JOURNAL OF ECONOMIC THEORY 7, 188–209 (1974)

Equilibrium Search and Unemployment

ROBERT E. LUCAS, JR. AND EDWARD C. PRESCOTT

Graduate School of Industrial Administration, Carnegie-Mellon University Pittsburgh, Pennsylvania 15213

Received November 27, 1972

1. INTRODUCTION

Thirty years after the Great Depression, economists have again worked up the nerve to ask an obvious question: Why is it that workers *choose* (under some conditions) to be unemployed rather than to take employment at lower wage rates? Soon after serious attention began to be focused on this question, a variety of models were advanced to illustrate how workers might rationally prefer some other activity to work at wage rates they perceive to be temporarily below normal.[1] A particularly interesting class of models arises when the alternate activity is taken to be job search: The worker is faced with a wage offer which he views as a drawing from a probability distribution; his choices are to accept the offer or to take another drawing.[2] (To be of interest, obviously, these choices must be mutually exclusive: One must be unable to search and work at the same time.)

Most contributors to this literature on search behavior subscribe to some form of the Friedman–Phelps notion that there exists a natural rate of unemployment which either cannot or should not (or perhaps both) be lowered (on average) by monetary and fiscal policies.[3] Yet while the language used in discussing this natural rate suggests that it may have the properties of a competitive equilibrium, there exist no theoretical models in which a nonzero equilibrium unemployment rate is determined

[1] A number of these are collected in Phelps *et al.* [10]. The central ideas can be traced at least back to Hicks [5].

[2] For example, the Mortensen and Gordon–Hynes chapters of [10], McCall [8], and Stigler [13]. It is perhaps necessary to emphasize that the class of models in which active job search is the *only* alternative to work by no means exhausts the class of models in which unemployment is viewed as a "rational" choice.

[3] In addition to the above references, see Friedman [4].

and its properties studied.[4] Normatively, this means that there is no framework within which important welfare issues such as those raised by Phelps [11, chapter 4] and Tobin [14] can be formulated and analyzed. Empirically, it means (for example) that there is no theoretical account as to why average rates of unemployment vary so widely from one advanced capitalist economy to another.

Clearly, one cannot hope to deal with these questions by the study of the optimal search behavior of a single agent in the face of a *given* probability distribution of wage offers. The issues are those of market equilibrium and must be met in a theoretical context in which employment behavior and wages are *simultaneously* determined. As the reader who proceeds into the body of this paper will discover, this problem is more difficult than it sounds. Let us try to indicate why in the remainder of this introduction.

In order for wage rates for a single type of labor to differ at a point in time, labor must clearly be exchanged in spatially distinct markets. (Otherwise, wages would be bid into equality in a period much too short to be of economic interest.) The distribution of wage rates governing the worker's decision problem, referred to above, must then be related to his knowledge of the likely outcome of searching over these distinct markets. On the other hand, the distribution of wages over markets will evidently be influenced by the mobility of labor suppliers. In short, optimal labor supply behavior and the wage distribution on which it is based must be simultaneously determined within a model of *market* (as opposed to individual) behavior.

While quite analogous to the problem of using supply and demand schedules to determine price and quantity in a single market, this simultaneity problem is analytically more difficult for at least two reasons. First, since movement in space takes time, labor market search must be studied in a dynamic context. Second, the outcome of the process at each point in time will be a probability distribution rather than simply a number. The *solution* of the model will then be a *stochastic process.*

The rest of this paper is devoted to the elaboration of a complete "search model" of this general type. To preserve simplicity, the treatment will be abstract and illustrative. Discussion of the relationship of the theory to observed labor market behavior will be deferred to the conclusion of the paper and will there be brief.

[4] Lucas [7] provides an equilibrium in which employment fluctuates with aggregate demand. In this model, however, "unemployment" as an activity is not differentiated from "leisure" or other nonwork alternatives.

2. Structure of the Model

We think of an economy in which production and sale of goods occur in a large number of spatially distinct markets.[5] Product demand in each market shifts stochastically, driven by shocks which are independent over markets (so that aggregate demand is constant) but autocorrelated within a single market. Output to satisfy current period demand is produced in the current period, with labor as the only input. Each product market is competitive.

There is a constant workforce which at the beginning of a period is distributed in some way over markets. In each market, labor is allocated over firms competitively with actual money wages being market clearing. Each worker may either work at this wage rate, in which case he will remain in this market into the next period, or leave. If he leaves, he earns nothing this period but enters a "pool" of unemployed workers which are distributed in some way over markets for the next period. In this way, a new workforce distribution is determined, new demands are "drawn," and the process continues.

In this process, all agents are assumed to behave optimally in light of their objectives and the information available to them. For firms, this means simply that labor is employed to the point at which its marginal value product equals the wage rate. For workers, the decision to work or to search is taken so as to maximize the expected, discounted present value of the earnings stream. In carrying out this calculation, workers are assumed to be aware of the values of the variables affecting the market where they currently are (i.e., demand and workforce) and of the true probability distributions governing the future state of this market and the present and future states of all others. That is, expectations are taken to be *rational.*[6]

The economic interpretation of this assumption of rational expectations is that agents have operated for some time in a situation like the current one and have therefore built up experience about the probability distributions which affect them. For this to have meaning, these distributions must remain stable through time. Mathematically, this means that we will be concerned only with *stationary distributions* of demand and workforce and with behavior rules under these stationary distributions. Although sequences tending toward these stationary distributions will be utilized analytically, these seem to have no counterpart in observed behavior.

[5] See Phelps' introductory chapter in [10] for the description of the "island economy" which is the direct ancestor of the present model.

[6] In the sense of Lucas and Prescott [6] and Muth [9].

The task of the following sections may now be outlined in more detail. In the next section, we study the determination of equilibrium employment and wages in a single market, with the expected return to workers of *leaving* that market taken as a parameter. In Section 4, the stationary joint distribution of demand and workforce in this market is determined. In Section 5, we aggregate the workforce over markets to obtain the total economy-wide workforce as a function of the parametric expected return. This relationship serves as an aggregate demand function for labor; given a fixed total workforce, the *equilibrium* expected return is then determined in the usual way. Finally, Section 6 discusses a certain kind of stability possessed by this equilibrium, and concluding remarks are given in Section 7.

