

Equilibrium Unemployment and Wage Formation with Matching Frictions and Worker Moral Hazard*

Guillaume Rocheteau[†]

DEEP - HEC, University of Lausanne

Revised version : September 28th, 2000

Abstract

This paper combines the shirking and the matching approaches of equilibrium unemployment in order to endogenize the wage formation process as a function of labour market conditions. The steady state equilibrium can take two forms depending on whether the no-shirking condition is binding or not. It is demonstrated that the efficiency wage approach is relevant when the unemployment rate is above a certain threshold. Furthermore, an efficiency wage is more likely when the disutility of effort is high, recruiting costs and workers' bargaining power are low, inspections are unlikely and the workers' productivity is weak.

JEL Classification: E24, J41, J64

Keywords: Equilibrium unemployment, Matching model, Moral hazard.

*Acknowledgments: I thank for their comments Gérard Ballot, Pierre Cahuc, François Langot, Dale Mortensen, Claude Vedel, Patrick Villieu and seminar participants at the University of Paris 2 (ERMES), Chung-Hua institution for Economic Research (CIER), National Taiwan University, Academia Sinica Institute of Economics, and University of Lausanne (DEEP). The revised version of this paper has benefited from the very helpful suggestions of an anonymous referee and Aleksander Berentsen.

[†]Address: DEEP-HEC, University of Lausanne, BFSH1, CH1015 Lausanne, Switzerland. E-mail : guillaume.rocheteau@hec.unil.ch

1 Introduction

A sound understanding of what determines the unemployment rate remains a major objective for economists. The explanation of a persistent level of involuntary unemployment requires a relevant description of both the wage formation and the matching process between workers and firms. The aim of this article is to endogenize the wage formation process as a function of market conditions and to characterize the circumstances under which the wage is used mainly as a motivation device. Furthermore, by combining the Pissarides (2000) and the Shapiro and Stiglitz (1984) models, our article makes a step toward a more complete theory of equilibrium unemployment.

Although our objective is not to explain a specific empirical observation, our model by incorporating both worker moral hazard and workers' bargaining power, can give insights into the answers of the following questions. Does the wage formation process differ in tight and thick labour markets? Is the wage more flexible when productivity is high? Is employers' monitoring more intensive in depressed labour markets? How are workers' earnings related to their productivity?

To explain the presence of involuntary unemployment, standard models depart from the Walrasian paradigm by introducing imperfect information or imperfect competition in the wage formation process.¹ According to the shirking model of Shapiro and Stiglitz (1984), in the presence of moral hazard, the wage is set by employers to motivate workers and boost their productivity. An alternative explanation of the equilibrium unemployment rate is provided by the search-matching approach.² In the presence of matching frictions, employees have some monopoly power which allows them to command a wage larger than their reservation wage.

The shirking problem is introduced in a natural way within the matching model by assuming that the monitoring technology is imperfect. An employer inspects his worker's effort from time to time only. Consequently,

¹See, for instance, Johnson and Layard (1986) and Lindbeck (1992).

²For a survey, see Mortensen and Pissarides (1999a).

a worker has the option to cheat his employer and to provide no effort at work. In contrast to the standard efficiency wage models, the wage is not set unilaterally by employers but is the outcome of a bargaining between each worker and the firm he is matched with.

In our model, employed workers benefit from a positive rent, resulting in involuntary unemployment, for two reasons: first, it is costly to fill a job and, second, the monitoring technology is imperfect. The bargaining process is affected by the relative size of these two rents. Hence, the equilibrium can take two forms depending on whether the no-shirking condition (NSC) is binding or not. The wage is called an *efficiency wage* (EW) if the NSC is binding and a *freely negotiated wage* (FNW) otherwise. The frontier between the two regimes is endogenous and varies with workers' bargaining power, advertising costs of vacancies, the frequency of inspection and the disutility of effort.

The main results of our paper are as follows. The FNW equilibrium occurs when the unemployment rate is below a certain threshold. Indeed, in a tight labour market, turnover costs are high and workers can obtain a wage superior to the EW. If, in contrast, the unemployment rate is high, the equilibrium wage is the EW. This result challenges the traditional view that a high unemployment rate acts to discipline workers and reduces moral hazard. Furthermore, it is demonstrated that an equilibrium with an EW is more likely when the disutility of effort is high, recruiting costs and workers' bargaining power are low, or inspections are unlikely.

We consider two extensions of the model. First, the monitoring activity is endogenized by allowing employers to choose the frequency of inspections. It is shown that there is less monitoring in the FNW equilibrium and that the wage is never superior to the level required to motivate workers. Indeed, when the FNW is superior to the EW, employers can relax the supervision of their employees and save some monitoring costs. Again, our model challenges the traditional view of efficiency wage models by predicting that there is less monitoring in a tight labour market.

Second, we introduce heterogeneity between workers to derive the shape of the wage distribution and to determine what types of workers are paid efficiency wages. In the presence of different skill lines, wages and exit rates out of unemployment vary across workers even if search is undirected. In particular, a low exit rate out of unemployment disciplines unskilled workers. Indeed, if a meeting occurs with a relatively low skilled worker, this worker is offered employment with some probability greater than zero but strictly less than one: discrimination against low skilled increases the threat of unemployment. For intermediary levels of skill, workers earn the same wage (an EW) and have the same exit rate out of unemployment. Finally, for high levels of skill, the NSC is not binding and workers receive a wage which is increasing with their productivity. We derive some comparative statics results in the case where there are only two types of workers, skilled workers and unskilled workers.

Basically, our paper considers the economic consequences of an additional constraint on wage bargaining. Obviously, there are other constraints such as minimum wages that could be incorporated. However, the NSC gives a minimum incentive-compatible wage that depends on the overall labour market conditions. This means that the constraint we focus on is endogenous. Furthermore, some of the results generated by the worker moral hazard problem, the wage distribution in the presence of heterogenous workers for instance, do not carry over to constraints such as minimum wages.

The most closely related literature is a range of papers which endogenize the choice of the wage policy as a function of labour market conditions. Ellingsen and Rosen (1997) resort to a search model with heterogenous workers to determine whether wages are posted by the firm or negotiated with each applicant individually. MacLeod and Malcomson (1998) model the choice between performance pay and efficiency wages to motivate employees.

Sattinger (1990) and Larsen and Malcomson (1999) introduce EW considerations within a matching model of unemployment. Mortensen (1989) and Mortensen and Pissarides (1999a, section 4) resort to the search equilib-

rium approach to study alternative wage determination mechanisms. However, none of these analyses introduce simultaneously a bargaining power for workers and an EW component. Lindbeck and Snower (1991) study the interactions between the EW and insider-outsider theories but do not formalize the recruiting costs as the result of matching frictions. Manning and Thomas (1997) build a partial equilibrium model based on job search theory in the presence of incentives problems.

This paper is organized as follows. Section 2 presents the model and its main assumptions. The steady-state equilibrium and the two regimes are studied in the third section. The model is extended in Section 4 to incorporate an endogenous monitoring technology. Section 5 introduces heterogeneity among workers, and Section 6 concludes.

2 The model

The basic framework is similar to that of Pissarides (2000, Chapter 1) except that unobservable shirking is allowed.

2.1 Main assumptions

Consider an economy composed of a continuum of infinitely-lived homogeneous workers of measure equal to one, and of a continuum of identical firms holding at most one vacancy. Time, denoted by t , is continuous.

Workers and firms are risk neutral. The utility function of an individual whose consumption and effort trajectories are $\{c(t), t \in \mathbb{R}^+\}$ and $\{e(t), t \in \mathbb{R}^+\}$ is:

$$\int_{\mathbb{R}^+} \exp(-rt) [c(t) - e(t)] dt,$$

where $r \in \mathbb{R}^+$ is the rate of time preference.

Following Shapiro and Stiglitz (1984), there are two levels of work intensity (\bar{e} or 0) and workers' effort is imperfectly observed by employers. The

productivity per unit of time of a job-worker match is y if the work intensity is \bar{e} , and 0 otherwise. Inspections obey a Poisson process with arrival rate $\lambda \in \mathbb{R}^+$ and the punishment for shirking is to be fired.

The labour market has matching frictions. The flow of job creations per unit of time is $m(u, v)$ where u is the number of unemployed, v the number of vacancies and $m(., .)$ exhibits constant returns to scale. Let θ be defined as v/u , which is commonly referred to as labour market tightness. Further, if q denotes the rate at which a vacancy fills, it is given by

$$q = \frac{m(u, v)}{v} = x\left(\frac{1}{\theta}, 1\right),$$

and hence $q = q(\theta)$ with $q'(\cdot) < 0$. As the exit rate out of unemployment is $m(u, v)/u$, note that it equals $\theta q(\theta)$.

