Efficiency Wages, Shirking Model, and Employer Size-Wage Differentials *

Makoto Masui †
Department of Economics
Soka University ‡

Abstract

We combine the Burdett-Mortensen wage-posting model with the shirking model derived from the efficiency wage hypothesis, and study how incorporating the shirking behavior of workers affects properties of wage differentials among identical agents. This paper shows that (i) we can characterize a unique equilibrium even when each worker's effort level is not perfectly monitored by employers, (ii) the threat of dismissals may not dissuade worker shirking if the existence of on-the-job search is assumed, (iii) we can specify a scenario in which the improvement of a firm's monitoring technology results in a less dispersed earning distribution.

keywords: wage-posting model, shirking behavior, wage dispersion

JEL numbers: D83, E24, J31, J64

1 Introduction

We extend the Burdett-Mortensen wage-posting model such that employed workers optimally choose their effort levels and firms cannot perfectly monitor these levels. In such a situation, we examine properties of wage dispersion among identical agents when employers are identical. Workers are also identical. Many studies indicate that the efficiency wage hypothesis is an important tool for explaining various kinds of wage differential. In particular, we take note of the employer size-wage differential, and examine how strategies firms employ to prevent employee shirking affect an equilibrium earning distribution. Then, we find that improvement of monitoring technology may reduce wage differentials among employees.

Groshen (1988) classifies various kinds of wage differentials and surveys the related literatures. Many studies focus in particular on inter-industry differentials, intra-industry differentials and employer size-wage differentials, both theoretically and empirically: see

*I am very grateful to Soka University for financial support.
†E-mail: mmasui@soka.ac.jp. Tel: +81-(0)426-91-8914
‡1-236 Tangi-cho, Hachiouji-city, Tokyo, 192-8577, Japan.

Krueger and Summers (1988), Blackburn and Neumark (1992), Fairris and Alston (1994), Grea and Grenier (1994), and Abe and Ohashi (2004). These studies examine the source of inter-industry wage differentials by focusing on the effect of efficiency wages. [1] However, some researchers, such as Groshen (1991), focus on intra-industry wage differentials. She suggests that most wage differentials can be explained by employer size and that the efficiency wage hypothesis and the rent sharing rule will be dependent on this factor. There are several reasons why firms pay efficiency wages (or wage premiums): to keep employees from leaving current jobs, to attract more qualified workers, or to prevent employees from shirking among others. The shirking behavior of workers is of great theoretical importance but it is difficult to treat empirically.

Brown and Medoff (1989), Troske (1997), Bayard and Troske (1999), and Oi and Idson (1999) examine employer size-wage differentials. It is well-known that larger firms pay higher wages, but there is no consensus about which factor accounts for this differentials. Bayard and Troske (1999), however, show the existence of establishment size-wage premiums and firm size-wage premiums across main industry classifications such as manufacturing, retail trade, and services. Furthermore, they find that establishment size premiums do not vary across industries even if indicators such as employer productivity and worker skill levels are added to the analysis. This result suggests that there exists wage differentials that are due to factors independent of industry-specific parameters; these differentials then, might be relevant to the intra-industry wage differentials identified in Groshen (1988)(1991).

When we construct a model generating dispersion using identical workers and firms, the framework of Burdett and Mortensen (1998) is useful. They established, analytically, a wage-posting model characterizing an equilibrium wage offer distribution. [2] In a wage posting model, firms post their wages before a search process begins. If wage offers are greater than the reservation wage of workers, a worker-firm match is created. The number of workers each employer can hire depends on the wage level as well as market conditions such as the offer arrival rate. Higher wages attract more applicants with lower profits per worker, while lower wages attract fewer applicants with higher profits per worker. This trade-off generates many alternatives for maximizing profit, and therefore, the model provides wage dispersion with identical agents. [3] However, there is a little research on the worker moral hazard problem with respect to the Burdett-Mortensen model, and the properties of the wage differentials that would result.

Some studies combine the shirking model with a matching framework other than the wage posting model (see Shapiro and Stiglitz (1984), Albrecht and Vroman (1992)(1998),

1) The following studies deal with wage differentials theoretically: Bulow and Summers (1986), Esfahani and Salehi-Isfahani (1989), Lang (1991), Ramaswamy and Rowthorn (1991), and Montgomery (1991).

2) Mortensen and Pissarides (1999) and Mortensen (2003) extend the model of Burdett and Mortensen (1998), producing many important results.

3) According to van den Berg and Ridder (1998), and Postel-Vinay and Robin (2002), the level of differential irrelevant to heterogeneity of workers or employers is approximately 20% in the former study and 50% in the latter. Although the heterogeneity of agents can evidently explain some portion of wage differential, it is important to emphasize that there exist differentials which are caused by market friction rather than heterogeneity.

and Macleod and Malcomson (1998)). Albrecht and Vroman (1998) extend the Shapiro-Stiglitz model by introducing the heterogeneity of worker type, and they characterize a continuous wage offer distribution in this complex model. But unfortunately, the distribution they have characterized was an implicit function and it was derived from numerical calculation; it is, then, difficult to understand the properties of this distribution analytically.

Considering the above-described research, we establish our model by incorporating the asymmetric information problem (worker moral hazard problem) into the Burdett-Mortensen model, and examine how information asymmetries affect wage differentials generated in equilibrium. The model allows for on-the-job search behavior and this establishes that the wage level required for no shirking is positively related to the observability by firms (denoted by the rate of detecting shirkers). Bulow and Summers (1986) and Esfahani and Salehi-Isfahani (1989) suggest that the employer size-wage differential arises from the substitutability between wages and the observability by firms. Their assertions seem to be intuitive and convincing. But, Neal (1993) and Walsh (1999) point out that the relationship between these elements is ambiguous, and counterintuitively, that they may be positively related to each other (it may be optimal for firms to adopt a high wage and frequent observation concurrently). The statement of our hypothesis is similar to that of Neal (1993) or Walsh (1999). In this regard, it is worth noting that equilibrium unemployment no longer works as a discipline device, as described in Shapiro and Stiglitz (1984), when on-the-job search is taken into consideration. This is because if workers avoid being unemployed, a rise in detection rate will give them an incentive to be more diligent.

We can also derive the characteristics of the wage dispersion at equilibrium. Given the relationship between wages and the detection rate of firms noted above, the variance (or the coefficient of variance) of the earnings distribution decreases as firms catch shirkers more accurately, when the arrival rate of wage offers is sufficiently high. Since these indicators are considered to be the factors representing the degree of wage dispersion, this result suggests that wage differential among homogeneous employers is negatively related to quality of the monitoring technology.

The organization of this paper is as follows. In Section 2, we describe basic frameworks of the model and characterize the Bellman equations for workers. In Section 3, the no shirking condition which is necessary for firms to deter shirking is derived. In Section 4, we compute an equilibrium wage offer distribution and its support, and at the same time, we show the existence of a wage which satisfies the no shirking condition with equality. In Section 5, we investigate properties of rents received by employees, and bring out the relationship between the detection rate and the wage level specified in the previous section. In Section 6, we examine how wage differentials respond to the rise in the detection rate by employers, and identify the situation in which the differentials expand. In Section 7, we show that workers choose to shirk when they receive the reservation wage that is the lowest offered wage in the standard Burdett-Mortensen model. In Section 8, we conclude.

2 Basic Frameworks

We assume a continuous time, and restrict our analysis to the steady state equilibrium. There are many identical workers and identical firms in the economy. Let m denote a measure of the number of workers, defined in proportion to employers. Both unemployed and employed workers engage in job search activity. A firm posts one wage offer to attract a workforce.

The potential employees are aware of a wage offer distribution, while the actual offer they will receive is not identified until a match is formed. Let $F(\cdot)$ be the distribution of wage offers defined on its support $[\,\underline{w},\,\bar{w}\,]$. If a job seeker receives an offer greater than his standard for acceptance, he accepts it and moves to the new workplace. Then, he must choose how much effort he will devote to work. For simplicity, we assume that the effort level is either high ($\bar{e}$) or low ($\underline{e}$) such that $\bar{e} > \underline{e} > 0$. An employer cannot perfectly observe the effort level made by employees, but she can detect shirking at the Poisson arrival rate γ. Workers who are detected shirking become unemployed and search for other opportunities for employment.

Workers become worn out from high effort and suffer disutility $\bar{v}$ if they make $\bar{e}$ or $\underline{v}$ if they make $\underline{e}$. At a firm with wage payment $\hat{w}$, they receive a flow utility $\hat{w} - v$ (v is either $\bar{v}$ or $\underline{v}$). The assumption $\underline{e} \neq 0$ does not mean that firms allow their employees to shirk. We set $\underline{e}$ so low that firms act to prevent employees from exerting any less effort. Therefore, employers must set a sufficiently high wage level to elicit high effort levels.

Bellman equations for employed workers receiving $\hat{w}$ can be written as follows: Let $V_E(\cdot)$ denote the value when workers exert high effort, and let $V_S(\hat{w})$ denote the value when shirking.

$$r\,V_E(\hat{w}) = \hat{w} - \bar{v} + \lambda\left\{\int \max\,[\,V_E(w),\,V_E(\hat{w})\,]\,dF(w) - V_E(\hat{w})\right\} + \delta\,[\,V_U - V_E(\hat{w})\,], \tag{2.1}$$

$$r\,V_S(\hat{w}) = \hat{w} - \underline{v} + \lambda\left\{\int \max\,[\,V_S(w),\,V_S(\hat{w})\,]\,dF(w) - V_S(\hat{w})\right\} + (\delta + \gamma)\,[\,V_U - V_S(\hat{w})\,], \tag{2.2}$$

where λ is the arrival rate of jobs, which is assumed to be common to all states of workers (unemployed or employed), and δ is an exogenous separation rate. The assumption for the job arrival rate simplifies the model. (2.1) shows that the value of workers with high effort is composed of the flow utility and the future expected value that results from a change of state by exogenous separation or worker turnover. (2.2) implies that shirkers may suffer additional loss by being detected and becoming unemployed. It is worth noting that making high (low) effort in this model implies workers exert high (low) effort at any wage contained in $[\,\underline{w},\,\bar{w}\,]$. In other words, we do not consider the case in which employees who are shirking at one job choose to make high effort at another job at some future time.

Many studies consider worker shirking behavior, such as Shapiro and Stiglitz (1984) and Albrecht and Vroman (1992)(1998). But these studies do not introduce on-the-job search into their models. So, it would be interesting to examine the properties of the no-shirking condition in a situation in which the shirking model and on-the-job search behavior are jointly considered. In such a case, searching while on-the-job not only generates wage dispersion, but also reveals various important behaviors of employers, and with respect to the equilibrium earning distribution.

