Econometrica, Vol. 58, No. 4 (July, 1990), 783–808

ESTIMATING A MARKET EQUILIBRIUM SEARCH MODEL FROM PANEL DATA ON INDIVIDUALS

BY ZVI ECKSTEIN AND KENNETH I. WOLPIN[1]

In this paper we demonstrate the feasibility of estimating a Nash labor market equilibrium model using only information on workers. The equilibrium model is adapted from Albrecht and Axell (1984) and is based on workers who are homogeneous in terms of market productivity and heterogeneous in terms of nonmarket productivity, and on firms which are heterogeneous in terms of productive efficiency. The equilibrium model is contrasted in terms of its fit to the data with an unrestricted version of the model which is based on a mixture of negative binomial distributions. The equilibrium model fails to conform to the data in exactly the dimension of its major focus, namely it implies that measurement error accounts for almost all of the dispersion in observed wages. The equilibrium model also does not do well in fitting the unemployment duration distribution compared to the unrestricted model. The problem is that the duration distribution itself does not support the existence of significant heterogeneity, as evidenced by the estimates of the unrestricted model. The paper also illustrates the use of such models for policy analysis by simulating the welfare effects of a minimum wage.

KEYWORDS: Search, equilibrium, unemployment.

1. INTRODUCTION

IN THE LAST DECADE a substantial effort has been devoted to estimating labor market models of search. The first attempts used the theory only loosely, as a guide to the specification of the estimation problem (e.g., Toikka (1976)). Somewhat later studies attempted to incorporate in an approximate fashion the restrictions embodied in the theory in terms of the relationship between unobserved reservation wages and the observed wage offers (e.g., Kiefer and Neumann (1979)). More recently, there have been methods developed to estimate job search models structurally, that is, taking into account the full set of restrictions of the theory obtained by rigorously solving the sequential decision rule. Examples of this approach are Flinn and Heckman (1982) who discuss issues in estimating continuous time structural search models, Miller (1984) who presents a method for estimating a job matching model, and Wolpin (1987) who estimates a standard finite horizon job search model. Other studies which have taken the structural approach to estimation of search models include Lancaster and Chesher (1983), Narendranathan and Nickell (1983), and Stern (1985).[2]

All of this work in the search setting addresses the problem of a price-taking individual; in the job search context individuals take the wage offer distribution

[1] We have benefited from useful discussions with Rao Aiyagari, Jim Heckman, Dale Mortensen, Randy Olsen, Steven Stern, and Shmuel Zamir as well as from comments by two anonymous referees. Research support is gratefully acknowledged from NSF Grant SES-8619181. This research was also supported by a grant from the United States-Israel Binational Science Foundation (BSF), Jerusalem, Israel. Hwei-Lin Chuang provided excellent research assistance.

[2] In addition, there is a growing literature on the structural estimation of sequential decision-making models outside of the job search context, e.g., Eckstein and Wolpin (1989), Gotz and McCall (1987), Pakes (1987), Rust (1987), and Wolpin (1984).

as given. Since Rothschild's (1973) criticism of the assumption of stable wage (price) distributions in partial equilibrium search models, much research effort has focused on developing equilibrium models which generate wage (price) dispersion among homogeneous workers (buyers) (e.g., Axell (1977), Burdett and Judd (1983), Albrecht and Axell (1984), and Rob (1985)).

One purpose of this paper is to develop estimation methods for equilibrium search models with homogeneous workers.[3] We discuss in detail such a methodology based on an extension of the Albrecht and Axell (1984) model. In that model individuals work in a market where firms differ in their productivity, although all workers would have the same productivity at the same firm, i.e., workers are homogeneous with respect to their market skills. Workers are, however, heterogeneous with respect to their nonmarket productivity (or preferences for leisure). Individual preferences and firm productivities are private information, although the distribution of preferences and productivities over individuals and firms, respectively, are known to all agents. A Nash equilibrium wage offer distribution is shown to exist which corresponds to the reservation wages of workers with the appropriately solved wage offer probability distribution. The wage offer distribution is a function of both the worker preference distribution for leisure and the firm productivity distribution. The equilibrium wage offer distribution implies an equilibrium unemployment rate and accepted wage offer distribution.[4]

The equilibrium outcome of the model is that the more productive firms will offer high wages, attracting high reservation wage types in addition to the low reservation types and the less productive firms offer low wages, attracting low reservation wage types. This implication has bearing on the recently revived literature on inter-industry wage differentials (for example, Krueger and Summers (1988)). The model also predicts that for otherwise homogeneous workers, those in larger firms and more profitable firms will enjoy higher wages. However, because the model assumes constant returns to scale in production, setting a minimum wage equal to the marginal product of the most productive firm will clearly maximize welfare. Not only will there be no (search) unemployment, but output will be maximized because all workers are employed at the most productive firm. We modify the Albrecht-Axell model in a manner which preserves the positive implication of wage heterogeneity, while altering this peculiar normative implication.[5] In particular, we allow for an endogenously determined offer probability which is assumed to depend on the number of active firms per worker, where an active firm is one whose marginal product exceeds the

[3] Flinn and Heckman (1982) develop an equilibrium model of sequential search based on an exogenous matching technology. The equilibrium wage distribution is thus nondegenerate because of real productivity differences in worker-firm matches, which is different in spirit from the class of equilibrium models we consider.

[4] There is no distinction between unemployment and out-of-the-labor force states. All individuals would accept an offer if the wage was high enough.

[5] Equivalently, in Rob's (1985) model of equilibrium price dispersion in the product market, a maximum price that is equal to marginal cost (zero) would minimize search and maximize welfare.

lowest reservation wage.[6] An endogenous increase in the lowest reservation wage will reduce the offer probability to all individuals. Thus, an exogenously set minimum wage above the lowest reservation age will also reduce the offer probability and, thus, tend to increase the unemployment rate. Whether an effective minimum wage is optimal becomes an empirical issue.

Given panel data information on unemployment durations, it is possible to specify the likelihood function for the equilibrium model and to estimate the parameters subject to identification conditions which depend on the degree of population heterogeneity in leisure preferences. Estimation involves solving for the equilibrium wage offer distribution at each evaluation of the likelihood function. It is important to recognize that from individual data on duration alone it is possible to recover the technology structure as well as preferences. Information on accepted wages also can be incorporated into the estimation and reduces the restrictiveness of the identification conditions at the cost of additional structure. In either case, estimating equilibrium labor market models does not require data on both workers and firms, although it does require assumptions about the structure of the distributions of preferences and technology.

An additional goal of this paper is to assess the ability of the equilibrium model to describe observed duration and wage data. The theoretical model yields an unemployment duration distribution that is a mixture of negative binomials of order equal to the number of types of individuals. Parameters of the wage offer distribution and the job offer probability can be estimated without imposing the restrictions of the equilibrium model. The equilibrium search model can then be compared in terms of fit to the data and to the unrestricted model.

We use data from the National Longitudinal Surveys of Labor Market Experience of Youth to follow several cohorts of high school graduates from the time they graduate until the time they accept their first full-time job or until the last survey date. The maximum observed duration of "unemployment" is 212 weeks. While the unrestricted model gives a fairly reasonable fit to the duration data the equilibrium model does very poorly. However, the unrestricted model reveals that in terms of acceptance probabilities, there are essentially two groups of individuals, one group who almost always accepts offers and one group who almost always rejects offers. The equilibrium model, because it mimics that result, cannot generate significant wage dispersion. Indeed, the parameter estimates of the equilibrium model imply that almost all of the observed wage dispersion is due to measurement error. Thus, the model fails to explain the phenomenon it was intended to explain. Nevertheless, to illustrate the value of estimating an equilibrium model, we simulate the welfare effects of a minimum wage. We find that an effective minimum wage is Pareto inferior in the sense that a noninterventionist policy maximizes per-capita utility. However, given the inadequate performance of the equilibrium model, this policy simulation should be viewed only as an illustration of the potential value of estimating equilibrium models.

[6] Flinn and Heckman (1982) make a similar assumption.

2. AN EQUILIBRIUM SEARCH MODEL

The model is an extension of Albrecht and Axell's (1984) labor market search equilibrium model. Although workers are homogeneous in productivity, wage dispersion arises due to heterogeneity in worker tastes for leisure (or alternative values of time) and heterogeneity in the efficiency of labor across firms. We extend the model by allowing for more than two types of workers and for an endogenous wage offer probability which depends on the number of active firms in the market.