Each filled job is destroyed according to a Poisson process with arrival rate $s \in \mathbb{R}^+$. As the flow out of unemployment is $m(u, v) = u\theta q(\theta)$, while the flow in is $s(1 - u)$, the steady state (which requires that these two flows are equal) is characterized by an unemployment rate equal to:

$$u = \frac{s}{s + \theta q(\theta)}. \tag{1}$$

2.2 The value functions

Since there is zero productivity while shirking, an active equilibrium (with positive production) is such that the incentive-compatibility constraint is always fulfilled.

A non-shirker is a worker who chooses not to shirk in all periods while attached to the current job. The lifetime expected utility of a non-shirker who earns wage w is denoted by $W_E(w)$ and obeys the following asset pricing equation:

$$rW_E(w) = w - \bar{e} + s(W_U - W_E(w)), \tag{2}$$

where W_U is the lifetime expected utility of an unemployed worker and w the real wage. If W_E represents the ‘‘asset’’ value of employment, eq. (2) simply

states that the opportunity cost of holding it, rW_E , is equal to the current income flow w minus the disutility of effort \bar{e} plus the expected capital loss flow, the third term of the RHS.

The lifetime expected utility of a currently employed worker who chooses to shirk during a spell of length dt , denoted by $W_E^S(w; dt)$ (S for *shirking*), satisfies:

$$\begin{aligned} W_E^S(w; dt) &= wdt + \exp(-rdt) \{ \mathbb{P} [\min(\tau_s, \tau_\lambda) \leq dt] W_U + \\ &+ (1 - \mathbb{P} [\min(\tau_s, \tau_\lambda) \leq dt]) W_E(w) \}, \end{aligned} \quad (3)$$

where τ_λ is the length of time until the next inspection (an exponential distribution of parameter $\lambda \in \mathbb{R}^+$) and τ_s the random duration of a job (an exponential distribution of parameter $s \in \mathbb{R}^+$).³ According to (3), during the time interval of length dt , the worker receives the real wage wdt and suffers no disutility; he loses his job if he is caught shirking or if his job is destroyed by an idiosyncratic shock. If neither of these two events occurs during the time interval of length dt , the employed worker stops shirking in all the subsequent periods: his lifetime expected discounted utility is equal to $W_E(w)$. After some manipulation, eq. (3) yields:

$$\begin{aligned} W_E^S(w; dt) &= wdt + (1 - rdt) \{ (s + \lambda)dtW_U + [1 - (s + \lambda)dt] W_E(w) \} + o(dt), \\ &= W_E(w) + \bar{e}dt - (1 - rdt)\lambda dt \{ W_E(w) - W_U \} + o(dt), \end{aligned}$$

with $\lim_{dt \rightarrow 0} o(dt)/dt = 0$. A worker who draws a benefit from cheating his employer over a period of time of length dt , chooses to shirk all the time. When dt approaches zero, the worker's optimal strategy is not to shirk if and only if $W_E^S(w; dt) - W_E(w) \simeq [\bar{e} - \lambda(W_E(w) - W_U)] dt \leq 0$, that is

$$W_E(w) - W_U \geq \frac{\bar{e}}{\lambda}. \quad (4)$$

To prevent the employed worker from shirking, he must get a rent at least equal to \bar{e}/λ . Indeed, a shirker saves the disutility of effort \bar{e} but bears a

³The random variable $\min(\tau_s, \tau_\lambda)$ is characterized by an exponential distribution of parameter $s + \lambda$.

capital loss $W_E(w) - W_U$ if he is dismissed (an event which occurs with an instantaneous probability λ).

As in Pissarides (2000), recruiting is costly. The instantaneous advertising costs are γ . Thus, the expected profits of a vacancy (W_V) satisfy the following Bellman equation in continuous time:

$$rW_V = -\gamma + q(\theta)(W_J - W_V), \quad (5)$$

where W_J is the value of a filled job.

Finally, the value functions of an unemployed worker (W_U) and of a filled job (W_J) obey the following asset pricing equations:

$$rW_U = b + \theta q(\theta)(W_E - W_U), \quad (6)$$

$$rW_J(w) = y - w + s(W_V - W_J(w)), \quad (7)$$

where b represents unemployment benefits and $\theta q(\theta)$ the exit rate out of unemployment. Note that, according to (6), an unemployed worker who finds a job becomes a non-shirker.

Under a free-entry condition, the expected profits of a vacant job are zero ($W_V = 0$). Thus, from (5), $W_J = \gamma/q(\theta)$. By substituting this expression into (7), we get the *vacancy supply condition* which gives a decreasing relation between labour market tightness ($\theta = v/u$) and the real wage:

$$y - w = (r + s) \frac{\gamma}{q(\theta)}. \quad (8)$$

2.3 The bargaining process

The wage flow w that will be paid to the worker during the lifetime of the job is determined through bilateral bargaining.⁴ The outcome of the bargaining process corresponds to the asymmetric Nash solution with threat points equal to the employer's and worker's respective values of continued search. The worker's bargaining power is denoted by $\beta \in (0, 1)$. Furthermore, the NSC

⁴We rule out upward-sloping age-earnings profile. A review of other devices to motivate employees is offered by Ritter and Taylor (1997).

gives the minimum wage that must be paid by an employer to prevent his worker from shirking. Indeed, an employer will never agree to pay a wage which does not satisfy the incentive constraint. The Nash program can be written as follows:

$$w = \arg \max_{w \in (0, y)} [W_E(w) - W_U]^\beta [W_J(w) - W_V]^{1-\beta},$$

subjected to (4)

— **Insert figure 1** —

The outcome of the bargaining is represented in figure 1. From (2) and (7), the Pareto frontier of the bargaining set is given by:

$$W_E(w) - W_U + W_J = \frac{y - \bar{e} - rW_U}{r + s}.$$

The Nash solution is efficient; it is the highest level of the asymmetric Nash product situated on the Pareto frontier. Two cases can be distinguished: a binding and a non-binding NSC. In Figure 1 (a), the NSC is not binding, and the asymmetric Nash solution is given by the tangency point between the Pareto frontier of the bargaining set and the highest Nash product curve (the locus of the pairs $(W_E - W_U, W_J)$ associated with the same Nash product). In Figure 1 (b), the NSC is binding, and the bargaining outcome is the point of intersection of the NSC and the Pareto frontier of the bargaining set.

- *The NSC is not binding*

The threat to shirk is not *credible* and the equilibrium wage, denoted by \tilde{w} , is the wage that would be obtained in the absence of unobservable shirking, that is:

$$(1 - \beta) [W_E(\tilde{w}) - W_U] = \beta [W_J(\tilde{w}) - W_V], \quad (9)$$

From (2) and (7), the wage expression is given by:

$$\tilde{w} = (1 - \beta)(rW_U + \bar{e}) + \beta y. \quad (10)$$

The FNW is a weighted mean of the workers' productivity and the workers' reservation wage.

- *The NSC is binding*

The FNW violates the incentive constraint: the threat to shirk is credible. Employers must use the wage as a motivation device. The EW, denoted by \hat{w} , satisfies (4) at equality. Using (2), we find:

$$\hat{w} = (r + s) \frac{\bar{e}}{\lambda} + \bar{e} + rW_U. \quad (11)$$

The EW has two components: the reservation wage ($\bar{e} + rW_U$) and a linear function of effort. It depends on workers' productivity only through rW_U .

Finally, note that the employees' rent ($W_E - W_U$) vanishes if the inspection rate is infinite and if workers have no bargaining power. The wage is then equal to the reservation wage.

$$\lim_{\lambda \rightarrow +\infty} \hat{w} = \lim_{\beta \rightarrow 0} \tilde{w} = \bar{e} + rW_U.$$

3 Steady-state equilibrium

The equilibrium can take two forms depending on whether the NSC is binding or not.