3. Equilibrium in a Single Market

In this section and the next, we study wage and employment determination in a single market, representing the impact of the rest of the economy on this market by certain given parameters. This impact takes three forms: first, product demand functions shift in an exogenously determined, stochastic manner; second, the outside economy offers alternative employment to workers; third, new workers arrive from the rest of the economy, augmenting the local work force. We discuss each effect in turn.

The individual market behaves as a Marshallian industry, faced with a demand function $p = D(s, Q)$, where p is price, Q is industry output, and s is a stochastic shift variable, *realized prior to trading*. Output is supplied by m identical firms, each with the production function $\varphi(n)$ depending on labor input only. The industry is competitive, so that the profit (and present value) maximizing policy for firms is to hire labor to the point at which the marginal value product of labor, $p\varphi'(n)$, equals the wage. When the product market is cleared, then, the function $R(s, n)$ defined by

$$R(s, n) = D(s, m\varphi(n))\, \varphi'(n)$$

gives the marginal value product of labor when demand is in state s and employment is n. Since R summarizes completely the demand side of the labor market, we shall discard the functions D and φ and place restrictions directly on R, as follows.

The function $R(s, n)$ is positive, differentiable, and bounded; its first derivatives satisfy

$$R_s(s, n) > 0, \qquad R_n(s, n) < 0. \tag{1}$$

For each fixed n,

$$\lim_{s \to 0} R(s, n) = 0. \tag{2}$$

For each fixed w, $0 < w \leqslant R(s, 0)$, the function $\hat{n}(s, w)$ defined by $R(s, \hat{n}(s, w)) = w$ satisfies

$$\lim_{s\to\infty} \hat{n}(s, w) = \bar{n}(w) < \infty. \tag{3}$$

The shift variable s follows a Markov process governed by

$$F(s', s) = \Pr\{s_{t+1} \leqslant s' \mid s_t = s\}.$$

For fixed s, F is a cumulative distribution function on $s' > 0$, with the continuous, strictly positive density $f(s', s)$. For fixed s', F is a strictly decreasing function of s on $s > 0$; further, if g is continuous,

$$\lim_{s\to 0} \int g(s') f(s', s)\, ds' = \lim_{s\to 0} g(s), \tag{4}$$

and if g is also positive and nondecreasing,

$$\lim_{s\to\infty} \int g(s') f(s', s)\, ds' \leqslant \lim_{s\to\infty} g(s). \tag{5}$$

The process defined by F is assumed to possess a unique stationary distribution.[7]

The demand shifts s are assumed to be independent *across* markets, and the number of markets is large.[8] Further, the total workforce of the economy is fixed. In consequence, once the workforce has settled down to a stationary distribution over markets, the expected present value of job search is a constant, say λ. In this section and the next, we treat λ as a given parameter; its equilibrium value will be determined in Section 5.

At the beginning of the period, each market has a fixed workforce, y, which serves as an upper bound on current period employment in that market. All currently employed workers remain into the next period; currently unemployed workers leave. In addition, new workers arrive in a stochastic fashion, the exact nature of which depends on the search process which is assumed. In the present paper, we shall impose a particular property on the *outcome* of this process, namely that unemployed

[7] For an example of a Markov process satisfying all these restrictions, including (4) and (5), let (ϵ_t) be a sequence of independent, normal variates, let $0 < r < 1$, and let s_t follow

$$\ln(s_{t+1}) = a + r \ln(s_t) + \epsilon_t .$$

[8] By large, we mean either a continuum of markets or a countable infinity. Economically, then, the assumption of independent demand shifts means that *aggregate* demand is taken to be constant through time.

workers are allocated over markets in such a way as to equate to the opportunity cost λ the expected return in each market receiving workers. The precise arrival rate which will guarantee this outcome will be specified below.[9]

To summarize, the *state* of a particular market is completely described by its state of demand, s, its beginning of period workforce, y, and the expected present value of search, λ. Of these three variables, only s and y vary from market to market; accordingly, we use (s, y) to index markets (referring, for example, to "market (s, y)"). Then for market (s, y), we seek equilibrium values of wages and employment, $w(s, y, \lambda)$ and $n(s, y, \lambda)$, as functions of the state of the market. An equilibrium must satisfy both the market clearing condition

$$w(s, y, \lambda) = R(s, n(s, y, \lambda)) \tag{6}$$

and the labor supply constraint

$$n(s, y, \lambda) \leqslant y. \tag{7}$$

Additional equilibrium conditions will be obtained by considering the present value maximizing work–search decision made by workers.

To study this choice, let $v(s, y, \lambda)$ be the expected present value of the wage stream for a worker who finds himself in (s, y) at the beginning of the period. In general, $v(s, y, \lambda)$ will equal the current wage plus the expected present value of the wage stream from next period on, discounted to the present by a constant factor β, $0 < \beta < 1$. Formally,

$$v(s, y, \lambda) = w(s, y, \lambda) + \beta E\{v(s', y', \lambda)\}$$

where the expectation is taken with respect to the distribution (as yet undetermined) of next period's state, (s', y') conditional on the information currently available to workers: (s, y, λ). The value of the terms on the right will vary with (s, y); it is convenient to consider three cases separately, as follows.

Case A. Some (or all) workers leave; some (or none) remain.

In this case, departing workers earn the expected return from search. Remaining workers earn no less, since they have the option to leave, and no more, since departing workers have the option to remain. Thus

$$v(s, y, \lambda) = \lambda. \tag{8a}$$

[9] In Eq. (19).

Case B1. All workers remain; no additional workers arrive next period.

In this case, current employment is the total workforce y and the current wage is, from (6), $R(s, y)$. Since the current workforce is maintained into the following period, next period's state is (s', y), with s' given probabilistically by $f(s', s)$. Thus

$$v(s, y, \lambda) = R(s, y) + \beta \int v(s', y, \lambda) f(s', s)\, ds'. \tag{8b1}$$

Case B2. All workers remain; some additional workers arrive next period.