Next, we derive the Bellman equation for unemployed workers. Let w_R denote the reservation wage that makes workers indifferent to being unemployed or employed. This wage level is one component of the equilibrium in the standard model. But, as has been noted, we suppose that employers have a desire to offer high enough wages to discourage employees from shirking. The infimum, then, of the support of $F(\cdot)$ is given by the minimum of these offered wages. If the reservation wage is greater than this minimum level (denote w_n), the model is essentially identical to the basic Burdett-Mortensen model. But in Section 6, we show that this never occurs. Therefore, we presume that $w_R \leq w_n$ does not generate any problems in the argument that follows. Rather, we must demonstrate the existence of w_n in an equilibrium.

The Bellman equation for unemployed workers is given as follows.

$$\begin{aligned} r\,V_U &= b + \lambda\left(\int_{\underline{w}}^{\bar{w}} \max\left[\,V_E(w),\,V_S(w)\,\right] dF(w) - V_U\right) \\ &= b + \lambda\left(\int_{w_n}^{\bar{w}} V_E(w)\,dF(w) - V_U\right), \end{aligned} \tag{2.3}$$

where b denotes unemployment compensation.

3 No Shirking Condition

Employers must offer a sufficiently high wage in order to elicit effort. At this wage, workers obtain greater utility from exerting $\bar{e}$ than from exerting $\underline{e}$. Firms actively construct their offers to satisfy this no shirking condition. Since a worker's utility increases with his wage, employers will pay at least at level w_n, which satisfies $V_E(w_n) = V_S(w_n)$ (hereafter, we call w_n the no shirking wage).

(2.1) and (2.2) enable us to describe this condition precisely. First, those equations yield

$$\begin{aligned} r\left[\,V_E(\hat{w}) - V_S(\hat{w})\,\right] \;=\; &\lambda\left\{\int_{\hat{w}}^{\bar{w}} \left[\,V_E(w) - V_S(w)\,\right] dF(w) - \;(1 - F(\hat{w}))\left[\,V_E(\hat{w}) - V_S(\hat{w})\,\right]\right\} \\ &- \delta\left[\,V_E(\hat{w}) - V_S(\hat{w})\,\right] - \gamma\left[\,V_U - V_S(\hat{w})\,\right] - (\bar{v} - \underline{v}). \end{aligned} \tag{3.1}$$

Arranging (3.1) leads to the following expression (see Appendix A for details).

$$
\begin{aligned}
&(r+\lambda+\delta)\,[\,V_E(\hat{w})-V_S(\hat{w})\,] \\
=\ & -(\bar{v}-\underline{v})+\frac{\gamma\,(\hat{w}-\underline{v})}{r+\delta+\lambda+\gamma}-\frac{\gamma(r+\lambda)}{r+\delta+\lambda+\gamma}V_U+\frac{\lambda}{r+\delta+\lambda+\gamma}V_S(\bar{w}) \\
& -\lambda\,\gamma\int_{\hat{w}}^{\bar{w}}\frac{F(w)}{[\,r+\delta+\lambda\,(1-F(w))\,]\,[\,r+\delta+\gamma+\lambda\,(1-F(w))\,]}\,dw \\
& -\frac{\lambda\,\gamma}{r+\delta+\lambda+\gamma}\int_{\hat{w}}^{\bar{w}}\frac{F(w)}{r+\delta+\gamma+\lambda\,(1-F(w))}\,dw. \qquad (3.2)
\end{aligned}
$$

Since the no shirking wage w_n satisfy $V_E(w_n)=V_S(w_n)$, (3.2) becomes zero at $\hat{w}=w_n$. That is,

$$
\begin{aligned}
0\ =\ & -(\bar{v}-\underline{v})+\frac{\gamma\,(w_n-\underline{v})}{r+\delta+\lambda+\gamma}+\frac{\lambda\,\gamma\,(\bar{w}-\underline{v})}{(r+\delta+\lambda)(r+\delta+\lambda+\gamma)}-\frac{\gamma\,r}{r+\delta+\gamma}V_U \\
& -\lambda\,\gamma\int_{w_n}^{\bar{w}}\frac{F(w)}{[\,r+\delta+\lambda\,(1-F(w))\,]\,[\,r+\delta+\gamma+\lambda\,(1-F(w))\,]}\,dw \\
& -\frac{\lambda\,\gamma}{r+\delta+\lambda+\gamma}\int_{w_n}^{\bar{w}}\frac{F(w)}{r+\delta+\gamma+\lambda\,(1-F(w))}\,dw, \qquad (3.3)
\end{aligned}
$$

where we have already developed $V_S(\bar{w})$ by using (2.2). Note that we cannot show the existence of w_n satisfying (3.3) at this time, because relationships between w_n and other endogenous variables have not been resolved yet. Thus, the existence and uniqueness of the no shirking wage are shown in the next section.

As Carmichael has suggested, employers can employ other devices to inhibit shirking, such as imposing entry fees or bonds on newly-employed persons. According to some related works, however, it is optimal for employers to pay efficiency wages when they cannot grasp the employment history of workers concerning the reason for a separation. [4)] Furthermore, an employer moral hazard problem may arise in this case because employers will have incentives to confiscate these bonds by intentionally regarding diligent workers as shirkers. It follows that we can state that such employers will benefit from paying efficiency wages in various situations.

4 Description of the Equilibrium

4.1 Wage Offer Distribution

In this section, we characterize an equilibrium wage offer distribution. It follows from the argument in the Burdett-Mortensen model that the equivalence of profits obtained by offering any offer in $[\,\underline{w},\ \bar{w}\,]$ provides a concrete shape to the distribution.

First, we compute the steady-state employment level that results from firms offering $\hat{w}$. Let u be the measure of unemployment. Then, the flow into the unemployment pool

4) See Macleod and Malcomson (1993)(1998), Farris and Alston (1994) and Ritter and Lowell (1994).

is $\delta\,(m-u)$, and the flow out of the pool is expressed by $\lambda\,[\,1-F(w_n)\,]\,u$. In the steady state, the equality of these flows yields the steady state unemployment level as follows:

$$u \;=\; \frac{\delta\, m}{\delta+\lambda\,[\,1-F(w_n)\,]}. \tag{4.1}$$

Also, let $L(\hat{w}\,|\,w_n,\,F)$ be the number of applicants for the job with $\hat{w}$. Except for the difference between w_n and w_R, we can apply the same derivation process in Burdett and Mortensen (1998). By using (4, 1), $L(\hat{w}\,|\,w_n,\,F)$ can be written as

$$L(\hat{w}\,|\,w_n,\,F) \;=\; \frac{m\,k}{[\,1+k\,(1-F(\hat{w}))\,]^2}, \tag{4.2}$$

where k is λ/δ. It follows from (4.2) that the expected profit of a firm offering $\hat{w}$ is

$$\pi(\hat{w}) \;=\; (\bar{y}-\hat{w}-\bar{C})\,L(\hat{w}\,|\,w_n,\,F) \;=\; \frac{m\,k\,(\bar{y}-\hat{w}-\bar{C})}{[\,1+k\,(1-F(\hat{w}))\,]^2}, \tag{4.3}$$

where $\bar{y}$ denotes the productivity of an employee with $\bar{e}$, and $\bar{C}$ denotes hiring cost per worker, such as advertising and training costs. We set $\bar{C}$ as constant.

In equilibrium, since each firm making a wage offer contained in the support of $F(\cdot)$ must obtain equal profits, the equivalence of $\pi(w_n)$ and $\pi(\hat{w})$ for $\forall\hat{w}\in[\,\underline{w},\,\bar{w}\,]$ yields the following expression: [5)]

$$\pi(\hat{w}) \;=\; \pi(w_n) \;\Rightarrow\; F(\hat{w}) \;=\; \frac{1+k}{k}\left(1-\sqrt{\frac{\bar{y}-\hat{w}-\bar{C}}{\bar{y}-w_n-\bar{C}}}\right). \tag{4.4}$$

Since the supremum of the support $\bar{w}$ satisfies $F(\bar{w})=1$, it is given by

$$\bar{w} \;=\; \bar{y}-\bar{C}-\frac{\bar{y}-w_n-\bar{C}}{(1+k)^2}. \tag{4.5}$$

Note that (4.4) and (4.5) depend on the no shirking wage w_n, while the reservation wage is one of the components of the wage-posting equilibrium in the standard Burdett-Mortensen model. The expressions above enable us to show the existence of w_n.

4.2 The Existence and Uniqueness of The No Shirking Wage

Since w_n affects the value of unemployment, we must identify the relationship between them. Developing V_U from (2.3) leads to

$$r\,V_U \;=\; b-\,\lambda\int_{w_n}^{\bar{w}}\frac{F(w)}{r+\delta+\lambda\,(1-F(w))}\,dw+\,\lambda\,V_E(\bar{w})-\,\lambda\,V_U,$$

$$\Rightarrow\;\; V_U \;=\; \frac{r+\delta}{r(r+\delta+\lambda)}\left[\,b+\frac{\lambda\,(\bar{w}-\bar{v})}{r+\delta}-\,\lambda\int_{w_n}^{\bar{w}}\frac{F(w)}{r+\delta+\lambda\,(1-F(w))}\,dw\,\right], \tag{4.6}$$

5) Burdett and Mortensen (1998) and Tudela (2004) prove other basic properties such as the fact that the distribution has no mass point, and its support is a connected set. These properties also hold in our model.

where $V_E(\bar{w})$ represents the value of employed workers receiving $\bar{w}$ with high effort.

As in the Burdett-Mortensen model, we now redefine λ and γ as follows for simplification: $r/\delta \to 0$, $\lambda/\delta \equiv k$, $\gamma/\delta \equiv \beta$, where we assume that the discount rate is zero. This assumption does not generate any serious problems for the results of the model. By using (4.6) and the new parameter values, (3.3) becomes

$$\begin{aligned} 0 \;=\; & -(\bar{v}-\underline{v}) + \frac{\beta\,(w_n-\underline{v})}{1+\beta+k} - \frac{\beta}{(1+\beta)(1+k)}\left[b + k\,(\bar{w}-\bar{v}) - \int_{w_n}^{\bar{w}} \frac{k\,F(w)}{1+k\,(1-F(w))}\,dw\right] \\ & - \int_{w_n}^{\bar{w}} \frac{\beta\,k\,F(w)}{[1+k\,(1-F(w))]\,[(1+\beta+k\,(1-F(w))]}\,dw \\ & - \frac{\beta\,k}{1+\beta+k}\int_{w_n}^{\bar{w}} \frac{F(w)}{1+\beta+k\,(1-F(w))}\,dw + \frac{\beta\,k}{(1+k)(1+\beta+k)}(\bar{w}-\underline{v}). \end{aligned} \tag{4.7}$$

The wage level satisfying this condition makes workers indifferent between being employed and unemployed. In other words, it is the minimum wage level in the offers that satisfy the no shirking condition.

