

RESIDUAL WAGE DISPERSION WITH EFFICIENCY WAGES

BY SUPHANIT PIYAPROMDEE*

December 13, 2016

Abstract

This paper extends a classic on-the-job search model of homogeneous workers and firms by introducing a shirking problem. Workers choose their effort levels and search on the job. Firms elicit effort through wages and monitoring; an inverse relationship between wages and monitoring rates is derived. Wages play a dual role by allocating labor supply and motivating employee effort. This gives rise to an equilibrium wage distribution that contrasts with existing literature in several aspects. In particular I show that a hump-shaped and positively skewed wage distribution, as observed empirically, can be derived even when firms and workers are respectively identical.

Keywords: Labor market frictions, wage dispersion, efficiency wages.

JEL codes: D83, J31

1. INTRODUCTION

It is well known that a large proportion of inter-industry and cross-employer wage differentials cannot be explained by observed differences in worker and firm characteristics (e.g. [Katz and Autor, 1999](#); [Idson and Feaster, 1990](#)). Moreover, aggregate earnings distributions are hump-shaped and right skewed. These qualitative properties remain true in the aggregate and within specific occupations, as depicted in Figure (1).

The theoretical literature has rationalized these empirical facts using ex-ante heterogeneity and market frictions. In particular, [Burdett and Mortensen \(1998\)](#), hereafter BM, show that wage dispersion can be obtained as an equilibrium outcome in a search framework even when workers and firms are homogeneous. However, in the homogeneous case, the densities of wages offered

*Department of Economics, University College London. Email: s.piyapromdee@ucl.ac.uk. I am grateful for invaluable guidance from John Kennan and Rasmus Lentz. I thank the editor Guido Menzio and three anonymous referees for their helpful reports. I also thank Andrew Anderson, Michael Choi, Tim Huegerich, Chenyan Lu, Kayuna Nakajima, Mariann Ollar, Fabien Postel-Vinay, Jean-Marc Robin, Nicolas Roys, Haomin Wang and Randy Wright as well as seminar participants at UW-Madison and Chicago Fed Search and Matching workshop for insightful comments.

and wages paid are monotonically rising. Therefore ex-ante heterogeneity across firms is required to match the observed shape of cross-firm wage distributions.

This paper introduces efficiency wages into the BM framework. The difference between this work and BM lies in the interplay between search frictions and the inability of firms to costlessly observe workers' effort. The novelty of this paper, in contrast to BM, is that under certain conditions a unique and hump-shaped wage distribution can be derived even when firms and workers are respectively identical. Moreover, in contrast to [Shapiro and Stiglitz \(1984\)](#) where identical firms offer the same wage and undertake a common and exogenous monitoring rate, firms in this present model post different wages and endogenously choose different monitoring rates.

The intuition behind the hump-shaped wage distribution is simple. In the original BM, where each worker yields the same marginal productivity, there is only one tradeoff between the wage and the labor supply faced by the firm. A high wage firm earns a smaller profit per worker, but attracts a larger labor force. In equilibrium, all firms make the same profit and the wage offer density has to be such that there is no profitable deviation. Consider a wage cut deviation. A high wage firm is more tempted to reduce its wage since this confers a larger cost reduction due to its labor force size. Since high wage firms have more incentive to reduce wages— and the only tradeoff is losing workers to other firms— these firms have to be more threatened by competition just below on the wage offer distribution. Therefore, the single tradeoff between wage and number of workers gives rise to the monotonically increasing wage offer density.

However, if wage reduction is associated with an additional tradeoff that is sufficiently strong, then the wage offer density need not be rising throughout its support. To illustrate this mechanism, I extend the homogeneous version of BM by allowing a firm's output to depend on its workers' effort. The firm cannot costlessly observe a worker's output. Unemployed workers randomly search for jobs while employed workers search for higher paying jobs and choose whether to exert effort. Each firm posts a wage and chooses a monitoring rate. The worker is fired if caught shirking. Using the standard efficiency wage argument, the inverse relationship between wage and monitoring rate is derived (e.g. [Shapiro and Stiglitz, 1984](#)). With the presence of on-the-job search, identical firms optimally choose different wages, and consequently undertake different monitoring rates. In equilibrium, all firms earn the same profit. A wage reduction in this model leads to smaller firm size as well as a higher monitoring cost per worker. Under the condition that a high wage firm is sufficiently threatened by a higher monitoring cost when lowering its wage, the threat of losing workers to other firms can be weaker. Thus the degree of firm competition, and hence the wage offer and equilibrium wage densities, can be declining over the upper range of the wage support.¹

[Mortensen \(2000\)](#) extends the BM framework by introducing match-specific investment and shows that the wage offer density can be unimodal when firms and workers are ex-ante identical. Using a directed search framework, several authors derive positively skewed wage distri-

¹The arguments here can be applied to other tradeoffs associated with wage reduction, such as job amenities. From the compensating differential point of view ([Rosen, 1974](#)), a firm with a clean and safe workplace may attract more workers even at lower wages. Thus it is possible to have identical firms offering different combinations of wages and amenities. Applying the arguments in this paper, an empirically plausible wage distribution can be derived using the tradeoff between wages and job amenities. This contrasts with [Hwang et al. \(1998\)](#) who introduce amenities into the [Burdett and Mortensen \(1998\)](#) framework. In their model, ex-ante identical firms choose the same amenity level since the demand for an amenity is independent of wage. Hence the resulting wage offer density is monotonically rising.

butions without assuming heterogeneity, for instance [Camera and Selcuk \(2004\)](#). As discussed in [Mortensen \(2000\)](#), even a wage equation residual term with an increasing density would be a departure from reality; thus an empirically consistent model should be able to deliver wage distributions with long right tails even in absence of firm and worker heterogeneity. One contribution of this paper is to provide an alternative mechanism to resolve the anomalous feature of [Burdett and Mortensen \(1998\)](#) using efficiency wages. The present framework is relevant for understanding wage dispersion, among equally productive firms, in many labor markets since both monitoring and job-to-job transitions are ubiquitous ([Neal, 1993](#); [Nagypál, 2008](#)). Further, a number of studies, e.g. [Brunello \(1995\)](#), find empirical support for the negative relationship between wages and monitoring intensity which is a key mechanism in this framework.²

This paper provides a note of caution for empirical analysis. Building on the BM framework, a number of empirical search papers appeal to heterogeneity in productivity to improve model fit (see for example, [Bontemps et al., 2000](#)). The identification of firm productivity is usually dependent on the one-to-one relationship between the productivity and wage densities. This approach is based on the BM implication that firm heterogeneity necessitates the shape of the empirically observed wage distributions. However, this paper shows that an empirically plausible wage distribution can be derived even in absence of heterogeneity.

Additionally, this paper contributes to the efficiency wage literature by endogenizing monitoring rates. Existing studies often assume firms to undertake a common and exogenous monitoring rate as in [Shapiro and Stiglitz \(1984\)](#), for instance [Strand \(1987\)](#); [Albrecht and Vroman \(1998\)](#); [Walsh \(1999\)](#). However, using the ratio of supervisors to workers as a proxy for monitoring intensity, empirical studies (e.g. [Brunello, 1995](#); [Neal, 1993](#)) find that these ratios vary across firms. Thus it is appealing to allow firms to choose different monitoring rates. [Albrecht and Vroman \(1998\)](#) combine a shirking problem with a wage posting random search model where the monitoring rate is exogenous and there is no on-the-job search. Under their assumption that workers are heterogeneous in the disutility of effort, shirking persists in equilibrium and adverse selection precludes the existence of a single wage outcome. Further, they provide a numerical example where a uniform distribution of effort costs can give rise to a downward sloping wage distribution. In my model, on the job search is what leads homogeneous firms to post different wages and consequently choose different monitoring rates. The tradeoff between wages and monitoring rates is what generates a declining wage offer density.

The rest of the paper is organized as follows. Section 2 describes the model and presents the problem of a worker taking the distribution of wages as given. Section 3 endogenizes the wage distribution and characterizes the equilibrium solution. Section 4 discusses the model's implications. Section 5 provides a numerical example and Section 6 concludes.

²See [Walsh \(1999\)](#) for a review of these studies.

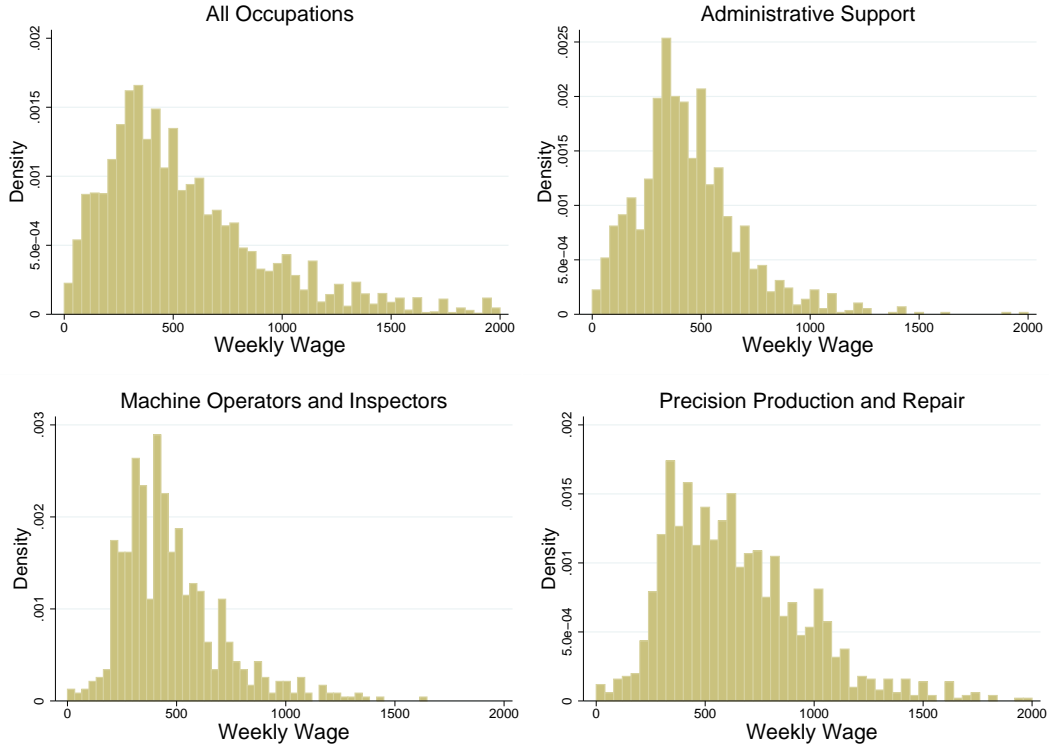


Figure 1: Wage Distributions from Current Population Survey, Year 2000

2. THE MODEL

The market consists of a continuum of infinite-lived, risk-neutral firms and workers. The measure of workers is indicated by N whereas the measure of firms is normalized to one, so that N is the ratio of workers to firms. All workers and firms are respectively identical and face the same discount rate r .

The search framework is similar to that of BM (Burdett and Mortensen, 1998). This model deviates from BM by allowing the output of a match to depend on a worker's effort level: $\{0, e\}$ where $e > 0$. For simplicity, it is assumed that each productive worker contributes one unit of output and zero otherwise.³ The output of an individual worker cannot be costlessly observed. Firms detect shirking by monitoring their workers at some cost. Each firm posts a wage w and a monitoring rate q .⁴ The firm pays w to all of its employees and hires any worker who is willing to accept w . For tractability, I maintain the assumption that firms post constant wages as in BM rather than setting wages as a function of the output history.⁵ It will be shown that the simple deviation

³This assumption rules out an equilibrium where some firms find it optimal to hire shirkers. If the output associated with the two effort levels were strictly positive, then using a similar argument as Albrecht and Axell (1984), it is possible for firms to offer wages for both effort levels.

⁴As will become clear later, the firm's choice of q is implied by its choice of w .

⁵Under the assumption that the output of an individual worker cannot be costlessly observed, an output contingent wage would still be tied to a monitoring problem since the firm has to first decide how often to check the output level

from BM considered here is sufficient to generate empirically observed wage distributions. Further, as discussed in [Board and Meyer-ter Vehn \(2015\)](#), non-stationary wage contracts may harm the worker without benefiting the firm.⁶

Workers are assumed to randomly search among firms; let $F(w)$ denote the distribution of wage offers from which workers sample. Consider the problem of a worker taking monitoring rates and the distribution of wages as given. At a moment in time, each worker is either employed or unemployed and searching. While unemployed, workers receive utility flow b per instant and receive job offers at rate λ_0 , each of which is a random draw from F . While employed, workers receive their wage w and choose whether to exert effort. Employed workers who exert an effort e obtain an instantaneous utility flow: $w - e$.⁷ Shirkers do not incur the effort cost, but at rate q they will be caught and lose their jobs. In both cases, employed workers receive new job offers from F at rate λ_1 and lose their jobs exogenously at rate δ . Workers cannot recall previous offers and must respond to offers as soon as they arrive.

