A list of Working Papers on the last pages

No. 267, 1990

DUAL LABOR MARKETS, EFFICIENCY WAGES, AND SEARCH

by

James W. Albrecht and Susan B. Vroman

November, 1990

Dual Labor Markets, Efficiency Wages, and Search

James W. Albrecht and Susan B. Vroman
Department of Economics
Georgetown University
Washington, DC 20057
May 1990

1. Introduction

In this paper we present an equilibrium model of a dual labor market in which firms endogenously separate into two sectors, even though the firms are identical *ex ante*. In the primary sector effort requirements are high, jobs are monitored, and workers are commensurately rewarded, while in the secondary sector effort requirements are low, monitoring is lax, and workers are poorly paid.

Our model is motivated in part by a resurgence of empirical interest in dualism in the labor market and in part by the recognition that an efficiency wage model might be able to provide a theoretical underpinning for dual labor markets, heretofore an almost purely empirical construct. Empirical work supporting a dual labor market view of the US economy is presented in Dickens and Lang [1985]; Heckman and Hotz [1986] take a position strongly to the contrary. Regardless of one's position on the controversy, the point is that this old question is once again being debated in the mainstream journals.

This empirical revival ultimately has its roots in theoretical developments. Proponents of efficiency wage models tend to view that theory as a useful tool for explaining the "rigidities" that used to be largely in the domain of old-style macroeconomists and "institutionalists." In that spirit Bulow and Summers [1986] constructed an efficiency wage theory of dual labor markets, a concept which, prior to their paper, was located squarely in the institutionalist bailiwick. In their model the technology associated with a particular job is specified exogenously. Assuming the existence of two types of jobs (secondary sector jobs, which are sufficiently menial as to not require supervision, and primary sector jobs in which supervision is required), they use efficiency wage theory to explain how an outcome in which equally skilled workers are paid different wages can persist as an equilibrium.

Our approach is complementary to that of Bulow and Summers. Rather than exploring the implications of an assumed technological duality, we instead generate that duality as an equilibrium outcome. Our approach thus addresses the obvious question that is begged in Bulow and Summers, namely, why there should be two as opposed to one or "many" or a continuum of job types. It is also consistent with the institutionalist dual labor market literature, which emphasizes the endogeneity of technology choice.

The key assumption behind our dual labor market outcome is a nonconvexity in the monitoring technology. It is this nonconvexity that allows firms, all identical in the sense of have access to the same technology, to be indifferent in equilibrium between offering the two job types. Specifically, we assume that a firm can observe costlessly whether its workers are exerting effort at an exogenously specified minimum level, but that effort above the minimum level can be monitored only imperfectly and at a cost. This is consistent with intuition: one can observe costlessly whether the receptionist is answering the telephone; observing whether he or she is more or less helpful to those who call requires the expenditure of time and effort.

Like Bulow and Summers, we use the Shapiro and Stiglitz [1984] model, one of the standards of the efficiency wage literature, as our starting point. However, our setup differs from that of Shapiro and Stiglitzz in one significant respect in that we allow for heterogeneity among workers. Workers are not assumed to be heterogeneous with respect to their abilities since we want to preserve the Bulow and Summers outcome that in equilibrium equally skilled workers are paid different wages on different jobs. Instead, we assume workers differ according to the value placed on leisure. This assumed heterogeneity leads to a search equilibrium with two realistic properties: not all workers will accept secondary sector jobs, and some primary sector workers will shirk.

The outline of the rest of the paper is as follows. In the next section we set out the decision problem faced by workers. A worker has two choices to make. First, if unemployed and offered a job, he or she must decide whether to accept the job or remain unemployed. Second, if employed on a job with an effort requirement, the worker must decide whether to shirk or work. A job is distinguished by its wage offer, its effort requirement and a match-specific nonpecuniary component. We show that the job acceptance decision depends on the wage and the nonpecuniary component, while the shirk/work decision also depends on the effort requirement. Both decisions depend on the worker's type.

Our analysis of worker decisions is used in Section 3 to characterize the probabilities entering the firm's decision problem, namely, the probability that its job offer will be accepted and the probability that a worker, having accepted the job, will meet its effort requirement. These probabilities depend on the distribution of wage/effort requirement packages extant in the market and on the distribution of worker types across the unemployed.

The firm's decision problem is discussed in Section 4. A firm can observe costlessly whether its workers are exerting effort at the exogenously specified minimum level, but effort above this minimum cannot be monitored perfectly. The monitoring technology is exogenous: the cost of monitoring and the rate at which shirkers are detected are parameters of the firm's problem. Workers who are monitored and found to be putting forth less than the required level of effort are fired. The firm's problem is to decide what wage to offer, whether to monitor or not, and if it does monitor, what effort requirement to set. We establish that if a firm monitors, it never sets an effort requirement at or arbitrarily close to the minimum effort level. This implies that an equilibrium in which some firms monitor while others do not involves "separation"; ie, such an equilibrium must exhibit "dualism."

We construct the equilibrium of the model in Section 5. An equilibrium is a distribution of wages and effort requirements across vacancies together with a corresponding distribution of individual types across unemployed workers. This equilibrium is explicitly Nash: wages and effort requirements are set optimally by firms in conjunction with optimal search and effort decisions by workers. We focus on symmetric equilibria, ie, those in which firms that choose to monitor offer a common wage/effort requirement package and firms that choose not to monitor offer a common wage. A symmetric equilibrium is characterized by four variables, the wage for non-monitoring firms, the wage and effort requirement for monitoring firms, and the fraction of vacancies arising in firms that monitor. Three types of symmetric equilibria can arise: (i) pure secondary sector equilibria in which no firms monitor, (ii) pure primary sector equilibria in which all firms monitor, and (iii) dual labor market equilibria in which both types of behavior are optimal. We prove the existence of symmetric equilibria and show that dual labor market equilibria will arise given appropriate values for the cost of monitoring and the rate at which shirkers can be detected. Such equilibria will involve separation: the effort requirement of primary sector firms will not be arbitrarily close to that of secondary sector firms.