Next, we show that there is only one w_n meeting (4.7). How the right hand side of (4.7) relates to w_n is revealed by differentiating it. Then,

$$\begin{aligned} &\text{Second term}\;:\;\frac{\beta}{1+\beta+k},\quad \text{Third term}\;:\;\frac{\beta\,k}{(1+\beta)(1+k)}\int_{w_n}^{\bar{w}}\frac{\partial\,S(w,\,w_n)}{\partial\,w_n}\,dw \\ &\text{Fourth term}\;:-\left[\frac{\beta\,k}{1+\beta}\,\frac{d\,\bar{w}}{d\,w_n} + \beta\,k\int_{w_n}^{\bar{w}}\frac{\partial\,T(w,\,w_n)}{\partial\,w_n}\,dw\right], \\ &\text{Fifth term}\;:-\left[\frac{\beta\,k}{(1+\beta)(1+\beta+k)}\,\frac{d\,\bar{w}}{d\,w_n} + \frac{\beta\,k}{1+\beta+k}\int_{w_n}^{\bar{w}}\frac{\partial\,R(w,\,w_n)}{\partial\,w_n}dw\right], \\ &\text{Sixth term}\;:\;\frac{\beta\,k}{(1+k)(1+\beta+k)}\,\frac{d\,\bar{w}}{d\,w_n}, \end{aligned}$$

where we have already used the facts $F(\bar{w}) = 1$ and $F(w_n) = 0$, and

$$\begin{aligned} &S(w,\,w_n) \equiv \frac{F(w)}{1+k\,(1-F(w))},\; T(w,\,w_n) \equiv \frac{F(w)}{[\,1+k\,(1-F(w))]\,[\,1+\beta+k\,(1-F(w))\,]}, \\ &R(w,\,w_n) \equiv \frac{F(w)}{1+\beta+k\,(1-F(w))}. \end{aligned}$$

Furthermore, simple calculations yield $d\bar{w}/dw_n$ and $\partial F(w)/\partial w_n$ from (4, 4) and (4.5).

$$\frac{d\,\bar{w}}{d\,w_n} > 0,\;\; \frac{\partial F(w)}{\partial w_n} < 0.$$

Now we give a description of the effect of w_n on the right-hand side of (4.7) in detail. First, organizing the coefficients of $d\,\bar{w}/d\,w_n$ results in

$$\text{与式} \;=\; \frac{\beta(\beta\,k^3 + 2\,\beta\,k^2 + 3\,\beta\,k + 2\,k + 1 + \beta)}{(1+\beta)(1+\beta+k)(1+k)^3} > 0. \tag{4.8}$$

Second, (4.7) indicates that w_n also affects the condition through the integrands appearing on the right-hand side. A sign of this effect is equivalent to that of the following expression:

$$\int_{w_n}^{\bar{w}} \left[\frac{\beta\, k}{(1+\beta)(1+k)} \frac{\partial S(w,\, w_n)}{\partial w_n} - \beta\, k \frac{\partial T(w,\, w_n)}{\partial w_n} - \frac{\beta\, k}{1+\beta+k} \frac{\partial R(w,\, w_n)}{\partial w_n} \right] dw. \tag{4.9}$$

This represents the effect of an increase in w_n on the effort-expenditure decision problem of workers through the distribution $F(\cdot)$. The result described in Appendix B shows the integrand parts in (4.9) to be negative. Since $\partial F(w)/\partial w_n$ is also negative, (4.9) is positive. We conclude, then, that the right-hand side of (4.7) is increasing in w_n.

Proposition 1
There exists only one w_n that satisfies the no shirking condition with equality, and an equilibrium in this model can be uniquely characterized.

Once w_n is determined, it specifies the distribution $F(\cdot)$ and its support. Since the unemployment rate given in (4.1) is independent of w_n, we can characterize the equilibrium under this w_n. Notice that the existence and the uniqueness of the equilibrium critically depends on whether w_n is uniquely determined or not in our model.

5 Properties of the Equilibrium

5.1 Rents Received by Employed Workers

In the previous section, we showed that equilibrium in this model was uniquely characterized by (4.1), (4.4), (4.5) and (4.7). In particular, properties of the equilibrium depend on how the no shirking wage affect it.

We first examine rents received by employed workers. They receive such a rent not only due to the existence of the equilibrium wage dispersion among identical firms, but also because of information asymmetry concerning employee effort levels. At some $\hat{w} \in [\, w_n,\, \bar{w}\,]$, employed workers obtain the rent expressed by $V_E(\hat{w}) - V_U$. This rent can be separated by $V_E(\hat{w}) - V_E(w_n)$ and by $V_E(w_n) - V_U$. Wage differentials generate the former premium while information asymmetry provides employees the latter. It follows from (2.1) that $V_E(w_n)$ is described as

$$V_E(w_n) \;=\; \frac{w_n - \bar{v}}{\delta+\lambda} + \frac{\lambda\,(\bar{w} - \bar{v})}{\delta(\delta+\lambda)} + V_U - \frac{\lambda}{\delta+\lambda} \int_{w_n}^{\bar{w}} \frac{F(w)}{\delta + \lambda\,[\,1 - F(w)\,]}\, d\,w.$$

Then, $V_E(\hat{w}) - V_E(w_n)$ and $V_E(w_n) - V_U$ are written as

$$\begin{aligned} V_E(\hat{w}) - V_E(w_n) \;&=\; \frac{\hat{w} - w_n}{\delta+\lambda} + \frac{\lambda}{\delta+\lambda} \int_{w_n}^{\hat{w}} \frac{F(w)}{\delta + \lambda\,[\,1 - F(w)\,]}\, d\,w, \\ V_E(w_n) - V_U \;&=\; \frac{w_n - \bar{v}}{\delta+\lambda} + \frac{\lambda\,(\bar{w} - \bar{v})}{\delta(\delta+\lambda)} - \frac{\lambda}{\delta+\lambda} \int_{w_n}^{\bar{w}} \frac{F(w)}{\delta + \lambda\,[\,1 - F(w)\,]}\, d\,w. \end{aligned}$$

As a result, partial derivatives of the two expressions with respect to w_n are

$$\frac{\partial\,[\,V_E(\hat{w}) - V_E(w_n)\,]}{\partial w_n} < 0, \quad \frac{d\,[\,V_E(w_n) - V_U\,]}{d\,w_n} > 0,$$

since $\partial F/\partial w_n$ are negative. These results indicate that an increase in the no shirking wage reduces the premium arising from the wage dispersion, and raises the premium arising from employer inability to observe effort. The width of the support described by $\bar{w} - w_n$ shrinks as w_n increases; this means that the absolute difference of wage payments between employees is small. Employed workers obtain less benefit from turnover in this situation. The former result reflects this context. On the other hand, the increase in w_n provides all employed workers with benefit. This expands the welfare difference between employed and unemployed workers, and the latter result reflects this circumstance.

The total rent represented by $V_E(\hat{w}) - V_U$ is given by the sum of $V_E(\hat{w}) - V_E(w_n)$ and $V_E(w_n) - V_U$:

$$V_E(\hat{w}) - V_U \;=\; \frac{\hat{w} - \bar{v}}{\delta + \lambda} + \frac{\lambda\,(\bar{w} - \bar{v})}{\delta(\delta + \lambda)} - \frac{\lambda}{\delta + \lambda}\int_{\hat{w}}^{\bar{w}} \frac{F(w)}{\delta + \lambda\,[\,1 - F(w)\,]}\,dw,$$

where $\hat{w}$ is an offer contained in $[\,\underline{w},\,\bar{w}\,]$. This allows us to calculate the partial derivative of $V_E(\hat{w}) - V_U$ with respect to w_n. We obtain

$$\frac{\partial\,[\,V_E(\hat{w}) - V_U\,]}{\partial w_n} \;=\; -\,\frac{\lambda}{\delta + \lambda}\int_{\hat{w}}^{\bar{w}} \frac{\delta + \lambda}{\{\,\delta + \lambda\,[\,1 - F(w)\,]\}^2}\,\frac{\partial F(w)}{\partial w_n}\,dw. \tag{5.1}$$

Again, it follows from $\partial F/\partial w_n < 0$ that

$$\frac{\partial\,[\,V_E(\hat{w}) - V_U\,]}{\partial w_n} > 0. \tag{5.2}$$

What is important in this result is that the sign observed in (5.2) never occurs in the standard Burdett-Mortensen model because firms appeared in the standard model to recognize the reservation wage as the minimum wage level that is necessary to attract a labor force. The premium received by employed workers in this situation corresponds to $V_E(\hat{w}) - V_E(w_n)$ in our model. However, they also obtain the additional premium $V_E(w_n) - V_U$ because firms must pay enough wages to elicit high effort levels. This changes the characteristic of the rent $V_E(\hat{w}) - V_U$.

Proposition 2
The increase in w_n raises the total rent $V_E(\hat{w}) - V_U$ received by employed workers while it diminishes the rent that is due to wage dispersion.

Proposition 2 provides the following implication: The facts described in (5.1) and (5.2) indicate that the disparity between employed workers and unemployed workers expands as w_n rises, and at the same time, the wage differentials between employed workers become smaller. We conclude that these differentials can not be resolved altogether through the change in w_n. [6)]

6) A Change of the Parameter having direct effects on w_n may remedy both of them.

5.2 Detection of Shirking and Work Incentives

In Proposition 2, we examined the effect of w_n on the rent received by employed workers. But w_n itself is an endogenous variable, and all effects through w_n should be induced by the change of parameters. There are many parameters in the model, but in particular, the detection rate of shirking β seems to be important. This is because this parameter has an influence on the effort decision of employed workers since this parameter characterizes the frequency of detection.

It seems that workers have strong incentives to shirk when firms monitor them less frequently. Such reasoning is valid if every firm offers the same wage level that is equal to the no shirking wage. But if identical firms offer different wages, unemployed workers who once received w_n could find better employment opportunities through the search process. This possibility will make firing less consequential for shirkers. Actually, we show that $\partial w_n/\partial\beta > 0$ for some parameter values in the following discussion. Note that this never occurs in the standard shirking models, such as Shapiro and Stiglitz (1984). [7)]

An effect of β on w_n is described by the partial derivative of (4.7) with respect to β. This effect can be classified as through integrants and otherwise. The former effect depends critically on the partial derivatives of integrands appearing in (4.7). This is expressed as follows:

$$\frac{k\,F(\hat{w})}{(1+k)(1+\beta)^2\,[\,1+k\,(1-F(\hat{w}))\,]} - \frac{k\,F(\hat{w})}{[\,1+\beta+k(1-F(\hat{w})\,]^2)} - \frac{k\,F(\hat{w})\,[\,(1+k)(1+k(1-F(\hat{w}))-\beta^2\,]}{(1+\beta+k)^2\,[\,1+\beta+k(1-F(\hat{w}))^2\,]} \;<\; 0, \tag{5.3}$$

if $k > \beta$.