Workers. We assume a labor market where each period new individuals, who live forever, enter and exit that market. The probability of leaving the market is constant over the individual's lifetime and is equal to τ. At each period in the market the individual is either working or (costlessly) searching for work. While searching, at most one wage offer is received per period, and if received may be accepted or rejected. If the wage offer is accepted, it is maintained as long as the individual remains in this market. If the individual leaves the market, there is no way of returning to this market and a lifetime income of R is received. Accepting a wage offer of w yields expected lifetime welfare (real income) equal to:

$$(1) \qquad V^a(w) = \frac{w}{1-(1-\tau)\delta} + \theta',$$

where

$$\theta' = \frac{\theta + \delta\tau R}{1-(1-\tau)\delta},$$

θ is income that is independent of the individual's market status, and δ is the objective discount factor.

There are $n+1$ types of individuals with many people of the same type. The types differ according to their per-period nonmarket value of time, z_j, $j = 0, 1, \ldots, n$. The value of not working for an individual of type j is

$$(2) \qquad V_j^r = z_j + b + \theta + \delta\left\{(1-\tau)\left[pE\max\left(V^a(w), V_j^r\right) + (1-p)\left(V_j^r\right)\right] + \tau R\right\},$$

where p, $0 < p \leqslant 1$, is the probability of receiving a positive wage offer and b is the per-period value of unemployment benefits. The solution to the worker optimization problem is characterized by a reservation wage, w_j^*, which satisfies

$$(3) \qquad V^a(w_j^*) = V_j^r = V_j^*$$

and the worker accepts any $w_j > w_j^*$. Anticipating the Nash equilibrium solution of the model, assume that wages are in discrete levels such that $w_{i-1} < w_i < w_{i+1}$ and w_i is sampled with probability γ_i, $i = 0, 1, 2, \ldots, n$. Furthermore, we conjecture that the set of wage offers is equal to the set of reservation wages, that is, $w_j^* = w_j$ for all j. Obviously there are as many wages as there are reservation wages. Let $\gamma^j = \sum_{i=0}^{j-1}\gamma_i$ be the probability of receiving an offer of $w < w_j$. Then,

$$(4) \qquad E\max\left[V^a(w), V_j^r\right] = \gamma^j V_j^r + \sum_{i=j}^{n} \gamma_i V^a(w_i).$$

Substituting (4) and (3) into (2) and equating to (1), we get the following solution to the (reservation) wage:

$$(5a) \qquad w_j = (z_j + b)\frac{1-\delta(1-\tau)}{1-\delta(1-\tau)(1-p+p\gamma^{j+1})} + \frac{\delta(1-\tau)p}{1-\delta(1-\tau)(1-p+p\gamma^{j+1})}\sum_{i=j+1}^{n}\gamma_i w_i,$$

$$(5b) \qquad w_n = (z_n + b).$$

Notice that w_j is independent of θ and R because both are independent of the individual's employment status and that w_j is a linear combination of z_j, $z_{j+1}, \ldots, z_n$, i.e., of one's own value of leisure and of all higher values of leisure.[7]

In order that equation (5) define an equilibrium wage function, consistent with our assumptions so far, it has to satisfy the restriction that $w_j < w_{j+1}$. This is established by the following lemma:

LEMMA 1: *For any given values of* δ, $\tau \in (0,1)$, $b \geqslant 0$, $0 \leqslant p \leqslant 1$, $z_j > 0$ *for all* $j = 0, \ldots, n$ *such that* $z_0 < z_1 \cdots < z_n$, *and probabilities* $\gamma_j \geqslant 0$, $\sum_{j=1}^{n}\gamma_j = 1$, *there exists a unique sequence of reservation wages* $w_j > 0$ *for all* $j = 0, \ldots, n$ *such that* $0 < w_0 < w_1 < \cdots w_n = z_n + b$.

PROOF: See Appendix.

Firms. We assume that there are F firms where F is large. Each firm faces the following linear production technology:

$$(6) \qquad y = \lambda l,$$

where y is output, l is the number of workers, and $\lambda \geqslant 0$ is a firm-specific productivity index.[8] The productivity index λ comes from a continuous distribution function $A(\lambda)$, such that $A(0) = 0$. Each firm's λ is the private information of that firm, e.g., the managerial capability of that firm. The firm's objective is to maximize profits at each discrete time of this economy. Profits are given by

$$(7) \qquad \pi(w, \lambda) = (\lambda - w)l(w),$$

where $l(w)$ is the supply of workers to the firm when wage w is offered.

[7] It is thus clear that income outside of this market, R, can be made type-specific without altering the reservation wage.

[8] Allowing for diminishing marginal product of labor adds one more parameter, but otherwise does not create conceptual difficulties.

At each point in time the firm meets many workers. The worker does not know the λ of the particular firm and the firm does not know the reservation wage of the particular worker. It is assumed that the firm's strategy is to make the same wage offer to all workers and that the worker's strategy is either to accept or decline the offer.[9] Firms know that there are only $n+1$ types of workers and, therefore, $n+1$ levels of wages. To maximize profits, a firm would employ as many workers as possible if $w < \lambda$. However, firms expect that in equilibrium the number of available workers per firm ($l(w)$) is decreasing as the wage is lower. Firms take the function $l(w)$ as given and offer a wage that would maximize their profits. Suppose that $w_0, w_1, \ldots, w_n$ are the equilibrium wages offered by the firms. Then, there will exist a critical value of λ, λ_j, which makes a firm indifferent between offering w_j or w_{j-1} ($w_{j-1} < w_j$), i.e., $(\lambda_j - w_{j-1})l(w_{j-1}) = (\lambda_j - w_j)l(w_j)$. The solution for this critical value is

$$\lambda_j = \frac{w_j l(w_j) - w_{j-1} l(w_{j-1})}{l(w_j) - l(w_{j-1})}. \tag{8}$$

A firm with λ between λ_j and λ_{j+1} will offer a wage w_j (the lower wage). Because w_0 is the lowest wage,

$$\lambda_0 = w_0 \tag{9}$$

is the lowest productivity associated with an active firm, i.e., a firm capable of earning positive profits. $A(w_0)$ is the proportion of potential firms that are not active because their productivity is lower than the least acceptable wage in the market.

The following lemma establishes the optimal wage offer of a given firm.

LEMMA 2: *If $l(w_j) > l(w_{j-1})$, then $\lambda_0 = w_0 < w_1 < \lambda_1$. Further, if $\lambda_j > \lambda_{j-1}$ for all $j = 2, 3, \ldots, n$, then a firm with λ between λ_j and λ_{j+1} maximizes profits by offering the wage w_j.*

PROOF: See Appendix.

The first part of the lemma establishes that $\lambda_1 > \lambda_0$. For $n > 1$, some additional restrictions on the parameters of the model are necessary in order that the λ_j's are rising, which is a necessary condition for the optimal wage offer strategy described in the lemma.

Having described both the worker and firm optimal behavior, we can now turn to characterizing the market equilibrium.

Equilibrium. In order to solve for the equilibrium, we have to find the supply of workers per firm offering a wage of w_j, $l(w_j)$ for all $j = 0, 1, 2, \ldots, n$. Let β_j be the proportion of workers of type j and k the total number of workers in the market.

[9] Perry (1986) has shown that in a bargaining situation with this type of incomplete information, if the firm has a lower cost of negotiation per meeting than does the worker, the firm's "take it or leave it" offer is an equilibrium bargaining solution.

Furthermore, it is assumed that $\tau k\beta_j$ new workers of type j join the market each period. Let $\mu(\lambda_0) = k/F(1 - A(\lambda_0))$ be the number of workers per active firm (F is the number of potential firms). Then the proportion of total workers per firm offering w_j is

$$(10)\qquad \begin{aligned} l(w_j) = {} & \sum_{r=0}^{j} \tau\mu\beta_r\left[1 + (1-\tau) + (1-\tau)^2 + \cdots\right] \\ & \times\left[p + ((1-p) + p\gamma^r)(1-\tau)p \right. \\ & \left. + ((1-p) + p\gamma^r)^2(1-\tau)^2 p + \cdots\right] \\ = {} & \sum_{r=0}^{j} \frac{\mu\beta_r p}{1 - ((1-p) + p\gamma^r)(1-\tau)} \end{aligned}$$

where $(1 - \tau)$ is the proportion of people who do not leave the market each period and where $\gamma^0 = 0$. The calculation here is made at the steady state of this economy where there are infinitely many past cohorts of new entrants to the market. As asserted earlier, (10) shows that $l(w_j)$ is increasing with j so that at higher wages there is greater market supply, although each individual supplies only one unit of labor inelastically. This result guarantees that λ_j in (8) is always positive.

The probability of receiving an offer p is assumed to be an increasing function of the number of active firms per worker, i.e.,

$$(11)\qquad p = G\left(\frac{1}{\mu(\lambda_0)}\right), \quad G' > 0,$$

where G is a technological function which maps from R to $[0,1]$. In this regard, we deviate significantly from the Albrecht and Axell formulation. The effect of this change is to make the demand for workers explicitly depend on the size of the market. Workers sample from the distribution of potential firms and only obtain an offer when they meet an active firm. The greater the proportion of firms that are active, the more likely is the worker to receive an offer.