In sum, efficiency wage and search considerations lead to an equilibrium in which firms having access to the same technology ex ante choose to produce output in two distinct sectors. Secondary sector firms offer lower wages and do not monitor, while firms in the primary sector monitor their workers and pay a higher wage in order to elicit greater productivity. This result is discussed in the final section.

2. Workers

We begin with the decision problem faced by workers. This decision problem has two aspects. First, if unemployed and offered a job paying a wage of w, providing a nonpecuniary benefit of ξ,[1] and requiring an effort of e, ie, a "(w,ξ,e) job," does the worker accept or reject the offer? Second, if the offer is accepted, does the worker shirk or meet the job's effort requirement? The decision rules determining these choices vary with the worker's value of leisure, which is the worker's private information. The value of leisure, θ, varies across workers according to the continuously differentiable density, $f(\theta)$. This density and $g(\xi)$, likewise a continuously differentiable density, are the fundamental exogenous elements of our model and are assumed to be common knowledge.

We examine worker behavior using three value functions: (i) the value of meeting the effort requirement on a (w,ξ,e) job, (ii) the value of shirking on a (w,ξ,e) job, and (iii) the value of unemployment. These value functions vary with the worker's type, θ. We establish four results in this section. First, we prove that the worker's problem is well-posed, ie, we verify that the three value functions are defined by contractions. Second, we show that the decision to accept or reject a job offer is independent of the job's effort requirement; ie, the job acceptance decision depends only on w and ξ and on the worker's type, θ. Third, for given w and θ, there is a critical value of ξ such that

---

[1]Upon receiving a job offer, the worker discovers the value of its nonpecuniary component, ξ. The random variable ξ is iid across matches, is independent of the worker's value of leisure, and its realization is the worker's private information. Its inclusion in the model is required for the existence of symmetric equilibrium; without ξ, the firms' payoff functions fail to be concave in the wage and effort requirement. This point is discussed in the context of equilibrium search theory in Albrecht and Vroman [1989].

jobs offering w with an associated ξ greater than or equal to this critical value are accepted by workers of type θ; jobs with a lower ξ are rejected. We show that this critical value is continuously differentiable, decreasing in w, and increasing in θ. Finally, for given w, e, and θ, there is an analogous critical value of ξ for the work/shirk decision. A worker of type θ meets the effort requirement on a (w,ξ,e) job if ξ is greater than or equal to this critical value; otherwise the worker shirks. We show that this critical value is also continuously differentiable in its arguments, decreasing in w, and increasing in e and θ.

The details of the model are as follows. Workers live forever. Time is continuous, and the future is discounted at the rate r. Utility is derived from the rate at which income and nonpecuniary benefits are received and disutility from the rate at which effort is expended. The rate at which effort is supplied is a choice variable, bounded below by the minimum level, which we normalize to 1;[2] the effort level is zero for the unemployed. A worker employed on a (w,ξ,e) job and meeting the effort requirement enjoys an instantaneous utility of $w + \xi - e$; ie, the worker's utility in an interval of time of length Δt is given by $[w + \xi - e]\Delta t + o(\Delta t)$. If the worker shirks on that job, his or her instantaneous utility is given by $w + \xi - e^*$, where $e^* < e$. Finally, unemployment generates an instantaneous utility of θ.

A worker on a (w,ξ,e) job must decide what effort rate to supply. A worker not meeting the effort requirement faces a separation risk of μ; ie, the probability of a separation in an interval of time of length Δt equals $\mu\Delta t + o(\Delta t)$. This separation risk is independent of how far below the requirement the

[2]We also assume that there is an upper bound on effort. This natural assumption ensures that the set of policies open to the firm is compact, as will be required by our existence proof.

worker's effort falls; so, if a worker decides to shirk, he or she never exerts more than the minimum level of effort, ie, $e^* = 1$. A worker meeting the effort requirement suffers a separation risk of $\delta < \mu$.

The worker's effort choice is then simply one of whether or not to shirk. The value of not shirking on a (w,ξ,e) job is:

$$V_N(w,\xi,e;\theta) = \frac{1}{1+r\Delta t}[(w+\xi-e)\Delta t + \delta\Delta t U(\theta) + (1-\delta\Delta t)V_N(w,\xi,e;\theta) + o(\Delta t)].$$

The non-shirking worker gets an instantaneous utility of $(w+\xi-e)\Delta t + o(\Delta t)$. With probability $\delta\Delta t + o(\Delta t)$ the worker loses the job, in which case he or she becomes unemployed with associated value $U(\theta)$; otherwise the value $V_N(w,\xi,e;\theta)$ is retained. Rearranging, dividing through by Δt, and taking the limit as $\Delta t \to 0$ yields:

(1) $$V_N(w,\xi,e;\theta) = \frac{w+\xi-e}{r+\delta} + \frac{\delta}{r+\delta} U(\theta).$$

The corresponding value of shirking is:

(2) $$V_S(w,\xi;\theta) = \frac{w+\xi-1}{r+\mu} + \frac{\mu}{r+\mu} U(\theta).$$

The job's effort requirement does not enter $V_S(\cdot)$. Note that if the firm sets $e = 1$, the shirk/no-shirk distinction disappears. The separation risk is δ, and the value of having the job is given by (1) with $e = 1$.

Next, consider an unemployed worker. Suppose job offers arrive at the rate α. This arrival rate is exogenous and reflects the underlying matching technology. Then the value of unemployment to a worker of type θ is:

(3) $$U(\theta) = \frac{\theta}{r+\alpha} + \frac{\alpha}{r+\alpha} \mathrm{E}\max[A(w,\xi,e;\theta),U(\theta)],$$

where

$$A(w,\xi,e;\theta) = \max[V_N(w,\xi,e;\theta),V_S(w,\xi;\theta)].$$

The expectation in (3) is taken with respect to the joint distribution of wages and effort requirements across all vacancies, $H(w,e)$, and with respect to the distribution of ξ across all matches, $G(\xi)$. The value of unemployment to a worker of type θ thus depends on all the (w,e) offers extant in the market.