In this case, the arriving workers, in common with *all* searchers, have an expected present value (discounted to the present) of λ. Thus, for them and for the workers remaining in (s, y), $\beta E\{v(s', y', \lambda)\}$ will have the common value λ, and

$$v(s, y, \lambda) = R(s, y) + \lambda. \qquad \text{(8b2))}$$

Evidently, these three cases divide the positive quadrant of the (s, y) plane into three mutually exclusive and exhaustive subsets.[10]

Now comparing cases B1 and B2, we observe that if no new workers are expected to arrive (case B1), it must be that expected rent in (s, y) is nonpositive with a future workforce of y, or that $\beta \int v(s', y, \lambda) f(s', s)\, ds' \leqslant \lambda$. Thus, (8b1) and (8b2) may be combined as

$$v(s, y, \lambda) = R(s, y) + \min \left[\lambda, \beta \int v(s', y, \lambda) f(s', s)\, ds'\right]. \tag{8b}$$

Finally, comparing cases A and B, we observe that remaining workers in either case have rejected the option to search, so that $v(s, y, \lambda) \geqslant \lambda$. Thus, (8a) and (8b) may combine to yield a single functional equation valid for all cases:

$$v(s, y, \lambda) = \max \left\{\lambda, R(s, y) + \min \left[\lambda, \beta \int v(s', y, \lambda) f(s', s)\, ds'\right]\right\}. \tag{8}$$

The relevant facts about (8) are given in:

PROPOSITION 1. *Equation* (8) *has a unique solution* $v(s, y, \lambda)$. *The function* v *is continuous in* (s, y, λ), *nondecreasing in* s *and* λ, *nonincreasing in* y, *and satisfies*

$$| v(s, y, \lambda_1) - v(s, y, \lambda_2)| < (1/\beta) \mid \lambda_1 - \lambda_2 \mid \tag{9}$$

[10] See Fig. 3 (which we do not at this point in the argument have enough information to draw) for this partioning of the positive quadrant.

for any λ_1, λ_2. *For each* y, λ,

$$\lim_{s \to 0} v(s, y, \lambda) = \lambda, \tag{10}$$

and for s *sufficiently large*,

$$v(s, y, \lambda) \leqslant R(s, y)/(1 - \beta). \tag{11}$$

Proof. Let T_λ, an operator which maps bounded continuous functions u on (s, y) into the same space, be defined by

$$T_\lambda u(s, y) = \max \left\{\lambda, R(s, y) + \min \left[\lambda, \beta \int u(s', y) f(s', s)\, ds'\right]\right\}.$$

The operator T_λ is monotonic: $u \geqslant v$ for all (s, y) implies $T_\lambda u \geqslant T_\lambda v$. For any constant c and function u, $T_\lambda(u + c) \leqslant T_\lambda u + \beta c$. By a slight modification of Theorem 5 of Blackwell [1], these two facts imply that T_λ is a contraction mapping. Thus, Eq. (8), $T_\lambda v = v$, has a unique, continuous solution and $\lim_{n \to \infty} T_\lambda u = v$ for any continuous u.

If $u(s, y)$ is increasing in s and decreasing in y, so is $T_\lambda u$, using (1). Hence, $v = \lim T_\lambda{}^n u$ is nondecreasing in s and nonincreasing in y.

Let $\lambda_1 > \lambda_2$. Clearly, $T_{\lambda_1} v(s, y, \lambda_2) \geqslant v(s, y, \lambda_2)$ for all (s, y). Since the operator T_{λ_1} is monotonic, we have

$$v(s, y, \lambda_1) = \lim_{n \to \infty} T_{\lambda_1}^n v(s, y, \lambda_2) \geqslant v(s, y, \lambda_2).$$

Hence v is nondecreasing in λ.

To verify (9), let $\lambda_1 > \lambda_2$ and define $u(s, y) = v(s, y, \lambda_2) + (\lambda_1 - \lambda_2)/\beta$. Then from the definitions of T_{λ_1} and $v(s, y, \lambda_2)$, we have, since $\beta < 1$,

$$T_{\lambda_1} u(s, y) = v(s, y, \lambda_2) + \lambda_1 - \lambda_2 < u(s, y).$$

Then by the monotonicity of T_{λ_1},

$$v(s, y, \lambda_1) = \lim_{n \to \infty} T_{\lambda_1}^n u(s, y) < u(s, y) = v(s, y, \lambda_2) + (\lambda_1 - \lambda_2)/\beta.$$

To prove (10), let $v_0 = 0$ and apply T_λ repeatedly, using (2) and (4) at each step.

To prove (11), let $v_0 = 0$ and apply T_λ repeatedly, using (5) at each step.

This proves Proposition 1.

With the value function v determined, we return to the determination of equilibrium employment and wages and of the equilibrium behavior of new arrivals. To determine employment, let $\tilde{n}(s, \lambda)$ be the employment that would occur in a market with demand s if the workforce constraint were not present. Thus, $\tilde{n}$ is the solution to

$$R(s, \tilde{n}(s, \lambda)) + \min \left[\lambda, \beta \int v(s', \tilde{n}(s, \lambda), \lambda) f(s', s)\, ds'\right] = \lambda.$$

Since R is positive, the solution cannot occur when the second term on the left is λ, so we may simplify to

$$R(s, \tilde{n}(s, \lambda)) + \beta \int v(s', \tilde{n}(s, \lambda), \lambda) f(s', s)\, ds' = \lambda. \tag{12}$$

Then, clearly, equilibrium employment is

$$n(s, y, \lambda) = \min[\tilde{n}(s, \lambda), y], \tag{13}$$

and equilibrium wages are found using (6). We summarize in

PROPOSITION 2. *For each fixed* (s, y, λ), *there exist unique equilibrium employment and wage functions* $n(s, y, \lambda)$ *and* $w(s, y, \lambda)$ *defined implicitly by* (6), (8), (12), *and* (13). *These functions are continuous in* (s, y, λ) *and satisfy the monotonicity properties*[11]

$$n_s \geqslant 0, \qquad n_y \geqslant 0, \qquad n_\lambda \leqslant 0, \tag{14}$$

$$w_s \geqslant 0, \qquad w_y \leqslant 0, \qquad w_\lambda \geqslant 0. \tag{15}$$

Also, for each fixed (y, λ),

$$\lim_{s \to 0} n(s, y, \lambda) = 0 \tag{16}$$

and

$$\lim_{s \to \infty} n(s, y, \lambda) = \min[\bar{n}(\lambda), y], \tag{17}$$

where $\bar{n}(\lambda)$ *is a finite bound, varying with* λ.