DEFINITION: A *Nash equilibrium for this model* is a probability density function such that w_j has probability γ_j, $\gamma_j \in [0,1]$, satisfying equation (5), the firm's strategy is to offer w_j if its $\lambda \in [\lambda_j, \lambda_{j+1})$, λ_j is determined by equations (8) and (9), and is increasing in j, $l(w_j)$ is determined by equation (10), p is given by equation (11), and the probabilities, γ's, are equal to the proportion of active firms offering each wage, that is,

$$(12)\qquad \gamma_j = \frac{A(\lambda_{j+1}) - A(\lambda_j)}{1 - A(\lambda_0)} \quad \text{for all} \quad j = 0, 1, \ldots, n-1,$$

and

$$\sum_{j=0}^{n} \gamma_j = 1.$$

PROPOSITION: *There exists a Nash equilibrium as defined above if $\lambda_j > \lambda_{j-1}$ for all $j = 1, \ldots, n$, p and γ_i, $i = 1, \ldots, n$.*

PROOF: See Appendix.

It is possible to solve for the equilibrium numerically when the conditions in the above proposition are met.[10] In the Appendix we discuss conditions for increasing values of the λ_j's. We have found that of several routines we have tried, ZSPOW, a modified Newton method within the IMSL package, is the most efficient in solving the equilibrium. We are usually able to obtain convergence with a degree of accuracy of 10^{-8} in terms of changes in each of the λ_j and p.[11]

Implications. Given that an equilibrium exists, the model generates an equilibrium distribution of unemployment durations for a given cohort of new entrants. If we let d be the duration of unemployment (or search), then the probability that a randomly selected individual will be hired in period $d+1$ is

$$f_{d+1} = \sum_{j=0}^{n} (1-\tau)^d \left[p\gamma^j + (1-p) \right]^d p(1-\gamma^j)\beta_j \qquad (d = 0, 1, 2, \cdots). \tag{13}$$

Equation (13) is the probability that an individual survives d periods, either does not receive an offer or rejects an offer for d periods, and receives and accepts an offer in period $d+1$, averaged over the $n+1$ types in the population. The probability given by (13) is a mixture (over types) of $n+1$ negative binomial distributions. The hazard rate in period $d+1$, i.e., the conditional probability of leaving unemployment in period $d+1$ given d periods of unemployment, is by definition

$$h_{d+1} = \frac{f_{d+1}}{(1-\tau)^d \prod_{k=1}^{d} (1-h_k)} \qquad (d = 0, 1, 2, \cdots). \tag{14}$$

It is a well-known result that the population hazard rate is decreasing in duration, i.e., it exhibits negative duration dependence, due to heterogeneity, even though each type has the constant hazard rate $p(1-\gamma^j)$.

[10] It is important to recognize that if the $A(\lambda)$ distribution is truncated from above, the equilibrium as described may not exist if there are some individuals with sufficiently high values of leisure. In this case, the equilibrium would be likely to include a group of non-participants in the sense that no existing wage would be high enough to induce them to work. This was pointed out to us by Jim Heckman.

[11] For some numerical examples we calculated the sequence of values of the simplex (Δ) using the mapping (F) as described in the Appendix starting from an initial vector in the simplex. It is evident from these examples that the mapping from the simplex to itself is not a contraction mapping. Therefore, we cannot use contraction mapping results to prove uniqueness of the equilibrium.

The equilibrium unemployment rate is

$$(15) \qquad s = \sum_{j=0}^{n} \frac{\tau\beta_j\left[p\gamma^j + (1-p)\right]}{1-(1-\tau)\left[p\gamma^j + (1-p)\right]}$$

and the mean accepted wage in the market is

$$(16) \qquad \bar{w} = \sum_{j=0}^{n} w_j \eta_j,$$

$$\eta_j = \left[p\gamma_j \sum_{r=0}^{J} \frac{\beta_r}{1-((1-p)+p\gamma^r)(1-\tau)}\right] \div \left[\sum_{j=0}^{n} \eta_j\right]$$

$$= \left[\gamma_j l(w_j)\right] \div \left[\sum_{j=0}^{n} \gamma_j l(w_j)\right]$$

where η_j is the probability that an individual receives and accepts a wage of w_j. The mean wage offer is obviously $\sum_{j=0}^{n} w_j \gamma_j$.

Because $\{\gamma_j\}$ and p are equilibrium outcomes, any comparative statics requires a full solution of the model. An exogenous change in the parameters would affect the entire solution in a complicated nonlinear way. Therefore, it is difficult to evaluate the properties of the equilibrium except through numerical calculations.

Albrecht and Axell demonstrate that there may be an optimal nonzero unemployment compensation level even if unemployment compensation increases unemployment, where the social objective is to maximize per-capita utility, i.e., per capita consumption plus the nonmarket value of time spent in unemployment. It is worth comparing the case we consider, where the offer probability is endogenous and possibly less than one, to their case where p is exogenously set at unity.

Total production (and consumption, C) is given by the product of the number of firms offering each wage $= F[A(\lambda_{j+1}) - A(\lambda_j)]$, the number of workers who will meet and accept each wage offer $= l(w_j)$, and the average productivity of firms offering each wage $= \int_{\lambda_j}^{\lambda_{j+1}} \lambda \, dA(\lambda)/[A(\lambda_{j+1}) - A(\lambda_j)]$ summed over each wage, i.e.,

$$(17) \qquad C = F \sum_{j=0}^{n} l(w_j) \int_{\lambda_j}^{\lambda_{j+1}} \lambda \, dA(\lambda).$$

The total value of nonmarket time while unemployed is

$$(18) \qquad L = k \sum_{j=0}^{n} \frac{z_j \tau\beta_j\left(p\gamma^j + (1-p)\right)}{1-\left[p\gamma^j + (1-p)\right](1-\tau)}.$$

It is tedious, but straightforward, to show that per capita utility can be expressed as

$$(19)\qquad u=\frac{C+L}{k}$$
$$=\sum_{j=0}^{n}\beta_j\int_{\lambda_j}^{\infty}\frac{\lambda\,dA(\lambda)}{1-A(\lambda_j)}+\sum_{j=0}^{n}\frac{\beta_j\tau\left[p\gamma^j+(1-p)\right]}{1-\left[p\gamma^j+(1-p)\right](1-\tau)}$$
$$\times\int_{\lambda_j}^{\infty}\frac{(z_j-\lambda)}{1-A(\lambda_j)}\,dA(\lambda).$$

The first term, as noted by Albrecht and Axell, is full employment per-capita output and the second is an adjustment due to unemployment, which, because $z_j<\lambda_j$, must be negative. If $p=1$, then this expression reduces to the $n+1$ type extension of the two type case considered by Albrecht and Axell. In the case where $p=1$, if unemployment compensation increases the unemployment rate, this direct effect is to reduce utility. However, an increase in unemployment compensation increases reservation wages, which increases the average productivity of active firms. Thus, full employment output increases. Albrecht and Axell demonstrate through simulation that there may be an optimal nonzero level of unemployment compensation.[12] However, a minimum wage intervention will always be optimal. Introducing a high enough minimum wage must decrease unemployment, e.g., a minimum wage set at w_n, and it also must increase full employment output. Indeed, a minimum wage equal to the maximum λ is optimal.

When p is endogenously determined by the proportion of active firms, a trade-off emerges with a minimum wage intervention, as it did with unemployment compensation. Because increasing the minimum wage will eventually increase the unemployment rate (as p falls), there is at least the possibility of the existence of an optimal *effective* minimum wage. However, it matters how p is related to $1-A(\lambda_0)$. For example, if $p=\alpha(1-A(\lambda_0))$ then $\mu p=(k/F)\alpha$, a constant, in which case the minimum wage equal to the maximum λ is still optimal. But, if $p=\alpha(1-A(\lambda_0))^2$ a trade-off emerges, and an optimal minimum wage below the maximum λ may exist.

Finally, the equilibrium implies that workers in larger firms will have higher wages simply because $l(w_j)$ is increasing in w_j. Less obviously, workers in more profitable firms will also have higher wages as shown in Lemma 2. These implications seem consistent with recent evidence on inter-industry wage differences (Krueger and Summers (1988)).

[12] Albrecht and Axell (1984) analytically show that an increase in unemployment compensation (b) could either increase or decrease unemployment in this model when there are only two types of people.

3. ESTIMATION ISSUES

The model generates a population distribution for unemployment durations and wages which can be compared to data. We now discuss how the parameters of the model, both supply and demand, can be estimated from panel data on individuals.