Next, the following expression is relevant to the latter effect (the effect of β on w_n, which is not through the integrants).

$$\frac{(1+k)(w_n-\underline{v})+k\,(\bar{w}-\underline{v})}{(1+\beta+k)^2} - \frac{b+k\,(\bar{w}-\bar{v})}{(1+k)(1+\beta)^2}. \tag{5.4}$$

But the sign of (5.4) is ambiguous at this time. We therefore illustrate some lemmas required for specifying the sign.

Lemma 1
The partial derivative of (5.4) *with respect to* $\underline{e}$ *is negative.*

We can prove this lemma by using the fact $\partial w_n/\partial\underline{e} < 0$. See Appendix D.

Lemma 2
The sign of (5.4) *is negative when it is evaluated at the extreme case* $\bar{v} = \underline{v}$, *and the offer arrival rate* k *is sufficiently large.*

---

7) In Appendix III, we develop a wage-posting model without on-the-job search, and we find that wage compensations and the monitoring frequency are substitutable for employers.

We have already shown that the condition for (5.3) to be negative is that k is sufficiently large compared to β. The condition described at Lemma 2 is consistent with this condition. Although the result in Lemma 2 is based on the situation $\bar{v} = \underline{v}$, there should be some $\underline{e}$ in the neighborhood of $\bar{e}$ such that (5.4) is negative at this $\underline{e}$.

We conclude then, that $\partial w_n/\partial\beta$ is positive under the specified condition described above. This high $\underline{e}$, however, may provide firms strictly positive profits by allowing their employees to shirk. If so, an equilibrium offer distribution and the no shirking wage will have different properties in comparison to the current model. We can avoid this possibility because we find $\bar{C}$ such that employers can not obtain nonnegative profits per employee with such high $\underline{e}$.

Lemma 3
For every $\bar{e}$ and $\underline{e}$ such that $\bar{e} > \underline{e}$, there exists $\bar{C}$ satisfying both $\bar{y} - \bar{w} - \bar{C} > 0$ and $\underline{y} - w_n - \bar{C} < 0$.

The existence of such $\bar{C}$ ensures that firms always prefer eliciting $\bar{e}$ from employees over making them shirk, even when $\underline{e}$ is high enough to result in $\partial w_n/\partial\beta > 0$. If $\bar{y} \neq \underline{y}$, that is, $\bar{e} \neq \underline{e}$, we find a certain $\bar{C}$ satisfying the above two inequalities.

Proposition 3
If the arrival rate k and the effort level $\underline{e}$ are both sufficiently large, the no shirking wage increases as β rises.

We can interpret this proposition as follows. First, when there exists wage dispersion, the equilibrium unemployment does not work as the discipline device indicated in Shapiro and Stiglitz (1984). In the standard shirking model with a unique equilibrium wage, shirkers suffer greater losses from being fired as the detection rate goes up. On the other hand, with wage differentials, they expect that they can find other good jobs during the unemployment period. In this case, the detection of shirking may result in a more favorable situation for shirkers. Then, firms must pay more wages to deter shirking at the higher detection rate. [8)]

Second, the increase in w_n due to the higher detection rate reduces the width of the support of $F(\cdot)$, and the distribution shifts to the right. In Albrecht and Vroman (1998), this shift arises from the improvement of the monitoring technology under more complex settings. We can say that the result $\partial w_n/\partial\beta > 0$ is not peculiar.

6 Variance in the Earnings Distribution

In this section, we investigate how the change in the detection rate affects wage dispersion. Although the wage offer distribution has been taken into account, wages that employed

8) Tudela (2004) characterizes two wage distributions depending on the states of workers, unemployed or employed. If firms could recognize why a worker had become unemployed, an equilibrium would have two equilibrium wage distributions (for laid-off workers and those who had been detected shirking and been fired).

workers actually receive must be represented by the earnings distribution; and we consider the variance (or the coefficient of variance) of this distribution as a measure of the wage differentials.

We first illustrate the structure of the earnings distribution. For any offer $\hat{w}$ contained in $[\underline{w},\, \bar{w}\,]$, let $H(\hat{w})$ be the ratio of employed workers receiving no more than the wage $\hat{w}$. The flows of these workers are described by a dynamic equation, as follows:

$$\dot{H}(\hat{w}) \;=\; \lambda\, F(\hat{w})\, u - \;\{\delta + \lambda\,[\,1 - F(\hat{w})\,]\}\, H(\hat{w}).$$

In a steady state, $\dot{H}(\hat{w}) = 0$. Thus, we obtain the stationary value of $H(\cdot)$.

Moreover, let $G(\hat{w})$ be the proportion of employed workers receiving no more than the wage $\hat{w}$ relative to all employees. Since this proportion can be defined as $G(\hat{w}) = H(\hat{w})/(m-u)$, using (4.4) yields

$$G(\hat{w}) \;=\; \frac{1}{1+k}\left(\sqrt{\frac{\bar{y} - w_n - \bar{C}}{\bar{y} - \hat{w} - \bar{C}}} - 1\right). \tag{6.1}$$

This is the earnings distribution of the model. Differentiating (6.1) with respect to $\hat{w}$ results in the density function $g(\cdot)$,

$$g(\hat{w}) \;=\; \frac{(\bar{y} - w_n - \bar{C})^{1/2}(\bar{y} - \hat{w} - \bar{C})^{-3/2}}{2(1+k)}, \tag{6.2}$$

and calculating the mean of this distribution gives

$$E(w) \;=\; \int_{w_n}^{\bar{w}} w\, dG(w) \;=\; \frac{k\,[\,k(\bar{y} - \bar{C}) + w_n\,]}{(1+k)^2}. \tag{6.3}$$

Let σ be the variance of the earning distribution. Then, following from the mathematical definition of the variance, σ is expressed as

$$\sigma = \int_{w_n}^{\bar{w}} w^2\, dG(w) - \left(\int_{w_n}^{\bar{w}} w\, dG(w)\right)^2 = \int_{w_n}^{\bar{w}} w^2\, g(w)\, d\, w - \left(\int_{w_n}^{\bar{w}} w\, g(w)\, d\, w\right)^2.$$

The derivative of σ with respect to w_n is given by

$$\begin{aligned}\frac{d\,\sigma}{d\,w_n} \;=\;& \bar{w}\, g(\bar{w})\,[\,\bar{w} - 2\,E(w)\,]\,\frac{d\,\bar{w}}{d\,w_n} + w_n\, g(w_n)\,[\,2\,E(w) - w_n\,] \\ &+ \int_{w_n}^{\bar{w}} w\;(w - 2\,E(w)\,)\,\frac{\partial g(w)}{\partial w_n}\,d\,w.\end{aligned} \tag{6.4}$$

By using (4.5), (6.2) and (6.3), the expression (6.4) can be rewritten as

$$\begin{aligned}&\frac{1}{12\,(1+k)^4(\bar{y} - w_n - \bar{C})}\,[\,-(6\,k^4 + 2\,k^3 - 24\,k^2 - 6\,k)(\bar{y} - \bar{C})^2 \\ &+ (10\,k^3 - 12\,k^2 + 30\,k)\, w_n(\bar{y} - \bar{C}) - (2\,k^3 + 18\,k - 6)\, w_n^2\,].\end{aligned} \tag{6.5}$$

That is, $d\,\sigma / d\,w_n$ is equal in sign to (6.5). Since (6.5) is a quadratic function with respect to w_n, we can ascertain whether this function takes positive value for some w_n or not. For that purpose, we define some new parameters $A \equiv -2\,k^3 - 18\,k + 6$, $B \equiv (5\,k^3 - 6\,k^2 + 15\,k)(\bar{y} - \bar{C})$, and $C \equiv (-6\,k^4 - \,k^3 + 24\,k^2 + 6\,k)(\bar{y} - \bar{C})^2$, and compute a discriminant of (6.5). Then, we obtain

$$B^2 - A\,C \;=\; (\bar{y} - \bar{C})^2\,[\,-12\,k^7 + 21\,k^6 - 120\,k^5 + 192\,k^4 + 264\,k^3 + 81\,k^2 + 72\,k\,]. \quad (6.6)$$

Concerning (6.6), it takes a positive sign for at least $k \in [\,0,\,2\,]$, but it is always negative for all $k \geq 3$. In the numerical analysis of van den Berg and Ridder (1998), the estimated value of the conditioned arrival rate (λ/δ) is almost always far greater than 3. In the first place, Proposition 3 says that it requires sufficiently high k for $\partial w_n / \partial \beta$ to be positive. Therefore, the solution to the quadratic equation described above does not have any real roots at such high k. It follows that (6.5) is always negative in this situation. [9]

Proposition 4
The higher no shirking wage makes the wage differentials among identical employees less dispersed when the arrival rate of job offers is sufficiently large.

Given the result suggested in Proposition 3, employers must pay additional wage premiums to deter shirking at the high detection rate. This implies that an improvement of the monitoring technology in some sector results in the relative contraction of wage differentials. On the other hand, this improvement also reduces the range of the support $[\,\underline{w},\,\bar{w}\,]$, which corresponds to the absolute wage differentials. Therefore, we conclude that there is a negative relationship between the monitoring technology of employers and the degree of wage differentials.

Our interpretation of this relationship follows. An increase in the no shirking wage means that employers must pay higher wages to deter shirking than before, because the improvement of monitoring technology would give other opportunities for matching better jobs under the condition of a sufficiently high arrival rate. Then, the employers can reduce the possibility of worker turnover by making offers that are less dispersed while they provide high no shirking wages; shirkers, then, have less prospects of getting favorable jobs if they are getting caught and being unemployed. Therefore, it is rational for firms to pay wages in the neighborhood of the mean in order to prevent employees from making less effort.

7 The No Shirking Wage and the Reservation Wage

We have so far considered the no shirking wage as the lowest wage offered. But whether w_n is actually the minimum offer must be explicitly proved. If the reservation wage exceeds w_n, the no shirking condition does not work as a constraint. Then, the model reduces to the

9) Note that the coefficient of variance (variance divided by mean) is also negative because the mean of the distribution $F(\cdot)$ is increasing with β.

standard Burdett-Mortensen model with minor differences. The purpose of this section, however, is to show that the reservation wage w_R is strictly less than w_n.