Let us define V_s, V_e and V_u respectively as the expected discounted lifetime income for an employed shirker, an employed nonshirker, and an unemployed individual. Following [Shapiro and Stiglitz \(1984\)](#), the worker chooses an effort level to maximize his expected discounted lifetime income by comparing the utility from shirking with the utility from working. The no-shirking condition (NSC) under which the worker chooses to exert effort is $V_e \geq V_s$ where shirking is more attractive when the firm monitors less frequently. In equilibrium, the NSC always binds, $V_e = V_s$ otherwise the firm could reduce q and increase its profit without affecting the worker's choice of effort.

Using standard arguments, the expected discounted lifetime income for an unemployed worker is given by

$$(1) \quad rV_u = b + \lambda_0 \left[\int \max \{V_u, V_e(x)\} dF(x) - V_u \right],$$

where acceptance of a job occurs when the value of employment exceeds the value of continued searching.⁸ Since all workers exert effort in equilibrium, there is no firing for shirking and employed workers move from lower to higher paying jobs when such offers arrive. Nevertheless, the expected discounted lifetime income for a shirker will be important in deriving the NSC. This can be written as

$$(2) \quad rV_s(w) = w + \lambda_1 \left[\int \max \{V_s(w), V_e(x)\} dF(x) - V_s(w) \right] + (\delta + q) [V_u - V_s(w)],$$

of each worker.

⁶[Board and Meyer-ter Vehn \(2015\)](#) studies a relational contract problem in a frictionless market where a firm chooses a sequence of wage and effort levels to maximize profit. Each firm has one worker and can perfectly observe the worker's effort. In equilibrium, identical firms are indifferent between a continuum of contracts. In their framework, on-the-job search only acts as a non-pecuniary job benefit as it allows employed workers to receive better offers, whereas in my paper on-the-job search also affects the firm's labor size.

⁷In the case of continuous effort levels, the qualitative results remain the same if the elasticity of effort disutility with respect to effort is increasing, see [Walsh \(1999\)](#) for further details.

⁸Generally, the integrand in the asset value of unemployment is $\max \{V_u, V_e(x), V_s(x)\}$. However, in equilibrium $V_e = V_s$, hence $V_s(x)$ is dropped in (1) as well as in the asset values of shirking and working for ease of exposition.

while for an employed nonshirker, the expected discounted lifetime income is

$$(3) \quad rV_e(w) = w - e + \lambda_1 \left[\int \max \{V_e(w), V_e(x)\} dF(x) - V_e(w) \right] + \delta [V_u - V_e(w)].$$

The difference between (2) and (3) is that the shirker obtains a higher instantaneous utility flow since there is no cost of effort, but faces a higher rate of separation.⁹

As mentioned earlier, in equilibrium the NSC must bind otherwise there is a profitable deviation for the firm. Using (2) and (3), the equality of $V_e(w)$ and $V_s(w)$ implies that

$$(4) \quad V_e(w) = \frac{e}{q} + V_u.$$

See Appendix A.1 for a more detailed derivation. Eq (4) shows that at a no-shirking wage where $V_e(w) = V_s(w)$, there is a rent associated with moving from unemployment to employment when monitoring is imperfect, i.e. $q < \infty$. The rent associated with a job is increasing in the effort level and decreasing in the monitoring intensity. As a lemma, we have

Lemma 1. *For $q < \infty$, any no-shirking wage gives the worker a surplus. The worker strictly prefers accepting a no-shirking wage offer rather than being unemployed.*

Let \underline{w} and \bar{w} denote the infimum and supremum of the support of F which will be endogenized in the next section. Setting $V_e(w) = V_s(w)$, and using the fact that

$$(5) \quad E_x \max \{V_i(w), V_i(x)\} = V_i(w) + \int_{\underline{w}}^{\bar{w}} V_i'(x) \bar{F}(x) dx,$$

where $i \in \{e, s\}$ and $\bar{F}(x) = 1 - F(x)$, then substituting $V_e'(x) = 1 / [r + \delta + \lambda_1 \bar{F}(x)]$, an equilibrium no-shirking wage can be expressed as

$$(6) \quad w = (r + \delta) \frac{e}{q} + e - \lambda_1 (B(\bar{w}) - B(w)) + rV_u$$

where

$$B(w) = \int_{\underline{w}}^w \frac{\bar{F}(x)}{\delta + \lambda_1 \bar{F}(x)} dx.$$

It can be shown that wages and monitoring rates are inversely related. Differentiating (6) with respect to q gives

$$(7) \quad \frac{dw}{dq} = -\frac{e}{q^2} (r + \delta + \lambda_1 \bar{F}(w)).$$

From the firm's point of view, (7) determines the substitution rate between wages and monitoring rates.¹⁰ The firm can motivate the worker to exert effort by either monitoring more intensely or setting a higher wage to increase the cost of losing the job. Thus firms offering lower wages choose higher monitoring rates. Let \bar{q} denote the highest monitoring rate, the rate associated with the lowest wage offered \underline{w} . Substituting (1) into (6), we have the following lemma

⁹Note that the firm's choice of q is implied by its choice of w , and so V_s only depends on w .

¹⁰It will be shown that in equilibrium F has no mass points and so $\frac{dw}{dq}$ is well defined.

Lemma 2. *In equilibrium, a no-shirking wage satisfies (8):*

$$(8) \quad w = \frac{(r + \delta)e}{q} + \frac{\lambda_0 e}{\bar{q}} + e + b + \lambda_0 B(\bar{w}) - \lambda_1 (B(\bar{w}) - B(w)).$$

See Appendix A.1 for a more detailed derivation of (8). Thus a no-shirking wage consists of a rent due to imperfect monitoring $(r + \delta)e/q + \lambda_0 e/\bar{q}$, the cost of effort e , the unemployment benefit b , and the difference between expected gains from searching while unemployed and employed which is represented by the last two terms. Taking F as given, (8) yields predictions, in line with [Shapiro and Stiglitz \(1984\)](#), that the firm has to pay a higher wage to persuade the worker to exert effort when (i) the disutility of effort e is higher; (ii) the unemployment benefit b is larger; (iii) the job destruction rate δ is higher; (iv) the discount rate r is higher; or (v) the offer arrival rate for unemployed workers λ_0 is higher. Additionally, the firm can reduce its wage when the offer arrival rate for employed workers λ_1 is higher. The relationship between λ_1 and the no-shirking wage deserves a comment. Given that workers move from lower to higher paying jobs, a larger value of λ_1 increases the asset value of employment. This in turn increases the cost of losing a job if caught shirking which allows the no-shirking wage to be lower.

Finally, it is convenient to define an inverse function that maps w to the q implied by the NSC, so that the firm's choice of q is given implicitly by its choice of w . Rearranging (8), the inverse function is given by

$$(9) \quad q = \phi(w) \\ := \frac{e(r + \delta)}{w - \lambda_0 \frac{e}{\bar{q}} - e - b - \lambda_0 B(\bar{w}) + \lambda_1 (B(\bar{w}) - B(w))}.$$

Conditional on F , ϕ is a one-to-one mapping since q is strictly decreasing in w as shown in (7).

Before solving for an equilibrium in the next section, let us define the distribution of wages paid to employed workers, denoted by $G(w)$. Let u denote the steady state number of unemployed workers. Since Lemma 1 implies that all no-shirking wages are acceptable by unemployed workers when monitoring is imperfect, the flow of workers into employment is given by $\lambda_0 u$ whereas the flow of workers into unemployment is $\delta(N - u)$. Setting these equal, one finds that in the steady state

$$(10) \quad u = \frac{\delta N}{\delta + \lambda_0}.$$

Then equating the flow of unemployed workers into firms with a wage no greater than w , $\lambda_0 F(w)u$, and the exodus of workers from employment at wages less than w , $[\delta + \lambda_1 \bar{F}(w)]G(w)(N - u)$ gives

$$(11) \quad G(w) = \frac{\delta F(w)}{\delta + \lambda_1 \bar{F}(w)}$$

where $G(w)$ coincides with the distribution of wages offered when there is no on-the-job search, i.e. $\lambda_1 = 0$.

3. STEADY STATE EQUILIBRIUM

In this section, I endogenize the wage offer distribution F . It is assumed that all firms receive the same flow of revenue generated per employed nonshirker p . Each firm faces a common per-worker monitoring cost $c(q)$ where $c(q)$ is strictly increasing, twice continuously differentiable and $\lim_{q \rightarrow \infty} c(q) \rightarrow \infty$.

Conditional on F , the firm posts a wage and a monitoring rate at which it commits to hire all the workers it contacts, subject to the NSC: $q = \phi(w)$.¹¹ The flows of employment into the firm consist of workers leaving unemployment and leaving lower paid jobs. The flows out of the firm consist of workers leaving exogenously and moving to higher paid jobs. Let $\ell(w|F)$ represent the steady state number of workers at a firm paying w . For simplicity, I assume that firms maximize steady-state profit. This can be understood as the limiting case of maximizing the present value of the profit flow when $r \approx 0$ (see Coles, 2001 and Burdett and Mortensen, 1998). Henceforth, it is assumed that $r \approx 0$. An optimal wage offer solves the following problem

$$(12) \quad \pi = \max_w (p - w - c(\phi(w))) \ell(w|F).$$

Further, I assume that each firm faces an entry cost. So the total profit π has to be strictly positive for firms to amortize their entry cost.¹² In equilibrium, all firms make the same maximal profit. Thus the following holds

$$(13) \quad \begin{aligned} \pi &= (p - w - c(\phi(w))) \ell(w|F) \quad \text{for all } w \in \text{supp}(F) \\ \pi &\geq (p - w - c(\phi(w))) \ell(w|F) \quad \text{otherwise.} \end{aligned}$$

Given that a higher wage offer decreases the flow of workers leaving for higher paid jobs and increases the acceptance rate by workers, a firm's labor size $\ell(w|F)$ is increasing on the support of F . Using the same argument as in BM, the following lemma rules out wage offer distributions with mass points.

Lemma 3. *F has no mass points.*

Proof. Suppose there exists a mass point at w' where $V_e(w') > V_u$ and the cost per worker is such that $w' + c(\phi(w')) < p$. Then any firm offering w' could earn a strictly greater profit by offering $w' + \varepsilon$ for some $\varepsilon > 0$. This is because such a firm attracts a significantly larger labor force since, given the mass point, it meets workers earning w' with positive probability. At the same time, a reduction in the profit per worker (if the wage increment exceeds the additional savings on monitoring) can be made arbitrarily small by taking $\varepsilon \rightarrow 0$. Hence there cannot be a mass point at w' where $w' + c(\phi(w')) < p$. Further, there can be no mass point at w where $w + c(\phi(w)) \geq p$ since no firm offers such a wage as it would imply a nonpositive profit. \square

¹¹The implicit assumption here is that wages and monitoring rates are written into a contract.

¹²The entry cost does not affect the equilibrium derived and therefore its notation is left out of the analysis. As discussed in Burdett and Mortensen (1998), one can impose free entry and endogenize the ratio of workers to firms N so that the profit net of entry cost is zero.

Next let us derive steady state firm size. The steady state number of workers earning a wage in the interval $(w - \varepsilon, w)$ is represented by $(G(w) - G(w - \varepsilon))(N - u)$ whereas the steady state number of firms offering a wage in this interval is $F(w) - F(w - \varepsilon)$. Thus, the number of workers at a firm paying w is given by

$$(14) \quad \begin{aligned} \ell(w|F) &= \lim_{\varepsilon \rightarrow 0} \frac{G(w) - G(w - \varepsilon)}{F(w) - F(w - \varepsilon)} (N - u) \\ &= \frac{\lambda_0 \delta N (\delta + \lambda_1)}{(\delta + \lambda_0) (\delta + \lambda_1 \bar{F}(w))^2}. \end{aligned}$$

To derive an equilibrium, I first determine the infimum of the wage support. Since there is no mass point, $F(\underline{w}) = 0$, (14) implies that

$$(15) \quad \ell(\underline{w}|F) = \frac{\lambda_0 \delta N}{(\delta + \lambda_0) (\delta + \lambda_1)}.$$

The number of workers at the lowest wage firm is independent of the wage offered. In BM, the lowest wage firm maximizes profit by setting \underline{w} at a point where the worker is just indifferent between working and being unemployed $V_e(\underline{w}) = V_u$. Any wage below \underline{w} would raise the profit per worker but attract no workers. In the current framework, however, for the worker to exert effort the continuation value associated with any offer has to be strictly greater than the continuation value of unemployment, $V_e(\underline{w}) > V_u$. So the firm, in this case, could reduce its wage below \underline{w} and still attract workers. However, to prevent workers from shirking the firm would also have to increase monitoring. This implies that for the lowest wage to be optimal, it must be that the gain from lowering the wage and the cost of increased monitoring are exactly offsetting. Consequently, the derivative of $\ell(w|F)$ and the wage offer density are zero at \underline{w} .

