Journal of Econometrics 18 (1982) 115–168. North-Holland Publishing Company

# NEW METHODS FOR ANALYZING STRUCTURAL MODELS OF LABOR FORCE DYNAMICS*

C. FLINN

*University of Wisconsin, Madison, WI 53706, USA*

J. HECKMAN

*University of Chicago, Chicago, IL 60637, USA*

The economic theory of decision-making under uncertainty is used to produce three econometric models of dynamic discrete choice: (1) for a single spell of unemployment; (2) for an equilibrium two-state model of employment and non-employment; (3) for a general three-state model with a non-market sector. The paper provides a structural economic motivation for the continuous time Markov (or more generally 'competing risks') model widely used in longitudinal analysis in biostatistics and sociology, and it extends previous work on dynamic discrete choice to a continuous time setting. An important feature of identification analysis is separation of economic parameters that can only be identified by assuming arbitrary functional forms from economic parameters that can be identified by non-parametric procedures. The paper demonstrates that most econometric models for the analysis of truncated data are non-parametrically under-identified. It also demonstrates that structural estimators frequently violate standard regularity conditions. The standard asymptotic theory is modified to account for this essential feature of many structural models of labor force dynamics. Empirical estimates of an equilibrium two-state model of employment and non-employment are presented.

## 1. Introduction

This paper presents new econometric methods for the analysis of labor force dynamics. The economic models we discuss assume that rational agents make choices about their employment and labor force activity in the face of uncertainty about key aspects of their labor market environment. A variety of such models has recently appeared in the literature [e.g., Burdett and Mortensen (1978, 1980), Lucas and Prescott (1974), Wilson (1980), Lippman and McCall (1976a,b, 1980)] and research in the area continues to flourish.

*This research was supported by NSF Grant SOC 77-27136 to NBER and Labor Department Contract No. 20-17-80-44 from ASPER-DOL to NORC. We thank Tom Coleman, B. Jovanovic, R. Miller, B. Payner and B. Singer for comments. Portions of this paper were presented in an invited lecture at the World Meeting of the Econometric Society, Aix-en-Provence, and at a seminar at Northwestern University in 1980. This paper (and its companion) were presented at the summer meetings of the North American Econometric Society, San Diego, June 1981, and the 2nd Annual meetings of the Latin American Econometric Society in Rio de Janiero, July 1981.

Theoretical results already in hand offer possible explanations for frictional unemployment, on the job wage growth, job turnover and wage dispersion.

To date there has been little systematic econometric analysis of these models although qualitative implications obtained from these models have sometimes been used as loose guides to interpreting unemployment duration regressions. The few attempts at developing econometric procedures to estimate this class of models are direct transcriptions of econometric models used to estimate reservation wages in the analysis of female labor supply.[1] A clear analysis of identification criteria and estimation methods for models of labor force dynamics does not appear in the literature.

This paper takes a first step toward developing econometric models that implement and extend the new theory. A key feature of the class of economic models that we consider is that they produce predictions about the dynamics of individual behavior. With the recent wide-scale availability of longitudinal labor market data with repeated observations on individuals it has become possible to test the implications of the new microdynamic models using microdynamic data. The goal of this paper is to develop procedures to enable analysts to apply the new theories to the new data.

This paper makes three contributions. First, economic theory is used to derive the appropriate econometric specification for three models of labor force dynamics. They are (a) a continuous time single-spell model of search unemployment, (b) an equilibrium continuous time two-state model of employment and non-employment, and (c) a three-state model of employment, unemployment and non-market activity.

Second, we present new identification criteria that must be satisfied in order to estimate the structural parameters of the three models considered in this paper. We distinguish structural parameters that can be estimated using non-parametric procedures from structural parameters that can only be identified if arbitrary functional forms are assumed. One identification condition ('recoverability') applies to all econometric models for the analysis of truncated data. An important conclusion of our analysis is that most econometric models for the analysis of truncated data are non-parametrically underidentified, and some are parametrically underidentified as well.

Our third contribution recognizes that maximum likelihood estimators for most of the models considered here are non-regular. We develop an appropriate asymptotic theory for the non-regular case that arises in the economic models analyzed in this paper, and in other econometric models derived from optimization theory.[2]

The plan of this paper is as follows. We first consider in detail a model for

[1]For a survey of such models, see Heckman and MaCurdy (1981). See also the discussion in appendix C which makes precise the statement in the text.

[2]For example, frontier production theory. See e.g. Aigner and Chu (1968), Green (1980) and Forsund, Lovell and Schmidt (1980).

a single spell of an unemployed worker's search from a known distribution of wage offers. Most of the new econometric points that arise in the analysis of the economically more interesting models that follow also appear in the simpler setting of the one state model of search.

The second part of the paper is devoted to analysis of a new equilibrium two-state model of employment and non-employment. Empirical estimates of this model are presented. An appendix presents a three-state model of labor force dynamics which introduces non-market activity into conventional search models and provides a structural economic motivation for the competing risks model widely used in social science.[3] The paper concludes with a summary.

## 2. Estimation of a one-state model of search unemployment

In this section, we consider estimation of a single-spell model of an unemployed worker's search from a known distribution of wage offers using longitudinal data on a random sample of agents. Despite the fact that this model is widely cited as a justification for much empirical research on unemployment and despite previous attempts at 'structural estimation' of this model, a number of important econometric aspects of this model have been neglected.

We first present a continuous time economic model for a stationary environment. We compare the continuous time model to previous discrete time formulations. We then consider issues of identification and estimation first under the assumption of no interpersonal variation in parameters ('no heterogeneity') and then with such variation ('heterogeneity'). In a concluding subsection, we consider econometric issues for a non-stationary search model.

### *2.1. A one-state model of search unemployment*

This model is well exposited in Lippman and McCall (1976a). Agents are assumed to be income maximizers (here and in the rest of the paper). If an instantaneous cost $c$ is incurred, job offers arrive from a Poisson process with parameter $\lambda$ independent of the level of $c$ ($c>0$). The probability of receiving a wage offer in time interval $\Delta t$ is $\lambda\Delta t+\mathrm{o}(\Delta t)$.[4] The probability of two or more job offers in interval $\Delta t$ is thus negligible.[5]

Successive wage offers are independent realizations from a known absolutely continuous wage distribution $F(x)$ with finite mean that is

[3]For much more extensive discussion of this model, see Coleman and Heckman (1981).

[4]$\mathrm{o}(\Delta t)$ is defined as a term such that $\lim_{\Delta t\to 0}\mathrm{o}(\Delta t)/\Delta t\to 0$.

[5]For one justification of the Poisson wage arrival assumption, see e.g. Burdett and Mortensen (1978).

assumed to be common to all agents (until section 2.5). Once refused, wage offers are no longer available. Jobs last forever and there is no on the job search. Workers live forever. (But see the discussion in section 2.5.3.) The instantaneous rate of interest is $r$ ($>0$).

$V$ is the value of search. Using Bellman's optimality principle for dynamic programming [see e.g. Ross (1970)], $V$ may be decomposed into three components plus a negligible component [of order $o(\Delta t)$],

$$V=-\frac{c\Delta t}{1+r\Delta t}+\frac{(1-\lambda\Delta t)}{1+r\Delta t}V+\frac{\lambda\Delta t}{1+r\Delta t}E\max[x/r;V]+o(\Delta t), \quad \text{for} \quad V>0, \tag{2.1}$$
$$=0 \quad \text{otherwise.}$$

The first term on the right of (2.1) is the discounted cost of search in interval $\Delta t$. The second terms is the probability of not receiving an offer $(1-\lambda\Delta t)$ times the discounted value of search at the end of interval $\Delta t$. The third term is the probability of receiving a wage offer $(\lambda\Delta t)$ times the discounted value of the expected value [computed with respect to $F(x)$] of the maximum of the two options confronting the agent who receives a wage offer: to take the offer (with present value $x/r$) or to continue searching (with present value $V$). Note that equation (2.1) is defined only for $V>0$. If $V=0$, we may define the agent as out of the labor force [see Lippman and McCall (1976a)]. From stationarity, once out the agent is always out. Sufficient to ensure the existence of an optimal reservation wage policy in this model is $E(|x|)<\infty$ [Robbins (1970)].

Collecting terms in (2.1) and passing to the limit, we reach the familiar formula [Lippman and McCall (1976a)]

$$c+rV=(\lambda/r)\int_{rV}^{\infty}(x-rV)\mathrm{d}F(x) \quad \text{for} \quad V>0, \tag{2.2}$$

where $rV$ is the reservation wage. It is implicitly determined from (2.2). For any offered wage $x\geqq rV$, the agent accepts the offer. The probability that an offer is unacceptable is $F(rV)$.

If $r=0$, an optimal policy still exists if $E(|l|)<\infty$ where $l$ is the lump sum value of the job offer over the infinite horizon. A reservation value $R$ exists. It is determined as the solution to

$$c=\lambda\int_{R}^{\infty}(l-R)\mathrm{d}G(l) \quad \text{if} \quad R>0, \tag{2.2'}$$

where $G(l)$ is the cdf of $l$.

To calculate the probability that an unemployment spell $T_u$ exceeds $t_u$, two ingredients are required. First, we must compute the probability that $j$ offers are received in time interval $t_u$. By assumption this is

$$\text{prob}(j \text{ offers} \mid t_u) = (\lambda t_u)^j \mathrm{e}^{-\lambda t_u}/j!, \qquad \lambda > 0. \tag{2.3}$$

We next need to compute the probability that none of the $j$ offers is acceptable. This is $[F(rV)]^j$. Assuming independence of arrival times and wage offers, the *survivor function* $P(T_u > t_u)$ is the product of these two probabilities summed over all $j$, i.e.,

$$P(T_u > t_u) = \sum_{j=0}^{\infty} ((\lambda t_u)^j/j!)\,\mathrm{e}^{-\lambda t_u}[F(rV)]^j = \mathrm{e}^{-\lambda(1-F(rV))t_u},^{6} \tag{2.4}$$

so that the density of $t_u$ is $f(t_u) = \lambda(1-F(rV))\,\mathrm{e}^{-\lambda(1-F(rV))t_u}$.