Definition 1 *A steady-state equilibrium is a quintuplet $(W_U, W_E, W_J, \theta, w) \in \mathbb{R}^{5+}$ that satisfies: (2), (6), (7), (8) and (11) if the NSC is binding; (2), (6), (7), (8) and (10) if the NSC is not binding.*

3.1 FNW equilibria

The equilibrium wage offered by each firm is \tilde{w} given by (10). According to the free-entry condition ($W_J = \gamma/q(\theta)$) and (9), the permanent income of unemployed workers is

$$rW_U = b + \frac{\beta}{1 - \beta} \gamma \theta$$

By substituting this expression into (10), we get:

$$\tilde{w} = \beta(y + \theta\gamma) + (1 - \beta)(\bar{e} + b). \quad (12)$$

The wage equation in the absence of incentive problems is similar to that of Pissarides (2000) with one additional term $(1 - \beta)\bar{e}$ reflecting the compensation for the disutility of effort. Furthermore, (12) indicates that each employed worker receives a fraction β of the recruiting costs per unemployed ($v\gamma/u$). Substituting (12) into (8) yields

$$(r + s)\frac{\gamma}{q(\theta)} = (1 - \beta)(y - \bar{e} - b) - \beta\theta\gamma. \quad (13)$$

Labour market tightness is a decreasing function of the disutility of effort (\bar{e}) and is independent of the inspection rate (λ).

The FNW regime occurs when $W_E(\tilde{w}) - W_U \geq \bar{e}/\lambda$. We find:

$$\frac{\beta}{(1 - \beta)}\frac{\gamma}{q(\theta)} \geq \frac{\bar{e}}{\lambda}. \quad (14)$$

According to (14), when the workers' rent arising from the moral hazard problem is smaller than the rent arising from the presence of search frictions, the equilibrium wage is a FNW.

3.2 EW equilibria

The equilibrium wage is \hat{w} given by (11). Substituting $W_E - W_U = \bar{e}/\lambda$ into (6) gives:

$$rW_U = b + \theta q(\theta)\frac{\bar{e}}{\lambda}.$$

Inserting this expression into (11) yields:

$$\hat{w} = b + \bar{e} \left(\frac{r + s + \theta q(\theta)}{\lambda} + 1 \right). \quad (15)$$

Productivity (y) and advertising costs (γ) influence wages only indirectly via the exit rate out of unemployment. By inserting (15) into (8), we can derive the equilibrium labour market tightness:

$$y - b - \frac{\bar{e}}{\lambda} \{r + s + \lambda + \theta q(\theta)\} = (r + s)\frac{\gamma}{q(\theta)}. \quad (16)$$

The EW regime occurs if the FNW is not sufficient to enforce effort, that is $W_E(\tilde{w}) - W_U < \bar{e}/\lambda$, or equivalently,

$$\frac{\beta}{(1 - \beta)}\frac{\gamma}{q(\theta)} < \frac{\bar{e}}{\lambda}. \quad (17)$$

3.3 The frontier between the two regimes

Let (y^*, θ^*) denote the pair of variables which satisfies the two following equations:

$$\frac{\beta\gamma}{(1-\beta)q(\theta^*)} = \frac{\bar{e}}{\lambda}, \quad (18)$$

$$(r+s)\frac{\gamma}{q(\theta^*)} = (1-\beta)(y^* - \bar{e} - b) - \beta\theta^*\gamma. \quad (19)$$

The pair (y^*, θ^*) also satisfies (16). From (16) if $y < y^*$, $\theta < \theta^*$ and (17) is fulfilled: the equilibrium wage is an EW. From (13), if $y \geq y^*$, $\theta \geq \theta^*$ and (14) is satisfied: the equilibrium wage is a FNW.

According to the shirking model of Shapiro and Stiglitz (1984), the presence of involuntary unemployment imposes a cost on laid-off workers and acts as a *discipline device*. Paradoxically, in our model, the wage is an EW when the unemployment rate is larger than a critical level $u^* = s/[s + \theta^*q(\theta^*)]$.

To understand this result, note that unemployment is high when the value of a match between a worker and a firm is low. As a consequence, the FNW is smaller than the EW, and the threat to shirk is credible. Conversely, in a tight labour market workers are in a strong position in the bargaining and the FNW is sufficiently large to motivate them. A consequence of this result is that wages become more rigid when the labour market is depressed. Indeed, a binding NSC prevents a sharp fall in wages.

These conclusions are related with those of Ellingsen and Rosen (1997) who find that, in tight labour markets, bargaining is more common than posting wages. In a different context, McLeod and Malcomson (1998) demonstrate that efficiency wages should be observed in high unemployment environments whereas bonus payments are used in low unemployment labour markets.

In Figure 2, the equilibrium of the labour market is represented for three values of workers' productivity ($y_0 < y^* < y_1$). The equilibrium solution for the pair (θ, w) lies at the unique intersection of a vacancy supply curve (VS) and a wage curve (WS_{EW} if the wage is an EW, and WS_{FNW} if the wage is

a FNW).⁵ In the grey region, the NSC is not satisfied.

— **Insert figure 2** —

Proposition 2 *The critical productivity y^* below which EW equilibria occur depends on the inspection rate (λ), unemployment benefits (b), workers' effort (\bar{e}), advertising costs (γ), and workers' bargaining power (β). Furthermore:*

$$\frac{\partial y^*}{\partial \lambda} < 0, \quad \frac{\partial y^*}{\partial b} > 0, \quad \frac{\partial y^*}{\partial \bar{e}} > 0, \quad \frac{\partial y^*}{\partial \gamma} < 0, \quad \frac{\partial y^*}{\partial \beta} < 0.$$

Proof. The pair (y^*, θ^*) obeys (18) and (19). In the (y^*, θ^*) -plane, eq. (18) is represented by a horizontal line and (19) by an upward sloping curve. If λ increases, the curve given by (18) moves downward and y^* decreases. When b increases, the curve given by (19) moves downward and y^* increases. When \bar{e} increases, the curve given by (19) moves downward and the curve given by (18) moves upward: thus, $\partial y^*/\partial \bar{e} > 0$. From (18) and (19), we get the following upward-sloping relationship between y^* and θ^* :

$$(r + s) \frac{\bar{e}}{\beta \lambda} = y^* - \bar{e} - b - \theta^* q(\theta^*) \frac{\bar{e}}{\lambda}.$$

This relationship is independent of γ . The curve given by (18) moves downward as γ increases: then $\partial y^*/\partial \gamma < 0$. Finally, the pair (y^*, θ^*) obeys (16) and (18). Only (18) depends on β : thus, $\partial y^*/\partial \beta < 0$. ■

When $y = y^*$, the level of the rent due to unobservable shirking is equal to the level of the rent due to workers' bargaining power. The first rent depends positively on the disutility of effort and negatively on the inspection rate, whereas the second rent depends positively on advertising costs and worker's bargaining power. As a consequence, a FNW is likely when the disutility of effort (\bar{e}) is low, or when advertising costs (γ), workers' bargaining power (β) and inspection rate (λ) are high.

⁵The curve WS_{EW} is analogous to the NSC of the Shapiro and Stiglitz model: it defines a region of the (w, θ) -plane such that the incentive constraint is met. Unlike Shapiro and Stiglitz's model, this curve is not represented in the plane (real wage - employment). Indeed, in presence of trade frictions a firm does not directly control its employment level: its choice variable is the number of vacant jobs it advertises.

4 An endogenous monitoring technology

Why should employers keep on monitoring their workers when the NSC is not binding? To answer this question, we endogenize the monitoring technology by allowing employers to choose the frequency of inspections.

4.1 The choice of the inspection rate

We introduce a monitoring cost, denoted by $c(\lambda)$, which is a strictly convex and increasing function of the inspection rate. The net production flow of each job is then equal to the worker's productivity (\bar{y}) minus monitoring expenses.

$$y = \bar{y} - c(\lambda)$$

The value function of a filled job satisfies a generalized version of (7):

$$rW_J = \max_{\lambda \in \mathbb{R}^+} \{ \bar{y} - c(\lambda) - w(\lambda) - sW_J \}, \quad (20)$$

where $w(\lambda)$, the wage paid to workers, depends on the inspection rate. Each employer faces a trade-off between wages and monitoring costs. In general, an employer saves some monitoring costs by decreasing the frequency of inspections but must offer a higher wage to motivate workers.

As it has been outlined in the preceding section, the wage received by workers is at least equal to the efficiency wage. Hence, $w(\lambda) = \max [\tilde{w}(\lambda), \hat{w}(\lambda)]$, with, from (10) and (11),

$$\tilde{w}(\lambda) = (1 - \beta)(rW_U + \bar{e}) + \beta(\bar{y} - c(\lambda)), \quad (21)$$

$$\hat{w}(\lambda) = (r + s) \frac{\bar{e}}{\lambda} + \bar{e} + rW_U \quad (22)$$

For rW_U given⁶, wages are decreasing in the intensity of the monitoring activity. According to (22), a higher frequency of inspections reduces the incentive to shirk; according to (21), it raises monitoring expenses and decreases the net production flow.

⁶Workers and firms take W_U as given. Of course, the value of W_U in equilibrium also depends on the inspection rate.