Let $f(w)$ denote the wage offer density. To see why f is zero at \underline{w} , differentiating (14) with respect to w and evaluating at \underline{w} gives

$$(16) \quad \ell'(w|F) |_{w=\underline{w}} = \frac{2\lambda_0 \delta N (\delta + \lambda_1) \lambda_1 f(\underline{w})}{(\delta + \lambda_0) (\delta + \lambda_1)^3},$$

The derivative of $\ell(w|F)$ is equal to zero if and only if the wage offer density at the lowest wage $f(\underline{w})$ is zero. It can be shown later in this section that if the gain from a wage reduction equals the cost of increased monitoring, i.e. $\frac{dc}{dq} = -\frac{dw}{dq}$, then $f(\underline{w}) = 0$.¹³ Given that $\ell'(w|F) = 0$ at \underline{w} , the lowest wage firm's problem is to choose a combination of \underline{w} and \bar{q} that induces the worker to exert effort as well as maximizing the profit per worker.¹⁴ It is more straightforward to solve this by first pinning down the profit maximizing \bar{q} , then given \bar{q} the lowest no-shirking wage is implied by (8).

Differentiating the lowest wage firm's profit per worker with respect to q and setting it to zero gives

¹³The expression of $f(w)$ is derived in (23).

¹⁴The restriction on the revenue per worker such that profit is strictly positive is derived in Appendix A.2.

$$(17) \quad \begin{aligned} \frac{dc}{dq} &= -\frac{dw}{dq} \\ &= \frac{e(\delta + \lambda_1 \bar{F}(w))}{q^2} \end{aligned}$$

where the LHS is the marginal monitoring cost and the RHS is the marginal benefit of monitoring. Using the fact that $F(\underline{w}) = 0$, the first order condition (FOC) determining \bar{q} in (17) can be written as

$$(18) \quad c'(\bar{q})\bar{q}^2 = e(\delta + \lambda_1).$$

The following assumption ensures that the solution to (18) is a unique maximizer.

Assumption 1. (i) $c'(q)q^2$ is strictly increasing, (ii) $\lim_{q \rightarrow 0} c'(q)q^2 < e(\delta + \lambda_1)$, (iii) $\lim_{q \rightarrow \infty} c'(q)q^2 \rightarrow \infty$.

Assumption 1.2 implies that the marginal benefit of monitoring exceeds the marginal cost on the very first unit of monitoring rate. This together with assumptions 1.1 and 1.3 imply that the marginal cost of monitoring is initially smaller than the marginal benefit and becomes larger as q rises. If the marginal cost was always lower than the marginal benefit then no firms would monitor. Therefore, assumptions 1.1-1.3 ensure that the solution \bar{q} is a unique maximizer and strictly positive. It is worth noting that assumption 1.1 is equivalent to the second order optimality condition.¹⁵ Further, assumption 1 is a weak requirement, for example, a power cost function $c(q) = q^s$ where $s > 0$ trivially satisfies these conditions. As a lemma, we have

Lemma 4. Under assumption 1, the optimal \bar{q} is the unique solution to (18). Given the profit maximizing \bar{q} , the lowest no-shirking wage implied by (8) is

$$(19) \quad \underline{w} = \frac{e(\delta + \lambda_0)}{\bar{q}} + e + b + (\lambda_0 - \lambda_1)B(\bar{w}).$$

Any no-shirking wage below \underline{w} leads to an increase in the monitoring cost such that the profit per worker would be reduced. In the special case where $\lambda_0 = \lambda_1$, \underline{w} only depends on a rent due to imperfect monitoring $e(\delta + \lambda_0)/\bar{q}$, the cost of effort e and the unemployment benefit b .

Definition 1. Given the model primitives $\{e, b, \lambda_0, \lambda_1, \delta, N, p, c(\cdot)\}$, the triple (F, ϕ, ℓ) defined on an interval $[\underline{w}, \bar{w}]$ is a steady state equilibrium if and only if

- (i) $f(\underline{w}) = 0$
- (ii) Wages and monitoring rates satisfy the no shirking condition $q = \phi(w)$ in (8)
- (iii) Firms maximize profits such that (13) holds and $q(\underline{w})$ satisfies (18)
- (iv) The steady state number of workers per firm ℓ satisfies (14).

¹⁵The second order condition for profit maximizing requires $c''(\bar{q})\bar{q}^3 > -2e(\delta + \lambda_1) \iff c''(\bar{q})\bar{q}/c'(\bar{q}) > -2$. This is equivalent to assuming that $c'(q)q^2$ is strictly increasing since $\frac{d(c'(q)q^2)}{dq} > 0 \iff c''(q)q/c'(q) > -2$.

Before completing the derivation of the equilibrium, it is worth discussing the connectedness of the supports of F . In standard models, for example BM, wage supports with gaps are trivially ruled out since a gap in the firms' wage strategy cannot be optimal. The argument is based on the fact that wages in those models play a single role: allocating labor supply. If there were a gap between w' and w'' where $w' < w''$, then any firm paying w'' can make a strictly greater profit by paying $w'' - \varepsilon$ for some $\varepsilon > 0$. This is because such a firm loses no more workers than it did previously while the profit per worker increases; hence there cannot be a gap. In the present model, however, a wage cut does not necessarily increase the profit per worker. Since wages also provide incentives for employee effort, a wage cut imposes a higher monitoring cost. The profit per worker could decrease if the penalty of increased monitoring cost exceeds the wage reduction in which case the wage strategy with a gap would be optimal.

Nonetheless, it is appealing to restrict attention to wage distributions with connected supports for two reasons: First, this is a more interesting case in terms of matching actual wage distributions. Second, qualitative features of the equilibrium are more transparent. The following proposition provides sufficient conditions under which F has a connected support.

Proposition 1. *If $c'(q)q^2$ is strictly increasing, then F has a connected support.*

Proof. Suppose w' and w'' are both part of an optimal wage strategy and there is a gap between w' and w'' where $\underline{w} < w' < w'' < \bar{w}$, but $c'(q)q^2$ is strictly increasing. Let $\psi(w) = -1 - c'(\phi(w))\phi'(w)$ denote the change in profit per worker with respect to wage. Then $\psi(w) = 0$ for $\forall w \in [w', w'']$ otherwise there would be a profitable deviation. Using (7), $\psi(w)$ can be expressed as $-1 + c'(q)q^2/e(\delta + \lambda_1\bar{F}(w))$. So $\psi(w) = 0$ means that $c'(q)q^2 = e(\delta + \lambda_1\bar{F}(w))$ for $\forall w \in [w', w'']$. Since F has a gap, $e(\delta + \lambda_1\bar{F}(w))$ would be constant for $\forall w \in [w', w'']$, but this implies that $c'(q)q^2$ would also have to be constant. This contradicts the assumption that $c'(q)q^2$ is strictly increasing. Therefore if $c'(q)q^2$ is strictly increasing, then F must have a connected support. \square

The intuition behind proposition 1 is that if $c'(q)q^2$ is strictly increasing, i.e. $c(q)$ is not too concave, then the penalty of increased monitoring cost would not exceed the wage reduction. Therefore, a wage cut would lead to a higher profit per worker, and hence F cannot have a gap. In what follows, I characterize wage offer distributions with connected supports. The connectedness of F implies that $\ell(w|F)$ is strictly increasing in w on the support of F , and hence the profit per worker must be strictly decreasing for all firms to make the same profit, i.e. $0 > c'(\phi(w))\phi'(w) > -1$ for any $w \in (\underline{w}, \bar{w})$. In any equilibrium, the indifference condition must hold such that every offer yields the same steady state profit. This means that

$$(20) \quad \pi = \frac{(p - \underline{w} - c(\phi(\underline{w})))\lambda_0\delta N}{(\delta + \lambda_0)(\delta + \lambda_1)} = (p - w - c(\phi(w)))\ell(w|F)$$

for all w on the support of F .¹⁶ Substituting (14) into (20), a candidate for $F(w)$ is

$$(21) \quad F(w) = \left(\frac{\delta + \lambda_1}{\lambda_1} \right) \left(1 - \left(\frac{(p - w - c(\phi(w)))}{(p - \underline{w} - c(\phi(\underline{w})))} \right)^{1/2} \right).$$

¹⁶The equilibrium converges to BM as the effort cost goes to zero.

Using the fact that $F(\bar{w}) = 1$, (21) yields

$$(22) \quad p - \bar{w} - c(\phi(\bar{w})) = \frac{(p - \underline{w} - c(\phi(\underline{w}))) \delta^2}{(\delta + \lambda_1)^2}.$$

Eq (22) implies that the support of the equilibrium candidate for F is nondegenerate when firms make a strictly positive profit, i.e. $p - \underline{w} - c(\phi(\underline{w})) > 0$. Profit will be positive when the revenue per worker exceeds the sum of the effort cost, the unemployment benefit, the rent associated with the NSC and the cost due to monitoring. It can be shown that if $p > e + b + e(\delta + \lambda_0)/\bar{q} + (\lambda_0 + \lambda_1)c(\bar{q})/\lambda_1$ where \bar{q} is a function of the primitives given in (18), then $p - \underline{w} - c(\phi(\underline{w})) > 0$ (see the proof in Appendix A.2.).

Since the wage offer distribution depends on a function of its integral in $\phi(w)$, (21) can be written as a system of differential equations. Differentiating (21) with respect to w gives

$$(23) \quad F'(w) = \left(\frac{\delta + \lambda_1}{2\lambda_1} \right) \left(\frac{(p - w - c(\phi(w)))^{-1/2}}{(p - \underline{w} - c(\phi(\underline{w})))^{1/2}} \right) (1 + c'(\phi)\phi'(w)),$$

where

$$\phi(w) = \frac{e\delta}{w - \lambda_0 \frac{e}{\bar{q}} - e - b - \lambda_0 B(\underline{w}) + \lambda_1 (B(\bar{w}) - B(w))},$$

and

$$(24) \quad \begin{aligned} \phi'(w) &= 1 / \frac{dw}{d\phi} \\ &= \frac{-\phi(w)^2}{e(\delta + \lambda_1 \bar{F}(w))}. \end{aligned}$$

So (23) can be expressed as,

$$(25) \quad \begin{aligned} F'(w) &= \left(\frac{\delta + \lambda}{2\lambda} \right) \left(\frac{(p - w - c(\phi(w)))^{-1/2}}{(p - \underline{w} - c(\phi(\underline{w})))^{1/2}} \right) \left(1 - \frac{c'(\phi)\phi(w)^2}{e(\delta + \lambda \bar{F}(w))} \right), \\ &\equiv \beta(w, F(w), B(w)). \end{aligned}$$

Recall that

$$B(w) = \int_{\underline{w}}^w \frac{\bar{F}(x)}{\delta + \lambda_1 \bar{F}(x)} dx,$$

differentiating $B(w)$ with respect to w gives

$$(26) \quad \begin{aligned} B'(w) &= \frac{\bar{F}(w)}{\delta + \lambda_1 \bar{F}(w)} \\ &\equiv \alpha(w, F(w)). \end{aligned}$$

Therefore, (25) and (26) give us the following system of differential equations

$$(27) \quad \begin{aligned} B'(w) &= \alpha(w, F(w)) \\ F'(w) &= \beta(w, F(w), B(w)) \end{aligned}$$

with the initial conditions $F(\underline{w}) = 0$ and $B(\underline{w}) = 0$. The existence of a unique solution $F(w)$ and $B(w)$ to (27) can be proved using Picard's theorem for systems of differential equations. The following proposition provides sufficient conditions under which an equilibrium exists and is unique.

Proposition 2. *If (i) $p > e + b + \frac{e(\delta + \lambda_0)}{\bar{q}} + \left(\frac{\lambda_0 + \lambda_1}{\lambda_1}\right) c(\bar{q})$ where \bar{q} is given by (18) and*

(ii) $c(q)$ satisfies assumption 1,

then there exists a steady state equilibrium. All steady state equilibria involve $F(w)$ with a connected support and $B(w)$ satisfying (27) for some \underline{w} . Further, if $\lambda_0 = \lambda_1$ then the steady state equilibrium is globally unique.

The proof is in Appendix A.3.