The reservation wage is the wage that makes workers indifferent between being unemployed and being employed. In our model, w_R satisfies $V_S(w_R) = V_U$. From (2.2), we obtain

$$w_R - \underline{v} - \lambda \int_{w_R}^{\bar{w}} \frac{F(w)}{r + \delta + \gamma + \lambda\,(1 - F(w))}\, dw + \lambda\, V_S(\bar{w}) \;=\; (r + \lambda)\, V_U. \tag{7.1}$$

Evaluating $V_S(\bar{w})$ at $\hat{w} = \bar{w}$ yields [10)]

$$V_S(\bar{w}) \;=\; \frac{\bar{w} - \underline{v}}{r + \delta + \gamma} + \frac{\delta + \gamma}{r + \delta + \gamma}\, V_U.$$

Substituting this into (7.1) gives

$$w_R - \underline{v} - \lambda \int_{w_R}^{\bar{w}} \frac{F(w)}{r + \delta + \gamma + \lambda\,(1 - F(w))}\, dw + \frac{\lambda(\bar{w} - \underline{v})}{r + \delta + \gamma} \;=\; \frac{r(r + \delta + \gamma + \lambda)}{r + \delta + \gamma} V_U. \tag{7.2}$$

Our purpose is to show $V_E(w_R) < V_S(w_R)$ by using (3.2) and (7.2). Since the coefficient of V_U in (3.2) is $\gamma\, r/(r + \delta + \gamma)$, we can derive the following expression from (7.2),

$$\begin{aligned} \frac{\gamma\, r}{r + \delta + \gamma}\, V_U \;=\; & \frac{\beta\,(w_R - \underline{v})}{1 + \beta + k} - \frac{\beta\, k}{1 + \beta + k} \int_{w_R}^{\bar{w}} \frac{F(w)}{1 + \beta + k\,(1 - F(w))}\, dw \\ & + \frac{\beta\, k\,(\bar{w} - \underline{v})}{(1 + \beta)(1 + \beta + k)}, \end{aligned} \tag{7.3}$$

where (7.3) is depicted by using the parameters k and β. (7.3) represents the value of unemployed workers when the lowest wage offer is w_R.

We put (7.3) into (3.2) instead of V_U with $\underline{w} = w_n$. Then,

$$\begin{aligned} V_E(w_R) - V_S(w_R) \;=\; & -\,(\bar{v} - \underline{v}) + \frac{\beta\, k\,(\bar{w} - \underline{v})}{1 + \beta + k} \left(\frac{1}{1 + k} - \frac{1}{1 + \beta} \right) \\ & - \int_{w_R}^{\bar{w}} \frac{\beta\, k\, F(w)}{[\,1 + k\,(1 - F(w))\,]\,[\,1 + \beta + k(1 - F(w))\,]}\, dw. \end{aligned}$$

If $k \geq \beta$, the second term in the right-hand side of (7.4) is non-positive, and the sign of (7.4) is negative. It follows that w_n is strictly greater than w_R, since w_n satisfies $V_E(w_n) = V_S(w_n)$ and $V_E(\hat{w}) - V_S(\hat{w})$ is increasing in $\hat{w}$ by (3.2).

Proposition 5
If the offer arrival rate is large enough to be greater than the rate of detecting shirkers, the reservation wage w_R does not meet the no shirking condition.

10) See the last expression in Appendix A.

8 Conclusions

In this paper, we have combined the shirking model with the Burdett-Mortensen wage-posting framework in order to examine how wage differentials among homogeneous firms are affected by information asymmetry associated with worker effort levels. When employers prefer diligent workers to shirkers, they must make wage offers high enough to give employees incentives to work hard. The model indicates that employers can no longer compensate workers by paying the reservation wage, but instead, they determine their offers by considering the no shirking condition which is one of the conditions characterizing a wage-posting equilibrium in this model. We have shown that there exists a unique wage level satisfying the no shirking condition with equality. This result means that the equilibrium can be uniquely characterized with a specified wage offer distribution.

We have obtained the following two main conclusions. First, under some parameters values, workers are more inclined to shirk as the rate of detecting shirking rises. At first glance, it is not intuitive. But when wages are dispersed among identical employers, receiving higher offers leads to higher utility for workers at a future date even if they got caught shirking and become unemployed. In this situation, the movement to the unemployment pool by firing is no longer a serious penalty for shirkers if employers cannot find out whether job seekers shirked or not in a previous workplace. This implies that equilibrium unemployment does not work as the discipline device described by Shapiro and Stiglitz.

Second, given the first result, we have shown that the variance in the earning distribution becomes less dispersed as the detection rate goes up. In other words, wage differentials among identical employers will contract when they become able to catch shirkers more accurately. That is, we conclude that the wage dispersion generated by the wage posting equilibrium and monitoring technology have a negative relationship, rather than the positive one suggested in the existing research.

Finally, we point out that the results obtained in this paper depend on the existence of on-the-job search, which is relevant not only to the existence of the equilibrium, but also to its properties. Therefore, the robustness of these results should be investigated in a situation where only unemployed workers engage in search activity. Then, the model would require other factors, such as the variation of monitoring technology among employers and the difference in worker characteristics, in order to generate wage dispersion in equilibrium. These are topics for future research.

Appendix I. Some Proof and Derivation Processes

A. The Derivation of (3.2)

To rewrite (3.2) by using (3.1), we have to know more about the following expressions:

$$(1)\ :\int_{\hat{w}}^{\bar{w}} [\, V_E(w) - V_S(w)\,]\, dF(w) \quad (2)\ : V_U - V_S(\hat{w}).$$

Concerning (1), simple calculations yield

$$
\begin{aligned}
&\int_{\hat{w}}^{\bar{w}} [\, V_E(w) - V_S(w) \,] \, dF(w) \\
&= [\, V_E(\bar{w}) - V_S(\bar{w}) \,] \; - \; [\, V_E(\hat{w}) - V_S(\hat{w}) \,] \, F(\hat{w}) - \int_{\hat{w}}^{\bar{w}} [\, V_E'(w) - V_S'(w) \,] \, F(w) \, dw, \\
&= [\, V_E(\bar{w}) - V_S(\bar{w}) \,] \; - \; [\, V_E(\hat{w}) - V_S(\hat{w}) \,] \, F(\hat{w}) \\
&\qquad - \int_{\hat{w}}^{\bar{w}} \frac{\gamma \, F(w)}{[\, r + \delta + \lambda \, (1 - F(w)) \,] \, [\, r + \delta + \gamma + \lambda \, (1 - F(w)) \,]} \, dw, \qquad \text{(A.1)}
\end{aligned}
$$

where $V_E'(w)$ and $V_S'(w)$ are calculated from (2.1) and (2.2) as

$$
V_E'(w) \; = \; \frac{1}{r + \delta + \lambda \, (1 - F(w))}, \;\; V_S'(w) \; = \; \frac{1}{r + \delta + \gamma + \lambda \, (1 - F(w))}.
$$

By using (2.2), the expression (2) can be written as

$$
\begin{aligned}
V_S(\hat{w}) = \; & \frac{\hat{w} - \underline{v} + \lambda \, V_S(\bar{w}) + (\delta + \gamma) \, V_U}{r + \delta + \lambda + \gamma} \\
& - \; \frac{\lambda}{r + \delta + \lambda + \gamma} \int_{\hat{w}}^{\bar{w}} \frac{F(w)}{r + \delta + \gamma + \lambda \, (1 - F(w))} \, dw. \qquad \text{(A.2)}
\end{aligned}
$$

Subtracting V_U from both sides of (A.2), and combined with (A.1) results in (3.2).

B. The Sign of (4.10)

We first note the following facts:

$$
\begin{aligned}
\frac{\partial S(w, \, w_n)}{\partial w_n} \; &= \; \frac{1 + k}{[\, 1 + k(1 - F(w)) \,]^2} \, \frac{\partial F(w)}{\partial w_n}, \\
\frac{\partial T(w, \, w_n)}{\partial w_n} \; &= \; \frac{(1 + \beta)(1 + k) + k + k^2 \, (1 - F(w)^2)}{[\, 1 + k(1 - F(w)) \,]^2 \, [\, 1 + \beta + k \, (1 - F(w)) \,]^2} \, \frac{\partial F(w)}{\partial w_n}, \\
\frac{\partial R(w, \, w_n)}{\partial w_n} \; &= \; \frac{1 + \beta + k}{[\, 1 + \beta + k \, (1 - F(w)) \,]^2} \, \frac{\partial F(w)}{\partial w_n}.
\end{aligned}
$$

Our purpose in this appendix is to examine the sign of (4.10). But it suffices to see the sign of its integrand parts. They can be written by

$$
\begin{aligned}
& \frac{\partial F(w)}{\partial w_n} \left\{ \frac{\beta \, k}{(1 + \beta)[\, 1 + k \, (1 - F(w)) \,]^2} \right. \\
& \left. - \frac{\beta \, k \, [\, (1 + \beta)(1 + k) + k + k^2 \, (1 - F^2) \,]}{[\, 1 + k(1 - F(w)) \,]^2 \, [\, 1 + \beta + k \, (1 - F(w)) \,]^2} - \frac{\beta \, k}{[\, 1 + \beta + k \, (1 - F(w)) \,]^2} \right\}. \qquad \text{(B.1)}
\end{aligned}
$$

The bracket in (B.1) is equal in sign to the following expression:

$$
- \, k \, (1 + \beta)^2 - \beta \, k^2 \, [\, 1 - F(\hat{w}) \,]^2 - (1 + \beta)(1 + k) - k^2 \, (1 + \beta) \, [\, 1 - F(\hat{w})^2 \,].
$$

Obviously, this has a negative sign. It implies (B.1) has also a negative sign. When an integrand is negative for every w, (4.10) is also negative.

C. The Derivation of $V_E(\hat{w}) - V_U$ in Section 5

Rewriting $V_E(\hat{w})$ lead to

$$r\, V_E(\hat{w}) = \hat{w} - \bar{v} + \lambda \left\{ [\, V_E(w)\, F(w)\,]_{\hat{w}}^{\overline{w}} - \int_{\hat{w}}^{\overline{w}} V_E'(w)\, F(w)\, dw - [\,1 - F(\hat{w})\,]\, V_E(\hat{w}) \right\} - \delta\, V_E(\hat{w}) + \delta\, V_U. \quad \text{(C.1)}$$

Remember that we have already shown the value $V_E'(\hat{w})$. By substituting this into (C.1) and arranging it, this results in

$$(r + \delta + \lambda)\, V_E(\hat{w}) = \hat{w} - \bar{v} + \lambda \left[V_E(\overline{w}) - \int_{\hat{w}}^{\overline{w}} \frac{F(w)}{r + \delta + \lambda\,[\,1 - F(w)\,]}\, dw \right] + \delta\, V_U. \quad \text{(C.2)}$$

Note that $V_E(\overline{w})$ is expressed as

$$V_E(\overline{w}) = \frac{\overline{w} - \bar{v}}{r + \delta} + \frac{\delta}{r + \delta}\, V_U.$$

We incorporate this into (C.2) and take limits as $r \longrightarrow 0$ for simplicity. Then

$$V_E(\hat{w}) = \frac{\hat{w} - \bar{v}}{\delta + \lambda} + \frac{\lambda\,(\overline{w} - \bar{v})}{\delta(\delta + \lambda)} + V_U - \frac{\lambda}{\delta + \lambda} \int_{\hat{w}}^{\overline{w}} \frac{F(w)}{\delta + \lambda\,[\,1 - F(w)\,]}\, dw.$$

This leads to $V_E(\hat{w}) - V_U$ which is described in Subsection 5.1.