Since $V_N(\cdot)$ and $V_S(\cdot)$ depend on $U(\cdot)$, they also depend on $H(w,e)$ and $G(\xi)$.

We now verify that the worker's decision problem is well-defined.

Proposition 1: For any joint distribution $H(w,e)$ across vacancies coupled with any $G(\xi)$ and for any offer arrival rate α, there exist unique value functions $V_N(w,\xi,e;\theta)$, $V_S(w,\xi;\theta)$, and $U(\theta)$.

Proof: Given in the Appendix.

The expression for $U(\theta)$ given by (3) incorporates the worker's job acceptance decision rule: a (w,ξ,e) job is accepted iff $A(w,\xi,e;\theta) \geq U(\theta)$. We can now show:

Proposition 2: The job acceptance decision is independent of the job's effort requirement.

Proof: Given in the Appendix.

A worker accepts a job if either $V_N(w,\xi,e;\theta) \geq U(\theta)$ or $V_S(w,\xi;\theta) \geq U(\theta)$. Since the condition $V_S(w,\xi;\theta) \geq U(\theta)$ is more easily satisfied than $V_N(w,\xi,e;\theta) \geq U(\theta)$, the acceptance decision is determined only by w and ξ. The intuition for this result is as follows. Consider a worker on the accept/reject margin. If the worker accepts the job, it is a matter of indifference to the worker whether the job is retained or lost. He or she therefore has no incentive to put forth the required effort.[3]

The "acceptance condition" (AC) is thus $V_S(w,\xi;\theta) \geq U(\theta)$, ie:

(4) AC: $w + \xi \geq rU(\theta) + 1$.

Let $\xi_A(w,\theta)$ be defined by:

(5) $\xi_A = rU(\theta) + 1 - w$;

that is, $\xi_A(w,\theta)$ is the acceptance value of ξ for a person of type θ

---

[3]This is similar to the "Dougal result," established in Burdett and Mortensen [1980]. They show that the layoff risk on a marginally acceptable job does not influence the chance that such a job is accepted.

considering a (w,ξ,e) job. The job is accepted if $\xi \geq \xi_A$ and rejected otherwise; equivalently, $\xi_A(w,\theta)$ satisfies $V_S(w,\xi_A(w,\theta);\theta) = U(\theta)$. The properties of $\xi_A(w,\theta)$ are crucial for the acceptance probability that enters the firm's problem.

Proposition 3: The critical value $\xi_A(w,\theta)$ is continuously differentiable, decreasing in w, and increasing in θ.

Proof: Given in the Appendix.

The second aspect of the worker's decision problem, the shirk/no-shirk decision, is characterized by an analogous critical value. The "no-shirk condition" (NSC) is $V_N(w,\xi,e;\theta) \geq V_S(w,\xi;\theta)$; or,

(6) NSC: $w + \xi \geq rU(\theta) + 1 + (\frac{r+\mu}{\mu-\delta})(e-1)$.

Let $\xi_N(w,e,\theta)$ be defined by:

(7) $\xi_N = rU(\theta) + 1 - w + (\frac{r+\mu}{\mu-\delta})(e-1)$.

This critical value, $\xi_N(w,e,\theta)$, is the value of ξ with the property that an individual of type θ is indifferent between meeting the effort requirement on a (w,ξ_N,e) job and shirking; ie, $V_N[w,\xi_N(w,e,\theta),e;\theta] = V_S[w,\xi_N(w,e,\theta);\theta]$.

We also need to examine the properties of $\xi_N(w,e,\theta)$, as this critical value is key to the second probability entering the firm's decision, the probability that a worker who accepts the job will meet its effort requirement.[4]

Proposition 4: The critical value $\xi_N(w,e,\theta)$ is continuously differentiable in its arguments, decreasing in w, and increasing in e and θ.

Proof: Given in the Appendix.

[4]Since $\xi_N(w,e,\theta) > \xi_A(w,\theta)$ for all $e > 1$, the shirk/no-shirk and acceptance decisions are independent in the sense that the unconditional probability that a worker will meet the effort requirement on an offered job is the same as the probability that he or she will meet that requirement conditional on acceptance.

3. The Acceptance and No-Shirk Probabilities

In the preceding section we developed two critical values to characterize worker decision rules. In this section we use these critical values to develop expressions for the probabilities entering the firm's problem. Let q(w) denote the probability that an applicant accepts a job offering a wage of w, and let p(w,e) be the probability that an applicant who accepts the job meets its effort requirement. The acceptance probability is given by one minus the distribution function of ξ evaluated at $\xi_A(w,\theta)$ and integrated against the density of θ among the unemployed, ie,

(8) $$q(w) = \int [1 - G(\xi_A)] f_U(\theta) d\theta = \int \{1 - G[rU(\theta)+1-w]\} f_U(\theta) d\theta.$$

The no-shirk probability is given by the same expression, evaluated at $\xi_N(w,e,\theta)$ and similarly integrated against $f_U(\theta)$. That is,

(9) $$p(w,e) = \int \{1 - G[rU(\theta)+1-w+(\frac{r+\mu}{\mu-\delta})(e-1)]\} f_U(\theta) d\theta.$$

The acceptance probability is a function of w alone, while the no-shirk probability depends on both w and e. The applicant's two decisions depend on ξ; however, q(w) and p(w,e) are probabilities viewed from the firm's point of view, and ξ is not under the firm's control. Indeed, the nonpecuniary component is the worker's private information.