Proposition 2 shows that for a given value of \underline{w} , there exists a unique local solution $F(w)$ and $B(w)$ to the system of differential equations (27); however, the solution can be extended to be globally unique on the relevant domain using boundedness of $F(w)$ and $B(w)$ (see proposition 5 in Appendix A.3. for more details). In the special case where $\lambda_0 = \lambda_1$, \underline{w} only depends on the exogenous parameters; the initial conditions $F(\underline{w}) = 0$ and $B(\underline{w}) = 0$ in this case are uniquely pinned down. Consequently, the solution $F(w)$ and $B(w)$ is globally unique. In the case where $\lambda_0 \neq \lambda_1$, then \underline{w} also depends on $F(w)$, and there may be more than one pair of $\{\underline{w}, F(w)\}$ that solves (27); thus the solution need not be unique. However, from an empirical point of view this is not a problem since the value of \underline{w} is observed.¹⁷

Finally, differentiating (21) twice, the slope of the wage offer density $f(w)$ is given by

$$(28) \quad f'(w) = \frac{\delta + \lambda_1}{2\lambda_1 (p - \underline{w} - c(\phi(\underline{w})))^{1/2}} \left(\frac{(c''(\phi)(\phi'(w))^2 + c'(\phi)\phi''(w))}{(p - w - c(\phi(w)))^{1/2}} + \frac{(1 + c'(\phi)\phi'(w))^2}{2(p - w - c(\phi(w)))^{3/2}} \right).$$

The novelty of this paper is that the density of wages offered need not be rising throughout its support even when workers and firms are respectively identical. Let $\eta^{c'} = \frac{c''(q)q}{c'(q)}$ denote the elasticity of the cost derivative. The sufficient conditions under which f declines involves an upper bound of $\eta^{c'}$, i.e. the monitoring cost cannot be too convex. The intuition why there has to be an upper bound of $\eta^{c'}$ is the following: in BM, the elasticity of profit per worker is falling as a firm increases its wage. For high wage firms to make the same profit as low wage firms, the number of additional workers poached from lower wage firms has to be increasingly rising to offset a reduction in the profit per worker. This implies that there has to be more firms offering higher wages, and hence the wage offer density is monotonically rising. In the present framework, there is an additional

¹⁷However, the possibility of global non-uniqueness when $\lambda_0 \neq \lambda_1$ may preclude comparative statics.

tradeoff between wages and monitoring rates. As a firm increases its wage, it attracts more workers and at the same time can lower its monitoring rate. If the monitoring cost was very convex and more responsive to high monitoring rates, then the gains from reduced monitoring cost by offering a higher wage would be more significant for relatively low wage firms. This would further increase the difference between the elasticity of profit per worker at high and low wage firms, reinforcing the necessity for the wage offer density to be monotonically rising to equalize profit. I discuss this intuition in more detail in the next section.

The mechanism which gives rise to a declining f in the general case where $\lambda_0 \neq \lambda_1$ shares the same insight as the special case where $\lambda_0 = \lambda_1$, but more algebraically involved. Thus I focus on the special case in the following proposition.

Proposition 3. Let $\eta^{c'} = \frac{c''(q)q}{c'(q)}$ denote the elasticity of $c'(q)$ and $\hat{\pi}$ denote $(p - \underline{w} - c(\bar{q}))$. If (i) $\lambda_0 = \lambda_1$ so that $\hat{\pi}$ is exogenous and

$$(ii) \eta^{c'} < -2 + \frac{e(\delta + \lambda_1)}{2\hat{\pi}q} - \frac{e^2(\delta + \lambda_1)^4}{2\hat{\pi}\delta^2 c'(q)q^3} - \frac{c'(q)q(\delta + \lambda_1)}{2\hat{\pi}\delta} \text{ for } \forall q \leq q^*$$

then the wage offer density satisfying (23) declines over the upper range of the distribution where $q^* = \bar{q} - \frac{(e\delta)^2(2\delta\lambda_1 + \lambda_1)^2\hat{\pi}}{\left(p - \frac{\lambda_0 e}{q} - e - b\right)^2 e(\delta + \lambda_1)^3}$.

The proof is in Appendix A.4.

Proposition 3 provides sufficient conditions under which f declines over the upper range of the distribution; this combined with the fact that $f(\underline{w}) = 0$ implies that there are regions of w where f increases. However, this does not necessarily imply that f is unimodal. The wage offer distribution may have multiple modes. Nonetheless, this should not be a major concern since empirical wage distributions are not always unimodal.

4. DISCUSSION

In contrast to the homogeneous workers and firms version of BM in which the wage offer density is monotonically rising, the wage offer density in this model can be declining. The main difference is that in BM, firms only face one tradeoff between the wage and the labor supply. In their unique equilibrium, all firms make the same profit. A firm's labor size is increasing in wage, and so a high wage firm earns a smaller profit per worker but is compensated by a larger labor force. To understand why the wage offer density in BM is rising throughout its support, let us consider a firm's steady state profit in the BM model,

$$(29) \quad \pi = (p - w)\ell(w|F).$$

where $\ell(w|F)$ has the same expression as in (14). Differentiating (29) with respect to w and setting the derivative equal to zero, one finds that

$$(30) \quad \frac{d\ell}{dw} = \frac{1}{p - w}.$$

Eq (30) shows that for all wages to yield the same steady state profit, the rate of change of the profit per worker (the RHS of (30)) has to be offset by the rate of change of labor size (the LHS) at all wages on the support. The wage offer density responds to the profitability at each wage. Since the profit per worker falls at an increasing rate as w approaches p , the rate of change of the labor size has to be increasing in wage to offset the reduction in the profit per worker. Consequently the wage offer density rises monotonically.

The intuition is that in the equal profit equilibrium, the wage offer density has to be such that no firm has an incentive to deviate. Consider a wage cut deviation. Clearly a high wage firm has more incentive to cut its wage since the total cost reduction is higher given its number of employees. To prevent that, high wage firms have to be more threatened by competition nearby in the sense that lowering wage leads to a significant loss of workers to firms just below on the wage offer distribution. Thus the wage offer density is rising in wage.

In the present model, firms face an additional tradeoff. Given the negative relationship between wages and monitoring rates, a firm faces two threats when cutting its wage: (i) a higher monitoring cost and (ii) losing its workers to other firms. The additional tradeoff allows the density of firm competition to be more flexible. Differentiating the firm's steady state profit in (13) with respect to w and setting the derivative equal to zero, one finds that the change in profit per worker, under this framework, is not necessarily increasing in w . Namely,

$$(31) \quad \frac{d\ell}{dw} = \frac{\left(1 + \frac{dc}{d\phi} \frac{d\phi}{dw}\right)}{p - w - c(\phi(w))}.$$

As shown in Eq (31), if the savings from decreased monitoring, $\left|\frac{dc}{d\phi} \frac{d\phi}{dw}\right|$ are rising in w sufficiently fast so that the rate of change of the profit per worker is declining (the RHS of (31)), then the rate of change of labor size (the LHS of (31)), and hence the wage offer density is falling.

The intuition here is analogous to the insight from BM in that the wage offer density has to be such that no firm has an incentive to deviate. A high wage firm can gain more from a wage reduction given its labor size. But if a high wage firm is more threatened by the higher monitoring cost associated with a wage reduction, then the threat of losing workers to other firms can be weaker. Hence the wage offer and equilibrium wage densities can be declining over the upper range of the wage support. This can happen if (i) the monitoring cost exhibits increasing returns so that a wage cut at a high wage firm leads to a more significant increase in monitoring cost; or (ii) the substitution rate between wages and monitoring rates is decreasing in wage, i.e. a wage reduction at a high wage firm entails a larger increase in the monitoring rate than at a low wage firm.

This is in line with Proposition 3 in that the wage offer density declines if the elasticity of $c'(q)$ is not too high. This is because if the elasticity of $c'(q)$ is too high, i.e. the monitoring cost is very convex and more responsive at high monitoring rates, then a wage cut at a high wage firm would lead to a smaller increase in the monitoring cost relative to a wage cut at a low wage firm. In this case, a high wage firm would be less threatened by the higher monitoring cost associated with a wage cut; therefore to prevent deviation, the threat of losing workers to other firms would have to be even more severe. Consequently, if the elasticity of cost derivative is too high, then the resulting wage offer density would have to be monotonically rising, as in BM, to equalize profit.

Before turning to numerical illustration, let us discuss some assumptions about monitoring technologies. There is little direct evidence on the shape of monitoring costs; however, a commonly assumed technology is that the production of monitoring uses supervisors as an input and exhibits decreasing returns (for example [Greenwood et al., 2010](#)). In reality the detection rate q would depend on the monitoring intensity chosen by the firm which includes a range of factors such as the output production technology, the number of supervisors, and the information technology such as computers and cameras. In particular, [Abowd and Kramarz \(2003\)](#); [Kramarz and Michaud \(2010\)](#) estimate firms' labor adjustment costs to be concave, especially for hires of managers. If the production of monitoring uses supervisors where an increase in monitoring rate involves hiring more supervisors, then the concavity in hiring costs could influence the shape of monitoring costs. Therefore, the monitoring technology need not exhibit decreasing returns when these factors are taken into account.

Finally, concavity in c is not a necessary condition for the wage offer density to be declining. In the next section, I illustrate numerically that when the substitution rate between wage and monitoring is more favorable for high wage firms, then the model can generate a declining wage offer density even when the monitoring cost is approximately linear.

5. NUMERICAL ILLUSTRATION

This section provides an example of a hump-shaped and right skewed equilibrium wage distribution and the associated cost function. The general idea is to specify an empirically plausible distribution of wages paid G . Then, given G , the associated cost function is approximated numerically.

The method is as follows: Choose G and the parameters $\Omega = (e, b, \lambda_0, \lambda_1, \delta, N, p, \pi, r)$. Given G , (11) implies $F(w)$ for $\underline{w} \leq w \leq \bar{w}$. Then at any point over the wage support, the equal profit condition (13) gives the associated cost c , and the NSC (9) yields the implied monitoring rate q . Plotting c against q gives a nonparametric cost profile.

[Mortensen \(2005\)](#) provides summary statistics on occupation-specific wage distributions using the Danish Integrated Database for Labor Market Research (IDA). In this example, G is set to resemble the reported distribution for unskilled workers which is hump-shaped and right skewed:

$$(32) \quad G(w) = \text{Beta} \left(\frac{w - \underline{w}}{\bar{w} - \underline{w}}; 2, 3 \right).$$

These qualitative properties of wage distributions are robust across occupations and countries (see [Mortensen \(2005\)](#) for other wage distributions in Denmark, [Postel-Vinay and Robin \(2002a\)](#) for wage distributions in France and [Figure \(1\)](#) for wage distributions in the US). The infimum and supremum of the wage support are set at 180,034 and 318,667 Danish kroner (DKK) per year, which correspond to the reported 10th and 90th percentiles. Annual labor-market transition parameters for unskilled manual workers, taken from [Postel-Vinay and Robin \(2002a\)](#), are $\lambda_0 = 1.529$, $\lambda_1 = 0.666$ and $\delta = 0.0989$. I set the revenue per worker $p = 400,000$, the unemployment benefit $b = 50,000$, and the profit flow $\pi = 50,000$. The ratio of workers to firms N is set to 50 which lies in the range of low skilled workers per firm provided in [Bagger et al. \(2011\)](#). The effort

cost $e = 224$ is chosen to satisfy the FOC in (18).¹⁸ Further, the discount rate r is set to zero in accordance with the analytic expressions.

Figure (2) depicts the graphical results. Panel (a) shows the densities of wages paid and wages offered, $g(w)$ and $f(w)$. The implied monitoring cost $c(q)$ is displayed in panel (b). The relationship between wages and monitoring rates implied by the NSC, $\phi(w)$, and the corresponding monitoring cost at each wage, $c(\phi(w))$, are plotted in panel (c) and (d), respectively. Panel (b) shows, in line with proposition 3, that the model can generate a declining wage offer density if the monitoring technology exhibits increasing returns. Concavity in c implies a more significant increase in monitoring cost from a wage cut at a high wage firm. The prospect of a greatly increased monitoring cost reduces the relative importance of firm competition. Consequently, the densities of wages offered and wages paid decline over the upper range of their support as depicted in panel (a). This numerical exercise has illustrated a case in which the increasing returns monitoring technology is sufficient to generate a hump-shaped and right skewed equilibrium wage distribution.

As discussed earlier, increasing returns in monitoring technology is not necessary for the equilibrium wage density to decline. To illustrate this, I reduce the offer arrival rate for unemployed workers, λ_0 by one-fifth to 0.306 and set e to 156 to satisfy the FOC while all other parameter values are unchanged. Figure (3) illustrates the graphical results. As can be seen in panel (c), the substitution rate of monitoring rate for wage is more favorable for high wage firms (i.e. $\frac{d^2\phi}{dw^2} < 0$). This is because when the offer arrival rate for unemployed workers λ_0 is lower, the cost of losing a job becomes relatively more severe for workers who are closer to the top of the wage ladder. So the penalty of getting fired if caught shirking is more significant at high wage firms. This, in turn, implies that a high wage firm can substitute wage for monitoring rate more proportionately than a low wage firm. The relative importance of firm competition is diminished even when the nonparametric monitoring cost profile is approximately linear with slight convexity. Hence the wage offer and equilibrium wage densities are declining over the upper range of the wage support as shown in panel (a). Finally, while the value of λ_0 in this case is low compared to the estimates in the literature (e.g. [Postel-Vinay and Robin, 2002a,b](#)), this example shows that strong increasing returns in c is not a necessary condition for the model to generate a hump-shaped wage density.

6. CONCLUSION

This paper combines two well-known frameworks: efficiency wages from [Shapiro and Stiglitz \(1984\)](#) and on-the-job search from [Burdett and Mortensen \(1998\)](#). An inverse relationship between

¹⁸In the numerical exercise, \bar{q} is implied by the NSC taking the wage offer distribution $F(w)$ as given. However, to be internally consistent, I choose the effort level e so that the FOC determining the optimal \bar{q} : $c'(\bar{q}) = \frac{e(\delta+\lambda_1)}{\bar{q}^2}$ as well as the SOC: $c''(\bar{q}) > -\frac{2e(\delta+\lambda_1)}{\bar{q}^3}$ hold where $c'(\bar{q})$ and $c''(\bar{q})$ are approximated numerically.