Accepted wages are truncated random variables with $rV$ as the lower point of truncation. The density of accepted wages is

$$f(x \mid x > rV) = \frac{f(x)}{1-F(rV)}, \qquad x \geqq rV. \tag{2.5}$$

Thus the one-spell search model has the same statistical structure for accepted wages as other models of self selection in labor economics [Lewis (1974), Heckman (1974), and Heckman and MaCurdy (1981)].

From the assumption that wages are distributed independently of wage arrival times, the joint density of duration times $t_u$ and accepted wages ($x$) is the product of the density of each random variable,

$$\begin{aligned} m(t_u, x) &= \{\lambda(1-F(rV))\exp - \lambda(1-F(rV))t_u\}\frac{f(x)}{1-F(rV)} \\ &= (\lambda \exp - \lambda(1-F(rV))t_u) f(x), \qquad x \geqq rV. \end{aligned}$$

The stationary search model generates a duration model with a constant hazard rate $h_u(t_u)$ where

$$h(t_u) = -\mathrm{d}\ln P(T_u > t_u)/\mathrm{d}t_u = \lambda(1-F(rV)). \tag{2.6}$$

Higher values of the hazard are associated with more rapid exit from the unemployment state. A model with a constant hazard is said to exhibit no duration dependence $[\mathrm{d}h(t_u)/\mathrm{d}t_u = 0]$.

[6]The power series expansion definition of the exponential is used in the final step.

Differentiation of (2.2) (and integration by parts) reveals that the reservation wage increases with positive translations of the wage offer distribution, decreases in the cost of search, increases in the rate of wage arrivals ($\lambda$), and decreases in the rate of interest. The hazard rate ($h(t_u)$) increases with increases in the cost of search, increases in the rate of interest, and with positive translations of the wage offer distribution. The effect of increases in the rate of job offer arrivals and increases in the variance of the wage offer distribution on the hazard rate is ambiguous. Thus, *ceteris paribus*, increases in the cost of search, the rate of interest and positive translations of the wage offer distribution decrease the mean length of unemployment and result in unemployment duration distributions that are stochastically dominated by the predisplacement distributions.

If $r=0$, straightforward manipulation of (2.2′) reproduces all of the preceding propositions except those for $r$ and translations of the wage offer distribution. It is well known that if $r=0$ unit translations of the wage offer distribution lead to unit increases in reservation wages. Thus translations of the wage offer distribution in this case produce no change in the hazard rate.

### 2.2. Relationship to previous discrete time models of discrete choice

Previous work in deterministic labor supply theory [e.g. Heckman and Willis (1977), Heckman and MaCurdy (1980), Heckman (1981)] estimates a discrete time analogue of the model described above. For a single spell of non-employment, the decision to remain non-employed is made by comparing a market wage offer $X$ to reservation wage $rV$ although a different theory is used to derive the reservation wage. Kiefer and Neumann (1979) have applied discrete time econometric models developed in the deterministic labor supply literature to the search unemployment problem.

The assumption used in previous work is that new wage offers (or in the labor supply theory, reservation wages, or both) arrive each time period. To focus on essential aspects of the comparison we assume that the reservation wage is non-stochastic, and the cdf of the independent wage offers is $F(x)$. We assume that $F(x)$ is absolutely continuous.

Previous models produce the probability that a non-employment spell terminates after $j$ periods as

$$P(j)=(F(rV))^{j-1}(1-F(rV)).$$

Accepted wages are truncated from below by $rV$ so that (2.5) is the density of these offers. Thus the joint density of duration times and accepted wages is

$$k(j,x)=[F(rV)]^{j-1}f(x), \qquad x \geqq rV.$$

This approach is somewhat arbitrary because it is time unit dependent. To use it one must assume that wage offers come in at the rate of exactly one per period. (This is a crucial assumption in the Kiefer–Neumann analysis; see appendix C.) The model developed in this paper assumes that wage offers arrive at random times. The cost of this generality is the introduction of parameters of the wage arrival function ($\lambda$).

### 2.3. Identification[7]

We assume the analyst has access to longitudinal data on $N$ independent spells of unemployment. At this point we assume that all individuals have common structural parameters. The sampling frame is a fixed known time frame of length $\bar{t}_u$. Some spells may be censored (so $t_u = \bar{t}_u$). At the end of the sampling period an unemployment spell may still be in progress. Thus for certain episodes we do not observe an accepted wage.

Our approach to identification is somewhat unusual in that it is non-parametric. We produce consistent estimators of the reservation wage and the distribution functions of accepted wages ($x$) and durations ($t_u$). We also ask under what conditions we can solve from conditional empirical distribution functions [which are pointwise consistently estimated except at points of discontinuity; see e.g. Billingsley (1968)] and the estimated reservation wage for the underlying structural parameters and distributions of interest. A crucial identification condition is a *recoverability condition*, not always satisfied, which enables the analyst to recover an untruncated distribution from a truncated distribution.

From a random sample of unemployment durations, possibly censored, it is possible to use a Kaplan–Meier estimator [see Kalbfleisch and Prentice (1980, pp. 10–16)] to consistently estimate the integrated hazard and hence the hazard [eq. (2.6)]. It is thus possible to consistently estimate $h = \lambda(1 - F(rV))$. A variety of alternative estimators could be used.[8] From duration data alone, it is not possible to separate $\lambda$ from $(1 - F(rV))$.

From data on accepted wage offers it is possible to directly estimate the reservation wage.[9] The density of accepted wage offers is given in eq. (2.5).

[7]Kiefer and Neumann (1979) have indirectly addressed this problem for the case of log normally distributed $X$ in a discrete time search model with one wage offer received each period. They address the problem of identification in these models using standard sample selection bias methods. The analysis considered in this section is more general because it permits wage offers to arrive randomly, it is not specific to a log normal wage offer distribution, and because it utilizes crucial information overlooked in their analysis. For example, their analysis apparently implies that $r > 0$ is a necessary condition for identification. See appendix C.

[8]We note that there are better behaved estimators for the hazard than the one mentioned here. We use the estimator in our identification analysis only to indicate that *in principle* it is possible to estimate $h$ non-parametrically. See the discussion in section 2.4.

[9]This obvious and essential point has not hitherto been noted in the literature.

The smallest accepted wage ($X_{\min}$) in a sample of size $N$ has density

$$q(x_{\min})=N\left[\frac{f(x_{\min})}{1-F(rV)}\right]\left[\frac{1-F(x_{\min})}{1-F(rV)}\right]^{N-1}, \qquad rV\leqq x_{\min}\leqq\infty. \tag{2.7}$$

The limiting distribution of $X_{\min}$ (obtained by letting $N\to\infty$) places point mass at $x_{\min}=rV$. $X_{\min}$ is a strongly consistent estimator of reservation wage $rV$. There are obviously many other strongly consistent estimators of $rV$ (e.g. the second smallest accepted wage, etc.). The fact that the range of $X$ depends on a parameter ($rV$) means that a standard regularity assumption is violated. We develop this point in the next section where specific estimation strategies for parametric models are discussed.

The empirical cdf of accepted wages converges pointwise (at all points of continuity) to the population cdf of accepted wages,

$$F(x\mid x\geqq rV)=\frac{F(x)-F(rV)}{1-F(rV)}, \qquad x\geqq rV.$$

In light of the discussion in the preceding paragraph, $rV$ can be consistently estimated. Provided that it is possible to recover the untruncated distribution $F(x)$ from the truncated distribution knowing the point of truncation, it is possible to estimate $F(x)$ from an estimate of $F(x\mid x\geqq rV)$ and $rV$.

We define a distribution $F(x)$ to be *recoverable* from a truncated distribution with known point of truncation ($rV$) if knowledge of $F(x|x\geqq rV)$ and $rV$ implies that $F(x)$ is uniquely determined.

Provided that $F(x)$ is recoverable, all of the parameters in eq. (2.2) are identified provided that $r$ is known. Thus if $F(x)$ and $rV$ are known, from the estimated hazard $\lambda$ is known. [See eq. (2.6).] With $\lambda$ known, $F$ and $rV$ can be inserted in (2.2) to estimate $c$.[10]

$F(x)$ is not recoverable without maintaining some functional form assumption, and it is not recoverable for all functional forms. Two absolutely continuous distributions that produce the same truncated distribution with the same known point of truncation have densities that differ only by a linear transformation. If densities $f(x)$ and $f^*(x)$ both generate the same truncated cdf, then $f^*(x)=bf(x)$ for $x\geqq rV$.

To show this note that if

$$F^*(x\mid x>rV)=\frac{F^*(x)-F^*(rV)}{1-F^*(rV)}=\frac{F(x)-F(rV)}{1-F(rV)}$$

$$=F(x\mid x>rV) \quad \text{for all} \quad x\geqq rV,$$

[10]If $c$ is known *a priori*, it is possible (with known $F$ and $rV$) to estimate $r$.

then from absolute continuity,

$$\frac{f^*(x)}{1-F^*(rV)}=\frac{f(x)}{1-F(rV)} \quad \text{for all} \quad x \geqq rV,$$

so the densities must be proportional at all points of continuity of the distributions, i.e., $f^*=bf$ for $x \geqq rV$ where $b=(1-F^*(rV))/(1-F(rV))$. Thus the distribution functions are related by $F^*(x)=a+bF(x)$ where $a=1-b$.[11] Without imposing a parametric structure on $f(x)$, it is not possible to determine the shape of $f(x)$ for values of $x$ below $rV$. It is possible to determine the shape of $f(x)$ for $x \geqq rV$ (up to a scalar multiplication) but not the mass below $rV$.