A. *Duration Data*

Consider a sample of individuals who enter the market at a point in time and are followed until a job offer is received and accepted. Let $d_i + 1$ be the period that individual i is observed to be working, i.e., d_i periods of unemployment. Let NWK indicate the state of not working and WK the working state. Because we do not know the type of any particular individual, the likelihood function using duration data only is given by

$$(20) \qquad L = \prod_{i=1}^{I}\left[\sum_{j=0}^{n} \Pr\left(NWK_1, NWK_2, \ldots, NWK_{d_i}, WK_{d_i+1}, type = j\right)\right], \\ = \prod_{i=1}^{I} f_{d_i+1}$$

where I is the number of observed individuals, d_i the duration of unemployment, and where f_{d+1} is given by (13). Note that if there are incomplete spells, the appropriate density function is divided by $(1-\tau)p(1-\gamma^j)$.

Writing the likelihood function (20) in terms of the model's parameters, letting $\pi_i = 1$ if the spell of length $d_i + 1$ is complete (*exactly* d_i periods of unemployment) and $\pi_i = 0$ if the spell of length d_i is incomplete (*at least* d_i periods of unemployment), we get after taking logs[13]

$$(21) \qquad \log L = \sum_i (1-\pi_i)\log\left[\sum_{j=0}^{n}(1-\tau)^{d_i-1}\left(p\gamma^j + (1-p)\right)^{d_i}\beta_j\right] \\ + \sum_i \pi_i \log\left[\sum_{j=0}^{n}\left[(1-\tau)^{d_i}\left(p\gamma^j + (1-p\right)^{d_i} p(1-\gamma^j)\beta_j\right]\right].$$

It should be understood that p and the $n+1$ γ's are themselves functions of the $n+1$ z's, δ, b, the $n+1$ β's, τ, the distribution parameters of $A(\lambda)$, and the parameters of the function G. Equation (18) can be considered as the unrestricted version of the equilibrium search model, i.e., with parameters τ, p, the γ's, and the β's. The equilibrium model imposes restrictions across the γ's and p.

Identification. Consider identification in the unrestricted model. Clearly, (21) is maximized at $\tau = 0$. The intuitive explanation for this is simply that we cannot distinguish leaving the market from incomplete spells, and so the true τ, i.e., the

[13] Actually there are two types of incomplete spells in the data, one type for whom there is no information after a particular period and another type who we know survived but not whether they were employed in the next period. In the latter case $1-\tau$ is raised to the d_i power.

one corresponding to the search model, is not identified in the unrestricted model. There are $2n+1$ free parameters (recall that the γ's and β's must each sum to one) still to be identified. The likelihood function for the unrestricted model is a finite mixture of negative binomial distributions. Identification of the parameters of this model has been established (see Blischke (1965) and Yakowitz and Spragins (1968)). Note that p is identified because $\gamma^0 = 0$. The question then is whether the parameters of the equilibrium search model, the "structural" parameters, can be identified from the "reduced form" parameters in (21). As noted, there are $2n+1$ reduced form parameters. If we assume that the number of parameters in the λ distribution is q and that there is one free parameter in the G function, then there are $2n+q+5$ structural parameters. A necessary condition for identification is therefore not satisfied without further structure. Clearly, b cannot be separately identified from the z's because it enters additively with the z's and only in the reservation wage equations (5a) and (5b). Further, without wage data the z's must be normalized so that one of them, say z_0, will be fixed *a priori*. Even if we fix the discount factor δ, there are still then $2n+q+2$ structural parameters. In order to satisfy the necessary conditions for identification, we assume that the z's are generated by

$$z_j - z_{j-1} = |\alpha_0 + \alpha_1 j + \alpha_2 j^2 + \cdots + \alpha_r j^r|.$$

Then, there are $n+(r+1)+q+2$ parameters to be estimated. Identification requires that $2n+1 \geqslant n+(r+1)+q+2$ or $n \geqslant r+q+2$. In the estimation we assume that λ is log normal so that $q=2$ and we restrict $r=2$, which implies that there must be at least seven types to identify the structural parameters.

B. *Duration and Wage Data*

So far we have made use only of the duration data and have ignored the information provided by knowledge of accepted wages. To make use of the wage data, however, requires further assumptions outside of the model. The model predicts that we should observe the same number of distinct wages as types of workers, i.e., $n+1$, while the wage is essentially a continuous variable in the data. A natural way to accommodate wage data is to assume that observed wages are measured with error generated from a parametric distribution. The likelihood function corresponding to (20) for a person with d periods of unemployment and observed accepted wage $\tilde{w}_i$ is

$$(22) \qquad L = \sum_{i=1}^{I}\left[\sum_{j=0}^{n}(1-\tau)^d\left(p\gamma^j + (1-p)\right)^d \times p\left(\sum_{k=j}^{n}\Pr(\tilde{w}_i|w_k, j)\Pr(w_k|j)\right)\beta_j\right]$$

where w_k is the true wage ($k=0,1,\ldots,n$). Recognizing that $\Pr(w_k|j)=\gamma_k$ and assuming that the conditional distribution of the observed wage is lognormal with mean $\ln w_k$ and standard deviation σ_k, (22) can be written as

$$(23) \qquad L = \sum_{i=1}^{I} \left[\sum_{j=0}^{n} (1-\tau)^{d} \left(p\gamma^{j} + (1-p) \right)^{d} \right.$$
$$\left. \times p \left(\sum_{k=j}^{n} \gamma_k \frac{1}{\sigma_k} \phi \left(\frac{\ln \tilde{w}_i - \ln w_k}{\sigma_k} \right) \right) \beta_j \right].$$

Incomplete spells, because no wage is observed, are treated in identical fashion as in (20), as are complete spells with missing wage information.

The unrestricted model with wage data clearly identifies the γ's, the β's, p, the w's, and the σ's. There are thus $4n+3$ reduced form parameters. With $b=0$, δ fixed, and $q=2$ as before, there are $3n+6$ structural parameters including the σ's. To identify all of the parameters n must be at least equal to 3, i.e., there must be four types. Thus, with wage data we can identify more parameters, in particular the structural form for the z's can be relaxed, with fewer types of individuals. The cost is a distributional assumption about measurement error in wages. In our view, an important advantage of ignoring the wage data in the estimation is that the wage data can then be used as an out-of-sample test of the equilibrium model. As we shall see, the point is rather moot given the performance of the equilibrium model in both cases.

4. THE DATA

The model is estimated using data from the National Longitudinal Survey of Labor Market Experience Youth Cohort (NLSY). The NLSY is a national sample of 12,686 youth who were 14 to 21 years old as of January, 1979. It contains a core random sample with oversamples of blacks, Hispanics, and disadvantages whites. These individuals have been surveyed annually since 1979 and detailed information on employment, schooling, family background, etc. has been collected. We have drawn a subsample consisting of all high school graduates in 1979, 1980, and 1981 who did not return to school through at least the 1983 interview. We do not believe the labor market to be the same for all schooling groups, although we do assume that the market is not segmented by race or sex. We define a "real" or full-time job to be one in which an individual worked at least 30 hours per week and which lasted at least three months. The duration of search to this job ("unemployment" duration) is the time it took to obtain the first real job after leaving school. We do not distinguish between out-of-the-labor force and "unemployment" states as is consistent with the model; unemployment and nonemployment are the same thing.

There are 1,624 individuals in the sample, 526 who graduated from high school in 1979, 550 in 1980, and 553 in 1981. Of the full sample, 987 had incomplete spells; 703 of them because hours worked was missing for the first job observed after graduation and so it could not be determined whether it was a "real" job, and 284 because no job was taken by the time of the 1983 interview. Of the 713

TABLE I

DESCRIPTIVE STATISTICS

	Mean	Standard Deviation
Duration of unemployment (weeks)[a]	45.0	49.1
Real weekly earnings: completed spells[b]	182	116
Tenure: completed spells	34.4	23.8

[a] The duration of unemployment is zero for those who obtained a job in the first week after graduation, one for those who obtained a job in the second week, etc.

[b] 1980 dollars.

incomplete spells of the first type, 285 occurred in the very first week after graduation, i.e., they had a job but did not report hours worked. Those observations contain no useful information and are thus dropped from the sample, leaving 1,339 individuals for analysis.

Table I provides descriptive statistics for the sample. The average duration of unemployment to the first real job is a little less than one year, 45 weeks, including complete and incomplete spells. Average tenure, restricted by the real job definition to be greater than 13 weeks, is 34 weeks, although this is an underestimate of completed tenure. The average real weekly wage is 182 dollars. As noted, we have combined three high school graduate cohorts (1979, 1980, and 1981). The steady state assumption of the model implies that all cohorts have the same distribution of non-market time. The average duration of unemployment is similar across the three cohorts, namely, 44, 48, and 44 weeks, respectively, as is the average real weekly wage, 167, 180, and 192 dollars in terms of 1980 prices, respectively.