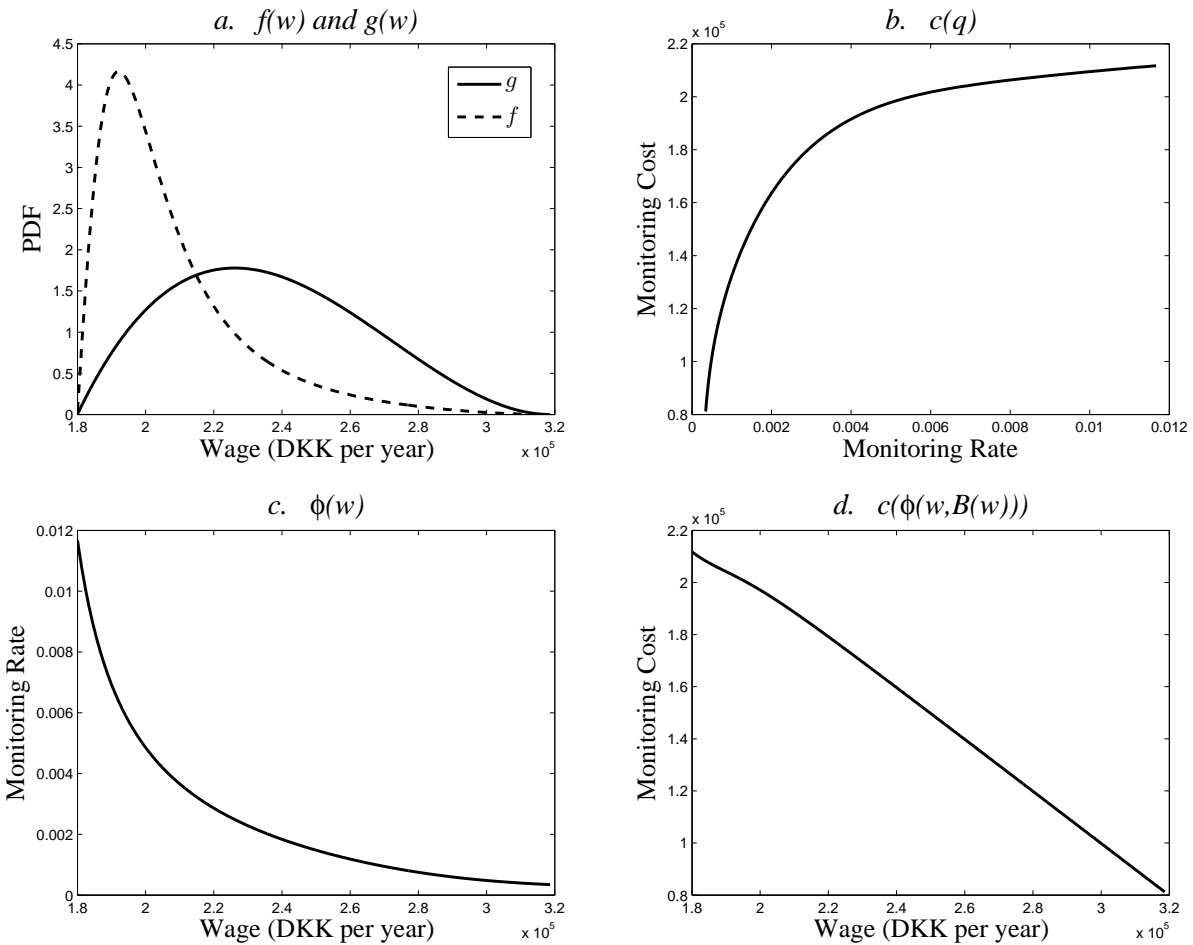


Figure 2: Benchmark Numerical Solution

wages and monitoring rates is derived which is consistent with a number of empirical studies, e.g. [Brunello \(1995\)](#). I show that this additional tradeoff can resolve the anomalous feature of [Burdett and Mortensen \(1998\)](#) which places more density on higher wages when workers and firms are homogeneous. The wage distributions implied by the model can exhibit hump shapes and long right tails as observed empirically. The present framework yields a contrasting result to previous work where ex-ante heterogeneity necessitates the right shape of wage distribution. The result that a hump-shaped wage distribution can be derived in absence of heterogeneity serves as a note of caution for empirical analyses that base identification of the productivity distribution on the assumption of a one-to-one relationship between the productivity and wage densities (e.g. [Bontemps et al., 2000](#)).

The estimated arrival rates provided in the literature imply a concave monitoring cost, in contrast to the commonly assumed decreasing returns technology (e.g. [Greenwood et al., 2010](#)). However, concavity is not a necessary condition for hump-shaped wage density. In reality the shape of monitoring cost is generally not observed. Monitoring may be done in a number of ways such as hiring supervisors and using information technology. [Abowd and Kramarz \(2003\)](#); [Kramarz](#)

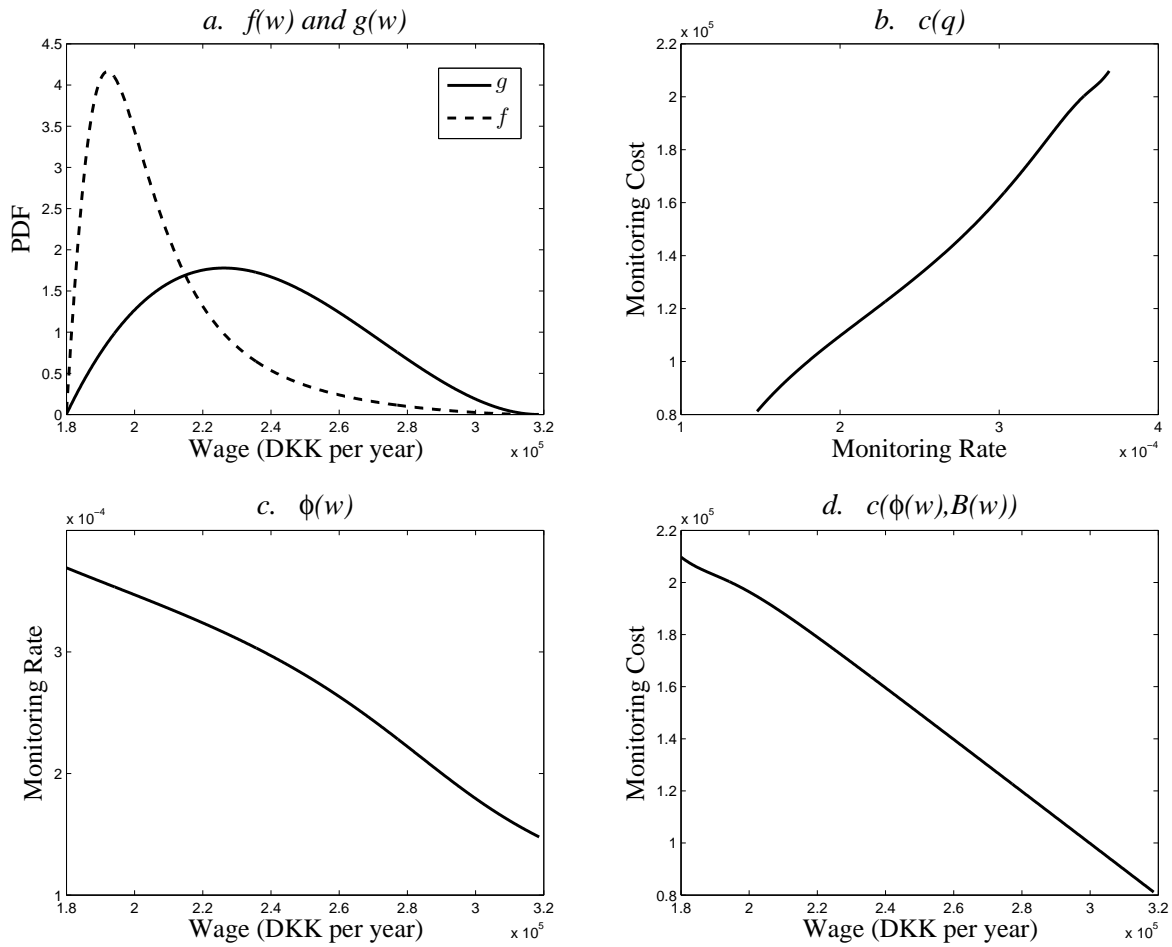


Figure 3: Numerical Solution: Low λ_0

and Michaud (2010) estimate firms' labor adjustment costs to be concave, particularly for hires of managers. If the production of monitoring uses supervisors, then the concavity in hiring costs would influence the shape of monitoring costs. So while the results do not depend on concavity, this type of monitoring technology is plausible.

One obvious extension of this work is to examine wage distributions when firms differ in productivity and analyze how productivity distribution can be identified. Additional extensions include using other means, such as job amenities, to derive an empirically plausible wage distribution when workers and firms are homogeneous.

APPENDIX

A.1 Derivation of NSC

To derive (4), rearrange (2) as follows

$$q[V_s(w) - V_u] = w + \lambda_1 \left[\int \max\{V_s(w), V_e(x)\} dF(x) - V_s(w) \right] + \delta[V_u - V_s(w)] - rV_s(w).$$

Using the fact that $V_s(w) = V_e(w)$ gives,

$$\begin{aligned} q[V_e(w) - V_u] &= w + \lambda_1 \left[\int \max\{V_e(w), V_e(x)\} dF(x) - V_e(w) \right] + \delta[V_u - V_e(w)] - rV_e(w) \\ &= e. \end{aligned}$$

Thus,

$$V_e(w) = \frac{e}{q} + V_u.$$

To derive (8), first note that

$$\begin{aligned} (33) \quad E_x \max\{V_i(w), V_i(x)\} &= F(w)V_i(w) + \bar{F}(w)E(V_i(x) | V_i(x) > V_i(w)) \\ &= F(w)V_i(w) + \bar{F}(w) \left[V_i(w) + \int_w^{\bar{w}} \frac{V_i'(x)\bar{F}(x)}{\bar{F}(w)} dx \right] \\ &= V_i(w) + \int_w^{\bar{w}} V_i'(x)\bar{F}(x) dx \end{aligned}$$

where $i \in \{e, s\}$ and $\bar{F}(w) = 1 - F(w)$. So the asset value of shirker can be written as

$$(34) \quad rV_s(w) = w + \lambda_1 \int_w^{\bar{w}} V_e'(x)\bar{F}(x) dx + (\delta + q)[V_u - V_s(w)].$$

Similarly the asset value of nonshirker can be expressed as

$$(35) \quad rV_e(w) = w - e + \lambda_1 \int_w^{\bar{w}} V_e'(x)\bar{F}(x) dx + \delta[V_u - V_e(w)].$$

Substituting (4) into (35) gives

$$(36) \quad w = (r + \delta) \frac{e}{q} + e - \lambda_1 \int_w^{\bar{w}} V_e'(x)\bar{F}(x) dx + rV_u.$$

Differentiating (35) with respect to w yields

$$\begin{aligned} (37) \quad rV_e'(w) &= 1 - \lambda_1 V_e'(w)\bar{F}(w) - \delta V_e'(w) \\ V_e'(w) &= \frac{1}{r + \delta + \lambda_1 \bar{F}(w)} \end{aligned}$$

Substituting (37) into (36), we have

$$(38) \quad w = (r + \delta) \frac{e}{q} + e - \lambda_1 \int_w^{\bar{w}} \frac{\bar{F}(x)}{r + \delta + \lambda_1 \bar{F}(x)} dx + rV_u.$$

Finally, by Lemma (1): $V_e(w) > V_u$ for any $w \in F(w)$, the asset value of unemployment can be expressed as

$$(39) \quad \begin{aligned} rV_u &= b + \lambda_0 \int_w^{\bar{w}} (V_e(x) - V_u) dF(x) \\ &= b + \lambda_0 \left(\frac{e}{q} + \int_w^{\bar{w}} V_e'(x) \bar{F}(x) dx \right) \\ &= b + \lambda_0 \left(\frac{e}{q} + \int_w^{\bar{w}} \frac{\bar{F}(x)}{r + \delta + \lambda_1 \bar{F}(x)} dx \right) \end{aligned}$$

where the second equality uses the fact that $V_u = V_e(\underline{w}) + \frac{e}{q}$. Substituting (39) into (38), the no-shirking wage is given by

$$\begin{aligned} w &= (r + \delta) \frac{e}{q} + \lambda_0 \frac{e}{q} + e + b + \lambda_0 \int_w^{\bar{w}} \frac{\bar{F}(x)}{r + \delta + \lambda_1 \bar{F}(x)} dx - \lambda_1 \int_w^{\bar{w}} \frac{\bar{F}(x)}{r + \delta + \lambda_1 \bar{F}(x)} dx \\ &= \frac{(r + \delta)e}{q} + \frac{\lambda_0 e}{q} + e + b + \lambda_0 B(\bar{w}) - \lambda_1 (B(\bar{w}) - B(w)) \end{aligned}$$