The *proof* of Proposition 2 is facilitated by reference to Fig. 1 which exhibits the left side of (12) as a function of n.

By (1) and Proposition 1, the curves in Fig. 1 are negatively sloped and shift to the right as s increases. As λ increases, these curves shift upward by Proposition 1 but, from (9), by an amount less than the increase in λ. Then, using (13), (14) is proved; (15) follows from (14) and (6).

[11] We use the usual notation for partial derivatives, recognizing that the monotonicity properties only imply that they exist almost everywhere.

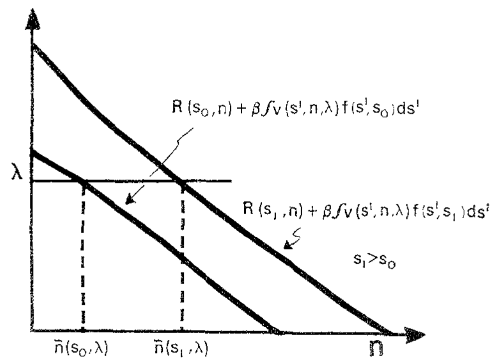

FIGURE 1

Applying (2), (4), and (10), (16) is proved.

Applying (3), (5), and (11), (17) is proved.

The results of this section may be illustrated on a conventional supply–demand diagram for labor. The demand curve is simply the marginal productivity condition (6), which shifts with the state of product demand, s. The curve SS is the relation between n and w implicit in (6) and (12): it is the locus of the wage–employment pairs which would be traced out as demand shifts if the workforce y did not constrain employment. The boldface curve is then the labor supply curve associated with the workforce y. The curve SS shifts up with increases in opportunity cost λ.

We remark that SS will not be flat, as would be the case if workers held a fixed "reservation wage" above which they accept employment and below which they do not. The reason this does not occur lies in the fact that as demand varies, wage and price changes convey information about future wage prospects as well as current earnings. Thus, as demand shifts to the left and employment declines, future prospects in (s, y) are affected

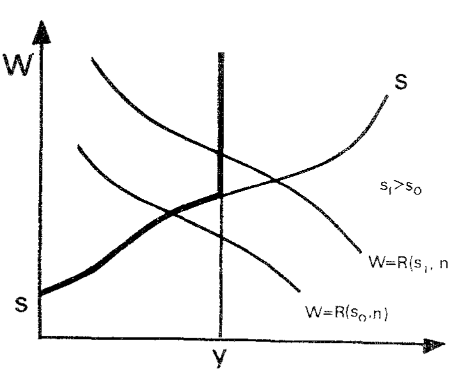

FIGURE 2

in two ways: first, lower demand this period increases the probability of a low demand next period as well; second, lower employment this period implies a lower workforce next period. These effects work in opposite directions, which is to say that on any interval, the curve SS may be either upward sloping (as drawn in Fig. 2) or downward sloping.

4. The Equilibrium Distribution of the Workforce

Implicit in the above discussion of equilibrium employment in a single market is the stochastic law relating a market's *next period* workforce to its current period demand and workforce, (s, y). In the present section, we make this law explicit, then develop its implications for the stationary joint distribution of demand and workforce.

In the preceding section, we postulated that all unemployed workers move toward markets with nonnegative expected rents. From the discussion of cases B1 and B2 above, it is clear that a market will attract new workers only if $n(s, y, \lambda) = y$ *and*

$$\beta \int v(s', y, \lambda) f(s', s)\, ds \geqslant \lambda. \tag{18}$$

If searchers were perfectly directed toward markets, each market satisfying (18) would receive exactly the arrivals a such that $y + a$ would satisfy (18) with equality. Equation (8) remains valid, however, under the somewhat weaker requirement that the search process eliminate rents on average. Specifically, let x be a positive random variable with the strictly positive density ϕ, c.d.f. Φ, and mean 1. We assume that each market (s, y) satisfying (18) receives $a(s, y, \lambda)x$ new workers, where the function $a(s, y, \lambda)$ is defined implicitly by

$$\beta \iint v(s', y + a(s, y, \lambda)\, x, \lambda) f(s', s)\, \phi(x)\, ds'\, dx = \lambda \tag{19}$$

if (s, y, λ) satisfies (18), and

$$a(s, y, \lambda) = 0 \tag{19a}$$

otherwise.[12]

[12] The arbitrariness in the search hypothesis (19) seems unavoidable, at least in the absence of a physically described process of search (e.g., the hypothesis that searchers follow a random walk over markets viewed as points in the plane). Our own attempts to formulate processes of the latter type have rapidly led to a complexity uncompensated by additional economic insight.

The hypothesis (19) seems roughly to capture the following sort of process. Unemployed workers are informed (by advertising, word of mouth, etc.) of which markets

The possible transitions from (s, y) are illustrated in Fig. 3. If (s, y) is in region I, current employment and next period's workforce is $n(s, \lambda)$, and $y - n(s, \lambda)$ workers enter the unemployed pool. Markets in region II neither contribute to nor receive from the unemployment pool, maintaining their current workforce into the next period. Markets in region III employ all their workforce and receive new workers for next period, as specified by (19).