Table II presents the empirical frequency for the duration of unemployment, while Table III presents the associated empirical hazard function. It is interesting to note from Table II that 11.1 percent of the sample were working at a real job in the first week after graduation, that *at least* 32.7 percent did not have a real job within one year of graduation, and *at least* 11.5 percent did not have a real job within two years of graduation. The observed hazard rates (Kaplan-Meier estimates; see Kalbfleisch and Prentice (1980)), as seen in Table III, decline with duration of nonemployment although not monotonically. Completed spells are fairly sparse after the sixth quarter, which may account for some of the lack of smoothness. The general (declining) shape of the hazard function is consistent with a pure heterogeneity model of nonemployment duration.

5. RESULTS

A. Duration Data

As baseline cases, we first present in Table IV estimates of the unrestricted model for two, seven, and ten types, i.e., $n = 1, 6, 9$ respectively, using duration data only. To reduce the number of parameters, we assume the β_j's are generated

TABLE II

DISTRIBUTION OF THE DURATION OF UNEMPLOYMENT

Number of Weeks	Frequency All Spells	Frequency Completed Spells	Percent All Spells	Cumulative Percent
0	148	148	11.1	11.1
1–12	334	167	24.9	36.0
13–25	202	79	15.1	51.1
26–38	124	50	9.3	60.4
39–51	92	67	6.9	67.3
52–64	45	36	3.4	70.7
65–77	53	33	4.0	74.7
78–90	104	10	7.8	82.5
91–103	80	15	6.0	88.5
104–116	30	13	2.2	90.7
117–129	8	8	0.5	91.2
130–142	40	4	3.0	94.2
143–155	33	4	2.5	96.7
156–168	8	1	0.5	97.2
169–181	3	2	0.2	97.4
182–211	35	0	2.6	100.0
	1339	637		

TABLE III

OBSERVED HAZARD RATES: KAPLAN-MEIER ESTIMATE

Week	1	2–13	14–26	27–39	40–52	53–65	66–78	79–91
Proportion employed by week i given not employed up to week i	.111	.152	.100	.083	.130	.083	.086	.033
Week	92–104	105–117	118–130	131–143	144–156	157–169	170–182	183–212
Proportion employed by week i given not employed up to week i	.081	.088	.063	.035	.069	.022	.053	0.0

from a binomial distribution with parameter p_β.[14] There are thus $n+2$ parameters estimated, $n\gamma_j$'s, p, and p_β.[15] The γ's can be interpreted in the unrestricted model as rejection probabilities. Then, for example, for seven types, the probability that type one rejects an offer is $\gamma^1 = .002E-4$, the probability that type two rejects is $\gamma^2 = .963$, the probability that type three rejects is $\gamma^3 = .993$, etc. Recall that type zero never rejects an offer, $\gamma^0 = 0$.

In all cases, the probability of receiving an offer, p, is very high; it is the lowest for $n=1$, 93 percent, and the highest for $n=9$, almost 100 percent. The β distribution for $n=6$ and $n=9$ are both peaked, rather than declining or increasing throughout. In addition, the γ_j distribution is also quite similar for $n=6$ and $n=9$, there being one type which rarely rejects (in addition to type zero) and the rest of the types who almost always reject. These results seem to

[14] $\beta_j = \binom{n}{j}(p_\beta)^j(1-p_\beta)^{n-j}$, $j=0,\ldots,n$. Note that this assumption is not relevant for the previous discussion of identification.

[15] We used the GQOPT package to perform all of the maximum likelihood estimation.

TABLE IV

MAXIMUM LIKELIHOOD ESTIMATES OF THE UNRESTRICTED MODEL
(Asymptotic Standard Errors in Parentheses)

	Two Types ($n = 1$)	Seven Types ($n = 6$) Duration Data	Seven Types ($n = 6$) Duration and Wage Data[a]	Ten Types ($n = 9$)
p	.926	.979	.941	.999
	(.035)	(.533 $E-3$)	(.112 $E-1$)	(.623 $E-2$)
p_β	.889	.510	.510	.367
	(.008)	(.298 $E-3$)	(.630 $E-2$)	(.184 $E-2$)
β_0	.111	.014	.014	.016
β_1	.889	.087	.087	.085
β_2		.225	.225	.197
β_3		.312	.312	.267
β_4		.243	.243	.233
β_5		.101	.101	.135
β_6		.018	.018	.052
β_7				.013
β_8				.002
β_9				.001 $E-1$
γ_0	.991	.162 $E-6$	.123 $E-6$	.282 $E-1$
	(.005 $E-1$)	(.760 $E-10$)	(.564 $E-6$)	(.138 $E-3$)
γ_1	.009	.963	.985	.936
		(.451 $E-3$)	(.098)	(.456 $E-2$)
γ_2		.301	.524 $E-2$	.274 $E-1$
		(.142 $E-4$)	(.335 $E-1$)	(.127 $E-3$)
γ_3		.862 $E-7$	.294 $E-7$	.215 $E-6$
		(.446 $E-10$)	(.091)	(.105 $E-8$)
γ_4		.666 $E-2$	.100 $E-1$	.877 $E-2$
		(.355 $E-5$)	(.092)	(.328 $E-4$)
γ_5		.735 $E-10$	.100 $E-72$	.220 $E-9$
		(.402 $E-13$)	(.117 $E-15$)	(.172 $E-12$)
γ_6		.732 $E-10$	.100 $E-72$	.178 $E-9$
				(.121 $E-12$)
γ_7				.146 $E-9$
				(.935 $E-12$)
γ_8				.208 $E-9$
				(.132 $E-11$)
γ_9				.787 $E-4$
Ln L	-3299.3	-3273.9	-3665.0	-3273.5

[a] The wage parameters and standard deviations of the measurement error associated with each wage are as follows (asymptotic standard errors in parentheses): $w_0 = 152.9$ (7.22), $w_1 = 161.3$ (3.51), $w_2 = 161.3$ (3.51), $w_3 = 167.0$ (2.77), $w_4 = 167.1$ (7.22), $w_5 = 169.3$ (3.24), $w_6 = 170.9$ (3.46), $\sigma_0 = .483\ E-1$ (.383 $E-1$), $\sigma_1 = .330$ (.301 $E-1$), $\sigma_2 = 1.75$ (2.32), $\sigma_3 = .915$ (1.73), $\sigma_4 = .336$ (.255 $E-1$), $\sigma_5 = .705\ E-4$ (.533 $E-4$), $\sigma_6 = .529\ E-1$ (.422 $E-1$).

indicate that if there are not exactly two types, there are at most two groups of individuals in the population.

The ln likelihood value declines considerably between $n = 1$ and $n = 6$, but is essentially unchanged between $n = 6$ and $n = 9$. A likelihood ratio test is, however, not appropriate because the restrictions are on the boundary of the parameter space, namely, some β_j's equal to zero. To get some feel for how much better the seven and ten types fit the data, Figure 1 graphs the implied survivor

FIGURE 1.—Survivor functions: unrestricted models.

function for $n = 1$ and 6, along with the Kaplan-Meier estimate. The fit is clearly superior for $n = 6$. As another indication, we calculated the χ^2 statistic for the comparison of predicted frequencies of unemployment durations and actual frequencies. Because there are many empty cells (out of the 213 periods there are 101 empty cells, periods in which no individual completed a spell), we calculated the χ^2 statistics in Table V using the classification scheme shown in Table III, approximately quarterly periods.[16] Although none of the specifications have χ^2 statistics below the critical value, the ordering is as suggested by the likelihood values and by the plot of the survivor functions. The second column of Table V considers the fit only for the first 50 weeks, taken individually, in which there are no empty cells. Again, none of the specifications pass the fit test; the ordering is consistent with the likelihood values.

Given the above results for the baseline unrestricted model, we estimated the equilibrium model for seven types only. As per our previous discussion, we adopted the following specification for the offer probability function: $p = \alpha(1 - A(\lambda_0))^2$. α is a free parameter estimated jointly with the other parameter of the equilibrium model. Because we treat p as a parameter, in the estimation the equilibrium model is numerically solved with p fixed. Recall that the necessary condition for identification using only duration data requires seven types, and the evidence is at least suggestive that the unrestricted model with seven types is appropriate. Table VI reports the estimates. We fixed the discount factor (δ) at .999 and z_0 was set equal to 100 as a scaling factor.

[16] The last quarterly period, because it is empty, is combined with the one preceding it.