A key point in deriving q(w) and p(w,e) is that the density function of θ among the unemployed, $f_U(\theta)$, and the corresponding population density, $f(\theta)$, are not the same. Individuals with higher values of θ are overrepresented among the unemployed since they are more likely to shirk (and be fired) and less likely to accept a given job. By definition:

(10) $$f_U(\theta) \equiv P[\Theta = \theta | \text{unemployed}].$$

By Bayes Rule:

(11) $$f_U(\theta) = P[\text{unemployed} | \Theta = \theta] \cdot P[\Theta = \theta] / P[\text{unemployed}]$$
$$= u(\theta) f(\theta) / u,$$

where $u(\theta)$ is the θ-specific unemployment rate and $u = \int u(\theta)f(\theta)d\theta$ is the overall unemployment rate.

The derivation of $u(\theta)$ is tedious and is given in the Appendix. We show that:

(12) $u(\theta) = \mu\delta/\{\mu\delta + \alpha\mu p^*(\theta) + \alpha\delta[q^*(\theta)-p^*(\theta)]\}$.

where $q^*(\theta) = \int\{1-G[\xi_A(w,\theta)]\}dH(w,e)$ and $p^*(\theta) = \int\{1-G[\xi_N(w,e,\theta)]\}dH(w,e)$.

We now have all the elements of the contaminated distribution of θ among the unemployed, and can show:

Proposition 5: The density $f_U(\theta)$ is continuously differentiable in θ.

Proof: Given in the Appendix.

From Propositions 3, 4, and 5, we have the following results:

Proposition 6: The acceptance probability $q(w)$ is continuously differentiable and increasing.

Proposition 7: The no-shirk probability $p(w,e)$ is continuously differentiable in both its arguments, increasing in w, and decreasing in e.

Proposition 7 shows that our model has a standard efficiency wage property, namely, higher wages reduce shirking.

4. Jobs

A firm must decide what wage to offer, whether to monitor or not, and, if it does monitor, what effort requirement to set. We first set up the maximization problem for a firm that chooses not to monitor; then we consider the analogous problem for a firm that does monitor. The choice of monitoring versus not monitoring is determined by comparing the values associated with the two strategies. We verify that the firm's problem is well-defined and establish a simple separation result.

A firm consists of a large number of jobs. A job is either occupied or vacant and entails a fixed cost at the rate c, whether occupied or vacant. If occupied, output equals the effort of the worker in the job. Effort is the only input. There is independence across jobs in the sense that a firm's aggregate output is just the sum of the outputs of its jobs. This means that we can treat the job as the basic unit of analysis. All jobs are identical *ex ante*, but production may vary *ex post* across jobs because different (w,e) packages in conjunction with the randomly drawn ξ can elicit variations in worker effort. Entry and exit costs are assumed to be zero; thus, new jobs will be created if the value of a vacancy is positive and eliminated if this value is negative.

The minimum effort level can be ensured costlessly in all occupied jobs. If a firm chooses to impose no effort requirement, it need only decide upon a wage offer. A wage offer of w will be accepted with probability q(w). If the wage offer is accepted, then the value to the firm of having a worker in that job is:

$$R(w) = \frac{1}{1+r\Delta t}[(1-w-c)\Delta t + \delta\Delta t\Pi + (1-\delta\Delta t)R(w) + o(\Delta t)].$$

This value is the sum of the instantaneous return, 1-w-c, realized over the interval of time Δt, and the future value. With probability $1-\delta\Delta t+o(\Delta t)$ the firm retains the worker and the associated value R(w). Otherwise, the firm loses the worker, and the job becomes vacant with associated value Π. Passing to the limit in the usual way gives:

$$(13)\quad R(w) = \frac{1-w-c}{r+\delta} + \frac{\delta}{r+\delta}\Pi.$$

With probability 1-q(w) the firm's wage offer is rejected. In this case the firm retains the value of a vacancy, Π. Thus, a firm that imposes no effort requirement chooses w to solve:

$$(14)\quad \max_{w}\ q(w)R(w) + [1-q(w)]\Pi.$$

Alternatively, a firm might attempt to achieve a higher level of effort on its jobs. In order to elicit a level of effort above the minimum, a firm must monitor. Monitoring entails a fixed cost incurred at the rate m while the job is occupied and enables the firm to detect shirking, albeit not necessarily immediately.[5] As discussed above, a shirking worker on a monitored job suffers a separation risk of $\mu > \delta$.

Consider a firm using the monitoring technology, offering a wage of w, and imposing an effort requirement of e. The firm's wage offer is accepted and the effort requirement is met with probability p(w,e). With probability q(w) - p(w,e) the job is accepted and the worker shirks. Finally, 1 - q(w) is the probability that the firm's offer is rejected.

The value of having a non-shirker on a (w,e) job is:

$$(15)\quad N(w,e) = \frac{e-w-c-m}{r+\delta} + \frac{\delta}{r+\delta}\Pi;$$

the value of having a shirker on the same job is:

$$(16)\quad S(w) = \frac{1-w-c-m}{r+\mu} + \frac{\mu}{r+\mu}\Pi.$$

The maximization problem faced by such a firm is thus:

$$(17)\quad \max_{w,e}\; p(w,e)N(w,e) + [q(w)-p(w,e)]S(w) + [1-q(w)]\Pi.$$

Finally, consider a vacant job. Suppose applicants for this job arrive at the rate λ.[6] Then the value of a vacancy is:

$$\Pi = \frac{1}{1+r\Delta t}[-c\Delta t + \lambda\Delta t B + (1-\lambda\Delta t)\Pi + o(\Delta t)],$$

where:

[5] We assume that each firm is sufficiently large as to preclude using its total output to infer whether a worker on a particular job is shirking or not.

[6] The arrival rate, λ, is endogenous. Its determination is discussed below.

$$(18)\quad B = \max\left[\max_{w}\{q(w)R(w) + (1-q(w))\Pi\},\right.$$
$$\left.\max_{w,e}\{p(w,e)N(w,e) + [q(w)-p(w,e)]S(w) + [1-q(w)]\Pi\}\right]$$

is the value to a firm of meeting a job applicant. The firm's decision of whether to monitor or not is incorporated in the value B. In the limit:

$$(19)\quad \Pi = \frac{-c}{r+\lambda} + \frac{\lambda}{r+\lambda}B.$$

We can now verify that the firm's decision problem is well-posed.