D. Proof of Lemma 1

The partial derivative of (5.4) with respect to $\underline{w}$ is

$$\frac{1+k}{(1+\beta+k)^2}\,\frac{\partial w_n}{\partial \underline{e}} - \frac{1+k}{(1+\beta+k)^2}\,\frac{\partial \underline{v}}{\partial \underline{e}} + \frac{k}{(1+\beta+k)^2}\,\frac{\partial \overline{w}}{\partial \underline{e}} - \frac{k}{(1+\beta+k)^2}\,\frac{\partial \underline{v}}{\partial \underline{e}} - \frac{k}{(1+k)(1+\beta)^2}\,\frac{\partial \overline{w}}{\partial \underline{e}}, \quad \text{(D.1)}$$

where we assume that $\underline{v}$ is increasing in $\underline{e}$. The sum of the first term, the third term and the fifth term, results in

$$\begin{aligned} & \frac{1+k}{(1+\beta+k)^2}\,\frac{\partial w_n}{\partial \underline{e}} - \frac{k^2\,(k+1-\beta^2)}{(1+k)^3(1+\beta)^2(1+\beta+k)^2}\,\frac{\partial w_n}{\partial \underline{e}}, \\ < \; & \frac{1+k}{(1+\beta+k)^2}\,\frac{\partial w_n}{\partial \underline{e}} - \frac{k^2\,(1+k)}{(1+k)^3(1+\beta)^2(1+\beta+k)^2}\,\frac{\partial w_n}{\partial \underline{e}}, \\ = \; & \frac{1+k}{(1+\beta+k)^2}\left[1 - \frac{k^2}{(1+k)^3(1+\beta)^2}\right]\frac{\partial w_n}{\partial \underline{e}}, \\ < \; & 0, \end{aligned}$$

where

$$\frac{\partial \overline{w}}{\partial \underline{e}} = \frac{1}{(1+k)^2}\,\frac{\partial w_n}{\partial \underline{e}} \quad \text{and} \quad \frac{\partial w_n}{\partial \underline{e}} < 0.$$

The last inequality holds because of $k > 0$. The proof is completed.

E. Proof of Lemma 2

We first consider (5.4) at $\underline{v} = \bar{v} \equiv v$ (or $\underline{e} = \bar{e}$). Then, it can be described as

$$\frac{(1+k)}{(1+\beta+k)^2}(w_n - v) - \frac{k^2\,(k+1-\beta^2)}{(1+k)(1+\beta)^2(1+\beta+k)^2}(\bar{w} - v) - \frac{b}{(1+k)(1+\beta)^2}. \tag{5.4'}$$

It follows from (5.4′) that it is negative if the sum of the first two terms is negative. Since $\bar{w} > w_n$, it remains to examine when the following result holds:

$$\frac{(1+k)}{(1+\beta+k)^2} - \frac{k^2\,(k+1-\beta^2)}{(1+k)(1+\beta)^2(1+\beta+k)^2} < 0 \implies (5.4') \ < 0.$$

The former part of this can be rewritten as

$$\begin{aligned} &\frac{(1+k)}{(1+\beta+k)^2} - \frac{k^2\,(k+1-\beta^2)}{(1+k)(1+\beta)^2(1+\beta+k)^2}, \\ &= \frac{-\,k^3 + 2\,\beta\,(1+\beta)\,k^2 + 2\,(1+\beta)^2\,k \ + (1+\beta)^2}{(1+k)(1+\beta)^2(1+\beta+k)^2}. \end{aligned} \tag{E.1}$$

In short, a sufficient high k ensures (E.1) to be negative. This is consistent with the condition required in Proposition 3 and Proposition 5.

Appendix II. Mathematical Supplementary concerning the Earning Distribution

The Mean of the Distribution $G(\cdot)$

It follows from the density function $g(\cdot)$, given by (6.2), that the mean of the earning distribution is

$$\begin{aligned} (6.3) \ &= \int_{w_n}^{\bar{w}} \frac{w\,(\bar{y} - w_n - \bar{C})^{1/2}(\bar{y} - w - \bar{C})^{-3/2}}{2(1+k)}\,d\,w, \\ &= \frac{(\bar{y} - w_n - \bar{C})^{1/2}}{2(1+k)} \int_{w_n}^{\bar{w}} (\bar{y} - w - \bar{C})^{-3/2}\,d\,w, \\ &= \frac{(\bar{y} - w_n - \bar{C})^{1/2}}{1+k} \int_{(\bar{y}-\underline{w}-\bar{C})^{-1/2}}^{(\bar{y}-\bar{w}-\bar{C})^{-1/2}} \left(\bar{y} - \bar{C} - \frac{1}{t^2}\right) dt, \\ &= \frac{(\bar{y} - w_n - \bar{C})^{1/2}(\bar{y} - \bar{C})\,[\,(\bar{y} - \bar{w} - \bar{C})^{-1/2} - (\bar{y} - \bar{w} - \bar{C})^{-1/2}\,]}{1+k} \\ &\qquad\qquad - \frac{(\bar{y} - w_n - \bar{C})^{1/2}}{1+k} \int_{(\bar{y}-\underline{w}-\bar{C})^{-1/2}}^{(\bar{y}-\bar{w}-\bar{C})^{-1/2}} \frac{1}{t^2}\,dt, \\ &= \frac{k\,(\bar{y} - \bar{C})}{1+k} - \frac{k\,(\bar{y} - w_n - \bar{C})}{(1+k)^2}, \\ &= \frac{k\,[\,k\,(\bar{y} - \bar{C}) + w_n\,]}{(1+k)^2}, \end{aligned}$$

where the second equality stems from the transformation of the variable $\hat{w}$ to t, such that $t = 1/\sqrt{\bar{y} - w - \bar{C}}$, and the fourth equality can be obtained from the following fact:

$$\bar{y} - \bar{w} - \bar{C} \; = \; \frac{\bar{y} - w_n - \bar{C}}{(1+k)^2}.$$

This is easily computed from (4.5).

Details of (6.4)

$w - 2\,E(w)$:

$$\begin{aligned}
\bar{w} - 2\,E(w) \; &= \; (\bar{y} - \bar{C}) \left[1 - \frac{1}{(1+k)^2} \right] + \frac{w_n}{(1+k)^2} - \frac{2\,k^2\,(\bar{y} - \bar{C})}{(1+k)^2} - \frac{(2\,k-1)\,w_n}{(1+k)^2}, \\
&= \; \frac{(\bar{y} - \bar{C})(2k - k^2)}{(1+k)^2} - \frac{w_n\,(2\,k-1)}{(1+k)^2},
\end{aligned}$$

$w_n - 2E(w)$:

$$w_n - 2E(w) \;\; = \;\; - \frac{2\,k^2\,(\bar{y} - \bar{C})}{(1+k)^2} + \frac{w_n\,(1+k^2)}{(1+k)^2},$$

Computation of the Integral given by (1):

$$\begin{aligned}
& 2\,E(w) \int_{w_n}^{\bar{w}} w\,\frac{\partial g(w)}{\partial w_n}\,d\,w, \\
&= - \frac{E(w)}{2\,(1+k)(\bar{y} - w_n\bar{C})^{1/2}} \left\{ (\bar{y} - \bar{C}) \left[\frac{1}{(\bar{y} - \bar{w} - \bar{C})^{1/2}} - \frac{1}{(\bar{y} - w_n - \bar{C})^{1/2}} \right] \right. \\
& \left. - \frac{k\,(\bar{y} - w_n - \bar{C})^{-1/2}}{1+k} \right\}, \\
&= - \frac{k\,E(w)}{2\,(1+k)(\bar{y} - w_n\bar{C})^{1/2}} \left[\frac{\bar{y} - \bar{C}}{(\bar{y} - w_n - \bar{C})^{1/2}} - \frac{(\bar{y} - w_n - \bar{C})^{1/2}}{1+k} \right], \\
&= - \frac{k\,E(w)}{2\,(1+k)} \left[\frac{\bar{y} - \bar{C}}{\bar{y} - w_n - \bar{C}} - \frac{1}{1+k} \right], \\
&= - \frac{k^2\,[\,k\,(\bar{y} - \bar{C}) + w_n\,]}{2\,(1+k)^3} \left[\frac{\bar{y} - \bar{C}}{\bar{y} - w_n - \bar{C}} - \frac{1}{1+k} \right], \qquad \text{(F.1)}
\end{aligned}$$

where

$$\frac{\partial g(w)}{\partial w_n} \; = \; - \frac{(\bar{w} - w_n - \bar{C})^{-1/2}(\bar{y} - w - \bar{C})^{-3/2}}{4\,(1+k)}.$$

Computation of the Integral given by (2):

$$\int_{w_n}^{\bar{w}} w^2\,\frac{\partial g(w)}{\partial w_n} d\,w \; = \; - \frac{(\bar{y} - w_n - \bar{C})^{1/2}}{4\,(1+k)} \int_{w_n}^{\bar{w}} w^2\,(\bar{y} - w - \bar{C})^{-1/2}\,d\,w, \qquad \text{(F.2)}$$

The integral appearing in (F.2) is

$$\int_{w_n}^{\bar{w}} w^2\,(\bar{y} - w - \bar{C})^{-1/2}\,d\,w \;=\; 2\,k\,(\bar{y} - \bar{C})^2\,(\bar{y} - w_n - \bar{C})^{-1/2}$$

$$- \frac{4\,k\,(\bar{y} - \bar{C})(\bar{y} - w_n - \bar{C})^{1/2}}{1+k} + \frac{2\,[\,(1+k)^3 - 1\,](\bar{y} - w_n - \bar{C})^{3/2}}{3(1+k)^3},$$

$$= \frac{2\,k(\bar{y} - \bar{C})\,[\,(1+k)(\bar{y} - \bar{C}) - 2(\bar{y} - w_n - \bar{C})\,]}{(1+k)(\bar{y} - w_n - \bar{C})^{1/2}} + \frac{2\,[\,(1+k)^3 - 1\,](\bar{y} - w_n - \bar{C})^{3/2}}{3(1+k)^3},$$

$$= \frac{2\,k(\bar{y} - \bar{C})\,[\,(k-1)(\bar{y} - \bar{C}) + 2\,w_n\,]}{(1+k)(\bar{y} - w_n - \bar{C})^{1/2}} + \frac{2\,[\,(1+k)^3 - 1\,](\bar{y} - w_n - \bar{C})^{3/2}}{3(1+k)^3}.$$

Substituting this result into (F.2) yields

$$(F.2) = -\frac{(\bar{y} - w_n - \bar{C})^{1/2}}{4\,(1+k)} \left\{ \frac{2\,k(\bar{y} - \bar{C})\,[\,(k-1)(\bar{y} - \bar{C}) + 2\,w_n\,]}{(1+k)(\bar{y} - w_n - \bar{C})^{1/2}} + \frac{2\,[\,(1+k)^3 - 1\,](\bar{y} - w_n - \bar{C})^{3/2}}{3(1+k)^3} \right\}, \tag{F.3}$$