TABLE V

χ^2 TEST STATISTICS COMPARING ESTIMATED FREQUENCIES TO ACTUAL FREQUENCIES OF UNEMPLOYMENT DURATION[a]

	All periods[b]	50 periods[c]
Unrestricted Model		
Two types	120.9	113.8
Seven types (Duration Data)	62.7	85.0
Seven types (Duration and Wage Data)	81.3	92.8
Ten types	64.9	84.0
Equilibrium Model		
Seven types (Duration Data)	643	543
Seven types (Duration and Wage Data)	1001	954

[a] The χ^2 statistic is defined as

$$\sum_{j=1}^{T} \frac{\left[j \prod_{k=1}^{j-1} (1-h_k))_{\text{data}} - \left\{ (1-\tau)^{j-1} \left[h_j \prod_{k=1}^{j-1} (1-h_k) \right] \right\}_{\text{estimate}} \right]^2}{\left[(1-\tau)^{j-1} h_j \prod_{k=1}^{j-1} (1-h_k) \right]_{\text{estimate}}} \cdot 1339$$

where h_j is the hazard rate in period j. For the data the hazard rate function is given by the Kaplan-Meier estimate, while for the unrestricted and equilibrium models, it is defined in (20) ($\tau = 0$ in the unrestricted model). In performing the tests, we adjust degrees of freedom by the number of estimated parameters. This is a conservative rejection criterion because the estimation of the frequencies in the models use all individual cells on which the test is conducted (see Kendall and Stuart (1979, p. 455)).

[b] This classification corresponds to Table III except that the last cell, which is empty, is combined with the preceding one. Reading down the column, the critical χ^2 values at the .05 significance level (12, 7, 7, 4, 7, 7 degrees of freedom, respectively) are 21.0, 14.1, 14.1, 9.5, 14.1, 14.1, respectively.

[c] The critical χ^2 values at the .05 significance level (47, 42, 42, 39, 42, 42 degrees of freedom, respectively) are 64, 58, 58, 54, 58, and 58, respectively. We use the approximation that $\sqrt{2\chi^2} - \sqrt{2n-1}$ is a normal deviate with unit variance, where n is the number of degrees of freedom.

Comparing Table IV and the first column of Table VI, it is clear that the ln likelihood of the unrestricted model is considerably larger than that of the equilibrium model. To get an idea of the fit of the model to the data, Figure 2 presents survivor functions. The equilibrium model seems to track the data for about the first 50 weeks, but is seriously divergent after that. Table V shows the χ^2 tests for the equilibrium model. The equilibrium model is resoundingly rejected using the quarterly data classification and is much worse than the unrestricted model. Consistent with Figure 2, the equilibrium model fits the data somewhat better in the first 50 weeks than over the entire period.

B. *Duration and Wage Data*

It is possible that the equilibrium model performs poorly relative to the unrestricted model because of the restrictions on the z_j generating function required for identification. Introducing wage data allows us to drop that restriction, at the cost of parametrically specifying a measurement error distribution to account for continuous wage observations. Before estimating the equilibrium

TABLE VI

MAXIMUM LIKELIHOOD ESTIMATES OF THE EQUILIBRIUM MODEL

	Seven Types[a,b] Duration Data	Seven Types[b] Duration and Wage Data
z_0	100	27.2
z_1	259	56.3
z_2	418	77.9
z_3	577	103
z_4	737	130
z_5	896	152
z_6	1055	154
λ	6.03	5.19
σ_λ	.950 $E-2$	.171 $E-1$
p_β	.332	.396
β_0	.089	.049
β_1	.265	.191
β_2	.329	.313
β_3	.218	.273
β_4	.082	.134
β_5	.016	.035
β_6	.001	.004
τ	1.94 $E-8$	.121 $E-17$
α	.625 $E-1$	.155 $E-1$
σ_0	—	1.53
σ_1	—	8.43
σ_2	—	11.3
σ_3	—	102
σ_4	—	.329
σ_5	—	.586 $E-7$
σ_6	—	1.393
$\ln L$	-3429.6	-3893.8

[a] $\delta = .999$, $z_0 = 100$, $z_j - z_{j-1} = |\alpha_0 + \alpha_1 z_{j-1} + \alpha_2 z_{j-1}^2|$; α_0, α_1 and α_2 are the estimated parameters ($\hat{\alpha}_1 = 159$, $\hat{\alpha}_2 = -.145\,E-3$, $\hat{\alpha}_3 = .955\,E-4$).
[b] $\delta = .999$.

model incorporating wage information, we estimated the unrestricted model for seven types augmented with the wage information. The results, reported in Table IV, reveal the unrestricted estimates to be qualitatively unchanged. There are again essentially only two types. It is interesting to note that the wage mass points (see the footnote to Table IV) are themselves not very disperse, the range from the lowest to highest being less than 20 dollars per week. The fit to the duration data of the unrestricted model estimated with the wage data is shown graphically in Figure 1 and in Table V. The fit to duration data is almost the same when the wage data is used in estimation.

The second column of Table VI reports the maximum likelihood estimates for the equilibrium model using the wage data. The only large differences arise in the estimates of τ, α, and the z's. Note that the value of leisure is estimated to range from 27 to 154 dollars per week, with an average value of 88 dollars. The mean offered wage is 148 dollars per week (see Table VIII). The χ^2 goodness of fit statistic for the duration data shown in the last row of Table V is significantly worse than the model fit only on duration data, as might have been anticipated.

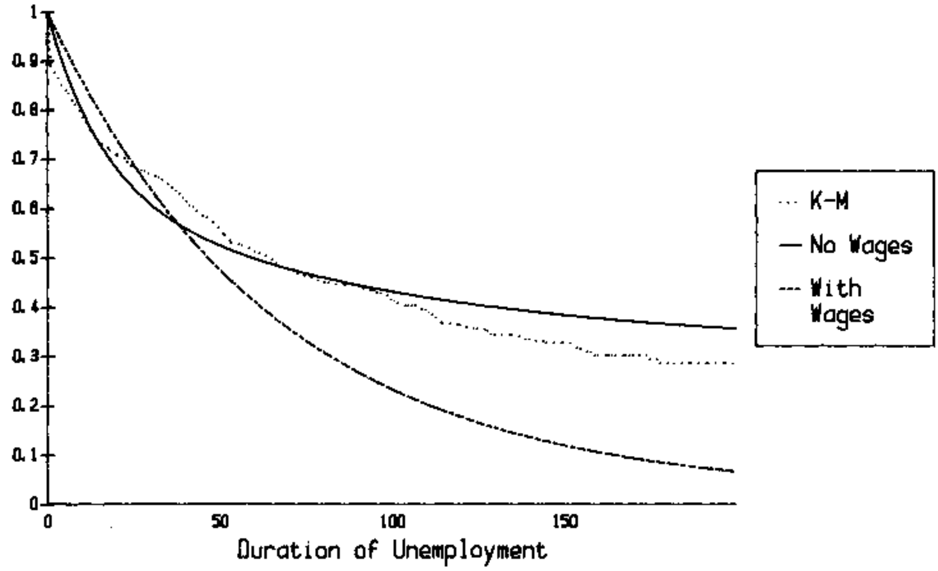

FIGURE 2.—Survivor functions: equilibrium models.

The equilibrium model using wage data does slightly better in fitting the first 50 periods than in fitting the entire 212 periods as did the equilibrium model, which did not use wage data. It does not do better than the estimated model without wage data. This is also evident in Figure 2. Thus, the restriction on the z-function does not do so much harm as to be misleading in deriving conclusions about the validity of the equilibrium model, at least conditional on the parametric assumption about measurement error.

As another indication of the fit between the data and the models, we calculated the expected duration of unemployment for the data and for the estimated unrestricted and equilibrium models. Table VII indicates that the expected

TABLE VII

COMPARISON OF EXPECTED DURATION OF UNEMPLOYMENT (IN WEEKS) FOR ALTERNATIVE MODELS[a]

	Duration $\leq$ 212	Duration $\leq$ 999
Actual Mean Duration (complete and incomplete spells)	45.0	—
Kaplan-Meier Estimate	47.5	—
Unrestricted Model		
Two types	66.8	109.8
Seven types (Duration Data)	54.2	102.0
Seven types (Duration and Wage Data)	59.5	83.6
Ten types	51.1	84.5
Equilibrium Model		
Seven types (Duration data)	39.5	52.8
Seven types (Duration and Wage data)	56.2	72.5

[a] $E(d|d \leq D) = \Sigma_{d=0}^{D} df_{d+1}/F_{D+1}$ where f_{d+1} is defined in (19) and $F_{D+1} = \Sigma_{d=0}^{D} f_{d+1}$.