Proposition 8: There exist unique value functions $R(w)$, $N(w,e)$, $S(w)$, and Π.

Proof: Given in the Appendix.

We can also establish that if a firm decides to monitor, it never sets an effort requirement at or arbitrarily close to the minimum level, $e = 1$.

Proposition 9: If a firm chooses to monitor, it sets an effort requirement no less than $1+m$.

Proof: Given in the Appendix.

Proposition 9 implies that if some firms are monitoring while others are not, then any equilibrium must involve "separation." Thus, equilibria in which there is a continuum of job types are ruled out. That is, even if a continuum of wage/effort requirement packages were offered in the primary sector, the lowest effort requirement would not be arbitrarily close to the minimum (secondary sector) effort requirement, ie, the equilibrium will exhibit dualism.

Thus far, our discussion of firm behavior has been limited to jobs "in the market." To complete the firm side of the model we invoke the free entry/exit condition. A firm creates jobs so long as the value of a vacancy, Π, is positive; a firm eliminates vacancies from the market if $\Pi < 0$. This free entry/exit condition implies that Π must be zero in equilibrium. From equation (19) this condition also determines the equilibrium applicant arrival rate, namely:

(20) $\lambda = c/B$.

5. Equilibrium

In the preceding sections we characterized optimal behavior for workers in the face of any arbitrary distribution H(w,e) of offers across vacancies and for firms in the face of any arbitrary distribution $F_U(\theta)$ of worker types across the unemployed. An equilibrium is a distribution H(w,e) together with a corresponding distribution $F_U(\theta)$ that is Nash in the sense that H(w,e) reflects the optimizing behavior of firms given $F_U(\theta)$, while at the same time $F_U(\theta)$ reflects the optimizing behavior of workers given H(w,e).

The most general equilibria to consider are those in which some firms monitor and some do not. The possibility that a variety of wage/effort requirement packages might be offered by primary sector firms and that such dispersion might be self-supporting in equilibrium is not ruled out *a priori*. However, we focus our attention on symmetric equilibria. A symmetric equilibrium is one in which those firms that choose to monitor offer a common wage/effort requirement package; likewise, firms that choose not to monitor offer a common wage. Given suitable restrictions on the underlying exogenous distributions, we prove the existence of symmetric equilibria.

We denote the common primary sector package by (w_P, e_P), the wage offered by secondary sector firms by w_S, and the fraction of vacancies arising in firms that monitor by φ. A symmetric distribution H(w,e) is thus characterized by four variables, w_S, w_P, e_P, and φ. Symmetric equilibria in which some firms monitor and some do not, ie, $0 < \varphi < 1$, are (symmetric) "dual labor market equilibria." Two degenerate cases can also arise. If $\varphi = 0$, we have "pure secondary sector equilibria," in which no firms monitor;

if $\varphi = 1$, we have "pure primary sector equilibria," in which all firms monitor.

To prove the existence of a symmetric equilibrium we construct a map that takes any initial symmetric distribution $H^0(w,e)$ into a new symmetric distribution $H^1(w,e)$. To use Brouwer's Theorem to show that this map has a fixed point, we must establish that the map is continuous and defined on a compact set. As we have assumed an upper bound on effort, the set of possible quadruples $\{w_S, w_P, e_P, \varphi\}$ is closed and bounded. Thus, to prove existence we need to prove continuity.

The map taking H^0 to H^1 has three basic components. First, optimal behavior by workers generates the contaminated distribution $F_U(\theta)$ and the probabilities q(w) and p(w,e) that enter firms' decisions. Second, given q(w) and p(w,e), firms compute the optimal secondary sector wage offer, w_S, and the optimal primary sector package (w_P,e_P). Finally, given the updated (w_S, w_P, e_P), firms optimally allocate vacancies across sectors, producing an updated value of φ.

The continuity of the map taking H^0 to H^1 is established by demonstrating the continuity of each step. Thus, we need to first demonstrate the continuity of $F_U(\theta)$ and of q(w) and p(w,e) in the variables comprising H^0. Second, we need to demonstrate the continuity of the optimal secondary sector and primary sector choices in the variables comprising H^0. To do this we show that these optimizing values are unique, so that the Maximum Theorem can be applied to establish continuity. Finally, given optimal behavior in both sectors, we need to show that optimal sectoral choice generates a unique φ. These results are given in the following three propositions.

<u>Proposition 10</u>: The distribution $F_U(\theta)$ and the probabilities q(w) and p(w,e) are continuous in the variables comprising H^0.

<u>Proof</u>: Given in the Appendix.

Proposition 11: Given suitable restrictions on $f(\theta)$ and $g(\xi)$, there exists a unique solution w_S to the secondary sector firm maximization problem and a unique solution (w_P, e_P) to the primary sector firm maximization problem.

Proof: Given in the Appendix.

Proposition 12: Given $\{w_S, w_P, e_P\}$, there is a unique φ reflecting optimal sectoral choice by firms.

Proof: Given in the Appendix.

This completes our characterization of the map from the set of quadruples (w_S, w_P, e_P, φ) into itself. This map is a continuous function defined on a compact set, so we can apply Brouwer's Fixed Point Theorem. Thus, we have proven the following proposition.

Proposition 13: Given suitable restrictions on $f(\theta)$ and $g(\xi)$, a symmetric equilibrium exists.