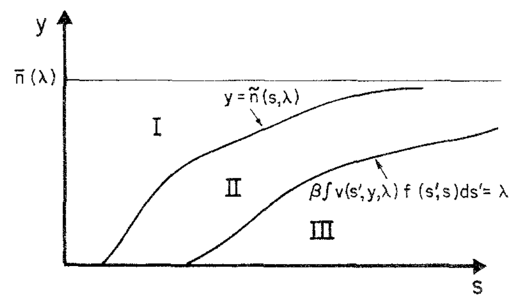

FIGURE 3

Analytically, the transitions from (s, y) are described by[13]

$$\begin{aligned}\Pr\{s_{t+1} \leqslant s', y_{t+1} \leqslant y' \mid s_t = s, y_t = y\} \\ = F(s', s) \Pr\{n(s, y, \lambda) + a(s, y, \lambda)\, x \leqslant y'\} \\ = F(s', s)\, \Phi\left(\frac{y' - n(s, y, \lambda)}{a(s, y, \lambda)}\right)\end{aligned}$$

These transition probabilities define an operator P on distribution functions $\Psi(s, y)$ as follows: Suppose that at a point in time, demand and workforce are distributed according to the c.d.f. Ψ; then the demand–workforce distribution next period is

$$P\Psi(s', y') = \iint F(s', s)\, \Phi\left(\frac{y' - n(s, y, \lambda)}{a(s, y, \lambda)}\right) \Psi(ds, dy). \tag{20}$$

We wish to show that the (s, y) process has a unique stationary distribution, or to prove.

need workers (are in region III of Fig. 3) and in which of these demand is greatest. All workers move toward a market in this class. Since the search is not coordinated, there is a stochastic element in the relationship between the actual "shortage" and arrivals of new workers.

[13] We use the convention that when $a(s, y, \lambda) = 0$, division of a positive (negative) number by $a(s, y, \lambda)$ yields $+ (-) \infty$. A c.d.f. evaluated at $+\infty$ is 1; evaluated at $-\infty$, it is 0.

PROPOSITION 3. *With P defined by* (11), *$P\Psi = \Psi$ has a unique solution Ψ^* (for each fixed λ) and $\lim_{n\to\infty} P^n\Psi = \Psi^*$ for all Ψ. Further, Ψ^* possesses a continuous density ψ^* which is strictly positive on $(s > 0, y \geqslant 0)$.*

The proof of Proposition 3 follows the treatment of Feller [3, pp. 264–268] or Doob [2, pp. 190–221]. The essential elements are the proofs of

LEMMA 1. *For any initial distribution Ψ, the sequence $\{P^n\Psi\}$ is stochastically bounded.*

and

LEMMA 2. *For any nondegenerate rectangle R in $(s > 0, y \geqslant 0)$ and any initial distribution Ψ, there is some m such that for all $n > m$, the distribution $P^n\Psi$ assigns positive probability to R.*

The second of these two lemmas specifies that the entire set $(s > 0, y > 0)$ is the ergodic set and contains no cyclically moving subsets; the first assures that most of the probability remains concentrated on a bounded subset of $(s > 0, y \geqslant 0)$. Together, these facts imply Proposition 3.

Proof of Lemma 1

For probabilities assigned by $P^n\Psi$, we use the notation $\Pr\{(s_n, y_n) \in A\}$. For arbitrary $\epsilon > 0$ and initial distribution Ψ, we wish to find $(\bar{s}, \bar{y})$ such that

$$(P^n\Psi)(\bar{s}, \bar{y}) = \Pr\{s_n \leqslant \bar{s}, y_n \leqslant \bar{y}\} \geqslant 1 - \epsilon$$

uniformly in n. Evidently, it will be sufficient to verify this inequality for $n \geqslant m$ for some m.

We have

$$\Pr\{s_n \leqslant \bar{s}, y_n \leqslant \bar{y}\} \geqslant 1 - \Pr\{s_n \geqslant \bar{s}\} - \Pr\{y_n \geqslant \bar{y}\}.$$

Since $\{s_n\}$ has a stationary distribution, $\bar{s}$ may be chosen so that $\Pr\{s_n \geqslant \bar{s}\} \leqslant \epsilon/2$ for n sufficiently large. Let $\bar{n}(\lambda)$ be the employment bound referred to in (17), Proposition 2. Choose $\bar{y}$ so that

$$\Pr\{x \geqslant (\bar{y} - \bar{n}(\lambda))/a(\bar{s}, 0, \lambda)\} = 1 - \Phi[(\bar{y} - \bar{n}(\lambda))/a(\bar{s}, 0, \lambda)] \leqslant \epsilon/2.$$

Then since $a(s, y, \lambda)$ is increasing in s and decreasing in y,

$$\Pr\{y_n \geqslant \bar{y}\} \leqslant \Pr\{\bar{n}(\lambda) + a(\bar{s}, 0, \lambda)x \geqslant \bar{y}\} \leqslant \epsilon/2.$$

This completes the proof of Lemma 1.

Proof of Lemma 2

We shall show that if the distribution is initially concentrated on an arbitrary point (s_0, y_0), then $\Pr\{\underline{s} \leqslant s_n \leqslant \bar{s},\ \underline{y} \leqslant y_n \leqslant \bar{y}\} > 0$ for all $n \geqslant 3$, provided $\underline{s} < \bar{s}$ and $\underline{y} < \bar{y}$.

Let $\tilde{y}$ satisfy $\beta \int v(s', \tilde{y}, \lambda) f(s', \underline{s})\, ds' = \lambda$ (so that $(\underline{s}, \tilde{y})$ is on the lower curve in Fig. 3), and let $\tilde{s}$ satisfy $\tilde{y} = \tilde{n}(\tilde{s}, \lambda)$ (so that $(\tilde{s}, \tilde{y})$ is on the upper curve of Fig. 3). Then since $f(s', s)$ is strictly positive,

$$\Pr\{n(s_1, y_1) \leqslant \tilde{y}, s_1 \leqslant \tilde{s} \mid (s_0, y_0)\} > 0$$

for any (s_0, y_0), and, therefore,

$$\Pr\{y_2 \leqslant \tilde{y}, \underline{s} \leqslant s_2 \leqslant \bar{s}\} > 0.$$

Then since $\phi(x)$ is strictly positive,

$$\Pr\{\underline{y} \leqslant y_3 \leqslant \bar{y}, \underline{s} \leqslant s_3 \leqslant \bar{s}\} > 0.$$

Evidently, the passage thus described may occur in any number of steps greater than three, so the proof is complete.