Without further information it is not possible to solve from the truncated distribution and the known point of truncation for a unique untruncated distribution. There are obviously infinitely many untruncated distributions which can produce the same truncated distribution at a given point of truncation. For this reason, a completely non-parametric approach to identification is impossible. Unless a recoverability condition is imposed, the structural duration model is not identified. (Precisely the same statement holds for all econometric models for the analysis of truncated data.)

It is sometimes possible to achieve recoverability by restricting the true distribution $F(x)$ to be a member of a certain class of parametric families. Thus if we require $F(x)$ and $F^*(x)$ to be normal distributions, $F(x)=F^*(x)$ ($a=0$ and $b=1$). (This restriction does not deny that some non-normal $F^*$ may also fit the truncated data.) Identification through the use of functional form is not always possible.

We present an example in which the recoverability condition is not met even if we impose the assumption that wage offers come from a distribution with known functional form. Let wages be Pareto distributed,

$$\mathrm{d}F(x)=\phi x^{\beta}\,\mathrm{d}x, \qquad c_2 \leqq x \leqq \infty, \qquad \beta \leqq -2,$$

where

$$\phi=\frac{-(\beta+1)}{(c_2)^{\beta+1}},$$

so

$$F(x \mid x>rV)=\frac{-(\beta+1)x^{\beta}}{(rV)^{\beta+1}}, \qquad x \geqq rV \geqq c_2.$$

[11]In a discrete time search model with one wage offer assumed to arrive each period, we can estimate $F(rV)$. In this case we can uniquely determine $F$ (for $F$ absolutely continuous) using the additional information that $F(rV)=k$, where $k$ and $rV$ are known constants. Since we know the denominator of the conditional density, we can uniquely determine $F$ and hence $f$. This is an essential feature of the Kiefer–Neumann analysis (see appendix C).

The parameter $\phi$ (or $c_2$) cannot be identified. Thus there are many Pareto distributions that fit the truncated wage data equally well.

To explore the consequences of the lack of recoverability in the Pareto case on the remaining parameters of the model, write (2.2) for the Pareto case as

$$c+rV=\left(\frac{\lambda\phi}{r}\right)\left(\frac{(rV)^{\beta+2}}{r(\beta+1)(\beta+2)}\right).$$

The hazard is

$$h=\lambda(1-F(rV))=\frac{-\lambda\phi}{(\beta+1)}(rV)^{\beta+1}:$$

Thus $rV$ is identified and so are $\beta$, $\lambda\phi$, and $c$. But obviously $\phi$ and $\lambda$ cannot be separately identified (we can use the inequality $rV \geqq c_2$ to produce an upper bound on $c_2$, an upper bound on $\phi$ and hence a lower bound on $\lambda$).

The fact that identification hinges on the functional form of the wage offer distribution is an unsettling result but accords with economic intuition. Without some restrictions imposed, it should not be possible to distinguish a model with a lot of the mass in the wage offer distribution near zero from a model with a low arrival rate for job offers, and with little mass in the wage offer distribution near zero. This point is discussed further in section 3.

If wages are normally distributed, the recoverability condition is satisfied. This fact is implicitly utilized in Amemiya (1973) and Pearson (1903). The condition is also satisfied for many commonly utilized distributions such as the exponential and log normal distributions.[12]

It is possible to test the concordance of any assumed functional form for $F(x)$ with the data. As in Heckman (1976a) it is possible to use a Kolmogorov–Smirnov (KS) test or chi-square test to test concordance between the empirical counterpart of $F(x \mid x \geqq rV)$ and a particular parametric cdf.[13] Thus the selection of a particular functional form need not be an entirely arbitrary process. But the fact that a given functional form is statistically concordant with the data does not imply that it is the only

[12]The identification analysis of Kiefer and Neumann (1979) based on a log normal distribution for wage offers apparently requires $r>0$. Their model satisfies the recoverability condition. The assumption $r>0$ is not needed to achieve model identification. Provided that $R$ [in eq. (2.2′)] is non-negative, it is possible to repeat the analysis in the text for $r=0$. Replacing $rV$ in the text equations with $R$, it is possible to use accepted wage data to consistently estimate $R$ using the minimum value of $X$ in the sample. From data on accepted wages it is possible to estimate the accepted wage distribution. Given the point of truncation, and satisfaction of the recoverability condition, it is possible to estimate $F(X)$, hence $F(R)$, and $\lambda$. Using eq. (2.2′) it is possible to consistently estimate $c$. For more discussion of the Kiefer–Neumann analysis, see appendix C.

[13]The KS test must be modified to correct for parameter estimation. See Durbin (1973, pp. 53–54).

functional form consistent with the data. *There is no statistical test for recoverability that can be based on the data generated by the model.*

### 2.4. Estimation

The preceding section generates identification criteria by way of producing consistent *non-parametric* estimators. No claim was made there about the efficiency or asymptotic distribution of those estimators. These topics are more readily addressed in a parametric framework. In this section we examine the maximum likelihood estimator of the one-spell search model in a parametric setting, and derive the appropriate asymptotic distribution.

Two new points arise. First, as is evident from eq. (2.7), estimation is being considered in a non-regular setting because the lower limit of the accepted wage offer distribution depends on a parameter of the model ($rV$). Thus standard asymptotic distribution arguments will not apply [Cox and Hinkley (1974, pp. 112)]. Second, because we utilize a parametric framework, it becomes essential to explicitly account for censoring of spells.[14] If the sampling plan of the longitudinal data is such that spells of length greater than $\bar{t}_u$ are not observed, the density of $t_u$ and $x$ *conditional* on $t_u \leqq \bar{t}_u$ is

$$m(t_u, x \mid t_u \leqq \bar{t}_u) = \left[\frac{\lambda(1-F(rV))\exp-\lambda(1-F(rV))t_u}{1-\exp-\lambda(1-F(rV)\bar{t}_u)}\right]\left[\frac{f(x)}{1-F(rV)}\right],$$

where the denominator of the first term in brackets is the probability that $t_u \leqq \bar{t}_u$.

Let $d$ denote an indicator function with

$$\begin{aligned} d &= 1 \quad \text{if} \quad t_u \leqq \bar{t}_u, \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

Then the joint density of $d$, $t_u$ and $x$ is

$$m(t_u, x, d) = [f(x)\lambda\exp-\lambda(1-F(rV))t_u]^d[\exp-\lambda(1-F(rV))\bar{t}_u]^{1-d}.$$ [15]

[14]The non-parametric Kaplan–Meier estimator can be applied to censored or uncensored data.

[15]In forming this density we use the fact that the probability $d=0$ is the same as the probability that $t_u$ exceeds $\bar{t}_u$ $(=\exp-(1-F(rV))\bar{t}_u)$. The probability that $t_u$ is less than $\bar{t}_u$ is $(1-\exp-(1-F(rV))\bar{t}_u)$.

The full economic model for $r>0$ implies the following restrictions:

$$x \geqq rV \geqq 0, \qquad rV+c=(\lambda/r)\int_{rV}^{\infty}(x-rV)\mathrm{d}F(x).^{16} \tag{2.8}$$

In the remainder of this section we write the distribution function of wage offers as $F(x\,|\,\theta)$ to explicitly demonstrate the dependence of $F(x\,|\,\theta)$ on a finite-dimensional vector of bounded parameters ($\theta$). We assume that $F(x\,|\,\theta)$ is recoverable (in the sense of section 2.3) and hence that $\theta$ is identified. We assume that $r$ is known.

The sample log likelihood function (with individual subscripts suppressed) for a longitudinal sample of $N$ independent individuals is

$$\begin{aligned}\pounds &= \sum \ln m(t_u, x, d)\\ &= (\textstyle\sum d)\ln\lambda + \sum d \ln f(x\,|\,\theta) - \lambda(1-F(rV\,|\,\theta))\sum t_u,\end{aligned} \tag{2.9}$$

where $t_u=\bar{t}_u$ for the censored observations.

Structural estimation of the model requires maximizing (2.9) incorporating the economic restrictions in (2.8). There are many possible parameterizations of the likelihood function. The most analytically convenient one treats $rV$ as an explicit parameter along with $\theta$ and $\lambda$ using (2.8) to solve for $c$. Note that for the model to make statistical sense, it is required that $\lambda>0$. This constraint is automatically imposed in the estimation as we demonstrate below.

The maximum likelihood estimator has a simple structure. Denoting the maximum likelihood estimator by $\hat{}$, $\widehat{rV}=x_{\min}$, where $x_{\min}$ is the minimum accepted wage observed in the sample of job takers (for whom $d=1$). That this is the mle estimator is verified by noting that $\pounds$ is monotonically increasing in $rV$ so that the first constraint in (2.8) is always effective at the upper limit.

This estimator is strongly consistent. Producing the asymptotic distribution theory for the estimator is a more delicate matter. As is well known [see Galambos (1978, p. 71)] for a given $F(x\,|\,\theta)$ there may exist *no* norming factors such that $a(n_1)X_{\min}+b(n_1)$ converges to a distribution where $a(n_1)$ and $b(n_1)$ are functions of $n_1$ $(=\sum d)$. If such factors exist, the asymptotic distribution is one of two possible functional forms [see Galambos (1978, pp. 56–57)].[17] The existence and functional form of the limiting distribution of $X_{\min}$ is critically dependent on the functional form of

[16]The second constraint for the case $r=0$ is eq. (2.2′). In that notation, the first constraint is $x \geqq R \geqq 0$.