The existence of a symmetric equilibrium does not guarantee *per se* the existence of a symmetric dual labor market equilibrium, ie, an equilibrium in which $0 < \varphi < 1$. However, such an equilibrium must exist for a range of values of the exogenous parameters of the model. The equilibrium quadruple (w_P, e_P, w_S, φ) depends on the underlying parameters of the monitoring technology, μ and m, in a continuous manner. Suppose $\varphi = 0$, so that all vacancies are in the secondary sector. If μ is increased and/or m is reduced by sufficient amounts then a dual market equilibrium arises. Similarly, if $\varphi = 1$, reducing μ and/or increasing m produces a dual market equilibrium.

6. Conclusion

In this paper we present a model in which a dual labor market arises as a Nash equilibrium. A key feature of our setup is that firms are identical *ex ante*. This distinguishes our model from that of Bulow and Summers [1986] in which the technology choices of firms are specified exogenously. While our approach complements that of Bulow and Summers, the endogeneity of technology choice makes our model consistent with an important theme of the institutionalist dual labor market literature.

Our equilibrium rests on both efficiency wage and equilibrium search considerations. This combination produces a useful extension of the basic Shapiro/Stiglitz model. In contrast to Shapiro/Stiglitz and Bulow/Summers, shirking by some, but not all, workers is a feature of our equilibrium. Of course, shirking in equilibrium follows in our model from the assumed heterogeneity of workers. Our model thus shows how some simple tools from equilibrium search theory can be used to incorporate worker heterogeneity into the Shapiro/Stiglitz model.

In proving existence we worked with symmetric equilibria. Whether asymmetric equilibria exist in this model is an open question. However, our focus on symmetric equilibria is not essential to the dual labor market result. Even if we were to allow for the possibility of a range of (w,e) combinations among "primary sector" firms, Proposition 9 establishes the required separation. The gap between the lowest primary sector effort requirement and the minimum effort level is ensured by the assumed nonconvexity in the monitoring technology.

This nonconvexity in the monitoring technology is the key assumption in our model. In the interest of making our point as cleanly as possible, the nonconvexity we used was a very stylized one. It is clear, however, that any fixed cost associated with setting up the monitoring technology could yield a dual outcome. Our model thus provides a theoretical basis for an essentially empirical construct that has been used by many labor economists. Although we have limited our attention to establishing the logical coherence of the setup, ie, to establishing the existence of equilibrium, the model could, with suitable modification, be used to analyze policy questions within a dual labor market context. For example, the effects of a minimum wage on the primary sector package and on the sizes of the two sectors could be examined.

Our model could be extended and refined in several directions. A particularly interesting possibility would be to explore the consequences of allowing firms to offer a menu of choices to workers, ie, to allow workers to self-select. However, even without extensions, the model performs its basic function of establishing the possibility of an endogenously generated dual labor market outcome.

References

1. J. Albrecht and S. Vroman, "Equilibrium in Search Models with Adverse Selection," Georgetown University Working Paper No. 89-8, April 1989.

2. J. Bulow and L. Summers, "A Theory of Dual Labor Markets with Application to Industrial Policy, Discrimination, and Keynesian Unemployment," Journal of Labor Economics, July 1986, part 1, pp 376-414.

3. K. Burdett and D. Mortensen, "Search, Layoffs, and Labor Market Equilibrium," Journal of Political Economy, August 1980, pp 652-72.

4. W. Dickens and K. Lang, "A Test of Dual Labor Market Theory," American Economic Review, September 1985, pp 792-805.

5. J. Heckman and V. Hotz, "An Investigation of the Labor Market Earnings of Panamanian Males: Evaluating the Sources of Inequality," Journal of Human Resources, Fall 1986, pp 507-42.

6. C. Shapiro and J. Stiglitz, "Involuntary Unemployment as a Worker Discipline Device," American Economic Review, June 1984, pp 433-44.

APPENDIX

Proof of Proposition 1: Substituting (1) and (2) into (3) gives a mapping of the form

$U(\theta) = T[U(\theta)]$. It is straightforward to check that $T(\cdot)$ is a contraction for each θ, ensuring the existence of a unique function $U(\theta)$. To establish the uniqueness of $V_N(\cdot)$ and $V_S(\cdot)$, substitute the unique $U(\cdot)$ into (1) and (2). QED.

Proof of Proposition 2:

$$V_N(w,\xi,e;\theta) = \frac{w+\xi-e}{r+\delta} + \frac{\delta}{r+\delta}U(\theta) \geq U(\theta) \text{ iff } w + \xi \geq rU(\theta) + e;$$
$$V_S(w,\xi;\theta) = \frac{w+\xi-1}{r+\mu} + \frac{\mu}{r+\mu}U(\theta) \geq U(\theta) \text{ iff } w + \xi \geq rU(\theta) + 1.$$

If either of these conditions is satisfied, the worker accepts. Since $V_S(w,\xi;\theta) \geq U(\theta)$ is more easily satisfied than $V_N(w,\xi,e;\theta) \geq U(\theta)$, the acceptance decision is determined only by w and ξ. QED.

Proof of Proposition 3: From (5) we have $\dfrac{\partial\xi_A(w,\theta)}{\partial w} = -1$ and $\dfrac{\partial\xi_A(w,\theta)}{\partial\theta} = rU'(\theta)$; thus, we need to show $rU'(\theta) > 0$. We can write $U(\theta)$ as:

$$U(\theta) = \frac{\theta}{r+\alpha} + \frac{\alpha}{r+\alpha}\{\int_{N(\theta)} V_N(w,\xi,e;\theta)dH(w,e)dG(\xi) + \int_{S(\theta)} V_S(w,\xi;\theta)dH(w,e)dG(\xi) + \int_{R(\theta)} U(\theta)dH(w,e)dG(\xi)\},$$

where $N(\theta)$ is the no-shirk region for an individual of type θ; ie,

$$N(\theta) = \{(w,\xi,e)\colon V_N(w,\xi,e;\theta) \geq V_S(w,\xi;\theta)\}.$$

The shirk region, $S(\theta)$, and the reject region, $R(\theta)$, are defined analogously. Using (1) and (2) and rearranging gives:

$$rU(\theta) = \theta + \alpha\{\int_{N(\theta)} [\frac{w+\xi-e-rU(\theta)}{r+\delta}]dH(w,e)dG(\xi) + \int_{S(\theta)} [\frac{w+\xi-1-rU(\theta)}{r+\mu}]dH(w,e)dG(\xi)$$