For each fixed λ, then, the (s, y) process has a unique stationary distribution, described by its c.d.f. $\Psi(s, y, \lambda)$ or its density $\psi(s, y, \lambda)$. In the remainder of this section, we study the behavior of mean values of functions of (s, y) taken with respect to Ψ as the parameter λ varies. The result of this examination is

PROPOSITION 4. *Let $\psi(s, y, \lambda)$ be the stationary density found in Proposition* 3, *and let $g(s, y)$ be continuous. Then if the integral*

$$h(\lambda) = \iint g(s, y)\, \psi(s, y, \lambda)\, ds\, dy$$

exists, it is a continuous function of λ.

The *proof* begins with the observation that one can always select a closed rectangle R, with the complement $\tilde{R}$ containing the (s, y) pairs with either very small or very large s-values, such that

$$\iint_{\tilde{R}} | g(s, y)|\, | \psi(s, y, \lambda_1) - \psi(s, y, \lambda_0)|\, ds\, dy \leqslant \delta$$

for any λ_0, λ_1 and $\delta > 0$. We shall be concerned, then, only with showing that the above integral taken over R tends to zero with $| \lambda_1 - \lambda_0 |$. We do so with heavy reliance on Fig. 3.

As λ increases (say from λ_0 to λ_1), the curves in Fig. 3 both shift down (by Propositions 1 and 2). This implies that $\Psi(s, y, \lambda_1)$ lies everywhere (on the y-axis) to the *left* of $\Psi(s, y, \lambda_0)$. (That is, high λ values are associated with low workforce levels.) Now since the functions $\tilde{n}(s, \lambda)$ and $v(s, y, \lambda)$ are continuous, there is a maximum absolute vertical shift, $c(\lambda_0, \lambda_1)$, of the two curves on R. Further, c tends to zero with $\lambda_1 - \lambda_0$.

By the argument used to prove Proposition 3, one can find the c.d.f. $\Psi(s, y, \lambda_0, c)$ implied by a *constant* shift of c in both curves of Fig. 3. Evidently, this c.d.f. lies everywhere to the left of $\Psi(s, y, \lambda_1)$, so that the *horizontal* distance between $\Psi(s, y, \lambda_1)$ and $\Psi(s, y, \lambda_0)$ is bounded from above by the horizontal distance between $\Psi(s, y, \lambda_0, c)$ and $\Psi(s, y, \lambda_0)$. But $\Psi(s, y, \lambda_0, c) = \Psi(s, y - c, \lambda_0)$, so this latter distance is simply c, which tends to zero with $\lambda_1 - \lambda_0$.

Since Ψ possesses a continuous density, this continuity property is sufficient to guarantee the continuity of $h(\lambda)$.

5. Economy-wide Equilibrium

Propositions 1–4 describe the determination of the stationary distributions of employment, workforce, and wages in a representative market, with the expected return from search, λ, treated as a given parameter. From an economy-wide viewpoint, however, it is the size of the workforce which is fixed and the "price" λ which adjusts to clear the market.

For given λ, the system described above would behave, in the aggregate, as an occupation with a membership elastically supplied at the expected present value λ. The distribution of the workforce over locations (indexed by (s, y)) would in this case be the same as the stationary distribution of (s, y) in any one market. (This follows from our assumptions that the number of markets is large and that demand shifts are independent across markets.) Then the total workforce demanded (per market) in this occupation, at the return λ, is

$$\iint y\Psi(s, y, \lambda)\, ds\, dy. \tag{21}$$

For each fixed λ, the integral (21) converges in view of the facts that employment is bounded for each fixed λ (Eq. (17), Proposition 2), that $a(s, y, \lambda)$ is bounded, and that the random variable x has a finite mean. By Proposition 4, the expression (21) is a continuous function of λ. As observed in Section 4, increases in λ shift the distribution function $\Psi(s, y, \lambda)$ to the left (along the y-axis), so that (21) is a decreasing function of λ. As $\lambda \to 0$, $E(y; \lambda) \to \infty$ since R is a positive strictly decreasing function of n; as $\lambda \to \infty$, $E(y; \lambda) \to 0$. The demand function is thus as shown in Fig. 4.

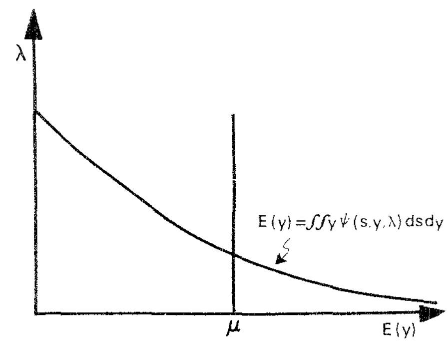

FIGURE 4

Now let μ denote the fixed workforce per market *supplied*. This vertical supply function together with the demand function just obtained gives the equilibrium λ: the solution to

$$\iint y\Psi(s, y, \lambda)\, ds\, dy = \mu. \tag{22}$$

We summarize in

PROPOSITION 5. *For all values of workforce-per-market μ, there is a unique positive equilibrium value of λ.*

Thus, Propositions 1–3 and 5 provide a full description of the equilibrium determination of wages, employment, and workforce in all markets of the economy.[14] By Proposition 3, there will always be some markets in region I of Fig. 3, where the workforce y exceeds the equilibrium

[14] Since the content of this paper consists as much in motivating and explaining a particular definition of equilibrium as in analyzing this equilibrium, we have intermingled definitions and results in a way which may be difficult for readers to disentangle. A different procedure would be to begin with the following (abbreviated)

DEFINITION. An *equilibrium* for the economy under study consists of a 5-tuple of nonnegative, continuous functions $n(s, y)$, $\tilde{n}(s, y)$, $w(s, y)$, $v(s, y)$, and $a(s, y)$, a c.d.f. $\Psi(s, y)$, and a nonnegative number λ such that (6), (8), (12), (13), (19), (20), and (22) are satisfied.