TABLE VIII

ESTIMATED DISTRIBUTION OF OFFERED AND ACCEPTED WAGES

Normalized Wages		Offer Wage Probabilities (γ_j)[a]		Accepted Wage Probabilities (η_j)[b]	
Duration Data	Duration and Wage Data (Estimated True Wages)	Duration Data	Duration and Wage Data	Duration Data	Duration and Wage Data
1.0	1.0 (141.0)	.145 $E-2$	.114 $E-37$	.188 $E-3$	.553 $E-39$
1.006	1.012 (142.7)	.829	.138 $E-29$	.430	.332 $E-30$
1.040	1.021 (144.0)	.170	.109 $E-52$	.570	.605 $E-53$
1.436	1.033 (145.6)	1.0 $E-20$	.304 $E-13$	.103 $E-13$	.251 $E-13$
1.832	1.044 (147.2)	1.0 $E-20$	.846	.142 $E-13$	.813
2.228	1.091 (153.8)	1.0 $E-20$	.011	.149 $E-13$	.013
2.624	1.095 (154.4)	1.0 $E-20$	.143	.150 $E-13$	.174

[a] Mean offered wage = 1.012 for duration data, = 1.052 for duration and wage data.
[b] Mean accepted wage = 1.025 for duration data, = 1.055 for duration and wage data.

duration in the data (column one), using the Kaplan-Meier estimate, is 47.5 weeks. The unrestricted models tend to overpredict relative to the Kaplan-Meier estimate, the prediction improving as the number of types increases. The equilibrium model using duration data underpredicts the within-sample expected duration, while the equilibrium model using duration and wage data overpredicts; however, the prediction within the former case is smaller in absolute value than in the latter. We also computed the expected duration conditional on duration less than 1000 weeks, which is well outside the sample data, for the unrestricted and equilibrium models. The equilibrium model's data predicts a much shorter average duration than the unrestricted models. The equilibrium model with duration and wage data, however, is closer to the unrestricted model.

The estimated offer wage distribution (the γ's) for the equilibrium models is shown in Table VIII.[17] The wage distribution is normalized so that the lowest wage equals unity. All other wages represent proportionate increases relative to the lowest wage. In terms of the offer probability, both equilibrium model estimates are very much smaller, .0625 and .0150, than the estimate from the respective unrestricted model, .979 and .941. Both equilibrium models also indicate the existence of two groups (within the seven types) although less markedly than the unrestricted model estimates. The mean offered wage in both cases, as calculated from the offer distribution, is only slightly higher than the lowest wage.

The accepted wage distribution, the η_j's in equation (23), is shown in the third column of Table VIII. The distribution for the equilibrium model using duration data is much less concentrated than the offer distribution and has a peak at the third lowest wage. The mean accepted wage is 2.6 percent higher than the

[17] The γ's are derived from (12). However, due to rounding error, when the first λ_j becomes greater than 19 standard deviations above the mean of the λ distribution, we assumed that γ_{j-1} was equal to $1-\gamma^{j-1}-(n-j-1)10^{-20}$ and all subsequent γ's $(\gamma_j, \ldots, \gamma_n)$ were set to 10^{-20}. None of the results are affected by using values different from 10^{-20}, e.g., 10^{-30}, 10^{-40}, or 10^{-50}. We also checked for nonuniqueness of the equilibrium around the estimated values. Small variations in starting values for the γ's did not alter the equilibrium values.

minimum offered wage. The highest wage is 2.62 times greater than the lowest wage. The distribution for the equilibrium model estimated with wage data has a peak at the third highest wage. The mean accepted wage is 5.5 percent higher than the minimum offered wage and the highest wage is only 9.5 percent greater than the lowest wage. As indicated in Table I, the mean real weekly wage is 1.6 times as large as its standard deviation. Comparable figures for the equilibrium models are 61 when duration data alone is used and 52 when both duration and wage data are used. The degree of measurement error necessary to account for the observed wage variation in neither case is credible.

6. SIMULATING MINIMUM WAGE EFFECTS

Given the parameter estimates of the equilibrium search model, it is possible to simulate the impact of alternative levels of the minimum wage on unemployment and wages. Because the equilibrium model does not fit either the duration or wage data very well, it is best to view this exercise as illustrative only. For this reason we do not present the results for both estimated equilibrium models. We choose the model estimated only with duration data because it fits the duration data better and does not generate a significantly lower degree of "true" wage dispersion.

As already noted, it is *a priori* unclear in this model whether the unemployment rate or the expected duration of unemployment will be adversely affected when a binding minimum wage is introduced or increased. Therefore, Table IX presents a number of simulations for the equilibrium model.[18] We take a

[18] Introducing a minimum wage alters the structure of the equilibrium model. In particular, letting $\underline{w}$ be the minimum wage, suppose $w_j < \underline{w} < w_{j+1}$. Equilibrium wages are now $\underline{w}, w_{j+1}, \ldots, w_n$. In addition,

$$l(\underline{w}) = \frac{\mu p \sum_{i=0}^{j} \beta_i}{1-(1-p)(1-\tau)},$$

$$l(w_k) = l(\underline{w}) + \sum_{r=j+1}^{k} \frac{\mu p \beta_r}{1-[(1-p)+p\gamma^r](1-\tau)} \qquad (k=j+1,\ldots,n),$$

$$\lambda_0 = \underline{\lambda} = \underline{w},$$

$$\lambda_k = \frac{w_k l(w_k) - w_{k-1} l(w_{k-1})}{l(w_k) - l(w_{k-1})} \qquad (k=j+2,\ldots,n),$$

$$\underline{\gamma} = \frac{A(\lambda_{j+1}) - A(\underline{\lambda})}{1-A(\underline{\lambda})},$$

$$\gamma_k = \frac{A(\lambda_{k+1}) - A(\lambda_k)}{1-A(\underline{\lambda})} \qquad (\kappa = j+1,\ldots,n),$$

$$p = \alpha(1-A(\underline{\lambda}))^2.$$

The equilibrium is obtained numerically using the same algorithm (ZSPOW) but with the above substitutions.

TABLE IX

MINIMUM WAGE SIMULATIONS

Minimum Wage (relative to w_0)	Expected Duration of Unemployment ($d \leq 212$)	Normalized Unemployment Rate	Mean Offered Wage	Mean Accepted Wage	Probability of Receiving an Offer (p)	Per Capita Utility (u)
1.01	39.6	1.0	1.015	1.028	.622 $E-1$	501.7
1.03	44.3	1.003	1.003	1.036	.367 $E-1$	485.8
1.05	103.8	1.003	1.050	1.050	.592 $E-3$	491.6
1.07	106.0	2.167	1.070	1.070	.165 $E-7$	453.7
1.09	106.0	3.151	1.090	1.090	.437 $E-15$	417.3
1.10	106.0	3.155	1.100	1.100	.539 $E-20$	417.3
Baseline (1.00)	39.5	1.003	1.012	1.026	.625 $E-1$	571.3

minimum wage at several different values between one percent and ten percent above the lowest wage.

As must be the case, as the minimum wage increases the offer probability falls, from .063 with no minimum wage to one hundredth of that at a five percent minimum wage. The expected duration of unemployment increases monotonically with an abrupt jump between a three and five percent minimum, while the unemployment rate is not very sensitive for minimum wages up to five percent but abruptly increases thereafter. At a ten percent minimum wage the offer probability is so close to zero that the unemployment rate is essentially unity. As the minimum wage increases, the wage offer distribution becomes almost single valued at the minimum wage so that both the mean offered and mean accepted wage are almost identical.

Table IX also shows how per capita utility (equation 26) varies with the minimum wage. As the minimum increases, the level of per capita utility declines, although not quite monotonically, and this is true even when the unemployment rate declines.[19] It appears that a zero minimum wage is optimal in the sense of maximizing per capita utility.

7. CONCLUSIONS

In this paper we have demonstrated the feasibility of estimating a Nash labor market equilibrium search model using only information on workers. The equilibrium model is based on workers who are homogeneous in terms of market productivity and heterogeneous in nonmarket productivity. Due to this heterogeneity, the frequency distribution of search duration is a mixture of negative binomial distributions. The equilibrium model places restrictions across the parameters of the mixture. We perform goodness-of-fit tests on both the unrestricted and equilibrium models. While the unrestricted specification seemed to fit the data reasonably well, the equilibrium model gave poor results. The equilibrium model did not fit the duration of unemployment distribution very well, and

[19] Notice that from (19) per capita utility may decline when the unemployment rate declines as long as the unemployment rate increases for some types.

the estimates imply that almost all of the wage dispersion is measurement error. Possibly a model which incorporates fluctuations in cohort size, or some form of true duration dependence would provide a better description of the data. Adequate equilibrium models of the labor market are needed in order to provide a more rigorous basis for quantitative policy analysis.

Department of Economics, Tel-Aviv University, Tel-Aviv, Israel, and Boston University, Boston, MA, U.S.A.

and

Department of Economics, University of Minnesota, Minneapolis, MN 55455, U.S.A.