A case that is worth mentioning is when workers have no bargaining power ($\beta = 0$). Then, the wage is always an EW. From (20) and (22), the profit-maximizing inspection rate is λ^* which satisfies:

$$c'(\lambda^*) = \frac{r+s}{\lambda^{*2}} \bar{e}$$

The marginal cost of the inspection technology is equal to the marginal gain in terms of the wage bill.

Proposition 3 *There are two types of equilibrium:*

- **EW equilibria** : the inspection rate is λ^* and $\hat{w}(\lambda^*) > \tilde{w}(\lambda^*)$.
- **FNW equilibria** : the inspection rate is inferior to λ^* and is such that $\hat{w}(\lambda) = \tilde{w}(\lambda)$.

Proof. Let $\lambda^{eq} > 0$ denote the value of the inspection rate in equilibrium. There are a priori three possible cases:

(i) $\hat{w}(\lambda^{eq}) < \tilde{w}(\lambda^{eq})$

It is impossible. Indeed, from (21)

$$\frac{\partial (c(\lambda) + \tilde{w}(\lambda))}{\partial \lambda} = (1 - \beta)c'(\lambda) > 0$$

Each firm bears a fraction $(1 - \beta)$ of the monitoring cost and has an incentive to reduce its inspection rate below λ^{eq} .

(ii) $\hat{w}(\lambda^{eq}) > \tilde{w}(\lambda^{eq})$.

Then, profits maximization implies $\lambda^{eq} = \lambda^*$.

(iii) $\hat{w}(\lambda^{eq}) = \tilde{w}(\lambda^{eq})$.

For profits to be maximized, employers must have no incentive either to increase or to decrease the inspection rate. For values of λ slightly below λ^{eq} , firms want to increase their inspection rate if both $\hat{w}(\lambda) > \tilde{w}(\lambda)$ and $\lambda < \lambda^*$. This requires $\lambda^{eq} \leq \lambda^*$, and so $c'(\lambda^{eq}) \leq -\hat{w}'(\lambda^{eq})$. Indeed,

$$c'(\lambda^{eq}) \leq -\hat{w}'(\lambda^{eq}) \Rightarrow \beta c'(\lambda^{eq}) < -\hat{w}'(\lambda^{eq}) \Rightarrow \hat{w}'(\lambda^{eq}) < \tilde{w}'(\lambda^{eq})$$

Under this last condition, if λ is slightly above λ^{eq} then $\widehat{w}(\lambda) < \widetilde{w}(\lambda)$ and, according to case (i), firms want to reduce their inspection rate. ■

Proposition 3 establishes that the wage expression in the two regimes is identical, and is given by eq. (22). When the NSC is not binding, workers have no incentive to cheat. Consequently, employers can relax supervision of employees and save some monitoring costs by decreasing the inspection rate below λ^* .

4.2 Steady-state equilibrium

Let us turn to the characterization of the steady-state equilibrium. Given that $W_E - W_U = \bar{e}/\lambda$, and from (6), the permanent income of unemployed workers is:

$$rW_U = b + \theta q(\theta) \frac{\bar{e}}{\lambda}. \quad (23)$$

By replacing rW_U by its expression given by (23) into (22), we obtain a wage equation identical to (15). Consequently, the labour market tightness in equilibrium obeys (16).

Definition 4 *A steady-state equilibrium is a sextuplet $(W_U, W_E, W_J, \theta, w, \lambda) \in \mathbb{R}^{6+}$ that satisfies: (2), (6), (20), (8), (22) and $\lambda = \lambda^*$ if the NSC is binding; (2), (6), (20), (8), (22) and $\widehat{w}(\lambda) = \widetilde{w}(\lambda)$ if the NSC is not binding*

Let us characterize the two types of steady-state equilibrium. For both types of equilibria, the labour market tightness (θ) obeys (16), that is:

$$\bar{y} - c(\lambda) - b - \frac{\bar{e}}{\lambda} \{r + s + \lambda + \theta q(\theta)\} = (r + s) \frac{\gamma}{q(\theta)} \quad (24)$$

- *FNW equilibria*

The inspection rate satisfies (14) at equality.

$$\frac{\beta}{1 - \beta} \frac{\gamma}{q(\theta)} = \frac{\bar{e}}{\lambda} \quad (25)$$

- *EW equilibria*

The inspection rate satisfies:

$$c'(\lambda) = \frac{r+s}{\lambda^2} \bar{e} \quad (26)$$

From (21), (22), (23) and (24), the condition $\widehat{w}(\lambda^*) > \widetilde{w}(\lambda^*)$ can be rewritten as:

$$\frac{\bar{e}}{\lambda^*} > \frac{\beta}{1-\beta} \frac{\gamma}{q(\theta)}. \quad (27)$$

In Figure 3, the vacancy supply condition (24) is represented by a curve labelled VS . The inspection rate schedule in the FNW regime, given by equation (25), is represented by a curve labelled \widetilde{IR} . Finally, the inspection rate schedule in the EW regime, which obeys equation (26), is represented by the curve \widehat{IR} . The curves VS , \widetilde{IR} and \widehat{IR} are respectively \cap -shaped, downward-sloping and vertical. Furthermore, it can be verified that VS reaches a maximum for a value of the inspection rate superior to λ^* .⁷

The equilibrium solution for the pair (λ, θ) lies at the unique intersection of VS and \widetilde{IR} , or VS and \widehat{IR} . The equilibrium is represented for three different values of the productivity ($\bar{y}_0 < \bar{y}^* < \bar{y}_1$). For instance, when $\bar{y} = \bar{y}_1$ the equilibrium pair (λ, θ) lies at the unique intersection of $VS(\bar{y}_1)$ and \widehat{IR} .

— **Insert Figure 3** —

Let (\bar{y}^*, θ^*) denote the pair of values (\bar{y}, θ) which satisfies (24) and (25) when $\lambda = \lambda^*$.

Proposition 5 *If $\bar{y} < \bar{y}^*$, the unique equilibrium is an EW equilibrium and $\lambda = \lambda^*$. If $\bar{y} \geq \bar{y}^*$, the unique equilibrium is an FNW equilibrium and $\lambda \leq \lambda^*$.*

⁷This reflects an externality in the choice of the inspection rate. Indeed, employers fail to internalize the effects of their monitoring decision on the permanent income of unemployed workers, and consequently on the wage paid by other firms.

Proof. According to (24), if $\bar{y} < \bar{y}^*$ and $\lambda = \lambda^*$ then $\theta < \theta^*$ and

$$\frac{\beta}{1 - \beta} \frac{\gamma}{q(\theta)} < \frac{\bar{e}}{\lambda^*}$$

Thus, the condition for an EW equilibrium is satisfied. Furthermore, equations (24) and (25) for a FNW equilibrium can only be satisfied for a value of the inspection rate above λ^* , which is impossible. Conversely, if $\bar{y} \geq \bar{y}^*$ and $\lambda = \lambda^*$, then $\theta \geq \theta^*$ and the condition (27) for an EW equilibrium is not satisfied. The equilibrium lies at the unique intersection of VS and \widetilde{IR} , and $\lambda \leq \lambda^*$. ■

Proposition 5 confirms and generalizes results from the preceding section. EW equilibria occur when productivity is low and unemployment is high; the inspection rate is then equal to λ^* . Conversely, in tight labour markets firms reduce their inspection rate below λ^* because they gain nothing by monitoring their workers more intensively than what is required to prevent shirking.

Again, our model generates a result which challenges the traditional view of EW models. Whereas in EW models the unemployment acts as a discipline device, our model predicts that employers' monitoring is more intensive in depressed labour markets.

5 Heterogeneity and wage formation

To derive the shape of the wage distribution and to determine what types of workers are paid efficiency wages, we assume from now on the presence of different skill lines. Jobs are identical but the production flow of a job-worker match increases with worker's qualification. Search is undirected and employers can meet either skill or unskilled workers.⁸ Moreover, when a match occurs there is no uncertainty about worker's characteristics.

⁸It is usually assumed in the search-matching literature that search is directed across skill lines. See, for instance, Mortensen and Pissarides (1999b). Under this assumption, worker's exit rate out of unemployment increases with his productivity. In our framework, this result holds even if search is undirected.

As in the previous sections, wages are determined by one-on-one negotiations. The bargaining outcome is given by the Nash solution subject to the NSC. The threat point of the worker is the value of continued search, which depends on his qualification. The threat point of the firm is the value of a vacancy (For some justifications, see the appendix).

This section proceeds as follows. First, we describe the links between the wage, the exit rate out of unemployment and the worker's productivity in equilibrium. Second, we determine the labour market tightness and the equilibrium unemployment rate in the presence of two types of workers, skilled and unskilled workers.