[17]The forms for the cdfs are (see Galambos)

$$L_2=1-\exp(-x^{\gamma}), \quad x>0, \quad \text{or} \quad L_3=1-\exp(-e^{x}).$$

$F(x \mid \theta)$ although the strong consistency of the estimator does not depend on this functional form.

The following theorem establishes a general condition under which $\sqrt{N}(X_{\min} - rV)$ has a degenerate distribution:

*Theorem 1. Define* $z = \sqrt{N}(X_{\min} - rV)$. *If and only if*

$$\lim_{N \to \infty} \left[ \frac{1 - F(rV + z/\sqrt{N} \mid \theta)}{1 - F(rV \mid \theta)} \right]^N \to 0 \quad \text{for} \quad z > 0, \tag{2.10}$$

$z = \sqrt{N}(X_{\min} - rV)$ *has a degenerate distribution at* $z = 0$.

*Proof.* From (2.7) and integration,

$$F(x_{\min}) = 1 - \left[ \frac{1 - F(x_{\min} \mid \theta)}{1 - F(rV \mid \theta)} \right]^N.$$

Thus the cdf of $z$, say $K(z)$, is

$$K(z) = 1 - \left[ \frac{1 - F(rV + z/\sqrt{N} \mid \theta)}{1 - F(rV \mid \theta)} \right]^N.$$

Then, we have $\lim K(z) = 1$, $z > 0$, if and only if

$$\lim_{N \to \infty} \left[ \frac{1 - F(rV + z/\sqrt{N} \mid \theta)}{1 - F(rV \mid \theta)} \right]^N \to 0, \qquad z > 0. \qquad \text{Q.E.D.}$$

For example, if $F(x \mid \theta) = 1 - e^{-\theta x}$, $\theta > 0$, the limit in (2.10) is achieved so that $\sqrt{N}(X_{\min} - rV)$ has a degenerate asymptotic distribution. In fact $N(X_{\min} - rV) \sim 1 - e^{-\theta(X_{\min} - rV)}$ and convergence in distribution is at rate $N$.

Violation of condition (2.10) is unusual as demonstrated in the next theorem:

*Theorem 2.* $F'(rV \mid \theta) = 0 = f(rV \mid \theta)$ *is a necessary condition for* (2.10) *to be violated for* $F$ *at least twice continuously differentiable with bounded second derivative.*

---

For the relevant discussion, see Galambos where the $a(n_1)$ and $b(n_1)$ functions are produced for specific functional forms.

*Proof*

$$N \ln\left[\frac{1-F(rV+z/\sqrt{N}\mid\theta)}{1-F(rV\mid\theta)}\right] \doteq N\left[1-\frac{f(rV\mid\theta)}{1-F(rV\mid\theta)}\frac{z}{\sqrt{N}} + \tfrac{1}{2}\frac{f'(rV\mid\theta)}{1-F(rV\mid\theta)}\frac{z^2}{N}+\mathrm{o}\left(\frac{1}{N}\right)\right].$$

If $f(rV\mid\theta)=0$,

$$\lim_{N\to\infty} N\left[\frac{1-F(rV+z/\sqrt{N}\mid\theta)}{1-F(rV\mid\theta)}\right]=\tfrac{1}{2}\frac{f'(rV\mid\theta)}{1-F(rV\mid\theta)}z^2,$$

so condition (2.10) is violated. Q.E.D.

The condition $f(rV\mid\theta)=0$ is not likely to be satisfied for most standard distributions.

The thrust of the two theorems is that in cases that are likely to arise in applied work, $\sqrt{N}(X_{\min}-rV)$ has a degenerate distribution. This, in turn, implies that if $\widehat{rV}$ is inserted in likelihood (2.9) we may treat $\widehat{rV}$ as if it is $rV$ in determining the $(\sqrt{N})$ asymptotic distributions of the remaining parameters.[18]

Conditional on $\widehat{rV}$, the likelihood equation for $\lambda$ is

$$\mathscr{L}_\lambda=0=(\textstyle\sum d)/\lambda-(1-F(rV\mid\theta))\sum t_u,$$

so

$$\hat{\lambda}=\frac{(\sum d)}{(1-F(\widehat{rV}\mid\hat{\theta}))(\sum t_u)}.$$

The equation for $\theta$ is

$$\mathscr{L}_\theta=0=\sum d\frac{f_\theta(x\mid\hat{\theta})}{f(x\mid\hat{\theta})}+\lambda F_\theta(\widehat{rV}\mid\hat{\theta})\sum t_u,$$

which can be simplified to

$$0=\sum d\frac{f_\theta(x\mid\hat{\theta})}{f(x\mid\hat{\theta})}+\frac{(\sum d)F_\theta(\widehat{rV}\mid\hat{\theta})}{1-F(\widehat{rV}\mid\hat{\theta})}.$$

[18]In a valuable paper on frontier production theory, Greene (1980) invokes conditions that ensure that (2.10) is violated and that standard asymptotic theory can be applied. He demonstrates that standard regularity conditions need not be satisfied in order to produce standard asymptotic distribution results. His results, while very useful in his context, are not of direct use here because to satisfy his conditions we would have to assume $f(x\mid\theta)=0$ over the economically interesting interval for $x$.

$\hat{\lambda}$ is always non-negative. Conditional on $\widehat{rV}$, the estimators are consistent and asymptotically normally distributed provided that the information matrix is of full rank. Proofs of the assertions are trivial and hence are deleted.

We state without proof:

*Theorem 3. If condition* (2.10) *is met,*

$$\sqrt{N}\begin{pmatrix}\hat{\lambda}-\lambda\\ \hat{\theta}-\theta\end{pmatrix}\sim N(0, I^{-1}(\widehat{rV})),$$

*where*

$$I(\widehat{rV})=-E\frac{\partial^2\mathscr{L}(\widehat{rV})}{\partial\psi\,\partial\psi'}=-E\frac{\partial^2\mathscr{L}(rV)}{\partial\psi\,\partial\psi'}\quad\text{with}\quad\psi=\begin{pmatrix}\lambda\\ \theta\end{pmatrix}.$$

With estimates of $rV$, $\theta$ and $\lambda$, we may solve for $c$ from the second equation of (2.8). We may use the delta method to obtain the sampling distribution for $c$ [Bishop, Feinberg and Holland (1975, p. 486)]. Assuming that $\sqrt{N}(X_{\min}-rV)$ is degenerate, $\widehat{rV}$ contributes only to the mean of the sampling distribution of $c$. Thus

$$\hat{c}=(\hat{\lambda}/r)\int_{\widehat{rV}}^{\infty}(x-\widehat{rV})\mathrm{d}F(x\mid\hat{\theta})-\widehat{rV},$$

and

$$\sqrt{N}(\hat{c}-c)\sim N(0,\sigma_c^2),$$

where $\sigma_c^2$ is obtained by the usual Taylor series development and the covariance of $\sqrt{N}(\hat{c}-c)$ with the other parameters is obtained in the usual way.

A simple example will help to clarify ideas. Suppose the sampling frame is such that there are no censored observations ($\bar{t}_u=\infty$). Suppose further that wages are exponentially distributed so

$$\mathrm{d}F(x\mid\theta)=\theta\mathrm{e}^{-\theta x}\mathrm{d}x,\qquad\theta>0. \tag{2.11}$$

With a zero rate of interest and Poisson arrival parameter $\lambda$, simple substitution in (2.2′) confirms that the reservation wage $R$ satisfies

$$R=(-1/\theta)\ln(C\theta/\lambda). \tag{2.12}$$

For the search problem to be economically meaningful we require $R>0$ so that $C\theta/\lambda<1$ since $\theta$, $\lambda$ and $C$ are all positive.[19]

[19]Note that $X$ is interpreted as a lump sum undiscounted value of the job.

In a sample of size $N$, the density of the minimum accepted wage offer $X_{\min}$ is

$$q(x_{\min}) = N\theta e^{-N(x_{\min}-R)\theta}, \qquad x_{\min} \geqq R, \tag{2.13}$$

with mean

$$E(X_{\min}) = (N\theta R + 1)/N\theta = R + 1/N\theta.$$

Thus $X_{\min}$ is an upward biased but consistent estimator for $R$.

The log likelihood function (2.9) with $\sum d = N$ is

$$\mathscr{L} = N \ln \theta - \theta \sum x + N \ln \lambda - \lambda e^{-\theta R} \sum t.$$

The maximum likelihood estimators are achieved from the following procedure. First estimate $R$ by $X_{\min}(\hat{R} = x_{\min})$. Then maximize the concentrated likelihood function $\mathscr{L}(\hat{R}) = N \ln \theta\lambda - \theta \sum x - \lambda e^{-\theta x_{\min}} \sum t$, where $x_{\min}$ is substituted for $R$ in the previous expression. The roots of the conditional maximum likelihood equations obtained from $\mathscr{L}(x_{\min})$ are

$$\frac{1}{\hat{\theta}} - \frac{\sum x}{N} + x_{\min} = 0, \qquad \frac{1}{\hat{\lambda}} - \frac{\sum t}{N} e^{-\hat{\theta} x_{\min}} = 0.$$

$\hat{\theta}$ and $\hat{\lambda}$ are consistent for $\theta$ and $\lambda$, respectively, and both are biased. The exact bias of $1/\hat{\theta}$ for $1/\theta$ can be computed since

$$E(1/\hat{\theta}) = 1/\theta + 1/N\theta.^{20}$$

The bias is of order $N^{-1}$. It can also be shown that the bias of $1/\hat{\lambda}$ for $1/\lambda$ is of order $N^{-1}$.