Manuscript received February, 1987; final revision received June, 1989.

APPENDIX

1. Proof of Lemma 1: Rewrite (5a) as

$$w_j = \left\{ [1-\delta(1-\tau)](z_j+b) + \delta(1-\tau)p\left[\gamma_{j+1}w_{j+1} + \sum_{i=j+2}^{n} \gamma_i w_i\right]\right\} \times \left\{\frac{1}{1-\delta(1-\tau)+\delta(1-\tau)p\sum_{i=j+1}^{n}\gamma_i}\right\}. \tag{A1}$$

Substituting z_{j+1} for z_j yields

$$w_j < \left\{ [1-\delta(1-\tau)](z_{j+1}+b) + \delta(1-\tau)p\sum_{i=j+2}^{n}\gamma_i w_i + \delta(1-\tau)p\gamma_{j+1}w_{j+1}\right\} \times \left\{\frac{1}{1-\delta(1-\tau)+\delta(1-\tau)p\sum_{i=j+1}^{n}\gamma_i}\right\}. \tag{A2}$$

The first two terms on the right hand side of (A2) can be replaced by (A1) for $j+1$, so that

$$w_j < \left\{\left[1-\delta(1-\tau)+\delta(1-\tau)p\sum_{i=j+2}^{n}\gamma_i\right]w_{j+1} + \delta(1-\tau)p\gamma_{j+1}w_{j+1}\right\} \times \left\{\frac{1}{1-\delta(1-\tau)+\delta(1-\tau)p\sum_{i=j+1}^{n}\gamma_i}\right\} = w_{j+1}. \tag{A3}$$

Q.E.D.

2. Proof of Lemma 2: From (8),

$$\lambda_j = w_j + \frac{(w_j - w_{j-1})}{l(w)-l(w_{j-1})} l(w_{j-1}).$$

Then $\lambda_1 > w_1 > w_0 = \lambda_0$ follows from Lemma 1 ($w_j > w_{j-1}$) and the assumption that $l(w_j) > l(w_{j-1})$.

To prove the second part of the lemma, define the critical profit level associated with offering wage w_j as

$$\pi_j = (\lambda_j - w_j)\,l(w_j) = (\lambda_j - w_{j-1})\,l(w_{j-1}). \tag{A4}$$

Thus, $\pi_{j-1} = (\lambda_{j-1} - w_{j-1})l(w_{j-1})$. But since $\lambda_j > \lambda_{j-1}$, it is immediate that $\pi_j > \pi_{j-1}$. Further, it is straightforward to show that $(\lambda - w_{j+1})l_{j+1} < (\lambda - w_j)l_j$. Hence, for any $\lambda \in [\lambda_j, \lambda_{j+1})$ the strategy that yields maximum profit is to offer w_j with profits equal to $(\lambda - w_j)l(w_j)$. *Q.E.D.*

3. Proof of the Existence Proposition: Let

$$\Delta = \left\{ \gamma = (\gamma_0, \gamma_1, \dots, \gamma_n, p) \,|\, 0 \leqslant p \leqslant 1, 0 \leqslant \gamma_i \leqslant 1, \sum_{i=0}^{n} \gamma_i = 1 \right\}$$

be a compact convex subset of R^{n+2}. Equation (5) in the text describes a continuous mapping (using Lemma 1) from Δ to R^{n+1}. For p approaching zero the limit of the ratio $l(w_{j-1}/l(w_j)$ is less than one and well defined; hence, λ_j is well defined by equation (8) for $p = 0$. Therefore equation (8) defines a continuous mapping from Δ to R^{n+1}, where $l(w_j)$ is given as in equation (10) and $l(w_j)$ is continuous on Δ. The above chain of continuous mappings shows that for $\lambda_j > \lambda_{j-1}$ for all j and γ that belong to Δ, equations (11) and (12) define a continuous mapping from Δ to Δ. Hence by Brouwer's fixed point theorem there exists a Nash equilibrium. *Q.E.D.*

4. Conditions for Increasing λ_j's: The existence of an equilibrium requires that the λ_j's be increasing in j. Recall that λ_j can be written as

$$\lambda_j = w_j + \frac{w_j - w_{j-1}}{l(w_j) - l(w_{j-1})}\, l(w_{j-1}).$$

Clearly, since w_j and $l(w_{j-1})$ are increasing in j, it is sufficient for λ_j to be increasing in j that $(w_j - w_{j-1})/[l(w_j) - l(w_{j-1})]$ be increasing in j. Now, from (A.1), (A.2), and (10) we get that

$$\frac{w_j - w_{j-1}}{l(w_j) - l(w_{j-1})} = \frac{\left[1 - (1-\tau)\left(1 - p\sum_{i=j}^{n}\gamma_i\right)\right][1 - \delta(1-\tau)]}{\left[1 - \delta(1-\tau)\left(1 - p\sum_{i=j}^{n}\gamma_i\right)\right]p\mu} \cdot \frac{z_j - z_{j-1}}{\beta_j}.$$

A sufficient condition for λ_j to increase is that $(z_j - z_{j-1})/\beta_j$ be increasing in j sufficiently. Thus, an equilibrium is assured if ever higher values of leisure are associated with ever lower proportions of individuals such that the ratio of the change in the value of leisure to the proportion of individuals with the higher value of leisure is increasing. Thus, some form of smoothness in the z's and β's, without large jumps up and down, is required.

REFERENCES

Albrecht, J. W., and B. Axell (1984): "An Equilibrium Model of Search Employment," *Journal of Political Economy*, 92, 824–840.

Axell, B. (1977): "Search Model Equilibrium," *Scandinavian Journal of Economics*, 79, 20–40.

Blischke, W. R. (1965): "Mixtures of Discrete Distributions," in *Classical and Contagious Discrete Distributions*, ed. G. P. Patil. Calcutta: Statistical Publishing Society.

Burdett, K., and K. Judd (1983): "Equilibrium Price Dispersion," *Econometrica*, 79, 955–970.

Eckstein, Z., and K. Wolpin (1989): "Dynamic Labor Force Participation of Married Women and Endogenous Work Experience," *Review of Economic Studies*, 56, 375–390.

Flinn, C., and J. Heckman (1982): "New Methods for Analyzing Structural Models of Labor Force Dynamics," *Journal of Econometrics*, 18, 115–168.

Gotz, G. A., and J. J. McCall (1984): "A Dynamic Retention Model for Air Force Officers," mimeo, Rand Corporation.

KALBFLEISH, J., AND R. PRENTICE (1980): *The Statistical Analysis of Failure Time Data*. New York: John Wiley and Sons.

KENDALL, M., AND A. STUART (1976): *The Advanced Theory of Statistics*, 4th ed., Vol. 2. London: Charles Griffin and Co.

KIEFER, N., AND G. NEUMANN (1979): "An Empirical Job-Search Model with a Test of the Constant Reservation Wage Hypothesis," *Journal of Political Economy*, 87, 89–108.

KRUEGER, A. B., AND L. H. SUMMERS (1988): "Efficiency Wages and the Inter-Industry Wage Structure," *Econometrica*, 56, 259–294.

LANCASTER, T., AND A. CHESHER (1983): "An Econometric Analysis of Reservation Wages," *Econometrica*, 51, 1661–1676.

MILLER, R. (1984): "An Estimate of a Job Matching Model," *Journal of Political Economy*, 92, 1096–1120.

NARENDRANATHAN, W., AND S. NICKELL (1983): "Modelling the Process of Job Search," mimeo, Center for Labor Economics.

PAKES, A. (1986): "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks," *Econometrica*, 54, 755–784.

PERRY, M. (1986): "An Example of Price Formation in Bilateral Situations: A Bargaining Model with Incomplete Information," *Econometrica*, 54, 313–321.

ROB, R. (1985): "Equilibrium Price Distributions," *Review of Economics Studies*, 52, 487–504.

ROTHSCHILD, M. (1973): "Models of Market Organization with Imperfect Information: A Survey," *Journal of Political Economy*, 81, 1283–1308.

RUST, J. (1984): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, 999–1034.

STERN, S. (1985): "Estimating a Nonsequential Search Model," mimeo, University of Virginia.

TOIKKA, R. (1976): "A Markovian Model of Labor Market Decisions by Women," *American Economic Review*, 66, 821–834.

WOLPIN, K. I. (1984): "An Estimable Dynamic Stochastic Model of Fertility and Child Mortality," *Journal of Political Economy*, 92, 852–874.

——— (1987): "Estimating a Structural Search Model: The Transition from School to Work," *Econometrica*, 55, 801–817.

YAKOWITZ, S., AND S. SPRAGINS (1968): "On the Identifiability of Finite Mixtures," *Annals of Mathematical Statistics*, 39, 209–214.