5.1 Wages and exit rates out of unemployment

In this subsection, we take the labour market tightness in equilibrium (θ) as given and we derive the distribution of wages and the exit rates of unemployment of workers with different productivities.

Consider a worker with productivity y . When a meeting with an employer occurs, the worker is recruited with an average probability $\Pi(y)$. Thus,

$$rW_{U,y} = b + \theta q(\theta) \Pi(y) \{W_{E,y} - W_{U,y}\}, \quad (28)$$

where $W_{U,y}$ is the value to be unemployed and $\theta q(\theta) \Pi(y)$ is the exit rate out of unemployment. The value to be employed, $W_{E,y}$, satisfies (2). The threat to shirk guarantees the worker for a rent at least equal to \bar{e}/λ .

$$W_{E,y} - W_{U,y} = \max \left(\frac{\bar{e}}{\lambda}, \frac{\beta}{1-\beta} W_{J,y} \right), \quad (29)$$

where $W_{J,y}$, the value of a filled job, is equal to $\frac{y-w(y)}{r+s}$.

Let us turn to firms' hiring strategy. We only consider symmetric Nash equilibria (NE). Each employer takes the strategy of other employers as given and chooses a probability $\pi(y)$ of recruiting the worker with productivity y in order to maximize his expected profits. Because in equilibrium the value of a vacancy is zero ($W_V = 0$), the best response function of an employer

satisfies the following rule:

$$\begin{array}{rcl} & > & = 1 \\ W_{J,y} < 0 & \Rightarrow \pi(y) = 0 & \\ & = & \in [0, 1] \end{array} \quad (30)$$

The employer accepts to recruit a worker with probability one if this worker generates positive profits for the firm. The NSC ($W_{E,y} - W_{U,y} \geq \bar{e}/\lambda$) gives a lower bound to the value of an acceptable match. Indeed, under (29),

$$\begin{array}{rcl} & > & > \frac{\bar{e}}{\lambda} \\ W_{J,y} < 0 & \Leftrightarrow W_{E,y} - W_{U,y} + W_{J,y} < \frac{\bar{e}}{\lambda} & \\ & = & = \end{array} \quad (31)$$

According to (30) and (31), $\Pi(y)$ sustains a symmetric NE if and only if:

$$\begin{array}{rcl} = 1 & & > \frac{\bar{e}}{\lambda} \\ \Pi(y) = 0 & \Rightarrow W_{E,y} - W_{U,y} + W_{J,y} < \frac{\bar{e}}{\lambda} & \\ \in [0, 1] & & = \end{array} \quad (32)$$

where, from (2) and (7), the total surplus of a match is:

$$W_{E,y} - W_{U,y} + W_{J,y} = \frac{y - \bar{e} - rW_{U,y}}{r + s}. \quad (33)$$

Two cases must be distinguished.

- $W_{E,y} - W_{U,y} + W_{J,y} \leq \bar{e}/\lambda\beta$.

Then, $W_{J,y} \leq \frac{(1-\beta)\bar{e}}{\lambda\beta}$ and, from (28)-(29), $rW_{U,y} = b + \Pi(y)\theta q(\theta)\bar{e}/\lambda$. The wage paid to the worker is an EW. Hence, (33) can be rewritten as follows:

$$W_{E,y} - W_{U,y} + W_{J,y} = \frac{y - \bar{e} - b - \Pi(y)\theta q(\theta)\bar{e}/\lambda}{r + s} \quad (34)$$

Let $\Pi(y)$ be the value of the recruiting probability that sustains a symmetric NE. According to (32) and (34):

$$\begin{array}{rcl} = 1 & & y - \bar{e} - b - \theta q(\theta)\bar{e}/\lambda > \\ \Pi(y) = 0 & \Leftrightarrow & y - \bar{e} - b < (r + s) \frac{\bar{e}}{\lambda} \\ \in [0, 1] & & y - \bar{e} - b - \Pi(y)\theta q(\theta)\bar{e}/\lambda = \end{array} \quad (35)$$

- $W_{E,y} - W_{U,y} + W_{J,y} > \bar{e}/\lambda\beta$.

Under this condition, $\Pi(y) = 1$ and the wage is a FNW. According to (28), (29) and (33), the total surplus of a match is:

$$W_{E,y} - W_{U,y} + W_{J,y} = \frac{y - \bar{e} - b}{r + s + \beta\theta q(\theta)} \quad (36)$$

These results yield the following proposition.

Proposition 6 *Assume that the labour market tightness in equilibrium is θ and consider a worker whose productivity is y :*

1. *If $y < b + \bar{e} + (r + s)\bar{e}/\lambda \equiv y_1$, $\Pi(y) = 0$.*
2. *If $y \geq y_1$ and $y \leq b + \bar{e} + (r + s + \theta q(\theta))\bar{e}/\lambda \equiv y_2$, the worker's wage income is equal to his productivity and $\Pi(y) \leq 1$.*
3. *If $y > y_2$ and $y < y^* \equiv b + \bar{e} + \{r + s + \beta\theta q(\theta)\}\bar{e}/\lambda\beta$, the worker receives an EW and $\Pi(y) = 1$.*
4. *If $y \geq y^*$, the worker earns a FNW and $\Pi(y) = 1$.*

Proof. Cases 1 and 2. Straightforward from (35). Because $W_{J,y} = \frac{y-w(y)}{r+s}$, $W_{J,y} = 0$ implies $w(y) = y$.

Cases 3 and 4. According to (35), for all $y > y_2$, $\Pi(y) = 1$. Furthermore, the critical productivity y^* which separates the EW and the FNW regimes satisfies $\frac{\beta}{1-\beta}W_{J,y^*} = \frac{\bar{e}}{\lambda}$. Hence,

$$W_{E,y^*} - W_{U,y^*} + W_{J,y^*} = \frac{\bar{e}}{\lambda\beta},$$

and from (28), (29) and (33):

$$\frac{y^* - \bar{e} - b - \theta q(\theta)\frac{\bar{e}}{\lambda}}{r + s} = \frac{\bar{e}}{\lambda\beta}.$$

In case 3, the wage satisfies $W_{E,y} - W_{U,y} = \frac{\bar{e}}{\lambda}$. Hence,

$$\hat{w} = \bar{e} + b + (r + s + \theta q(\theta))\frac{\bar{e}}{\lambda} = y_2, \quad \forall y \in (y_2, y^*) \quad (37)$$

In case 4, the wage satisfies the following splitting-the-surplus rule, $W_{E,y} - W_{U,y} = \frac{\beta}{1-\beta} W_{J,y}$. Hence, after some calculation:

$$\tilde{w}(y) = \frac{\beta(r+s+\theta q)y + (r+s)(1-\beta)(\bar{e}+b)}{(r+s) + \beta\theta q} \quad (38)$$

It can be verified that $\tilde{w}(y^*) = \hat{w}$. ■

Proposition 6 offers new insights about the distribution of wages and exit rates out of unemployment in the presence of worker's bargaining power and moral hazard. Three properties are worthwhile to mention.

First, if the worker's productivity lies between y_1 and y_2 , this worker receives an EW which is equal to his productivity, and he is discriminated against by employers (Case 2 of proposition 6). Indeed, if $\Pi(y)$ were equal to one, the non-shirking wage would be superior to the worker's productivity. Conversely, if $\Pi(y)$ were equal to zero, the worker would generate positive profits for his employer. As a consequence, employers adopt a mixed strategy. The decrease of the exit rate out of unemployment can be interpreted as a *discipline device* for less productive workers.⁹

Second, there is a "middle class" of workers who receive the same wage and have the same exit rate out of unemployment irrespective of their productivity (Case 3 of proposition 6). Indeed, for all $y \in (y_2, y^*)$ the worker's productivity is high enough to generate positive profits, but the FNW, $\tilde{w}(y)$, is always inferior to the EW, \hat{w} .

Third, workers receive a FNW which is increasing in their productivity (Case 4 of proposition 6). This suggests that the only workers who are able to negotiate their wage with their employer are the more qualified workers. For all other workers, the wage is just set at the level that prevents shirking. These results are recapitulated in Figure 4.

— **Insert Figure 4** —

To conclude, note that our results would not hold without assuming worker moral hazard, or simply by assuming the existence of a minimum

⁹For a similar argument, see Weiss (1991, p.80).

wage. In the absence of worker moral hazard ($\lambda = +\infty$), the first three cases of Proposition 6 would not be relevant. Furthermore, our model can explain the decrease in the exit rate out of unemployment for less qualified workers without assuming directed search across skill lines. Finally, the wage distribution in the presence of a minimum wage would simply be a truncated version of the distribution of the freely-negotiated wages.