Computation of $\bar{w}\,g(\bar{w})\,[\,\bar{w} - 2\,E(w)\,]\,d\,\bar{w}\,/\,d\,w_n$:

$$\bar{w}\,g(\bar{w})\,[\,\bar{w} - 2\,E(w)\,]\,\frac{d\,\bar{w}}{d\,w_n}$$

$$= \frac{\bar{w}\,(\bar{y} - w_n - \bar{C})^{1/2}(\bar{y} - \bar{w} - \bar{C})^{-3/2}\,[\,(2k - k^2)(\bar{y} - \bar{C}) - (2k-1)\,w_n\,]}{2\,(1+k)^5},$$

$$= \frac{\bar{w}(1+k)^3\,[\,(2k - k^2)(\bar{y} - \bar{C}) - (2k-1)\,w_n\,]}{2\,(1+k)^5\,(\bar{y} - w_n - \bar{C})},$$

$$= \frac{\bar{w}\,[\,(2k - k^2)(\bar{y} - \bar{C}) - (2k-1)\,w_n\,]}{2\,(1+k)^2\,(\bar{y} - w_n - \bar{C})}, \tag{F.4}$$

Computation of $w_n\,g(w_n)\,[\,2\,E(w) - w_n\,]$:

$$w_n\,g(w_n)\,[\,2\,E(w) - w_n\,] \;=\; \frac{w_n\,[\,2k^2(\bar{y} - \bar{C}) - (1+k^2)\,w_n\,]}{2\,(1+k)^3(\bar{y} - w_n - \bar{C})^{-1}}, \tag{F.5}$$

where

$$g(w_n) \;=\; \frac{(\bar{y} - w_n - \bar{C})^{1/2}(\bar{y} - w_n - \bar{C})^{-3/2}}{2(1+k)} \;=\; \frac{(\bar{y} - w_n - \bar{C})^{-1}}{2\,(1+k)}.$$

$(F.4) - (F.1)$:

$$
\begin{aligned}
(F.4) - (F.1) = &\ \frac{\bar{w}\,[\,(2k - k^2)(\bar{y} - \bar{C}) - (2k-1)\,w_n\,]}{2\,(1+k)^2\,(\bar{y} - w_n - \bar{C})} \\
&+ \frac{k^2\,[\,k\,(\bar{y} - \bar{C}) + w_n\,]}{2\,(1+k)^3} \left[\frac{k\,(\bar{y} - \bar{C}) + w_n}{(1+k)(\bar{y} - w_n - \bar{C})} \right], \\
= &\ \frac{\bar{w}\,[\,(2k - k^2)(\bar{y} - \bar{C}) - (2k-1)\,w_n\,]}{2\,(1+k)^2\,(\bar{y} - w_n - \bar{C})} \\
&+ \frac{k^2\,[\,k\,(\bar{y} - \bar{C}) + w_n\,][\,k\,(\bar{y} - \bar{C}) + w_n\,]}{2\,(1+k)^4(\bar{y} - w_n - \bar{C})}, \\
= &\ \frac{[\,(k^2 + 2k)(\bar{y} - \bar{C}) + w_n\,]\,[\,(-k^2 + 2k)(\bar{y} - \bar{C}) - (2k-1)w_n\,]}{2\,(1+k)^4\,(\bar{y} - w_n - \bar{C})} \\
&+ \frac{k^2\,[\,k\,(\bar{y} - \bar{C}) + w_n\,]^2}{2\,(1+k)^4(\bar{y} - w_n - \bar{C})}, \\
= &\ \frac{1}{2\,(1+k)^4\,(\bar{y} - w_n - \bar{C})}\,\{-(k^4 - 4k^2)(\bar{y} - \bar{C})^2 \\
&- (2k^3 + 4k^2 - 4k)w_n\,(\bar{y} - \bar{C})\; - \;(2k-1)w_n^2 + k^4(\bar{y} - \bar{C})^2 \\
&+ 2k^3\,w_n(\bar{y} - \bar{C}) + k^2\,w_n\}, \\
= &\ \frac{1}{2\,(1+k)^4\,(\bar{y} - w_n - \bar{C})}\,\{(\bar{y} - \bar{C})\,[\,(\bar{y} - \bar{C})\,[\,4\,k^2(\bar{y} - \bar{C}) - (4k^2 - 4k)\,] \\
&+ (k-1)^2\,w_n^2\,\},
\end{aligned} \tag{F.6}
$$

$(F.3) + (F.5)$:

$$
\begin{aligned}
(F.3) + (F.5) = &\ \frac{w_n\,[\,2k^2(\bar{y} - \bar{C}) - (1+k^2)\,w_n\,]}{2\,(1+k)^3(\bar{y} - w_n - \bar{C})^{-1}} - \frac{2k(\bar{y} - \bar{C})\,[\,(k-1)(\bar{y} - \bar{C}) + 2\,w_n\,]}{4(1+k)^2(\bar{y} - w_n - \bar{C})} \\
&- \frac{[\,(1+k)^3 - 1\,](\bar{y} - w_n - \bar{C})}{6(1+k)^4}, \\
= &\ \frac{4k^2 w_n(\bar{y} - \bar{C}) - 2(1+k^2)\,w_n^2 - (2k^3 - 2k^2)(\bar{y} - \bar{C})^2 - (2k + 2k^2)w_n(\bar{y} - \bar{C})}{4(1+k)^3(\bar{y} - w_n - \bar{C})} \\
&- \frac{[\,(1+k)^3 - 1\,](\bar{y} - w_n - \bar{C})}{6(1+k)^4}, \\
= &\ \frac{(2k^2 - 2k)w_n(\bar{y} - \bar{C}) - (2k^3 - 2k^2)\,(\bar{y} - \bar{C})^2}{4(1+k)^3(\bar{y} - w_n - \bar{C})} \\
&- \frac{[\,(1+k)^3 - 1\,](\bar{y} - w_n - \bar{C})}{6(1+k)^4}, \\
= &\ \frac{1}{12(1+k)^4(\bar{y} - w_n - \bar{C})}\{(1+k)(6k^2 - 6k)w_n(\bar{y} - \bar{C}) \\
&- (1+k)(6k^3 - 6k^2)(\bar{y} - \bar{C})^2 - 2\,[(1+k)^3 - 1\,]\,(\bar{y} - w_n - \bar{C})^2\}, \\
= &\ \frac{1}{12(1+k)^4(\bar{y} - w_n - \bar{C})}\,[\,(10k^3 + 12k^2 + 6k)w_n(\bar{y} - \bar{C}) \\
&- (6k^4 + 2k^3 - 6k)(\bar{y} - \bar{C})^2 - 2(k^3 + 3k^2 + 3k)\,w_n^2\,],
\end{aligned} \tag{F.7}
$$

The sum of $(F.6)$ and $(F.7)$:

$$
\begin{aligned}
&(F.6)+(F.7)\\
&= \frac{1}{12(1+k)^4(\bar{y}-w_n-\bar{C})}\,[\,(10k^3+12k^2+6k)w_n(\bar{y}-\bar{C})\\
&-\;(6k^4+2k^3-6k)(\bar{y}-\bar{C})^2-2(k^3+3k^2+3k)\,w_n^2+24k^2(\bar{y}-\bar{C})^2\\
&-\;(24k^2-24k)w_n\,(\bar{y}-\bar{C})+6\,(k-1)^2\,w_n^2\,],\\
&= \frac{1}{12(1+k)^4(\bar{y}-w_n-\bar{C})}\,[\,-(6k^4+2k^3-24k^2-6k)(\bar{y}-\bar{C})^2\\
&+\;(10k^3-12k^2+30k)w_n(\bar{y}-\bar{C})-(2k^3+18k-6)\,w_n^2\,].
\end{aligned}
$$

This is (6.5).

The sign of (6.5):
It is worth noting that $d\,\sigma/d\,w_n$ is equal in sign to (6.5) and it is a quadratic (and concave) function with respect to w_n. We define new parameters A, B, and C such that

$$
\begin{aligned}
&A\equiv -2k^3-18k+6,\\
&B\equiv (5k^3-6k^2+15k)(\bar{y}-\bar{C}),\\
&C\equiv (-6k^4-2k^3+24k^2+6k)(\bar{y}-\bar{C})^2.
\end{aligned}
$$

We compute the discriminant of the quadratic equation which is given by taking (6.5) equal to zero. This results in

$$
\begin{aligned}
B^2-AC &= (\bar{y}-\bar{C})^2\,[\,(5k^3-6k^2+15k)^2-(2k^3+18k-6)(6k^4+2k^3-24k^2-6k)\,],\\
&= (\bar{y}-\bar{C})^2\,[\,(25k^6-60k^5+180k^4-180k^3+225k^2)\\
&\qquad -\;(12k^7+4k^6+60k^5-12k^4-444k^3+144k^2-72k)\,],\\
&= -(\bar{y}-\bar{C})^2\,(\,12\,k^7+21\,k^6-120\,k^5+192\,k^4+264\,k^3+81\,k^2+72\,k). \qquad \text{(F.8)}
\end{aligned}
$$

(F.8) is equivalent to (6.6) in Section 6.

Appendix III. A Case without On-The-Job Search

Suppose now that only unemployed workers engage in searching for jobs. Then, we will show that only one wage is determined in the equilibrium. In the standard wage posting model, it is well-known that the model without on-the-job search generates Diamond paradox (no worker participates in the search process in equilibrium). On the other hand, employers in our model intend to pay high wages to deter shirking, and so, a unique wage determined in this case satisfies the no shirking condition rather than the reservation property, and workers are willing to be employed even at this wage level. There exists an equilibrium with positive employment. We will show this below.

It is worth noting that since only unemployed workers seek jobs, the firms' hiring level depends on the reservation wage, and not on the offer they actually post. It follows that hiring levels can be described as

$$L(w_n,\, F) \;=\; \frac{m\,k}{1 + k\,(1 - F(\hat{w}_n))}.$$

An equilibrium wage will be determined uniquely, and it must satisfy the no shirking condition with equality. Then, we have an interest in the value of this wage (denote it $\hat{w}_n$) and examine whether this wage exceeds the reservation wage. The employment level of a firm, therefore, depends on $\hat{w}_n$.