A consistent estimator of $c$ is obtained by substituting $x_{\min}$ for $R$ and $\hat{c}$ and $\hat{\theta}$ for $c$ and $\theta$, respectively, in (2.12). Thus

$$\hat{c} = (\hat{\lambda}/\hat{\theta}) e^{-\hat{\theta} x_{\min}}.$$

A norming factor of $N$ is required to produce a non-degenerate asymptotic distribution for $X_{\min}$. With this factor

$$N(X_{\min} - R) \sim \theta e^{-\theta\psi}, \qquad \psi \geqq 0.$$

[20]An unbiased estimator of $1/\theta$ is

$$1/\hat{\hat{\theta}} = N(\textstyle\sum x/N - x_{\min})/(N-1),$$

so that

$$(2N-1)((1/\hat{\hat{\theta}})/(1/\theta)) \sim \chi^2 \quad \text{with } 2(N-1) \text{ degres of freedom.}$$

Because $\sqrt{N}(X_{\min}-R)$ is a degenerate random variable,

$$\sqrt{N}\begin{pmatrix}\hat{\theta}-\theta\\ \hat{\lambda}-\lambda\end{pmatrix}\sim N(0,\mathscr{I}^{-1}),$$

where

$$\mathscr{I}^{-1}=\begin{pmatrix}\theta^2 & R\theta^2\lambda\\ R\theta^2\lambda & \lambda^2+\lambda^2\theta^2R^2\end{pmatrix}=\begin{pmatrix}1/\theta^2+R^2 & -R/\lambda\\ -R/\lambda & 1/\lambda^2\end{pmatrix}^{-1}.$$

Thus $\hat{\theta}$ and $\hat{\lambda}$ are asymptotically positively correlated. The information matrix may be consistently estimated replacing parameters with maximum likelihood estimates. The asymptotic distribution for $\sqrt{N}(\hat{c}-c)$ can be derived from the delta method using the joint asymptotic distribution for $\hat{\theta}$ and $\hat{\lambda}$ and the fact that $\sqrt{N}(X_{\min}-R)$ is degenerate.[21] For a comparison between the approach to structural estimation based on maximum likelihood which we use and other approaches, see appendix C.

### 2.5. *Accounting for population dispersion in the parameters of the model*

We have thus far assumed that all agents in our samples have identical values of $\lambda$, $c$, $r$ and $\theta$. There are plausible reasons why values of those parameters might vary among members of the population. In many economic models of equilibrium price distributions, dispersion in $c$ (or some other parameter) is required to produce equilibrium. Failure to account for such dispersion can produce badly biased estimates of the parameters of structural duration models [Silcock (1954), Blumen, Kogan and McCarthy (1955)]. This is the problem of 'heterogeneity'.

In principle, each parameter of the one state search model can be written as a function of observed and unobserved variables. The parameters associated with the observed variables in these functions can be estimated. Unobserved variables can be integrated out using distributions whose parameters can be estimated. This approach is used in many studies of labor force dynamics [Heckman and Willis (1975, 1977), Nickell (1979), Heckman (1981)].

A practical problem with this solution is the selection of particular functional forms for use in estimation. Except for occasional qualitative

[21]An estimator such as $\widehat{rV}$ that is based on an extreme observation in a sample is obviously quite sensitive to measurement error in the $x$ data. Since any fixed number of extreme observations (say the first up to the $k$th) are consistent estimators for $rV$, a consistent estimator for $rV$ that is less sensitive (in an imprecisely defined way) to measurement error is an average of the lowest $k$ observations. Greene's estimator (1980) in a related problem is also worthy of exploration: compute $\widehat{rV}$ from adjacent values of the lowest observations of $x$ that are 'close'. See the discussion of measurement error in section 2.5.2 below.

guidelines, economic theory offers little guidance. It is silent on the topic of the correct specification of functional forms for the distributions of unobservables.

This would not be a serious problem if estimates of key parameters of the model were 'relatively' insensitive to assumptions made about specific functional forms. In fact, the available evidence [Heckman and Singer (1981a,b)] suggests the opposite. Estimates of key parameters are very sensitive to the choice of particular functional forms.

The situation is not entirely hopeless. Heckman and Singer (1981a,b) present a computationally feasible non-parametric estimator for the distribution functions of unobservables, so that at least one aspect of arbitrariness in model specification can be eliminated with their methods. Nonetheless, a considerable amount of arbitrariness remains, and this arbitrariness vitally affects the inference from the structural models.

These problems are more troublesome than recent analyses indicate is the case. In this subsection we make a first attempt at clarifying the issues that must be addressed in order to account for population dispersion in parameters. We also offer some solutions. What follows is only the first and (we hope) not the last word on these problems.

It is analytically useful to distinguish two cases: (1) the case in which population heterogeneity arises solely from dispersion in observed variables, and (2) the case in which the heterogeneity arises from unobserved variables. We treat these two cases, starting with the first. In both cases we assume a parametric distribution of wage offers $F(x|\theta)$ that satisfies the recoverability condition. A fully general analysis of both cases is not attempted, but we indicate the gaps that remain to be filled (of which we are aware).

### 2.5.1. Observed heterogeneity

We first write $\lambda=\lambda(Z_\lambda\beta_\lambda)$, $\theta=\theta(Z_\theta\beta_\theta)$ and $c=c(Z_c\beta_c)$ to indicate the dependence of $\lambda$, $\theta$ and $c$ on explanatory variables. From eq. (2.8), $rV$ is a function of $Z$ and all the $\beta$ vectors. The density for $t_u$, $x$ and $d$ conditional on $Z=(Z_\lambda Z_\theta Z_c)$ is

$$m(t_u, x, d \mid Z)=[f(x \mid \theta)\lambda]^d \exp-[\lambda(1-F(rV \mid \theta))t_u], \tag{2.14}$$

where

$$x \geqq rV \geqq 0 \quad \text{for each value of } Z, \tag{2.15}$$

and

$$c=(\lambda/r)\int_{rV}^{\infty}(X-rV)\,\mathrm{d}F(x|\theta)-rV, \tag{2.16}$$

where it is understood that $\lambda$, $\theta$ and $c$ are functions of exogenous variables,

and $t_u$ denotes the duration of all measured spells including the censored spells; $r$ is assumed to be known.

From eq. (2.16) we may solve for $rV$ as a function of $c$, $\lambda$, $r$ and $\theta$,

$$rV = rV(Z_\lambda\beta_\lambda, Z_\theta\beta_\theta, Z_c\beta_c, r).$$

The reservation wage is monotonic in the cost of search $c$ $(=c(Z_c\beta_c))$.

Using the inequalities in (2.15) that must be satisfied for each observation, the log likelihood is (suppressing subscripts)

$$\begin{aligned}\mathscr{L} = &\sum d\ln\lambda + \sum d\ln f(x|\theta) - \sum \lambda(1-F(rV|\theta))t_u \\ &+ \sum \mu(x - rV(\lambda,\theta,c,r)) + \sum \phi(rV(\lambda,\theta,c,r)),\end{aligned} \tag{2.17}$$

where $\theta = \theta(Z_\theta\beta_\theta)$, $\lambda = \lambda(Z_\lambda\beta_\lambda)$ and $c = c(Z_c\beta_c)$ and where the multipliers $\mu$ and $\phi$ are defined for each observation.[22] The final two sets of constraints arise as a consequence of inequality (2.15) that must be satisfied for each observation. These constraints are a vital source of identification in this model and cannot be ignored in securing structural estimates.

Provided that the multipliers are well defined (see the footnote below) the first-order conditions for $\beta_\theta$, $\beta_\lambda$ and $\beta_c$ are

$$\begin{aligned}\mathscr{L}_{\beta_\theta} = 0 = \sum \Bigg\{ &\left( d\frac{\partial \ln f(x|\theta)}{\partial\theta} \right) + \lambda\left( \frac{\partial F(rV|\theta)}{\partial rV}\frac{\partial rV}{\partial\theta} + \frac{\partial F(rV|\theta)}{\partial\theta} \right) t_u \\ &+ (\phi-\mu)\frac{\partial rV}{\partial\theta} \Bigg\} \frac{\partial\theta}{\partial\beta_\theta},\end{aligned} \tag{2.18a}$$

$$\begin{aligned}\mathscr{L}_{\beta_\lambda} = 0 = \sum \Bigg\{ &\frac{d}{\lambda} - (1-F(rV|\theta))t_u + \lambda\frac{\partial F(rV|\theta)}{\partial rV}\frac{\partial rV}{\partial\lambda} \\ &+ (\phi-\mu)\frac{\partial rV}{\partial\lambda} \Bigg\} \frac{\partial\lambda}{\partial\beta_\lambda},\end{aligned} \tag{2.18b}$$

$$\mathscr{L}_{\beta_c} = 0 = \sum \left\{ \lambda\frac{\partial F(rV|\theta)}{\partial rV} t_u + (\phi-\mu) \right\} \frac{\partial rV}{\partial c}\frac{\partial c}{\partial\beta_c}. \tag{2.18c}$$

A critical feature of this optimization problem is revealed in condition

[22]The existence of these multipliers and the quasi-saddle-point characterization of the optimum used in the text follows if $rV$ and $x - rV$ as functions of $\beta$ satisfy the Arrow–Hurwicz–Uzawa conditions [see Takayama (1974, pp. 93–94)]. Their rank condition is likely to be satisfied in 'large enough' samples. We have not investigated the restrictions these conditions place on our choice of $c$, $\lambda$ and $\theta$ functions although this would be a potentially fruitful topic.