5.2 A seemingly-segmented labour market

In this subsection, we endogenize the labour market tightness and we provide some results of comparative statics in the case where the economy is composed of low skilled and high skilled workers in proportion p and $1 - p$. The productivity of low (high) skilled workers is y_L ($y_H > y_L$).

We focus on a tight labour market where $y_H > y^*$ and $y_1 < y_L < y_2$.¹⁰ Firms discriminate against unskilled workers ($\Pi(y_L) < 1$) and high skilled workers receive a FNW (See figure 4). Jobs occupied by low skilled workers generate no profit. This case is of particular interest because our economy is similar to an economy with two segmented markets: a market for unskilled workers where the wage is an EW and a market for skilled workers where the wage is a FNW. We call this market a *seemingly-segmented labour market*.

The equality of worker flows in and out of unemployment yields the steady state unemployment rates of skilled (u_H) and unskilled workers (u_L).¹¹

$$u_H = \frac{s}{s + \theta q(\theta)}, \quad u_L = \frac{s}{s + \Pi(y_L) \theta q(\theta)}, \quad (39)$$

where, from (35),

$$\Pi(y_L) \theta q(\theta) = \frac{\lambda(y_L - b - \bar{e})}{\bar{e}} - r - s \quad (40)$$

Obviously, the unemployment rate of low skilled workers is higher than the unemployment rate of high skilled ones. The total unemployment rate in the

¹⁰These conditions will be fulfilled if the proportion of skilled workers (p) is sufficiently high, and if the productivity of unskilled workers (y_L) is close to y_1 .

¹¹The unemployment rate for each category of workers is computed as the incidence of unemployment divided by the sum of the incidence of unemployment and the exit rate out of unemployment.

economy is:

$$u = pu_H + (1 - p)u_L \quad (41)$$

Let μ denote the fraction of unemployed workers that are high skilled. Then $\mu = pu_H/u \equiv \mu(\theta, p, y_L, \lambda)$. It can be verified that:

$$\frac{\partial \mu}{\partial \theta} < 0, \quad \frac{\partial \mu}{\partial p} > 0, \quad \frac{\partial \mu}{\partial y_L} > 0, \quad \frac{\partial \mu}{\partial \lambda} > 0.$$

Because the exit rate out of unemployment is higher for skilled workers, the fraction of skilled workers among unemployed is decreasing with the number of vacancies. The positive impact of p on μ is straightforward. Finally, if for given θ unskilled workers become more productive or if the moral hazard problem is less severe, firms raise the probability to recruit unskilled workers and the fraction of unskilled among unemployed decreases.

The value of a vacant job in equilibrium satisfies the following Bellman equation:

$$rW_V = -\gamma + q(\theta)\mu \{W_{J,y_H} - W_V\} + q(\theta)(1 - \mu)\Pi(y_L) \{W_{J,y_L} - W_V\}, \quad (42)$$

According to (42), firms meet skilled workers with the instantaneous probability $q(\theta)\mu$ and unskilled workers with the instantaneous probability $q(\theta)(1 - \mu)$. From the free-entry condition, the value of a vacancy is zero ($W_V = 0$). Furthermore, we only consider equilibria such that y_L is between y_1 and y_2 in equilibrium. Hence, $W_{J,y_L} = 0$. Then, equation (42) yields:

$$W_{J,y_H} = \frac{\gamma}{q(\theta)\mu(\theta, p, y_L, \lambda)}. \quad (43)$$

According to (43), the value of a job occupied by a high skilled worker is equal to the average advertising costs borne by firms to find a high skilled worker.

The wage received by high-skilled workers is a FNW that satisfies the following rent sharing rule:

$$W_{J,y_H} = (1 - \beta)(W_{E,y_H} - W_{U,y_H} + W_{J,y_H}) \quad (44)$$

From (28) and (33), (44) can be rewritten as follows::

$$W_{J,y_H} = (1 - \beta) \frac{y_H - \bar{e} - b - \theta q(\theta) \frac{\beta}{1-\beta} W_{J,y_H}}{r + s} \quad (45)$$

Inserting W_{J,y_H} by its expression given by (43) into (45) gives the labour market tightness in equilibrium (θ^{eq}) as a function of p , y_L and λ :

$$(1 - \beta) (y_H - \bar{e} - b) - \beta \frac{\theta^{eq} \gamma}{\mu(\theta^{eq}, p, y_L, \lambda)} = \frac{(r + s) \gamma}{q(\theta^{eq}) \mu(\theta^{eq}, p, y_L, \lambda)} \quad (46)$$

The RHS of (46) is increasing in θ^{eq} whereas the LHS of (46) is decreasing in θ^{eq} : thus, the equilibrium is unique.

The comparative statics of the model are as follows:

$$\frac{\partial \theta^{eq}}{\partial p} > 0, \quad \frac{\partial \theta^{eq}}{\partial y_L} > 0, \quad \frac{\partial \theta^{eq}}{\partial \lambda} > 0$$

When the proportion of skilled workers increases, when unskilled workers become more productive, or when the moral hazard problem is less severe, then firms are incited to open vacancies and the labour market tightness increases. Indeed, in all these cases it is less costly for a firm to recruit a skilled worker. For instance, if the monitoring technology becomes more efficient, the exit rate out of unemployment of unskilled workers is higher and their fraction in the unemployment pool is lower. As a consequence, it is easier for an employer to find a skilled worker who is unemployed. One can also check from (39) and (41) that in all these three cases the overall equilibrium unemployment rate in the economy (u^{eq}) decreases.

$$\frac{\partial u^{eq}}{\partial p} < 0, \quad \frac{\partial u^{eq}}{\partial y_L} < 0, \quad \frac{\partial u^{eq}}{\partial \lambda} < 0$$

Finally, we determine the effects of an increase in p , y_L or λ on the fraction of unemployed workers that are high skilled in equilibrium (μ^{eq}). There are two opposite effects: a direct effect that increases μ and an indirect effect, through the increase in the labour market tightness, that goes in the opposite direction. To determine the overall effect, we rewrite (46) as follows:

$$(1 - \beta) (y_H - \bar{e} - b) = \beta \frac{\theta^{eq} q(\theta^{eq}) \gamma}{q(\theta^{eq}) \mu^{eq}} + \frac{(r + s) \gamma}{q(\theta^{eq}) \mu^{eq}} \quad (47)$$

Following an increase in p , y_L or λ , the labour market tightness increases. Consequently, according to (47), the probability to fill a vacancy with a skilled worker, $q(\theta^{eq})\mu^{eq}$, also increases. Because $q'(\theta) < 0$, we have:

$$\frac{\partial \mu^{eq}}{\partial p} > 0, \quad \frac{\partial \mu^{eq}}{\partial y_L} > 0, \quad \frac{\partial \mu^{eq}}{\partial \lambda} > 0$$

An increase in the fraction of skilled workers (p), in the productivity of low skilled workers (y_L) or in the inspection rate (λ) leads to an increase in the fraction of unemployed workers that are high skilled in equilibrium.

6 Conclusion

This paper has combined the shirking and the matching approaches in order to endogenize the wage formation process as a function of labour market conditions. We have assumed the presence of a moral hazard problem and that workers have some bargaining power in the wage bargaining. It has been shown that the shirking approach is relevant when the unemployment rate is above a certain threshold. Indeed, in a depressed labour market, the worker's rent due to unobservable shirking is larger than the rent generated by the presence of recruiting cost. Consequently, the threat to shirk is credible and employers must pay an efficiency wage. An efficiency wage is more likely when the disutility of effort is high, recruiting costs and workers' bargaining power are low and when inspections are unlikely. Conversely, in a tight labour market, workers are strong in the bargaining and they receive a wage, the so-called freely-negotiated wage, which is higher than the efficiency wage.

In many respects, this paper challenges the traditional view of efficiency wage models that emphasizes the role of unemployment as a discipline device. In our model, the need to motivate workers and to pay them an efficiency wage occurs when the unemployment rate is sufficiently high. Furthermore, employers' monitoring needs to be higher when the labour market is depressed.

Our framework also gives new insights about the wage distribution and the exit rates out of unemployment. We find that the wage paid to less

qualified workers is equal to their productivity, but that these workers have less chance than others to find a job. Workers with a productivity which is neither too high nor too low receive the same efficiency wage and have the same exit rate out of unemployment. These workers compose what we call a “middle-class”. Finally, the most productive workers receive a wage which is freely-negotiated with the firm and which increases with workers performances.