The Bellman equations for employed workers are

$$r\,V_E(\hat{w}) = \hat{w} - \bar{v} + \delta\,[\,V_U - V_E(\hat{w})\,], \tag{G.1}$$

$$r\,V_S(\hat{w}) = \hat{w} - \underline{v} + (\delta + \gamma)\,[\,V_U - V_S(\hat{w})\,], \tag{G.2}$$

for a wage $\hat{w}$. It follows from (G.1) and (G.2) that the no shirking constraint is written as

$$(r + \delta)\,[\,V_E(\hat{w}) - V_S(\hat{w})\,] \;=\; -\bar{v} + \underline{v} + \gamma\,[\,V_S(\hat{w}) - V_U\,] \geq 0. \tag{G.3}$$

Furthermore, (G.2) yields

$$V_S(\hat{w}) - V_U \;=\; \frac{\hat{w} - \underline{v}}{r + \delta + \gamma} - \frac{r\,V_U}{r + \delta + \gamma}, \tag{G.4}$$

and substituting (G.4) into (G.3) results in

$$(r + \delta)\,[\,V_E(\hat{w}) - V_S(\hat{w})\,] \;=\; -\bar{v} + \underline{v} + \gamma\,\left[\frac{\hat{w} - \underline{v}}{r + \delta + \gamma} - \frac{r}{r + \delta + \gamma} V_U\right] \geq 0. \tag{G.5}$$

On the other hand, the value of unemployed workers is

$$r\,V_U \;=\; b + \,\lambda\,[\,V_E(\hat{w}_n) - V_U\,]\,.$$

Since $V_E(\hat{w}_n)$ is derived from (G.1), V_U can be rewritten by

$$V_U \;=\; \frac{r + \delta}{r\,(r + \lambda + \delta)}\,\left[b + \frac{\lambda}{r + \delta}\,(\hat{w}_n - \bar{v})\right]. \tag{G.6}$$

Substituting (G.6) into (G.5) results in

$$\begin{aligned}
&(r + \delta)\,[\,V_E(\hat{w}) - V_S(\hat{w})\,] \\
&= -\bar{v} + \underline{v} + \gamma\,\left\{\frac{\hat{w} - \underline{v}}{r + \delta + \gamma} - \frac{r + \delta}{r\,(r + \lambda + \delta)}\,\left[b + \frac{\lambda}{r + \delta}\,(\hat{w}_n - \bar{v})\right]\right\}.
\end{aligned} \tag{G.7}$$

At $\hat{w} = \hat{w}_n$, (G.7) must be zero by definition of $\hat{w}_n$. Since (G.7) is increasing with $\hat{w}$, it suffices for employers to pay $\hat{w}_n$ in order to prevent workers from shirking. We can solve for $\hat{w}_n$:

$$\hat{w}_n \;=\; \frac{(1 + k)(1 + \beta)\,\bar{v} - (1 + k + \beta\,k)\,\underline{v} \,+ \beta\,b}{\beta}, \tag{G.8}$$

where we have derived (G.8) by using the parameters k, β instead of λ, δ. (G.8) is essentially same as the wage characterized in Shapiro and Stiglitz (1984).

The partial derivative of (G.8) leads to $\partial\hat{w}_n/\partial\beta < 0$ because of $\bar{v} > \underline{v}$. This result indicates the complementarity between wages and the ability of detecting shirking. Now, workers take great risk for losing a high wage job. This risk encourages employed workers to work hard. Thus, employers pay less wages as the detection rate tends to decrease. But these arguments do not hold in the model with on-the-job search.

Finally, we show that $\hat{w}_n$ is greater than the reservation wage. It follows from (G.4) and (G.6) that we obtain

$$V_S(\hat{w}) - V_U = \frac{\hat{w} - \underline{v}}{r+\delta+\gamma} - \frac{(r+\delta)\,b}{(r+\delta+\gamma)(r+\delta+\lambda)} - \frac{\lambda\,(\hat{w}_n - \bar{v})}{(r+\delta+\gamma)(r+\delta+\lambda)}. \qquad \text{(G.9)}$$

By the definition of the reservation wage, (G.9) becomes zero at this wage (denote it w_r). Solving for w_r yields

$$w_r = \frac{b + k\,(\hat{w}_n - \bar{v}) + (1+k)\,\underline{v}}{1+k}. \qquad \text{(G.10)}$$

Note that this wage, w_r, results in Diamond paradox if we consider search costs to be positive. But, comparing w_r and $\hat{w}_n$, by using (G.8) and (G.10), leads to

$$\begin{aligned} \hat{w}_n \geq w_r &\iff -(1+k+\beta)\,\bar{v} + (1+\beta+k+2\,\beta\,k)\,\underline{v} \leq 0, \\ &\iff \frac{\bar{v}}{\underline{v}} \geq \frac{1+\beta+k+2\,\beta\,k}{1+\beta+k}. \end{aligned}$$

That is, the condition required for $\hat{w}_n > w_r$ is that $\bar{e}$ is sufficiently greater than $\underline{e}$. This is obviously satisfied if we assume $\underline{e} = 0$.

We conclude that there is a unique wage-posting equilibrium characterized by the wage satisfying the no shirking condition, even when only unemployed workers engage in seeking jobs. In other word, the last argument in this appendix indicates that workers participate in the labor market since they can obtain enough compensation to exert high effort levels, that is strictly greater than the reservation wage.

References

Abe, M., and I. Ohashi (2004), "Inter-Industry and Firm Size Effects on Wage Differentials and Efficiency Wages in Japan." *Hitotsubashi University Research Unit for Statistical Analysis in Social Sciences,* Discussion Paper Series, No 25.

Albrecht, J. W., and S. B. Vroman (1992), "Dual Labor Markets, Efficiency Wages and Search." *Journal of Labor Economics*, Vol. 10, No. 4, pp. 438 – 461.

Albrecht, J. M., and S. B. Vroman (1998), "Nash Equilibrium Efficiency Wage Distributions." *International Economic Review,* Vol. 39, No. 1, pp. 183 – 203.

Bayard, A. K., and K. R. Troske (1999), "Examining the Employer-Size Wage Premium in the Manufacturing, retail Trade, and Service Industries Using Employer-Employee Mathced Data," *American Economic Review*, Vol. 89, No. 2, pp. 99 – 103.

Bontemps, C., Lean-Marc. Robin and G. J. van den Berg (2000), "Equilibrium Search with Continuous Productivity Dispersion: Theory and Nonparametric Estimation." *International Economic Review,* Vol. 41, No. 2, pp. 305 – 358.

Brown, C., and L. Medoff (1989), "The Employer Size-Wage Effect," *Journal of Political Economy*, Vol. 97, No. 5, pp. 1027 – 1059.

Burdett, K., and D. T. Mortensen (1998), "Wage Differentials, Employer Size, and Unemployment." *International Economic Review,* Vol. 39, No. 2, pp. 257 – 273.

Blackburn, M., and D. Neumark (1992), "Unobserved Ability, Efficiency Wages, and Interindustry Wage Differentials." *Quarterly Journal of Economics,* Vol. 107, No. 4, pp. 1421 – 1436.

Esfahani,H. S., and Salehi-Isfahani, D (1989), "Effort Observability and Worker Productivity: Towards an Explanation of Economic Dualism." *Economic Journal*, Vol.99, No. 397, pp. 818 – 836.

Fairris, D., and L. J. Alston (1994), "Wages and the Intensity of Labor Effort: Efficiency Wages versus Compensating Payments." *Southern Economic Journal,* Vol. 61, pp. 149 – 160.

Gera, S., and G. Grenier (1994), "Interindustry Wage Differentials and Efficiency Wages: Some Canadian Evidence." *The Canadian Journal of Economics,* Vol. 27, No. 1, pp. 81 – 100.

Groshen, E. L. (1988), "Why Do Wages Vary among Employers ?" Economic Review, pp. 19 – 38.

Groshen, E. L. (1991), "Sources of Intra-Industry Wage Dispersion: How Much Do Employers Matter ?" *Quarterly Journal of Economics,* Vol. 106, No. 3, pp. 869 – 884.

Krueger, A. B., and L. H. Summers (1988), "Efficiency Wages and Inter-Industry Wage Structure." *Econometrica,* Vol. 56, No. 2, pp. 259 – 293.

Lang, K. (1991), "Persistent Wage Dispersion and Involuntary Unemployment." *Quarterly Journal of Economics,* Vol. 106, No. 1, pp. 181 – 202.

Macleod, W. B., and J. M. Malcomson (1993), "Wage Premiums and Profit Maximization in Efficiency Wage Models." *European Economic Review,* Vol. 37, No. 6, pp. 1223 – 1249.

Macleod, W. B., and J. M. Malcomson (1998), "Motivation and Markets." *American Economic Review,* Vol. 88, No. 3, pp. 388 – 411.

Montgomery, J, D. (1991), “Equilibrium Wage Dispersion and Interindustry Wage Differentials.” *Quarterly Journal of Economics,* Vol. 106, No. 1, pp. 163 – 179.

Mortensen, D. T. (2003), *Wage Dispersion: Why are Similar Workers Paid Differently ?*: The MIT Press.

Mortensen, D. T., and C. A. Pissarides (1999), “New Developments in Models of Search in the Labor Market.” In O. Ashenfelter and D. Card, eds. *Handbook of Labor Economics,* Vol. 3B, pp. 2567 – 2624. Amsterdam: Elsevier Science.

Neal, D. (1993), “Supervision and Wages Across Industries,” *The Review of Economics and Statistics,* Vol. 75, No. 3, pp. 409 – 417.

Oi, W. Y., and T. L. Idson (1999), “Firm Size and Wages,” In O. Ashenfelter and D. Card, eds. *Handbook of Labor Economics*, Vol. 3B, pp. 2165 – 2214. Amsterdam: Elsevier Science.

Postel-Vinay, F., and Jean-Marc. Robin (2002), “Equilibrium Wage Dispersion with Worker and Employer Heterogeneity,” *Econometrica*, Vol. 70, No. 6, pp. 2295 – 2350.

Ramaswamy, R., and R. E. Rowthorn (1991), “Efficiency Wages and Wage Dispersion.” *Economica,* Vol. 58, No. 2 pp. 501 – 514.

Ritter, J. A., and J. T. Lowell (1994), “Workers as Creditors: Performance Bonds and Efficiency Wages.” *American Economic Review,* Vol. 84, No. 3, pp. 694 – 704.

Shapiro, C., and J. E. Stiglitz (1984), “Equilibrium Unemployment as a Worker Discipline Device.” *American Economic Review*, Vol. 74, No. 4, pp. 433 – 444.

Troske, K. (1997), “Evidence on the Employer Size-Wage Premium From Worker-Establ ishment Matched Data.” Working Papers from U.S. Census Bureau, Center for Economic Studies.

Tudela, C. C. (2004), “Recruitment Policy When Firms Observe Workers’ Employment Status: An Equilibrium Search Approach.” No 584, Economics Discussion Papers from University of Essex, Department of Economics.

Van den Berg, G. J., and Ridder, G (1998), “An Empirical Equilibrium Search Model of the Labor Market.” *Econometrica*, Vol.66, No. 5, pp. 1183 – 1222.

Walsh, C. F. (1999), “A Multisector Model of Efficiency Wages”, *Journal of Labor Economics*, Vol. 17, No. 2, pp. 351 – 376.