A.2. Positive Profit Condition

Lemma 5. *If $p > e + b + \frac{e(\delta + \lambda_0)}{q} + \left(\frac{\lambda_0 + \lambda_1}{\lambda_1} \right) c(\bar{q})$ then $p - \underline{w} - c(\phi(\underline{w})) > 0$, profit is strictly positive and F has a nondegenerate support.*

To show that $\pi > 0$ if $p > \frac{e(\delta + \lambda_0)}{q} + e + b + \left(\frac{\lambda_0 + \lambda_1}{\lambda_1} \right) c(\bar{q})$, let us consider the following cases:

Case1) Suppose $\lambda_0 > \lambda_1$. From (19) we have

$$(40) \quad \underline{w} = \frac{e(\delta + \lambda_0)}{q} + e + b + \frac{(\lambda_0 - \lambda_1)}{\lambda_1} \int_w^{\bar{w}} \frac{\lambda_1 \bar{F}(x)}{\delta + \lambda_1 \bar{F}(x)} dx$$

which cannot be expressed in terms of the primitives due to the complexity of the last term. However, note that

$$(41) \quad \begin{aligned} \int_w^{\bar{w}} \frac{\lambda_1 \bar{F}(x)}{\delta + \lambda_1 \bar{F}(x)} dx &< \int_w^{\bar{w}} \frac{\lambda_1 \bar{F}(x)}{\lambda_1 \bar{F}(x)} dx \\ &= \int_w^{\bar{w}} 1 dx \\ &= \bar{w} - \underline{w}. \end{aligned}$$

Replacing the last term in (40) with (41), we have

$$\underline{w} < \frac{e(\delta + \lambda_0)}{\bar{q}} + e + b + \frac{(\lambda_0 - \lambda_1)}{\lambda_1} (\bar{w} - \underline{w}).$$

Since $\bar{w} < p$,

$$(42) \quad \underline{w} < \frac{e(\delta + \lambda_0)}{\bar{q}} + e + b + \frac{(\lambda_0 - \lambda_1)}{\lambda_1} (p - \underline{w}).$$

Rearranging (42) gives,

$$(43) \quad \begin{aligned} \underline{w} &< \left(\left(\frac{e(\delta + \lambda_0)}{\bar{q}} + e + b \right) + \frac{(\lambda_0 - \lambda_1)}{\lambda_1} p \right) / \left(1 + \frac{(\lambda_0 - \lambda_1)}{\lambda_1} \right) \\ 0 &< p - \underline{w} + \left(\frac{e(\delta + \lambda_0)}{\bar{q}} + e + b - p \right) \frac{\lambda_1}{\lambda_0}. \end{aligned}$$

Given that $\frac{e(\delta + \lambda_0)}{\bar{q}} + e + b + \left(\frac{\lambda_0 + \lambda_1}{\lambda_1} \right) c(\bar{q}) < p \iff \frac{e(\delta + \lambda_0)}{\bar{q}} + e + b - p < - \left(\frac{\lambda_0 + \lambda_1}{\lambda_1} \right) c(\bar{q})$, (43) implies

$$\begin{aligned} 0 &< p - \underline{w} - \left(\frac{\lambda_0 + \lambda_1}{\lambda_0} \right) c(\bar{q}) \\ &< p - \underline{w} - c(\bar{q}). \end{aligned}$$

Finally, replacing $\phi(\underline{w}) = \bar{q}$, we have

$$0 < p - \underline{w} - c(\phi(\underline{w})).$$

Case 2) Suppose $\lambda_0 \leq \lambda_1$, then (19) implies that

$$(44) \quad \begin{aligned} \underline{w} &\leq \frac{e(\delta + \lambda_0)}{\bar{q}} + e + b \\ 0 &\leq \frac{e(\delta + \lambda_0)}{\bar{q}} + e + b - \underline{w}. \end{aligned}$$

Since $\frac{e(\delta + \lambda_0)}{\bar{q}} + e + b + \left(\frac{\lambda_0 + \lambda_1}{\lambda_1} \right) c(\bar{q}) < p \iff \frac{e(\delta + \lambda_0)}{\bar{q}} + e + b < p - \left(\frac{\lambda_0 + \lambda_1}{\lambda_1} \right) c(\bar{q})$, (44) implies

$$(45) \quad \begin{aligned} 0 &< p - \underline{w} - \left(\frac{\lambda_0 + \lambda_1}{\lambda_1} \right) c(\bar{q}) \\ &< p - \underline{w} - c(\bar{q}). \end{aligned}$$

Using the fact that $\phi(\underline{w}) = \bar{q}$, we have

$$0 < p - \underline{w} - c(\phi(\underline{w})).$$

Therefore the steady state profit: $\pi = (p - \underline{w} - c(\phi(\underline{w}))) \ell(\underline{w}|F) > 0$ when $p > \frac{e(\delta + \lambda_0)}{\bar{q}} + e + b + \left(\frac{\lambda_0 + \lambda_1}{\lambda_1} \right) c(\bar{q})$.

A.3. Proof of Proposition 2: Existence of Equilibrium Solution

Before proving existence of equilibrium, let us rewrite (21) as a system of differential equations. Recall that

$$(46) \quad B(w) = \int_{\underline{w}}^w \frac{\bar{F}(x)}{\delta + \lambda_1 \bar{F}(x)} dx,$$

Eq (9) can be written explicitly as a function of w and $B(w)$ as

$$(47) \quad \phi(w, B(w)) = \frac{e\delta}{w + \chi - \lambda_1 B(w)},$$

with $\chi = -\lambda_0 \frac{e}{q} - e - b - \lambda_0 B(\bar{w}) + \lambda_1 B(\bar{w})$. Note that in the main text, I have dropped $B(w)$ as an argument of ϕ to reduce clutter in notation. Differentiating (21) with respect to w gives

$$(48) \quad F'(w) = \left(\frac{\delta + \lambda_1}{2\lambda_1} \right) \left(\frac{(p - w - c(\phi(w, B(w))))^{-1/2}}{(p - \underline{w} - c(\phi(\underline{w})))^{1/2}} \right) (1 + c'(\phi) \phi'(w)),$$

where

$$(49) \quad \begin{aligned} \phi'(w) &= 1 / \frac{dw}{d\phi} \\ &= \frac{-\phi(w, B(w))^2}{e(\delta + \lambda_1 \bar{F}(w))}. \end{aligned}$$

So (48) can be expressed as

$$(50) \quad \begin{aligned} F'(w) &= \left(\frac{\delta + \lambda_1}{2\lambda_1} \right) \left(\frac{(p - w - c(\phi(w, B(w))))^{-1/2}}{(p - \underline{w} - c(\phi(\underline{w})))^{1/2}} \right) \left(1 - \frac{c'(\phi) \phi(w, B(w))^2}{e(\delta + \lambda_1 \bar{F}(w))} \right), \\ &\equiv \beta(w, F(w), B(w)). \end{aligned}$$

Differentiating (46) with respect to w gives

$$(51) \quad \begin{aligned} B'(w) &= \frac{\bar{F}(w)}{\delta + \lambda_1 \bar{F}(w)} \\ &\equiv \alpha(w, F(w)). \end{aligned}$$

Therefore, (50) and (51) give us the following system of differential equations

$$(52) \quad \begin{aligned} B'(w) &= \alpha(w, F(w)) \\ F'(w) &= \beta(w, F(w), B(w)) \end{aligned}$$

with the initial conditions: $F(\underline{w}) = 0$ and $B(\underline{w}) = 0$. For a given value of \underline{w} , the existence and local uniqueness of $F(w)$ and $B(w)$ that solve (52) can be proved using Picard's theorem for systems of differential equations. Further, the boundedness of $F(w)$ and $B(w)$ allow us to extend the solution to be globally unique as discussed later in this section.

Proposition 4. Let $\alpha(w, F(w))$ and $\beta(w, F(w), B(w))$ be continuous functions defined on a convex open region $U \subseteq \mathbb{R}^3$. If $\alpha(F(w))$ and $\beta(w, F(w), B(w))$ satisfy the Lipschitz condition, i.e. there exists a constant L such that

$$|\alpha(w, F_1) - \alpha(w, F_2)| \leq L(|F_1 - F_2| + |B_1 - B_2|)$$

and

$$|\beta(w, F_1, B_1) - \beta(w, F_2, B_2)| \leq L(|F_1 - F_2| + |B_1 - B_2|),$$

for any (w, F_1, B_1) and (w, F_2, B_2) in U , then for a given $(\underline{w}, \underline{F}, \underline{B}) \in U$, the system of differential equations (52) has a unique solution $F(w)$ and $B(w)$ such that $F(\underline{w}) = \underline{F}$ and $B(\underline{w}) = \underline{B}$ for $|w - \underline{w}| \leq h^*$.

Proof. See Theorem 2, Chapter 4 of [Murray and Miller \(1976\)](#). \square

Since U is an open region, let us restrict our attention to the relevant closed neighborhood of U of the initial points $(\underline{w}, \underline{F}, \underline{B})$,

$$V = \{(w, F, B) : |w - \underline{w}| \leq j, |F - \underline{F}| \leq j, |B - \underline{B}| \leq j\}$$

on which $\alpha(w, F(w))$ and $\beta(w, F(w), B(w)) \leq M$. A range of w on which the solution $F(w)$ and $B(w)$ exist is given by $|w - \underline{w}| \leq h^*$ where $h^* = \min\left\{j, \frac{j}{M}\right\}$ (see Lemma 2, Chapter 4 of [Murray and Miller \(1976\)](#) for the derivation of h^*).

Since $0 \leq F \leq 1$, we begin by choosing $j = 1$ and first prove the existence of a locally unique solution. In equilibrium, firms always offer wages at least $e + b$ otherwise no workers would work, and no firms offer wages more than p otherwise profit would be negative. Therefore if $[\underline{w} - h^*, \underline{w} + h^*] \supset (e + b, p)$ then we have a globally unique solution on the relevant domain. However, if $[\underline{w} - h^*, \underline{w} + h^*] \subset (e + b, p)$ then one needs to show that the solution can be extended to $\{w : e + b < w < p\}$ to establish global uniqueness.

To prove existence, I will show that if (i) $p > e + b + \frac{e(\delta + \lambda_0)}{\bar{q}} + \left(\frac{\lambda_0 + \lambda_1}{\lambda_1}\right)c(\bar{q})$, i.e. firms make a strictly positive profit, and (ii) $c(q)$ satisfies assumption 1, then α and β are continuous functions which satisfy the Lipschitz condition. By proposition 4, there exists a unique solution $F(w)$ and $B(w)$ for a given value of \underline{w} .

First, note that α is continuous since it is the quotient of two continuous functions on V (and the denominator $\neq 0$). Second, provided that profit is strictly positive and c is continuously differentiable, β is continuous on V . It remains to show that α and β satisfy the Lipschitz condition.

Consider (w, F_1, B_1) and $(w, F_2, B_2) \in V$ with w fixed. Without loss of generality, assume $F_1 < F_2$. Take the absolute difference of (51) at these two points,

$$\begin{aligned} |\alpha(w, F_1) - \alpha(w, F_2)| &= \left| \frac{1 - F_1}{\delta + \lambda_1(1 - F_1)} - \frac{1 - F_2}{\delta + \lambda_1(1 - F_2)} \right| \\ &< \left| \frac{1 - F_1}{\delta + \lambda_1(1 - F_2)} - \frac{1 - F_2}{\delta + \lambda_1(1 - F_2)} \right| \\ &= \frac{1}{\delta + \lambda_1(1 - F_2)} |F_2 - F_1|. \end{aligned}$$

Since $|F - \underline{F}| \leq 1$ and $F(\underline{w}) = \underline{F} = 0$, we have $F \leq 1$ for all $F \in V$. So,

$$\begin{aligned}
|\alpha(w, F_1) - \alpha(w, F_2)| &< \frac{1}{\delta} |F_2 - F_1| \\
&< \frac{1}{\delta} (|F_2 - F_1| + |B_2 - B_1|) \\
(53) \qquad \qquad \qquad &= L_\alpha (|F_2 - F_1| + |B_2 - B_1|)
\end{aligned}$$

where the constant $L_\alpha = \frac{1}{\delta}$. Hence α is Lipschitz continuous.

To show that β satisfies the Lipschitz condition, it is sufficient (and less algebraically involved) to show that β is Lipschitz continuous in B given F , and in F given B using the following lemma.