(2.18c). Note that

$$(\partial F(rV \mid \theta)/\partial rV)(\partial rV/\partial c)$$

is always non-positive and for most parametric problems is negative.[23] If constraint (2.15) is ignored in solving for $\beta_c$ (so we assume $\phi=\mu=0$), the likelihood function possesses no interior maximum with respect to $\beta_c$ unless we impose an arbitrary nonlinearity that guarantees that $c$ is not monotonic in $\beta_c$ and that bounds $c$ from below. For example, if $c$ is a constant, a maximum likelihood estimator of $c$ that ignores (2.15) sets $c$ at its lowest possible value.

A general analysis of identification and distribution theory is a task for the future. We make some progress by assuming that the number of distinct $Z$ vectors is finite and equal to $K$ and that the population covariance matrix of $Z$ is of full rank. This is consistent with 'fixed in repeated samples' sampling schemes or with the $Z$ vectors selected by a finite $K$ state irreducible Markov chain so that all states are recurrent. This assumption is critical in the ensuing analysis. If it is not invoked, we are unable to produce consistent estimators for the model although we speculate below on how this might be done. We denote the population proportion of the exogenous variables of type $k$ by $P_k$, $0<P_k<1$.

By analogy with the analysis in section 2.4 we may estimate $rV_k$ for each $k=1,\ldots,K$. For $N\to\infty$ so $P_kN\to\infty$, these estimators are strongly consistent. Insert these estimators into

$$\tilde{\tilde{\mathscr{L}}}=\sum d(\ln\lambda+\ln f(x\mid\theta))-\sum\lambda(1-F(rV\mid\theta))t_u,$$

in place of $\widehat{rV}$ (associating $rV_k$ with observations of type $k$) and maximize with respect to $\beta_\theta$ and $\beta_\lambda$. Provided the matrix of second partials of $\tilde{\tilde{\mathscr{L}}}$ (with respect to $\beta_\lambda$ and $\beta_\theta$) is non-singular, these equations may be solved uniquely for $\beta_\lambda$ and $\beta_\theta$ and provided condition (2.10) is met for each $k$,

$$\sqrt{N}\begin{pmatrix}\hat{\beta}_\lambda-\beta_\lambda\\ \hat{\beta}_\theta-\beta_\theta\end{pmatrix}\sim N(0,\mathscr{I}^{-1}),$$

where $\mathscr{I}^{-1}$ is minus the inverse matrix of the expectation of the second partials of the log likelihood. Estimates of $\beta_c$ are obtained by solving $c_k$ from eq. (2.16) and estimating a nonlinear regression of the $\hat{c}_k$ on $c(Z_c\beta_c)$. If the model is of full rank (the $Z_c$ covariance matrix is positive definite), $\beta_c$ is identified and its asymptotic distribution can be produced by generating the asymptotic distribution of $c_k$ by the delta method, as before.

[23] $\partial F(rV|\theta)/\partial rV$ is non-negative and $\partial rV/\partial c$ is negative. The last assertion can be proved by differentiating (2.16) with respect to $c$.

The key requirement is that the matrix of second partials of $\tilde{\tilde{\mathscr{L}}}$ with respect to $\beta_\lambda$ and $\beta_\theta$ be of full rank. Verifying this condition in practice requires a case by case analysis for each choice of $f(x \mid \theta)$.

For the sake of specificity assume that $f(x \mid \theta) = \theta \mathrm{e}^{-\theta x}$. We adopt the specification of the example presented in section 2.4 with $r=0$. As is done there, we ignore censored observations. We further assume that

$$c = \exp(Z_c \beta_c), \qquad \theta = \exp(Z_\theta \beta_\theta), \qquad \lambda = \exp(Z_\lambda \beta_\lambda). \tag{2.19}$$

This parametrization imposes the required positivity for these parameters.

The log likelihood for this model is

$$\begin{aligned} \mathscr{L} = &\sum d(Z_\lambda \beta_\lambda) + \sum d(Z_\theta \beta_\theta - \exp(Z_\theta \beta_\theta) x) \\ &- \sum \exp(Z_\lambda \beta_\lambda)(\exp(-rV_k \mathrm{e}^{Z_\theta \beta_\theta}) t_u), \end{aligned} \tag{2.20}$$

where

$$rV_k = -\exp(-Z_{\theta_k} \beta_\theta)(Z_{c_k} \beta_c + Z_{\theta_k} \beta_\theta - Z_{\lambda_k} \beta_k) \geqq 0, \tag{2.21}$$

and where $Z_{c_k}$, $Z_{\theta_k}$ and $Z_{\lambda_k}$ are values of the $Z$ vectors for population type $k$.

For each type $k$ we estimate the reservation wage by $\widehat{rV}_k = \min \{x_{ik}\}_{i=1}^{N_k}$, where $x_{ik}$ is the $i$th observation on accepted wages for type $k$. These estimators are strongly consistent for $rV_k$ and $\lim_{N \to \infty} \sqrt{P_k N}(\widehat{rV}_k - rV_k)$ is a degenerate random variable.

The roots of $\tilde{\tilde{\mathscr{L}}}$ with respect to $\beta_\theta$ and $\beta_\lambda$ are

$$0 = \partial \tilde{\tilde{\mathscr{L}}} / \partial \beta_\theta = \sum Z_\theta - \sum Z_\theta \mathrm{e}^{Z_\theta \beta_\theta} x + \sum (\mathrm{e}^{Z_\lambda \beta_\lambda} t_u) \widehat{rV}_k \mathrm{e}^{-(\exp Z_\theta \beta_\theta) \widehat{rV}_k} \mathrm{e}^{Z_\theta \beta_\theta},$$

$$0 = \partial \tilde{\tilde{\mathscr{L}}} / \partial \beta_\lambda = \sum Z_\lambda - \sum Z_\lambda (\mathrm{e}^{Z_\lambda \beta_\lambda} \mathrm{e}^{-(\exp Z_\theta \beta_\theta) \widehat{rV}_k}) t_u.$$

As $N \to \infty$, the expectation of these first partials is zero at the true parameter values.

Conditional on $\widehat{rV}_k$, minus the expectation of the matrix of second partials of the log likelihood is [letting $\beta = (\beta_\theta \beta_\lambda)$]

$$\begin{aligned} \mathscr{I} &= E\left( -\frac{\partial^2 \tilde{\tilde{\mathscr{L}}}(rV_k)}{\partial \beta \partial \beta'} \right) \\ &= \begin{pmatrix} \sum Z_\theta Z_\theta' (1 + (Z_c \beta_c + Z_\theta \beta_\theta - Z_\lambda \beta_\lambda)^2) & \\ \sum Z_\lambda Z_\theta' (Z_c \beta_c + Z_\theta \beta_\theta - Z_\lambda \beta_\lambda) & \sum Z_\lambda Z_\lambda' \end{pmatrix}. \end{aligned} \tag{2.22}$$

Note that $Z_c \beta_c + Z_\theta \beta_\theta - Z_\lambda \beta_\lambda \leqq 0$ since $rV > 0$ [see eq. (2.15)].

Even if $Z_\theta = Z_\lambda = Z_c = \bar{Z}$, this matrix is non-singular provided that $\sum \bar{Z}\bar{Z}'$ is non-singular. This follows from the matrix Cauchy–Schwartz inequality. Thus, even if the same variables appear in the $c$, $\lambda$ and $\theta$ functions, it is possible to identify $\beta_\lambda$, $\beta_\theta$ and $\beta_c$.

To estimate $\beta_c$, take the estimates $\hat{\beta}_\theta$, $\hat{\beta}_\lambda$ and $\widehat{rV_k}$ and form the left-hand side of the function

$$-\widehat{rV_k}\exp(Z_{\theta_k}\hat{\beta}_\theta) - Z_{\theta_k}\hat{\beta}_\theta + Z_{\lambda_k}\hat{\beta}_\lambda = Z_{c_k}\beta_c.$$

The distribution of the left-hand side can be derived from the delta method. Use GLS to estimate $\beta_c$. To identify $\beta_c$ uniquely, it is required that the rank of $\sum Z_c Z_c'$ exceed the number of elements in $\beta_c$. A necessary condition is that $K$ not be less than the number of elements in $\beta_c$.

Note that this procedure does not in general maximize the likelihood (2.17). Moreover it is wasteful in information. If the number of parameters in $\beta_c$ is less than $K$, in general fewer than $K$ of the estimated reservation wages are required to estimate $\beta_c$. A second-round iteration of the procedure just presented produces estimators with a covariance matrix no larger than the one produced by the method just presented.

The proposed iteration sequence operates by taking estimates of $\beta_c$, $\beta_\lambda$ and $\beta_\theta$ from the first-round procedure and forming $rV_k$ for each type $k$ using eq. (2.16). For each $k$, use the minimum of $rV_k$ so formed and $\widehat{rV_k}$. Insert these values into $\tilde{\tilde{\mathscr{L}}}$ and obtain estimates of $\beta_\lambda$ and $\beta_\theta$ as before. Using the second-round estimated reservation wages and the estimates of $\beta_\lambda$ and $\beta_\theta$ obtained on the second round, repeat the non-linear regression procedure to estimate $\beta_c$. Continue iteration until convergence. In the case in which the number of parameters in $\beta_c$ is the same as $K$, in general the iteration cycle just described will produce the same final estimates of $\beta$ as used in the initial estimates because none of the reservation wage information is redundant.

For the case of a general distribution for $Z$, the likelihood and its roots are given in (2.17) and (2.18). The derivation of the asymptotic properties of the maximum likelihood estimator in the general case is a task for the future. We offer some conjectures on this problem in appendix A where we make some progress on the problem for certain cases of interest.