The content of Propositions 1–3 and 5, then, is that a unique equilibrium in the sense of the definition exists. Of course, these propositions also contain information useful in characterizing this equilibrium.

employment level $\tilde{n}(s, \lambda)$. This means that labor market equilibrium necessarily involves positive unemployment.

Numerical calculations of the equilibrium pictured in Fig. 4 are provided in the appendix to this paper.

6. Stability of Equilibrium

The equilibrium obtained for this model economy provides a complete description of the *time paths* of all variables involved, both at an aggregate and the individual market level. Since provision of such a description is frequently thought to be the task of "stability theory" (in the sense of, for example, Samuelson [12]), one may ask whether the latter theory has any applicability to the present model. The answer, we think, is "yes," provided one raises the stability question in its most fundamental sense of determining whether if an equilibrium approximately describes the economy at a point in time, it will continue to do so in the future.

In the present context, this approximation question is particularly pertinent, since we have provided no account as to how workers arrive at the state of perfect knowledge of the probability distributions relevant to their decision problem. Ultimately, this is a question for psychological rather than economic theory, so we do not apologize for framing it here in *ad hoc* "adaptive" terms.

The distributions F and Φ refer to variables exogenous to the markets under study; presumably, they are learned by processing observed frequencies in some sensible fashion, "Bayesian" or otherwise, which has the property that the "true" distributions become "known" after enough time has passed. The distribution $\Psi(s, y, \lambda)$, on the other hand, depends on the behavior of workers, so that as worker perceptions change, so does the "true" Ψ which is being learned. This could, in general, raise insuperable analytical difficulties, but in the present context it does not, since the *only* feature of Ψ which is relevant to worker decisions is the parameter λ. We must describe, then, how the economy operates when the λ perceived by workers differs from the equilibrium value λ^* (say) and how, under this circumstance, perceptions are revised.

For specificity, suppose $\lambda > \lambda^*$. Then the number of workers entering the pool exceeds the number which can be reassigned at an average return of λ. One could modify the reallocation mechanism in many ways, but suppose in particular that the mean of the random variable x varies so as to equate the total number of workers reallocated to the size of the pool. Then both searchers and workers who remain on the job will be disappointed (on average) in their wage expectations. Presumably, this will lead them to revise their perceived λ downward, slowly relative to the passage of trading time t. Thus, we assume

$$d\lambda/dt = g(\lambda - \lambda^*),$$

where g is a decreasing function vanishing at zero. Clearly (from Fig. 4), the equilibrium is stable. Equally clearly, this stability result can have no relevance to the dynamic response to regularly recurring shocks.[15]

7. Concluding Remarks

Although there are (by assumption) no *aggregate* dynamics in the model developed above, it should be obvious that the mechanism we have described is consistent with the now familiar account of the observed Phillips curve in terms of expectations. Thus, an unanticipated change in aggregate demand (a change in $E(s' \mid s)$) will move unemployment and wage changes in opposite directions. Of course, if aggregate demand changes were a recurrent event, as they are in reality, this fact would become incorporated into the maximum problem facing workers and would result in different equilibrium functions $w(s, y)$ and $n(s, y)$. We leave this nontrivial development for future research.

The implications one *can* draw from the model as it stands are of a comparative static nature, both positive and normative. As an example of the former, suppose a lump-sum cost is imposed on leaving one's market to search, so that the right side of (12) becomes $\lambda - c$ rather than simply λ. This will raise the curve $y = \tilde{n}(s, \lambda)$ in Fig. 3 and shift the "demand curve" in Fig. 4 downward. The result is a decrease in unemployment and a decrease in the equilibrium present value of wages, λ. (This example also shows that lower average unemployment is not, in general, associated with higher welfare for workers.) It may well be, though one could hardly demonstrate it at this level of abstraction, that differences of this sort in the actual or perceived costs of changing jobs can help to account for the observed differences in average unemployment across occupations and among countries.

We can also examine Tobin's normative concern [14, p. 8] that "the external effects [of search] are the familiar ones of congestion theory. A worker deciding to join a queue or to stay in one considers the probabilities of getting a job, but not the effects of his decision on the probabilities that others face." Now one could *add* congestion in the usual sense to the search model we have developed (say, by assuming that searching workers

[15] For reasons developed by Gordon and Hynes in [10].

travel on a congested route).[16] But it should be clear that congestion of this sort is not a necessary component of an equilibrium search model. In our scheme, the injury a searching worker imposes on his fellows is of exactly the same type as the injury a seller of any good imposes on his fellow sellers: the equilibrium expected return λ from job search serves the function of any other equilibrium price of signalling to suppliers the correct social return from an additional unit supplied.

The question of whether there exist important external effects in *actual* labor markets remains, of course, to be settled. However this may turn out, it is surely a major advance even to be discussing unemployment from the point of view of the usual (in better developed areas of economics) standard of allocative efficiency. Our intention in this paper has been to indicate the general kind of framework within which such discussions can be conducted and to begin to develop suitable analytical methods.

APPENDIX: EXAMPLES

Several examples were analyzed numerically to determine the workforce demand and unemployment rate as a function of the market parameter λ. In order to compute these solutions, it was necessary to assume a finite number of market demand states and to permit only integer values for the workforce. In addition, we assumed that x had a degenerate distribution concentrated at one.