Lemma 6. *Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. If $f(x, y)$ is Lipschitz continuous in x given y , and in y given x then $f(x, y)$ is Lipschitz continuous in (x, y) .*

Proof. First, fix y then

$$\begin{aligned}
|f(x_1, y) - f(x_2, y)| &\leq L \|(x_1, y) - (x_2, y)\| \\
&= L(|x_1 - x_2| + |y - y|) \\
(54) \qquad \qquad \qquad &= L|x_1 - x_2|.
\end{aligned}$$

Second, fix x then

$$\begin{aligned}
|f(x, y_1) - f(x, y_2)| &\leq L \|(x, y_1) - (x, y_2)\| \\
&= L(|x - x| + |y_1 - y_2|) \\
(55) \qquad \qquad \qquad &= L|y_1 - y_2|.
\end{aligned}$$

So

$$\begin{aligned}
|f(x_1, y_1) - f(x_2, y_2)| &= |f(x_1, y_1) - f(x_2, y_1) + f(x_2, y_1) - f(x_2, y_2)| \\
&\leq |f(x_1, y_1) - f(x_2, y_1)| + |f(x_2, y_1) - f(x_2, y_2)|
\end{aligned}$$

Using (54) and (55), the above inequality becomes

$$\begin{aligned}
|f(x_1, y_1) - f(x_2, y_2)| &\leq L|x_1 - x_2| + L|y_1 - y_2| \\
&= L \|(x_1, y_1) - (x_2, y_2)\|.
\end{aligned}$$

□

Applying Lemma 6, let us first show that β is Lipschitz continuous in F given B . Take (w, F_1, B) and $(w, F_2, B) \in V$, the absolute difference of (50) at these two points is

$$|\beta(w, F_1, B) - \beta(w, F_2, B)| = A_0 \left| \left(1 + \frac{dc}{d\phi} \frac{d\phi}{dw} \Big|_{F_1} \right) - \left(1 + \frac{dc}{d\phi} \frac{d\phi}{dw} \Big|_{F_2} \right) \right|$$

where $A_0 = \frac{\delta + \lambda_1}{2\lambda_1(p - \underline{w} - c(\phi(\underline{w})))^{1/2}(p - \underline{w} - c(\phi(w, B)))^{1/2}}$. Substituting (49) in gives

$$\begin{aligned} |\beta(w, F_1, B) - \beta(w, F_2, B)| &= A_0 \left| \left(\frac{-c'(\phi)\phi(w, B)^2}{e(\delta + \lambda_1(1 - F_1))} + \frac{c'(\phi)\phi(w, B)^2}{e(\delta + \lambda_1(1 - F_2))} \right) \right| \\ &= A_0 c'(\phi) \frac{\phi(w, B)^2 \lambda_1}{e} \left| \frac{F_2 - F_1}{(\delta + \lambda_1(1 - F_1))(\delta + \lambda_1(1 - F_2))} \right|. \end{aligned}$$

Without loss of generality, suppose $F_2 < F_1$ then

$$|\beta(w, F_1, B) - \beta(w, F_2, B)| < A_0 c'(\phi) \frac{\phi(w, B)^2 \lambda_1}{e(\delta + \lambda_1(1 - F_1))^2} |F_2 - F_1|.$$

The Lipschitz constant in this case is $A_0 c'(\phi) \frac{\phi(w, B)^2 \lambda_1}{e(\delta + \lambda_1(1 - F_1))^2}$ which is strictly positive since c is strictly increasing and $A_0 > 0$ since profit is strictly positive.

Next, let us take (w, F, B_1) and $(w, F, B_2) \in V$ where w and F are fixed, then

$$|\beta(w, F, B_1) - \beta(w, F, B_2)| = A_1 \left| \frac{\left(1 + \frac{dc}{d\phi} \frac{d\phi}{dw} \Big|_{B_1}\right)}{(p - w - c(\phi(w, B_1)))^{1/2}} - \frac{\left(1 + \frac{dc}{d\phi} \frac{d\phi}{dw} \Big|_{B_2}\right)}{(p - w - c(\phi(w, B_2)))^{1/2}} \right|$$

where $A_1 = \frac{\delta + \lambda_1}{2\lambda_1(p - \underline{w} - c(\phi(\underline{w})))^{1/2}}$. Substituting in (49) gives

$$\begin{aligned} |\beta(w, F, B_1) - \beta(w, F, B_2)| &= A_1 \left| \frac{\left(1 - \frac{c'(\phi(w, B_1))\phi(w, B_1)^2}{e(\delta + \lambda_1(1 - F))}\right)}{(p - w - c(\phi(w, B_1)))^{1/2}} - \frac{1 - \frac{c'(\phi(w, B_2))\phi(w, B_2)^2}{e(\delta + \lambda_1(1 - F))}}{(p - w - c(\phi(w, B_2)))^{1/2}} \right| \\ &= A_2 \left| \frac{\left(e(\delta + \lambda_1(1 - F)) - c'(\phi(w, B_1))\phi(w, B_1)^2\right)}{(p - w - c(\phi(w, B_1)))^{1/2}} \right. \\ &\quad \left. - \frac{\left(e(\delta + \lambda_1(1 - F)) - c'(\phi(w, B_2))\phi(w, B_2)^2\right)}{(p - w - c(\phi(w, B_2)))^{1/2}} \right| \end{aligned}$$

where $A_2 = \frac{A_1}{e(\delta + \lambda_1(1 - F))} > 0$. Without loss of generality, suppose $\phi(w, B_1) < \phi(w, B_2)$. Since $c(q)$ is strictly increasing and by assumption 1.1: $c'(q)q^2$ is strictly increasing, we have $c'(\phi(w, B_1))\phi(w, B_1)^2 < c'(\phi(w, B_2))\phi(w, B_2)^2$ and $c(\phi(w, B_1)) < c(\phi(w, B_2))$. Hence,

$$\begin{aligned} |\beta(w, F, B_1) - \beta(w, F, B_2)| &< \frac{A_2}{(p - w - c(\phi(w, B_2)))^{1/2}} \left| c'(\phi(w, B_2))\phi(w, B_2)^2 - c'(\phi(w, B_1))\phi(w, B_1)^2 \right| \\ &< \frac{A_2}{(p - w - c(\phi(w, B_2)))^{1/2}} \int_{\phi(w, B_1)}^{\phi(w, B_2)} |c''(x)x^2 + c'(x)2x| dx. \end{aligned}$$

Since c is twice continuously differentiable, $c''(x)x^2 + c'(x)2x$ is continuous on V . Since V is compact, there exists a constant J such that $c''(x)x^2 + c'(x)2x < J$ for all $\phi(w, B) \in V$. So,

$$\begin{aligned} |\beta(w, F, B_1) - \beta(w, F, B_2)| &< \frac{A_2}{(p - w - c(\phi(w, B_2)))^{1/2}} \int_{\phi(w, B_1)}^{\phi(w, B_2)} |J| dx \\ &< \frac{A_2 |J|}{(p - w - c(\phi(w, B_2)))} |\phi(w, B_1) - \phi(w, B_2)| \end{aligned}$$

Substituting in (47), we get

$$\begin{aligned} |\beta(w, F, B_1) - \beta(w, F, B_2)| &< \frac{A_2 |J|}{(p - w - c(\phi(w, B_2)))} \left| \frac{e\delta}{w + \chi - \lambda_1 B_1} - \frac{e\delta}{w + \chi - \lambda_1 B_2} \right| \\ &= \frac{A_2 |J| e\delta \lambda_1}{(p - w - c(\phi(w, B_2)))} \left| \frac{(B_1 - B_2)}{(w + \chi - \lambda_1 B_1)(w + \chi - \lambda_1 B_2)} \right|. \end{aligned}$$

Without loss of generality, suppose $B_1 < B_2$ then

$$|\beta(w, F, B_1) - \beta(w, F, B_2)| < \frac{A_2 |J| e\delta \lambda_1}{(p - w - c(\phi(w, B_2)))(w + \chi - \lambda_1 B_2)^2} |B_1 - B_2|.$$

The Lipschitz constant in this case is strictly positive since profit is strictly positive. So β is Lipschitz continuous in B given F , and in F given B . By Lemma 6, β is Lipschitz continuous in (F, B) .

Since α and β are Lipschitz continuous, there exists a unique solution $F(w)$ and $B(w)$ such that $F(\underline{w}) = \underline{F}$ and $B(\underline{w}) = \underline{B}$ for $|\underline{w} - \underline{w}| \leq h^*$. Further, by Proposition 1, $F(w)$ has a connected support. The remaining task is to show that the solution $F(w)$ and $B(w)$ can be extended to a range of w on which $e + b < w < p$ if $[\underline{w} - h^*, \underline{w} + h^*] \subset (e + b, p)$. The following proposition provides necessary and sufficient conditions under which a continuation of solution exists.

Proposition 5. *Let $F(w)$ and $B(w)$ be a solution to (52) defined on $|\underline{w} - \underline{w}| \leq h^*$ where $\underline{w} - h^* < e + b$ and $\underline{w} + h^* < p$. There exists a continuation of $F(w)$ and $B(w)$ to a larger interval $I = \{w : e + b < w < p\}$ if and only if the limits of $F(w)$ and $B(w)$ are finite on I .*

Proof. See proposition 1.17 in Markley (2004). □

By construction, $0 \leq F(w) \leq 1$ so $F(w)$ has finite limits for all w on I . Since $B(w)$ is continuous and increasing in w , we only need to check whether the limits as $w \rightarrow p^-$ and $w \rightarrow e^- + b^-$ are finite. The lower limit is

$$\lim_{w \rightarrow e^- + b^-} \int_{\underline{w}}^w \frac{\bar{F}(x)}{\delta + \lambda_1 \bar{F}(x)} dx = \lim_{w \rightarrow e^- + b^-} - \int_w^{\underline{w}} \frac{\bar{F}(x)}{\delta + \lambda_1 \bar{F}(x)} dx > \lim_{w \rightarrow e^- + b^-} - \int_w^{\underline{w}} \frac{1}{\delta} dx = \frac{-\underline{w} + e + b}{\delta}.$$

The upper limit is

$$\lim_{w \rightarrow p^+} \int_{\underline{w}}^w \frac{\bar{F}(x)}{\delta + \lambda_1 \bar{F}(x)} dx < \lim_{w \rightarrow p^+} \int_{\underline{w}}^w \frac{1}{\delta} dx = \frac{p - \underline{w}}{\delta}.$$

Therefore for a given value of \underline{w} , boundedness of $F(w)$ and $B(w)$ guarantees that the solution is globally unique on the relevant domain of the problem.

A.4. Proof of Proposition 3: Declining Wage Offer Density

Using (28), $f'(w) < 0$ iff

$$(56) \quad \left(c''(\phi) (\phi'(w))^2 + c'(\phi) \phi''(w) \right) < -\frac{(1 + c'(\phi) \phi'(w))^2}{2(p - w - c(\phi(w)))},$$

which depends on endogenous variables. However, by tightening the bounds, (56) can be expressed in terms of the exogenous parameters.

First, since the profit per worker is decreasing, i.e. $0 < (1 + c'(\phi) \phi'(w)) \leq 1$, we can tighten the bound by replacing $(1 + c'(\phi) \phi'(w))$ with 1. So (56) can be written as,

$$(57) \quad \left(c''(\phi) (\phi'(w))^2 + c'(\phi) \phi''(w) \right) < -\frac{1}{2(p - w - c(\phi(w)))}.$$

Using (7), we have that

$$(58) \quad \phi'(w) = -\frac{q^2}{e(\delta + \lambda_1 \bar{F}(w))}.$$

Differentiating (58) with respect to w gives

$$(59) \quad \phi''(w) = \frac{2q^3 - e\lambda_1 f(w) q^2}{(e(\delta + \lambda_1 \bar{F}(w)))^2}.$$

Substituting (58) and (59) into (57), we have

$$\begin{aligned} \left(c''(q) \frac{q^4}{(e(\delta + \lambda_1 \bar{F}(w)))^2} + c'(q) \left(\frac{2q^3 - e\lambda_1 f(w) q^2}{(e(\delta + \lambda_1 \bar{F}(w)))^2} \right) \right) &< -\frac{1}{2(p - w - c(q))}, \\ c''(q) q^4 + c'(q) q^3 &< -\frac{(e(\delta + \lambda_1 \bar{F}(w)))^2}{2(p - w - c(q))} \\ &+ e\lambda_1 f(w) c'(q) q^2. \end{aligned}$$

Substituting (23) into $f(w)$ gives

$$\begin{aligned} c''(q) q^4 + 2c'(q) q^3 &< -\frac{(e(\delta + \lambda_1 \bar{F}(w)))^2}{2(p - w - c(q))} \\ &+ e\lambda_1 c'(q) q^2 \left(\left(\frac{\delta + \lambda_1}{2\lambda_1} \right) \left(\frac{(p - w - c(q))^{-1/2}}{(p - \underline{w} - c(\bar{q}))^{1/2}} \right) (1 + c'(\phi) \phi'(w)) \right). \end{aligned}$$

Since the profit per worker of the lowest wage firm $(p - \underline{w} - c(\bar{q}))$ is higher than other firms' profits per worker, we can tighten the bound by replacing $(p - w - c(\phi(w)))$ in the second term of RHS with $(p - \underline{w} - c(\bar{q}))$. This yields

$$c''(q) q^4 + 2c'(q) q^3 < -\frac{(e(\delta + \lambda_1 \bar{F}(w)))^2}{2(p - w - c(q))} + \left(\left(\frac{c'(q) q^2 e(\delta + \lambda_1)}{2(p - \underline{w} - c(\bar{q}))} \right) (1 + c'(\phi) \phi'(w)) \right).$$