Differentiating and collecting terms gives:

$$rU'(\theta) = 1/[1 + \frac{\alpha}{r+\delta}\int_{N(\theta)} dH(w,e)dG(\xi) + \frac{\alpha}{r+\mu}\int_{S(\theta)} dH(w,e)dG(\xi)] > 0.$$ QED

Proof of Proposition 4: From (7) we have:

$$\frac{\partial\xi_N(w,e,\theta)}{\partial w} = -1, \quad \frac{\partial\xi_N(w,e,\theta)}{\partial e} = \frac{r+\mu}{\mu-\delta}, \text{ and } \frac{\partial\xi_N(w,e,\theta)}{\partial\theta} = rU'(\theta_N).$$ In the proof of Proposition 3 we showed that $rU'(\theta) > 0$ for all θ. QED.

Derivation of $u(\theta)$

We use steady-state flow conditions to derive the θ-specific unemployment rates. For each θ, (i) the rates of flow of nonshirkers into and out of unemployment must be equal and (ii) the corresponding rates for shirkers must also be equal. Let $n(\theta)$ denote the probability that an individual of type θ is employed and not shirking; let $s(\theta)$ be the probability that he or she is employed and shirking. The rate of flow into unemployment of non-shirkers of type θ is $\delta n(\theta)$; the corresponding rate of flow for shirkers of type θ is $\mu s(\theta)$.

The flows out of unemployment of workers of type θ consist of new hires. The flow of offers to unemployed workers of type θ is $\alpha[1-n(\theta)-s(\theta)]$. To compute the flow rates of new hires this offer arrival rate needs to be multiplied by the relevant acceptance probability. Let $q^*(\theta)$ denote the probability that an unemployed worker of type θ accepts a random offer. This probability is:

(A1) $q^*(\theta) = \int\{1-G[\xi_A(w,\theta)]\}dH(w,e)$.

Similarly, let $p^*(\theta)$ be the probability that a worker accepts a job and chooses to meet its effort requirement. This probability is:

(A2) $p^*(\theta) = \int\{1-G[\xi_N(w,e,\theta)]\}dH(w,e)$.

The rate of flow of workers of type θ into jobs on which they will not shirk is thus $\alpha[1-n(\theta)-s(\theta)]p^*(\theta)$. In steady-state this must equal $\delta n(\theta)$. The flow of new hires who shirk is the difference between the total outflow from unemployment and the outflow of nonshirkers; thus, the flow of workers of type θ into jobs on which they will shirk is $\alpha[1-n(\theta)-s(\theta)][q^*(\theta)-p^*(\theta)]$. This expression must equal $\mu s(\theta)$ in equilibrium.

Equating flow rates into and out of unemployment for workers of type θ gives:

$$\delta n(\theta) = \alpha[1-n(\theta)-s(\theta)]p^*(\theta)$$

$$\mu s(\theta) = \alpha[1-n(\theta)-s(\theta)][q^*(\theta)-p^*(\theta)].$$

These two steady-state conditions, plus the identity $n(\theta)+s(\theta)+u(\theta) = 1$, yield:

$$u(\theta) = \mu\delta/\{\mu\delta + \alpha\mu p^*(\theta) + \alpha\delta[q^*(\theta)-p^*(\theta)]\}.$$

Proof of Proposition 5: The population density $f(\theta)$ is continuously differentiable by assumption. To show that $f_U(\theta)$ is continuously differentiable, we need to show that $q^*(\theta)$ and $p^*(\theta)$ have this property. By Propositions 3 and 4, ξ_A and ξ_N are continuously differentiable in θ; by assumption, G is continuously differentiable in its argument. The continuous differentiability of $q^*(\theta)$ and $p^*(\theta)$ then follows directly from equations (A1) and (A2). QED

Proof of Proposition 8: Substituting the expression for B into (19) gives an expression of the form $\Pi = T(\Pi)$. It is straightforward to check that $T(\cdot)$ is a contraction, ensuring the existence of a unique value Π. To establish the corresponding properties for R(w), N(w,e) and S(w), substitute the unique Π

into (13), (15), and (16). QED.

Proof of Proposition 9: Suppose that a firm is monitoring and its optimal effort requirement is e. The value to the firm of a nonshirker is $N(w,e) = \frac{e-w-c-m}{r+\delta} + \frac{\delta}{r+\delta}\Pi$. The value of a worker to a firm that does not monitor is $R(w) = \frac{1-w-c}{r+\delta} + \frac{\delta}{r+\delta}\Pi$. A minimum requirement for monitoring to be profitable is that $N(w,e) > R(w)$ since the monitoring firm will also have some shirking workers. For a given w, $(r+\delta)[N(w,e) - R(w)] =$

$e - 1 - m$. That is, for $N(w,e) - R(w) > 0$ it must be the case that $e > 1 + m$. In fact, the monitoring firm will have to set w higher than the non-monitoring firm or all its workers will shirk, so that e can never be less than 1 + m. QED.