The method of solution used the T_λ operator, defined in Section 3, to determine the value function $v(s, y, \lambda)$. The initial approximation was $v_0(s, y, \lambda) = \lambda$. The nth approximation $v_n(s, y, \lambda)$ was $T_\lambda v_{n-1}(s, y, \lambda)$. With the assumed discount factor $\beta = 0.9$, the convergence to $v(s, y, \lambda)$, the unique fixed point of T_λ, was rapid. Equation (12) was then solved to determine $\tilde{n}(s, \lambda)$, and Eqs. (19) and (19a) were used to determine $a(s, y, \lambda)$. Next period's workforce, given $x = 1$, will be

$$y' = \min[\tilde{n}(s, \lambda), y] + a(s, y, \lambda). \tag{23}$$

The workforce will be bounded, which along with the previous assumptions implies a finite number of possible market states (s, y). Thus, the stochastic process for a market is a finite-state Markov chain with some transition probability matrix, say, P. This matrix whose ijth element specifies the probability that state j will occur next period given current

[16] Phelps [11, Chapter 4, pp. 103–105] also discusses congestion problems, but in a way which makes it clear that these problems arise under nonwage rationing of jobs (i.e., under *dis*equilibrium prices) as opposed to being externalities in the usual equilibrium sense.

state i is determined by (23) and the transition probability matrix of the s-process.

Let u be a function (represented by a vector) defined on the possible market states. Using the analysis of Feller [3, pp. 264–268], the expected value of u with respect to the stationary distribution implied by P can be determined by computing

$$\lim_{n\to\infty} P^n u.$$

The limiting vector has elements all of which are equal to the expected value of u. This was the procedure we used to compute

$$\sum_{s,y} y\psi(s, y, \lambda);$$

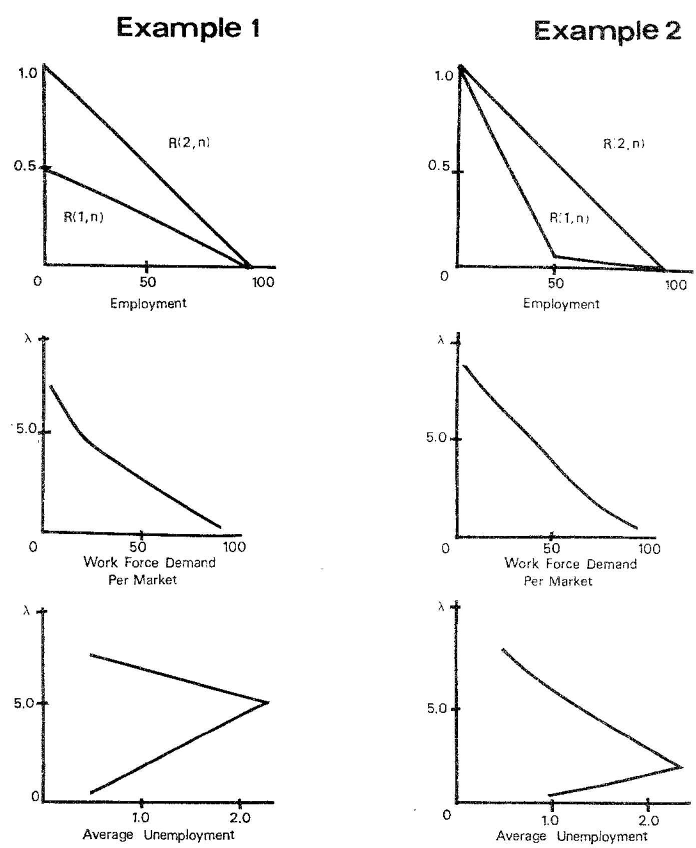

FIGURE 5

the average workforce per market, and

$$\sum_{s,y} a(s, y, \lambda)\, \psi(s, y, \lambda),$$

the average unemployment per market.[17]

Two of the examples considered had the marginal revenue schedules depicted in Fig. 5. There are but two demand states: $s = 1$ or $s = 2$. The transition probability matrix for the s process was

$$\begin{bmatrix} .9 & .1 \\ .1 & .9 \end{bmatrix},$$

so there was a strong persistence in demand. The discount factor β was 0.9.

As the theory predicts, the labor demand curve, pictured in Fig. 5, is downward sloping. On the other hand, the unemployment level, also pictured in Fig. 5, is not monotonic, having maxima. Overall, we found for low and high persistence in demand that unemployment rates were low. In the former case, there was little gained by reallocating workers, while in the latter reallocation occurred infrequently. As expected, the greater the variability of demand, holding the degree of persistence fixed, the greater the level of unemployment. This result is reasonable for more workers should be reallocated when demand conditions change.

Acknowledgment

We thank E. S. Phelps for his helpful general comments and, in particular, for pointing out an error in an earlier draft in the discussion of Fig. 2.

References

1. D. Blackwell, Discounted dynamic programming, *Ann. Math. Statist.* **36** (1965), 226–35.
2. J. L. Doob, "Stochastic Process," Wiley, New York, 1953.
3. W. Feller, "An Introduction to Probability Theory and Its Applications," Vol. ii, Wiley, New York, 1966.
4. M. Friedman, The role of monetary policy, *Amer. Econ. Rev.* **58** (1968), 1–17.
5. J. R. Hicks, "Value and Capital," Clarendon Press, Oxford, England, 1939.
6. R. E. Lucas, Jr. and E. C. Prescott, Investment under uncertainty, *Econometrica* **39** (1971), 659–81.
7. R. E. Lucas, Jr., Expectations and the neutrality of money, *J. Economic Theory* **4** (1972), 103–124.

[17] The computer program used for these calculations is available upon request.

8. J. McCall, The economics of information and optimal stopping rules, *J. Business* **38** (1965), 300–17.
9. J. F. Muth, Rational expectations and the theory of price movements, *Econometrica* **29** (1961).
10. E. S. Phelps *et al.* "Microeconomic Foundations of Employment and Inflation Theory," Norton, New York, 1969.
11. E. S. Phelps, "Inflation Policy and Unemployment," Norton, New York, 1972.
12. P. A. Samuelson, "Foundations of Economic Analysis," Harvard University Press, Cambridge, 1947.
13. G. J. Stigler, The economics of information, *J. Political Economy* **69** (1961), 213–35.
14. J. Tobin, Inflation and unemployment, *Amer. Econ. Rev.* **62** (1972), 1–18.