Similarly, since the profit per worker of the highest wage firm $(p - \bar{w} - c(q))$ is lower than other firms' profits per worker, we can tighten the bound by replacing $(p - w - c(q))$ in the first term of RHS with $(p - \bar{w} - c(q))$,

$$c''(q)q^4 + 2c'(q)q^3 < -\frac{(e(\delta + \lambda_1 \bar{F}(w)))^2}{2(p - \bar{w} - c(q))} + \left(\left(\frac{c'(q)q^2 e(\delta + \lambda_1)}{2(p - \underline{w} - c(\bar{q}))} \right) (1 + c'(\phi)\phi'(w)) \right),$$

and using (22): $(p - \bar{w} - c(q)) = \frac{(p - \underline{w} - c(\bar{q}))\delta^2}{(\delta + \lambda_1)^2}$ we have

$$c''(q)q^4 + 2c'(q)q^3 < -\frac{(e(\delta + \lambda_1 \bar{F}(w)))^2 (\delta + \lambda_1)^2}{2(p - \underline{w} - c(\bar{q}))\delta^2} + \left(\left(\frac{c'(q)q^2 e(\delta + \lambda_1)}{2(p - \underline{w} - c(\bar{q}))} \right) (1 + c'(\phi)\phi'(w)) \right).$$

Substituting (58) in and simplifying gives

$$\begin{aligned} c''(q)q^4 + 2c'(q)q^3 < & -\frac{(e(\delta + \lambda_1 \bar{F}(w)))^2 (\delta + \lambda_1)^2}{2(p - \underline{w} - c(\bar{q}))\delta^2} + \left(\frac{c'(q)q^2 e(\delta + \lambda_1)}{2(p - \underline{w} - c(\bar{q}))} \right) \\ & - \left(\frac{(c'(q)q^2)^2 (\delta + \lambda_1)}{(\delta + \lambda_1 \bar{F}(w)) 2(p - \underline{w} - c(\bar{q}))} \right). \end{aligned}$$

Using the fact that $F(w) \in [0, 1]$, we can tighten the bound as

$$c''(q)q^4 + 2c'(q)q^3 < -\frac{(e(\delta + \lambda_1))^2 (\delta + \lambda_1)^2}{2(p - \underline{w} - c(\bar{q}))\delta^2} + \left(\frac{c'(q)q^2 e(\delta + \lambda_1)}{2(p - \underline{w} - c(\bar{q}))} \right) - \left(\frac{(c'(q)q^2)^2 (\delta + \lambda_1)}{2\delta(p - \underline{w} - c(\bar{q}))} \right).$$

Let $\hat{\pi}$ denote $(p - \underline{w} - c(\bar{q}))$ and simplify the expression gives

$$c''(q)q^4 + 2c'(q)q^3 - c'(q)q^2 \left(\frac{e(\delta + \lambda_1)}{2\hat{\pi}} \right) < -\frac{(e(\delta + \lambda_1))^2 (\delta + \lambda_1)^2}{2\hat{\pi}\delta^2} - \left(\frac{(c'(q)q^2)^2 (\delta + \lambda_1)}{2\delta\hat{\pi}} \right).$$

Finally, let $\eta^{c'} = \frac{c''(q)q}{c'(q)}$ denote the elasticity of $c'(q)$ then the above expression gives an upper bound of $\eta^{c'}$ as

$$\begin{aligned} c'(q)q^3 \left(\eta^{c'} + 2 - \frac{e(\delta + \lambda_1)}{2\hat{\pi}q} + \frac{c'(q)q(\delta + \lambda_1)}{2\hat{\pi}\delta} \right) & < -\frac{e^2(\delta + \lambda_1)^4}{2\hat{\pi}\delta^2}. \\ (60) \quad \eta^{c'} & < -2 - \frac{e^2(\delta + \lambda_1)^4}{2\hat{\pi}\delta^2 c'(q)q^3} + \frac{e(\delta + \lambda_1)}{2\hat{\pi}q} \\ & & - \frac{c'(q)q(\delta + \lambda_1)}{2\hat{\pi}\delta}. \end{aligned}$$

Finally, assuming $\lambda_0 = \lambda_1$ then $\hat{\pi}$ only depends on exogenous parameters. If the monitoring cost satisfies (60) for all $q \leq q^*$ where $q^* \in (q, \bar{q})$, then f is declining over the upper range of the

distribution. The condition derived here ensures that f has a declining right tail. The fact that $f(w) = 0$ combined with f declines over the upper range of the support implies that there are regions of w where f increases, however this does not necessarily imply that f is unimodal. The wage offer distribution may have multiple modes.

The remaining task is to derive $q^* \in (\underline{q}, \bar{q})$ in terms of exogenous parameters. Since \underline{q} depends on endogenous variables, q^* is derived as some arbitrary interior $q \in (\underline{q}, \bar{q})$ in terms of exogenous parameters. I express q^* as $\bar{q} - D$ where D depends only on exogenous parameters and $0 < D < \bar{q} - \underline{q}$ so that $q^* = \bar{q} - D \in (\underline{q}, \bar{q})$. Using (8) and assuming that $\lambda_0 = \lambda_1$, we have

$$\begin{aligned} \bar{w} - \underline{w} &= \frac{\delta e(\bar{q} - \underline{q})}{\bar{q}\underline{q}} + \lambda_0 B(\bar{w}) \\ (61) \quad \frac{\delta e(\bar{q} - \underline{q})}{\bar{q}\underline{q}} &= \bar{w} - \underline{w} - \lambda_0 B(\bar{w}) \end{aligned}$$

Now

$$\begin{aligned} \bar{w} - \underline{w} - \lambda_0 B(\bar{w}) &= \int_{\underline{w}}^{\bar{w}} 1 dx - \int_{\underline{w}}^{\bar{w}} \frac{\lambda_1 \bar{F}(x)}{\delta + \lambda_1 \bar{F}(x)} dx \\ &= \int_{\underline{w}}^{\bar{w}} \frac{\delta}{\delta + \lambda_1 \bar{F}(x)} dx \\ &> \int_{\underline{w}}^{\bar{w}} \frac{\delta}{\delta + \lambda_1} dx \\ (62) \quad &= \frac{\delta(\bar{w} - \underline{w})}{\delta + \lambda_1}. \end{aligned}$$

Substituting (62) into (61) gives

$$\begin{aligned} \frac{\delta e(\bar{q} - \underline{q})}{\bar{q}\underline{q}} &> \frac{\delta(\bar{w} - \underline{w})}{\delta + \lambda_1} \\ \bar{q} - \underline{q} &> \frac{\bar{q}\underline{q}(\bar{w} - \underline{w})}{e(\delta + \lambda_1)}. \end{aligned}$$

Replacing \bar{q} with \underline{q} on RHS to find a smaller distance between \bar{q} and \underline{q} , we have

$$(63) \quad \bar{q} - \underline{q} > \frac{\underline{q}^2(\bar{w} - \underline{w})}{e(\delta + \lambda_1)}.$$

Now \bar{w} depends on endogenous variables, but we can express its lower bound as follows. First, use (22), we have

$$\begin{aligned} \bar{w} &= p - c(\underline{q}) - \frac{(p - \underline{w} - c(\bar{q}))\delta^2}{(\delta + \lambda_1)^2} \\ (64) \quad &> p - c(\bar{q}) - \frac{(p - \underline{w} - c(\bar{q}))\delta^2}{(\delta + \lambda_1)^2}. \end{aligned}$$

Substituting (64) into (63) gives

$$(65) \quad \bar{q} - \underline{q} > \frac{q^2 (2\lambda_1 \delta + \lambda_1^2) \hat{\pi}}{e(\delta + \lambda_1)^3}$$

where $\hat{\pi} = (p - \underline{w} - c(\bar{q}))$. Finally, using (9) we can express \underline{q} as

$$\begin{aligned} \underline{q} &= \frac{e\delta}{\bar{w} - \frac{\lambda_0 e}{\bar{q}} - e - b - \lambda_0 B(\bar{w})} \\ &> \frac{e\delta}{p - \frac{\lambda_0 e}{\bar{q}} - e - b}. \end{aligned}$$

So (65) can be expressed as

$$\bar{q} - \underline{q} > \frac{(e\delta)^2 (2\lambda_1 \delta + \lambda_1^2) \hat{\pi}}{\left(p - \frac{\lambda_0 e}{\bar{q}} - e - b\right)^2 e(\delta + \lambda_1)^3}.$$

The expression for q^* at which f declines $\forall q \geq q^*$ where $q^* \in (\underline{q}, \bar{q})$ can be given by

$$q^* = \bar{q} - \frac{(e\delta)^2 (2\lambda_1 \delta + \lambda_1^2) \hat{\pi}}{\left(p - \frac{\lambda_0 e}{\bar{q}} - e - b\right)^2 e(\delta + \lambda_1)^3}.$$

References

- ABOWD, J. M. AND F. KRAMARZ, “The costs of hiring and separations,” *Labour Economics* 10 (2003), 499–530.
- ALBRECHT, J. W. AND B. AXELL, “An Equilibrium Model of Search Unemployment,” *Journal of Political Economy* 92 (October 1984), 824–40.
- ALBRECHT, J. W. AND S. B. VROMAN, “Nash equilibrium efficiency wage distributions,” *International Economic Review* (1998), 183–203.
- BAGGER, J., B. CHRISTENSEN AND D. MORTENSEN, “Wage and Productivity Dispersion: The Roles of Rent Sharing, Labor Quality, and Capital Intensity,” Technical Report, working paper, 2011.
- BOARD, S. AND M. MEYER-TER VEHN, “Relational contracts in competitive labour markets,” *The Review of Economic Studies* 82 (2015), 490–534.
- BONTEMPS, C., J.-M. ROBIN AND G. J. VAN DEN BERG, “Equilibrium Search with Continuous Productivity Dispersion: Theory and Nonparametric Estimation,” *International Economic Review* 41 (2000), 305–358.

- BRUNELLO, G., “The Relationship between Supervision and Pay: Evidence from the British New Earnings Survey,” *Oxford Bulletin of Economics and Statistics* 57 (1995), 309–21.
- BURDETT, K. AND D. T. MORTENSEN, “Wage Differentials, Employer Size, and Unemployment,” *International Economic Review* 39 (May 1998), 257–73.
- CAMERA, G. AND C. SELCUK, “Price Dispersion with Directed Search,” Purdue University Economics Working Papers 1173, Purdue University, Department of Economics, December 2004.
- COLES, M., “Equilibrium Wage Dispersion, Firm Size, and Growth,” *Review of Economic Dynamics* 4 (2001), 159 – 187.
- GREENWOOD, J., J. M. SANCHEZ AND C. WANG, “Financing Development: The Role of Information Costs,” *American Economic Review* 100 (September 2010), 1875–91.
- HWANG, H.-S., D. T. MORTENSEN AND W. R. REED, “Hedonic Wages and Labor Market Search,” *Journal of Labor Economics* 16 (October 1998), 815–47.
- IDSON, T. L. AND D. J. FEASTER, “A Selectivity Model of Employer-Size Wage Differentials,” *Journal of Labor Economics* 8 (January 1990), 99–122.
- KATZ, L. F. AND D. H. AUTOR, “Changes in the Wage Structure and Earnings Inequality,” in O. C. Ashenfelter and D. Card, eds., *Handbook of Labor Economics* volume 3, Part A, chapter 26 (Elsevier, 1999), 1463 – 1555.
- KRAMARZ, F. AND M.-L. MICHAUD, “The shape of hiring and separation costs in France,” *Labour Economics* 17 (2010), 27–37.
- MARKLEY, N. G., *Principles of Differential Equations* (John Wiley & Sons, Inc., 2004).
- MORTENSEN, D., *Wage dispersion: why are similar workers paid differently?* (MIT press, 2005).
- MORTENSEN, D. T., “Equilibrium unemployment with wage posting: Burdett-Mortensen meet Pissarides,” *CONTRIBUTIONS TO ECONOMIC ANALYSIS* 243 (2000), 281–292.
- MURRAY, F. J. AND K. S. MILLER, *Existence theorems for ordinary differential equations* (Robert E. Krieger Publishing Company, 1976).
- NAGYPÁL, É., “Worker reallocation over the business cycle: The importance of employer-to-employer transitions,” *Manuscript, Northwestern Univ* (2008).
- NEAL, D., “Supervision and wages across industries,” *The Review of Economics and Statistics* 75 (1993), 409–17.
- POSTEL-VINAY, F. AND J.-M. ROBIN, “Equilibrium Wage Dispersion with Worker and Employer Heterogeneity,” *Econometrica* 70 (November 2002a), 2295–2350.

- , “The Distribution of Earnings in an Equilibrium Search Model with State-Dependent Offers and Counteroffers,” *International Economic Review* 43 (2002b), 989–1016.
- ROSEN, S., “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition,” *Journal of Political Economy* 82 (1974), pp. 34–55.
- SHAPIRO, C. AND J. E. STIGLITZ, “Equilibrium Unemployment as a Worker Discipline Device,” *American Economic Review* 74 (June 1984), 433–44.
- STRAND, J., “Unemployment as a discipline device with heterogeneous labor,” *The American Economic Review* 77 (1987), 489–493.
- WALSH, F., “A Multisector Model of Efficiency Wages,” *Journal of Labor Economics* 17 (April 1999), 351–76.