Proof of Proposition 10: The continuity of $F_U(\theta)$ and of q(w) and p(w,e) in the variables comprising H^0 all depend on the continuity of $U(\theta)$ in those variables. In the case of a symmetric initial H, $U(\theta)$ is implicitly defined by:

$$\text{(A3)}\quad rU(\theta) - \theta - \alpha\Bigg[\varphi\int_{\xi_N}^{\infty}\Big[\frac{w_P+\xi-e_P-rU(\theta)}{r+\delta}\Big]g(\xi)d\xi + \varphi\int_{\xi_P}^{\xi_N}\Big[\frac{w_P+\xi-1-rU(\theta)}{r+\mu}\Big]g(\xi)d\xi + (1-\varphi)\int_{\xi_S}^{\infty}\Big[\frac{w_S+\xi-1-rU(\theta)}{r+\delta}\Big]g(\xi)d\xi\Bigg] = 0,$$

where $\xi_P \equiv \xi_A(w_P,\theta)$ and $\xi_S \equiv \xi_A(w_S,\theta)$. Differentiating with respect to w_S, w_P, e_P, and φ, and applying the Implicit Function Theorem gives the desired result. QED

Proof of Proposition 11: The optimal wage for the firm, should it choose not to monitor, maximizes $q(w)R(w) + [1 - q(w)]\Pi$, where $R(w) = \frac{1-w-c}{r+\delta} + \frac{\delta}{r+\delta}\Pi$. In considering the maximization problem we impose the long-run equilibrium condition $\Pi = 0$ in advance. Thus, the optimal wage w_S is unique if $q(w)R(w)$ is concave.

The first-order condition for this problem can be written:

$$q_w(w)[1-w-c] - q(w) = 0,$$

and the second-order condition is:

$$q_{ww}(w)[1-w-c] - 2q_w(w) < 0.$$

From the first-order condition, $1-w-c = q(w)/q_w(w)$; thus, the second-order condition can be written as:

$$q_{ww}(w)q(w) < 2q_w(w)^2.$$

The acceptance probability is:

$$q(w) = \int f_U(\theta)\{1 - G[rU(\theta)+1-w]\}d\theta.$$

Thus:

$$q_w(w) = \int f_U(\theta)g[rU(\theta)+1-w]d\theta$$
$$q_{ww}(w) = -\int f_U(\theta)g'[rU(\theta)+1-w]d\theta.$$

By Proposition 5, $f_U(\theta)$ is continuously differentiable; so, the condition on $q(w)$ required for the existence of a unique maximizing value w_S can be satisfied by suitably restricting $f(\theta)$ and/or $g(\xi)$. For example, the condition is satisfied if ξ is distributed as a uniform random variable.

Should the firm choose to monitor, its optimal wage/effort requirement package maximizes:

$$p(w,e)[N(w,e)-S(w)] + q(w)S(w).$$

Again, we impose the long-run equilibrium condition $\Pi = 0$ in advance, so $N(w,e) = \frac{e-w-c-m}{r+\delta}$ and $S(w) = \frac{1-w-c-m}{r+\mu}$. Sufficient conditions for the existence of a unique optimal wage/effort requirement package are:[7]

$$q_{ww}(w)q(w) < 2q_w(w)^2$$
$$p_{ww}(w,e)p(w,e) < 2p_w(w,e)^2$$
$$p_{ee}(w,e)p(w,e) < 2p_e(w,e)^2.$$

The no-shirk probability is:

$$p(w,e) = \int f_U(\theta)\{1 - G[rU(\theta)+1-w+(\tfrac{r+\mu}{\mu-\delta})(e-1)]\}d\theta.$$

The sufficient conditions for the existence of a unique (w_P, e_P) can be satisfied by suitably restricting $f(\theta)$ and/or $g(\xi)$. Again, taking ξ to be a uniform random variable suffices. QED

Proof of Proposition 12: In a dual labor market equilibrium, φ equates the value of meeting an applicant for a non-monitoring firm to the analogous value for a monitoring firm; ie, φ solves:

(A4) $q(w_S;\varphi)R(w_S) - p(w_P,e_p;\varphi)N(w_P,e_P) - [q(w_P;\varphi)-p(w_P,e_P;\varphi)]S(w_P) = 0.$

For this equation to yield a unique value of φ it is sufficient that the above equation be continuous and monotonic in φ. We have already shown that $q(\cdot)$ and $p(\cdot)$ are continuous in φ; thus, monotonicity is the property we need to verify.

Monotonicity of (A4) in φ is established if the derivative of the LHS of (A4) with respect to φ is of the same sign for all φ. This derivative is:

$$\frac{\partial q(w_S)}{\partial\varphi}[(r+\mu)(1-w_S-c)] - [\frac{\partial q(w_P)}{\partial\varphi} - \frac{\partial p(w_P,e_P)}{\partial\varphi}][(r+\delta)(1-w_P-c-m)]$$
$$- \frac{\partial p(w_P,e_P)}{\partial\varphi}[(r+\mu)(e_P-w_P-c-m)].$$

Since firms take $F_U(\theta)$ as given, the partials of $q(\cdot)$ and $p(\cdot)$ with respect to φ depend on $\frac{\partial U(\theta;\varphi)}{\partial\varphi}$. That is, variations in φ affect the acceptance and

[7]The fact that these 3 inequalities are sufficient to ensure that the second-order conditions are satisfied follows from the relationship $p_e(w,e) = -(\frac{r+\mu}{\mu-\delta})p_w(w,e)$. It is this equality that causes the cross-derivative terms to cancel.

no-shirk probabilities by affecting the value $U(\theta)$ among the unemployed. Differentiating (A3) with respect to φ, we find that the sign of this derivative depends on the difference between the value of an acceptable primary sector offer and the value of an acceptable secondary sector offer, which must be positive. That is, $\frac{\partial U(\theta;\varphi)}{\partial\varphi} > 0$. Given $\frac{\partial U(\theta;\varphi)}{\partial\varphi} > 0$, the partials of $q(\cdot)$ and $p(\cdot)$ with respect to φ depend only on the form of $g(\xi)$. With minimal restrictions on the distribution of ξ, (A4) is monotonic in φ. For example, if ξ is a uniform random variable, the above derivative is positive for all φ.

If there is no $\varphi \in [0,1]$ solving (A4), then we have a degenerate solution. If the LHS of (A4) is positive for all φ in this range, we have a pure secondary sector solution, ie,
$\varphi = 0$. If the LHS of (A4) is negative for all $\varphi \in [0,1]$, we have a pure primary sector solution, ie, $\varphi = 1$. QED