To conclude, the shirking-matching framework we have proposed provides a useful tool to analyze many issues related to unemployment and wages. As an example, we have studied in a companion paper (Rocheteau, 2000) the consequences of a work-sharing policy on unemployment and wages within a search model with worker moral hazard.

Appendix : The wage formation

In the presence of a worker moral hazard problem, the value of a match is negative when the wage is inferior to the non-shirking wage. Consequently, the wage formation in our model is formalized by an asymmetric Nash bargaining subjected to a constraint on the set of feasible outcomes; namely, the no-shirking condition.

Binmore *et al.* (1986) demonstrated that the issue of the bargaining game with alternate offers tends to the axiomatic Nash solution when the length of the game period approaches zero. The asymmetric Nash solution is chosen if players value time differently or if the amount of time that elapses between a rejection and an offer is different for the worker and for the firm (Osborne and Rubinstein, 1990, chapter 4).

Consequently, it would have been equivalent to formalize in our model the wage process by the mean of a strategic game with alternate offers. At each period, one of the two players (the worker, with a probability β , or the employer, with a probability $1 - \beta$) announces a value for the wage. This offer can be accepted or rejected by the second player. If it is rejected, each player may be matched to a new partner; otherwise, the same game is played again at the following period. If it is accepted, the worker must decide to shirk or to provide the high level of effort. The last two nodes of the game (shirking/no shirking) can be reduced to a single node. Indeed, the possibility of shirking does not affect the negotiation process itself but only the set of feasible outcomes. Moreover, the choice of the statu quo points in the Nash bargaining (the employer's and worker's respective values of continued search) is justified by the assumption that the two players go on searching for new partners while bargaining. This generates an endogenous risk of breakdown.

One round of the bargaining game is shown in figure 5a. At each node of the game, the player who has to make a move is identified by a letter within a circle: W for the worker and F for the firm. At each round, the first move is a chance move. With probability β (resp. $1 - \beta$) the worker (resp. the

firm) proposes a wage $w_W \in [rW_U + \bar{e}, y]$ (resp. $w_F \in [rW_U + \bar{e}, y]$): there is a continuum of choices indicated by the triangle attached to the decision node. Each proposal by a player leads to a decision to accept (“Yes”) or to reject (“No”) the proposal by the other. If the offer is accepted, the worker must choose to shirk (“*shirking*”) or not to shirk (“*no-shirking*”). If it is rejected, the worker (resp. the employer) goes on searching for a partner and meets a vacancy (resp. an unemployed) with a probability $\theta q(\theta)\Delta$ (resp. $q(\theta)\Delta$) where Δ is a period length. The amount of advertising costs borne by firms on a period of time of length Δ is $\gamma\Delta$. This justifies the choice of the statu quo points, W_U for the employee and W_J for the employer.

The offers that would incite workers to cheat are represented by a grey area. Such offers will never be proposed or accepted by employers. In figure 5b, the same game is represented in the case where the set of agreement is restricted to incentive-compatible wages ($w \in [\hat{w}, y]$). Then, once an offer is accepted, it always implies that the worker chooses not to shirk. One could check that the outcome of this bargaining game when Δ goes to zero is an immediate agreement on w that satisfies:

$$w = \arg \max_{w \in [\hat{w}, y]} [W_E(w) - W_U]^\beta [W_J(w) - W_V]^{1-\beta}$$

— **Insert figure 5** —

The choice of the employers’ statu quo point, W_V , is still relevant in the presence of heterogenous workers. To see this, consider the situation of a firm between two rounds of the negotiation. The firm meets an unskilled worker with probability $(1 - \mu)q(\theta)\Delta$ and a skilled worker with probability $\mu q(\theta)\Delta$. There is no uncertainty about the quality of the worker, and the value to be matched with an unskilled worker is 0, whereas the value to be matched with a skilled worker is W_{J,y_H} . Furthermore, searching is costly: the amount of advertising cost on a period of time of length Δ is $\gamma\Delta$. Consequently, the expected value of a firm if an agreement is never reached is \widetilde{W}_V that satisfies the following Bellman equation:

$$\widetilde{W}_V = -\gamma\Delta + \mu q(\theta)\Delta W_{J,y_H} + (1 - r\Delta)(1 - \mu q(\theta)\Delta)\widetilde{W}_V$$

It can be checked that \widetilde{W}_V is equal to W_V when Δ approaches 0. This gives the statu quo point of the firm in the bargaining.

REFERENCES

- Binmore, K., Rubinstein, A., and A. Wolinsky, 1986, The Nash bargaining solution in economic modelling, *Rand Journal of Economics* 17, 176–188.
- Ellingsen, T. and Rosén, A., 1997, Fixed or flexible? Wage setting in search equilibrium, *Stockholm School of Economics Discussion Paper*.
- Johnson, G.E. and P. R. Layard, 1986, The natural rate of unemployment: explanation and policy, in: O. Ashenfelter and R. Layard, eds, *Handbook of labor economics*, chapter 16 (Elsevier Science).
- Larsen, J. D. and Malcomson, J. M. (1999). Efficiency wages, matching, and unemployment: an analysis for four OECD countries. Communication at 1999 meeting of the Society for Economic Dynamic, Alghero.
- Lindbeck, A., 1992, Macroeconomic theory and the labor market, *European Economic Review* 36, 209–235.
- Lindbeck, A. and D. J. Snower, 1991, Interactions between the efficiency wage and insider-outsider theories, *Economics Letters* 37, 193–196.
- MacLeod, W. and Malcomson, J., 1998, Motivation and markets, *The American Economic Review*, 88, 388-411.
- Manning, A. and J. Thomas, 1997, A simple test of the shirking model, *Mimeo* (Centre for Economic Performance).
- Mortensen, D. T., 1989, The persistence and indeterminacy of unemployment in search equilibrium, *Scandinavian Journal of Economics* 91, 347–370.
- Mortensen, D.T. and C. A. Pissarides, 1999a, New developments in models of search in the labour market, Chapter 39 in Orley Ashenfelter and David Card (eds.) *Handbook of Labor Economics*, Vol. 3B (North-Holland, Amsterdam).
- Mortensen, D.T. and C. A. Pissarides, 1999b, Unemployment response to ‘skill-biased’ technology shocks: the role of labour market policy, *The Economic Journal* 109, 242-265.
- Osborne, M. J. and A. Rubinstein, 1990, *Bargaining and Markets* (Academic Press, New York).

Pissarides, C. A., 2000, *Equilibrium unemployment theory*, 2nd edition (MIT Press, Cambridge).

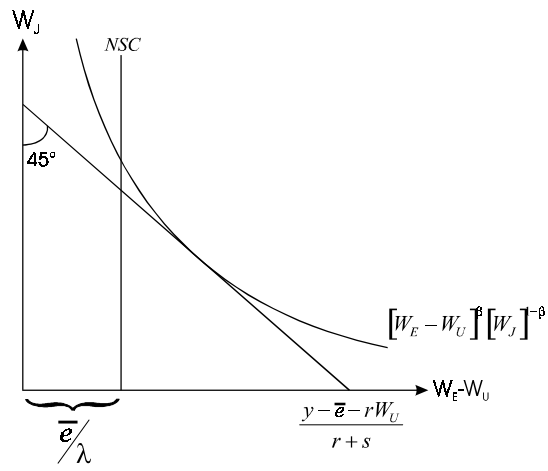
Ritter, J. A. and L. J. Taylor, 1997, Economic models of employee motivation, *Federal Reserve Bank of St Louis Review*, September-October, 3-21.

Rocheteau, G., 2000, Working time regulation in a search economy with worker moral hazard, DEEP Discussion Paper No 00.06 , University of Lausanne.

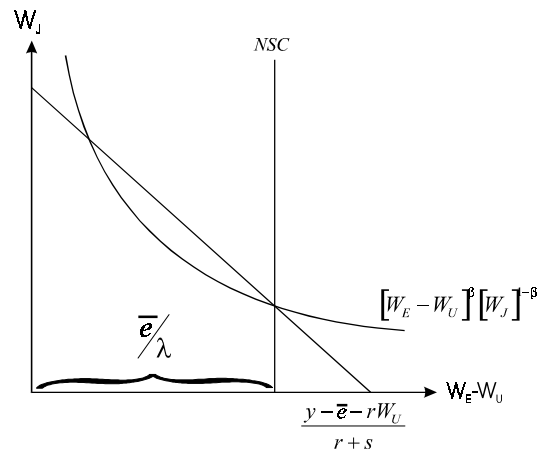
Sattinger, M., 1990, Unemployment, the market for interviews, and wage determination, *Journal of Political Economy* 98, 356–371.