### 2.5.2. *Unobserved heterogeneity*[24]

Extending the analysis to a case in which population heterogeneity also arises from unobserved variables is in principle straightforward. When the models are identified the appropriate asymptotic distribution theory is much

[24]Conversations with T. Coleman and B. Payner clarified the analysis presented in this section.

simpler than in the models of the preceding subsection. The practical problem is to decide where in the model to introduce the population heterogeneity. In principle we should allow for such heterogeneity in the $c$, $\lambda$ and $\theta$ functions so $c=c(Z_c\beta_c,\varepsilon_c)$, $\lambda=\lambda(Z_\lambda\beta_\lambda,\varepsilon_\lambda)$, $\theta=\theta(Z_\theta\beta_\theta,\varepsilon_\theta)$ where $\varepsilon=(\varepsilon_c,\varepsilon_\lambda,\varepsilon_\theta)$ is a vector of random variables with cdf $G(\varepsilon\,|\,\eta)$ where $\eta$ is a vector of parameters. For simplicity we assume that $c$, $\lambda$ and $\theta$ are monotone increasing functions of $\varepsilon_c$, $\varepsilon_\lambda$, and $\varepsilon_\theta$, respectively.

The reservation wage function is defined as

$$rV=rV(c,\theta,\lambda,r), \tag{2.23}$$

where $c$, $\lambda$, $\theta$ are functions of $\varepsilon_c$, $\varepsilon_\lambda$ and $\varepsilon_\theta$, respectively. In the ensuing analysis it is convenient to work with a normalized cdf,

$$G^*(\varepsilon\,|\,\eta)=G(\varepsilon\,|\,\eta)\Big/\int_{\{\varepsilon|rV\geqq 0\}}\mathrm{d}G(\varepsilon\,|\,\eta),\qquad \{\varepsilon\,|\,rV>0\},$$

where $\{\varepsilon\,|\,rV\geqq 0\}$ is the set of values of $\varepsilon$ such that the agent participates in the labor force. The probability that a spell lasts more than $t_u$ periods conditional on $c$, $\lambda$, $\theta$ is

$$P(T_u>t_u\,|\,c,\lambda,\theta)=\exp-\lambda[1-F(rV\,|\,\theta)]t_u.$$

Integrating out $\varepsilon=(\varepsilon_c,\ \varepsilon_\lambda$ and $\varepsilon_\theta)$, using the information that $rV\geqq 0$, yields

$$P(T_u>t_u)=\int_{\{\varepsilon|rV\geqq 0\}}[\exp-\lambda[1-F(rV\,|\,\theta)]t_u]\,\mathrm{d}G^*(\varepsilon|\eta). \tag{2.24}$$

To reduce notational complexity we suppress the conditioning set of exogenous variables. The density of accepted wages conditional on $c$, $\lambda$, $\theta$ is

$$f(x\,|\,\theta)/(1-F(rV\,|\,\theta)),\qquad x\geqq rV.$$

Integrating out $\varepsilon$, the density of accepted wages is

$$\int_{\{\varepsilon|x\geqq rV\geqq 0\}}f(x\,|\,\theta)/(1-F(rV\,|\,\theta))\mathrm{d}G^*(\varepsilon\,|\,\eta). \tag{2.25}$$

The joint density of $x$, $t_u$ given $d=1$ and $c$, $\lambda$, $\theta$ is, for a sampling frame of length $\bar{t}_u$,

$$m(t_u,x\,|\,d=1)=\frac{\lambda f(x\,|\,\theta)\exp-\lambda[1-F(rV\,|\,\theta)]t_u}{1-P(T_u<\bar{t}_u\,|\,c,\lambda,\theta)}.$$

Integrating out $\varepsilon$, the joint density of $x$ and $t_u$, given $d=1$, is

$$m^*(t_u, x \mid d_u = 1) = \frac{\int_{\{\varepsilon \mid x \geqq rV \geqq 0\}} f(x \mid \theta)[\exp - [1 - F(rV \mid \theta)]t_u] \, \mathrm{d}G^*(\varepsilon \mid \eta)}{1 - P(T_u < \bar{t}_u)}. \tag{2.26}$$

Using (2.24) and (2.26), the joint density of $t_u$, $x$ and $d$ is

$$m^*(t_u, x, d) = \left\{ \int_{\{\varepsilon \mid x \geqq rV \geqq 0\}} \lambda f(x \mid \theta)[\exp - \lambda[1 - F(rV \mid \theta)]t_u] \, \mathrm{d}G^*(\varepsilon) \right\}^d$$
$$\times \left\{ \int_{\{\varepsilon \mid rV \geqq 0\}} [\exp - \lambda(1 - F(rV))\bar{t}_u] \, \mathrm{d}G^*(\varepsilon \mid \eta) \right\}^{1-d}. \tag{2.27}$$

The likelihood function for duration of unemployment and accepted wages is defined as the product of densities (2.27).[25] Provided the model is identified, maximum likelihood estimators of $\beta$ and $\eta$ are consistent and asymptotically normally distributed provided standard conditions on the regressors ($Z$) are assumed [see e.g. Jennrich (1969)]. Because $rV$ is a random variable, the appropriate asymptotic distribution theory is much simpler than the distribution theory required to analyze the models presented in previous sections of this paper. This is so because in the present model the *range* of the random variable $X$ does not depend on a parameter of the model.

The crucial identifying role played in the preceding section by the restriction that $X$ must exceed the deterministic reservation wage [eq. (2.21)] is played in the current model by a restriction on the range of unobservables,

$$\{\varepsilon \mid 0 \leqq rV(c, \lambda, \theta, r) \leqq x\}, \tag{2.28}$$

that must be satisfied for all observations for which a completed unemployment spell and accepted wage are observed.

To demonstrate how (2.28) is used to achieve identification, assume $\varepsilon_\lambda = \varepsilon_\theta = 0$. Further assume $c = c(Z_c\beta_c + \varepsilon_c)$. Denote by $G_c^*(\varepsilon_c \mid \eta_c)$ the cdf of $\varepsilon_c$ which depends on parameter vector $\eta_c$. To simplify the argument assume the sampling frame is long enough that there are no censored observations ($d=1$).

[25]Note that we could in principle exploit the information that individuals are out of the labor force. Density (2.27) is conditional on labor force participation. The modification of the analysis to incorporate information on non-participants is straightforward. Obviously, use of the additional information will improve the efficiency of the estimators.

The log likelihood for the model is

$$\mathscr{L} = \sum \int_{\{\varepsilon_c | x \geqq rV \geqq 0\}} \lambda f(x|\theta)[\exp - \lambda(1 - F(rV|\theta))t_u] \, \mathrm{d}G_c^*(\varepsilon_c|\eta_c). \tag{2.29}$$

The likelihood is not monotonic in $Z_c\beta_c$. This is confirmed by direct differentiation of (2.29). The intuition underlying this result is given here rather than the direct formal argument which is both tedious and unilluminating. The key idea is that $\beta_c$ shifts both $rV$ *and the admissible range of the unobservable* $\varepsilon_c$.

The kernel of the integral in (2.29) is monotone decreasing in $Z_c\beta_c$ because $rV$ is monotone decreasing in $c$ (and hence $Z_c\beta_c$) and because the kernel is monotone increasing in $rV$. For each value of $x$, $\lambda$, $\theta$ and $\eta_c$, (2.28) defines an admissible region of $\varepsilon_c$. In general, both the location and length of the admissible region change as $Z_c\beta_c$ and the other parameters are varied. For the special functional form assumed here $[c = c(Z_c\beta_c + \varepsilon_c)]$ for each $x$ (and for fixed values of $\theta$ and $\lambda$) the admissible region for $\varepsilon_c$ is connected and of the same length for different values of $Z_c\beta_c$. However, the position of the interval varies as $Z_c\beta_c$ varies, shifting to the left as $Z_c\beta_c$ increases. For fixed values of the other parameters (including $\eta_c$), increasing $Z_c\beta_c$ shifts the admissible $\varepsilon_c$ interval to the left. As $Z_c\beta_c$ increases the likelihood must eventually decrease. In the limit as $Z_c\beta_c \to \infty$, holding the other parameters fixed, the likelihood diverges to minus infinity as the integrals inside the logarithms of (2.29) vanish. A parallel argument demonstrates that $Z_c\beta_c \to -\infty$ is another inadmissible value of the maximum likelihood estimator.[26]

The key point to extract from this heuristic discussion is that changing a parameter in $\beta_c$ alters the likelihood function in two ways: directly through its effect on $rV$ and indirectly through its effect on the range of the unobservable $\varepsilon_c$. It is the second effect which substitutes for the identifying information achieved from the order statistics in the models presented in the preceding sections.

A more concrete example may prove helpful. Assume the sampling frame is such that $d = 1$. Let $f(x|\theta) = \theta \mathrm{e}^{-\theta x}$. Assume $c$ is a constant and let $\ln \lambda$ be

$$\ln \lambda = \mu + \varepsilon_\lambda,$$

where $\varepsilon_\lambda$ is a mean zero normal random variable with variance $\sigma^2$. Assuming $r = 0$, the reservation wage is [see eq. (2.12)]

$$R = \frac{1}{\theta}(\ln \lambda - \ln c\theta) \geqq 0.$$

[26]This argument is only an heuristic interpretation of first- and second-order conditions which the reader can supply. It is predicated on the assumption that the likelihood is not a monotonic function of $\eta_c$, and that $\eta_c$ can be identified.

The probability that a person engages in search is

$$P(R \geqq 0) = P(\varepsilon_\lambda/\sigma \geqq (\ln c\theta - \mu)/\sigma) = \Phi((\mu - \ln c\theta)/\sigma),$$

where $\Phi$ is the cdf of a unit normal distribution.