Shapiro, C. and J. Stiglitz, 1984, Equilibrium unemployment as a discipline device, *American Economic Review* 74, 433–444.

Weiss, A., 1991, *Efficiency wages: models of unemployment, layoffs, and wage dispersion* (Clarendon Press, Oxford).



(a)



(b)

Figure 1 : The bargaining solution

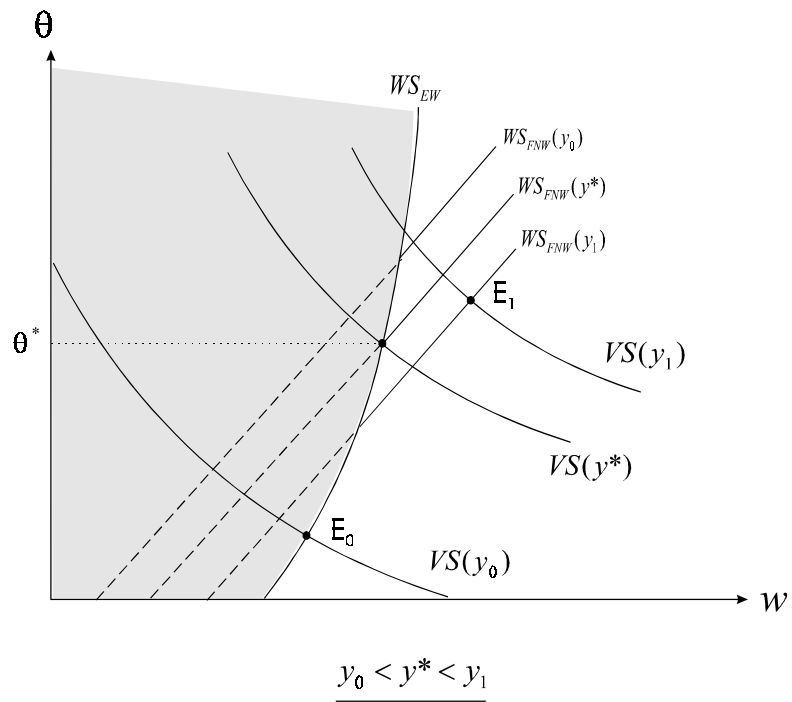


Figure 2 : Equilibrium determination

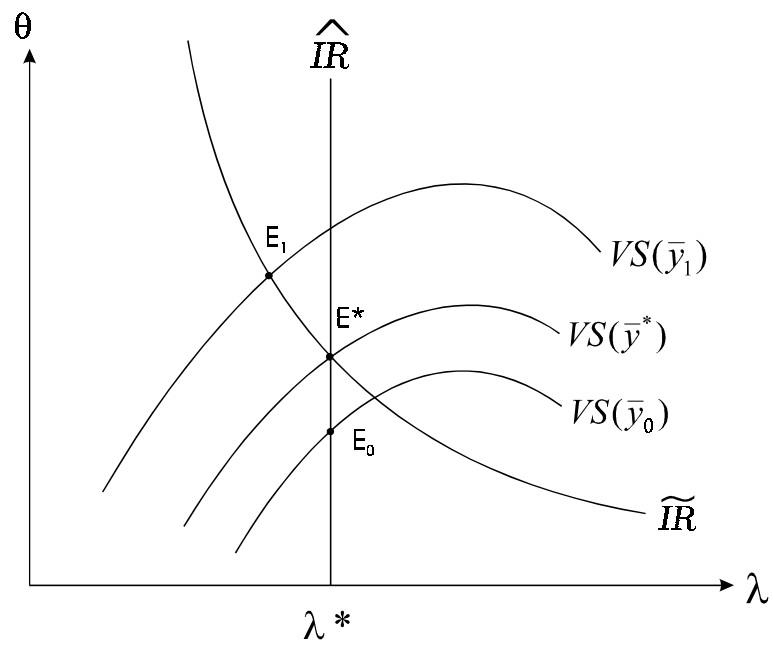


Figure 3 : Equilibrium with endogenous monitoring technology

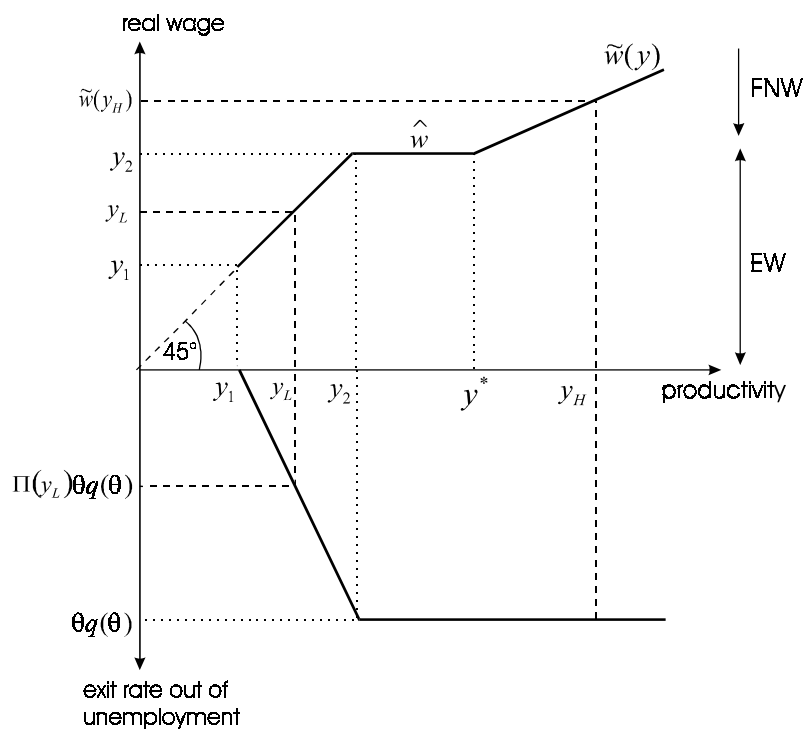


Figure 4 : Heterogeneity among workers

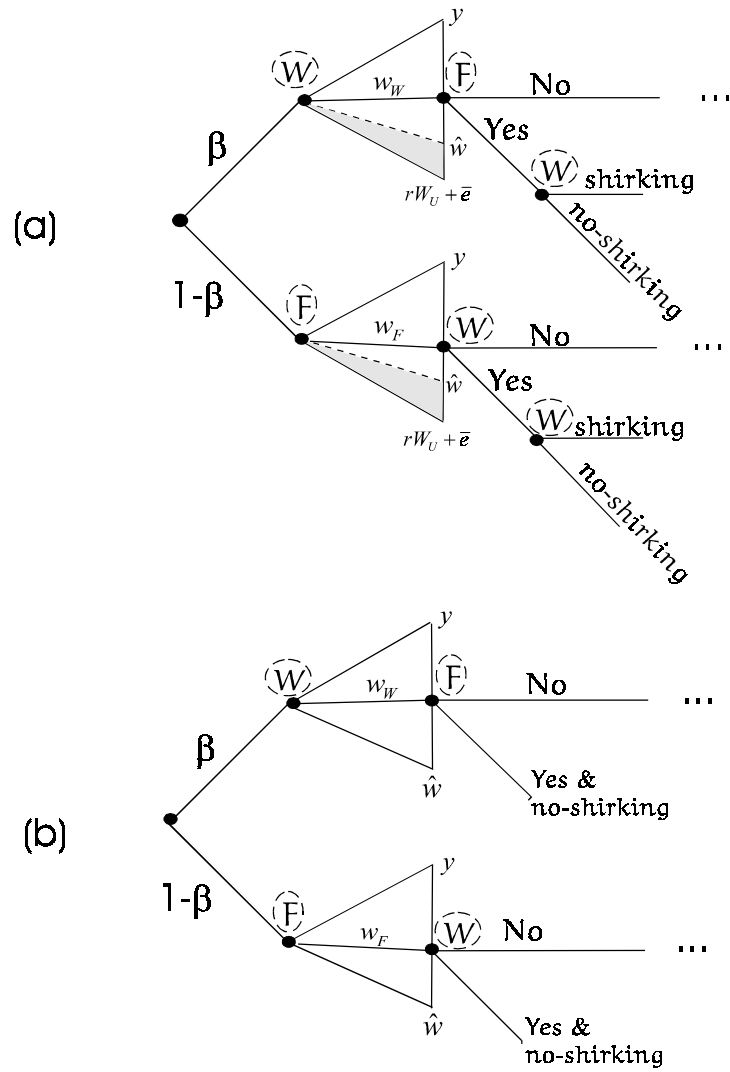


Figure 5 : A round of the bargaining game