In this example, the restriction on the unobservable $\varepsilon_\lambda$ [eq. (2.28)] is

$$\{\varepsilon_\lambda \mid 0 \leqq (1/\theta)(\ln \lambda - \ln c\theta) \leqq x\}.$$

Substituting $\ln \lambda = \mu + \varepsilon_\lambda$ produces a more explicit restriction on the set of admissible $\varepsilon_\lambda$, $\{\varepsilon_\lambda \mid 0 \leqq (1/\theta)(\mu + \varepsilon_\lambda - \ln c\theta) \leqq x\}$. The density for $x$, $t_u$ is [substituting for $R$ in (2.27) assuming $d = 1$]

$$\begin{aligned} m^{**}(x, t_u) &= \left\{ \theta\, \mathrm{e}^{-\theta x - c\theta t_u + \mu} \int_{(\ln c\theta - \mu)/\sigma}^{(\theta x + \ln c\theta - \mu)/\sigma} (2\pi)^{-\frac{1}{2}}\, \mathrm{e}^{-\frac{1}{2}(\varepsilon_\lambda/\sigma)^2} \mathrm{e}^{\varepsilon_\lambda/\sigma}\, \mathrm{d}(\varepsilon_\lambda/\sigma) \right\} \\ &\quad / \Phi((\ln c\theta - \mu)/\sigma) \\ &= \theta\, \mathrm{e}^{-\theta x - c\theta t_u + \mu + (\sigma^2/2)} \\ &\quad \times \left[ \frac{\Phi((\theta x - \mu + \ln c\theta - \sigma^2)/\sigma) - \Phi((\ln c\theta - \mu - \sigma^2)/\sigma)}{\Phi((\ln c\theta - \mu)/\sigma)} \right]. \end{aligned} \tag{2.30}$$

The second line follows from standard results on the moment generating function of truncated normal random variables. Direct differentiation of the likelihood formed from (2.30) reveals that $\theta$, $c$, $\mu$ and $\sigma$ are identified.[27]

[27]An informal approach to identification notes that the hazard for exit to employment ($h$) is independent of $\lambda$. Thus $h = c\theta$. Thus from the duration data, it is possible to estimate $c\theta$. The density of accepted wages is [see (2.25)]

$$c^{-1}\, \mathrm{e}^{-\theta x + \mu + (\sigma^2/2)} \left[ \frac{\Phi((\theta x - \mu + \ln c\theta - \sigma^2)/\sigma) - \Phi((\ln c\theta - \mu - \sigma^2)/\sigma)}{\Phi((\ln c\theta - \mu)/\sigma)} \right].$$

Heuristically we can estimate $\theta$ from the linear term in $x$ of a regression of the log empirical density of accepted wages. Departures from linearity in $x$ identify $\theta/\sigma$ and hence $\sigma$. From the intercept it is possible to identify $\mu$. Since $c\theta$ is known, $\theta$ is known and hence $c$ is known. Note that there is additional information available from knowledge that a person is a labor force participant [which occurs with probability $\Phi((\ln c\theta - \mu)/\sigma)$].

Note further that little modification of the analysis in the text is required if we cannot distinguish individuals who are unemployed from those out of the labor force. Just note that the probability that a spell exceeds $t$ is

$$P(T_u > t_u) = \Phi((\mu - \ln c\theta)/\sigma) + \mathrm{e}^{-(c\theta t_u}\Phi((\ln c\theta - \mu)/\sigma),$$

so now there is duration dependence in the survivor function. (This is an example of the mover–stayer problem). From the survivor data it is clearly possible to estimate $c\theta$ and we can proceed with identification as before.

Moreover, standard asymptotic theory can be used to derive the asymptotic distribution theory of the maximum likelihood estimators.

A complete analysis of identification in a model with unobserved heterogeneity requires knowledge of the functional forms for $G(\varepsilon \mid \eta)$ and $c$, $\lambda$ and $\theta$. A non-parametric identification analysis of this model is impossible and for this reason the model is non-parametrically underidentified.[28]

Even when specific parametric functional forms are assumed identification analysis is difficult for high-dimensional $G$. Moreover, the analysis of Heckman and Singer (1981, 1982) demonstrates that parameter estimates obtained from nonlinear models of the sort described here are very sensitive to choices of functional forms for the distribution of unobservables, and so it is necessary to try a variety of functional forms for $G$ to be sure that estimates and inferences are robust to changes in the assumed functional form of the distribution of unobservables. Because of computational cost, high-dimensional $G$ distributions are unlikely to be used in practice. Selecting which sources of heterogeneity to model and which to ignore injects a further source of arbitrariness into the empirical procedure. For these reasons we claim that the greater simplicity in the asymptotic distribution theory for this model is offset by the arbitrariness inherent in 'controlling for' unobserved heterogeneity.

We note that a procedure analogous to (but not identical to) the one presented in this section is appropriate if $x$ is measured with error. Thus if $x^* = x + \tau$ where $\tau$ is mean zero measurement error, and if $\varepsilon_c = \varepsilon_\lambda = \varepsilon_\theta = 0$, the range of $\tau$ is defined implicitly by

$$\{\tau \mid rV \leqq x^* - \tau\}.$$

In place of (2.27) the density of $x^*$, $t_u$ and $d$ is

$$m^{**}(t_u, x^*, d) = \left\{ \int_{\{u \mid rV \leqq x^* - \tau\}} \lambda f(x^* - \tau \mid \theta)[\exp - \lambda(1 - F(rV \mid \theta))t_u] \, \mathrm{d}G(\tau) \right\}^d$$

$$\times \{\exp - \lambda(1 - F(rV \mid \theta))t_u\}^{1-d},$$

for $rV \geqq 0$, where $G(\tau)$ is the cdf of $\tau$. The likelihood is formed in the usual way. Assuming the model is identified, maximum likelihood estimators are consistent and asymptotically normally distributed.

### 2.5.3. Non-stationarity

There are many interesting and empirically relevant factors that produce a non-stationary model; finite lives for individuals [DeGroot (1970)],

[28]This conclusion is not special to a continuous time search model.

unemployment insurance available for a limited duration, etc. Non-stationarity raises two difficult problems: (A) The econometrician must be much more careful about the initial conditions of the process than has been required in the models analyzed up to this point [Flinn and Heckman (1982)], and (B) properly controlling for non-stationarity raises formidable computational problems [Heckman and Singer (1982)].

## 3. An equilibrium model of sequential search

The one-state model of sequential search analyzed in section 2 is widely cited as the theoretical basis for empirical analysis of unemployment durations. Its value as a guide to data has been questioned because without modification it does not support a non-degenerate equilibrium wage distribution [Diamond (1971), Rothschild (1973)]. This now widely echoed objection has caused many applied economists to reject search theory as an interesting source of hypotheses on unemployment durations. Many would argue that developing the econometrics of an economically uninteresting model puts the cart before the horse.

In this section we develop an equilibrium two-state model of sequential search. The model is closely related to recent work on matching equilibria by Diamond and Maskin (1979) and Jovanovic (1979). We demonstrate that the econometric framework presented in section 2 can be applied, with minor modification, to the equilibrium model analyzed here.

The simplicity of the model proposed and estimated here may offend the sensibilities of readers accustomed to regression analysis with numerous 'control variables'. The simplicity of our model is characteristic of the theory on which we draw. As discussed in Heckman and Singer (1982) small departures from the stationary search model affect the predictions from these models, the existence of optimal solutions and the nature of the appropriate econometric model. For these reasons we are reluctant to 'throw in' numerous control variables in performing our empirical analysis, especially variables that control for non-stationarity.

### *3.1. An equilibrium sequential search model*

Labor market equilibrium is interpreted as the outcome of a steady state sorting and matching process. Pools of unattached risk-neutral workers and firms are in the market seeking to form partnerships to perform tasks. Before a worker and firm meet they do not know the potential rate of productivity of a match. Thus *ex ante* all firms look alike to all workers and all workers look alike to all firms. Upon encounter the productivity rate is perfectly revealed. The time required to complete a task is an exponential random variable. The distribution of the completion time of a task is beyond the control of either party. Each member of a potential partnership gets half of

the total productivity of the match if it is consummated. There is a distribution of match productivity rates in the population known to both workers and firms. Values of potential matches are distributed independently across workers and firms and are distributed independently across potential matches for the same workers and firms. The outcome of one match is of no value in predicting the outcome of another.

The probability that a worker meets a firm in a small time interval $\Delta t$ is less than one and the probability of two or more such encounters is negligible. The probability of encountering a potential partner is assumed to follow the Poisson probability law. The Poisson parameter of encounters depends on the number of firms searching in the market. For simplicity we assume equal numbers of unattached firms and workers ($L$) and that the rate of arrivals to a given worker or firm is the same linear function of $L$. The symmetry of the position of workers and firms is one possible justification for the 50–50 splitting rule assumed here.

Each party to an encounter has two options: (a) to accept the offered terms of the match, or (b) to continue searching. The task production technology is characterized by extreme indivisibilities such that partners cannot search while working. If the prospective partners continue search they incur an instantaneous flow rate of costs of $c$.

Search is with recall. If two prospective partners decide to continue searching they can consummate their match at a later time. However, it is known from the theory of search in stationary environments [Lippman and McCall (1976a)] that they never consummate their match if they ever reject it. The intuitive reason for this result is that since the environment is stationary the attractiveness of any forsaken option is the same tomorrow as it is today.

This model generates equilibrium sequential search which is supported by a non-degenerate equilibrium wage distribution. It shares many features in common with the discrete outcome model of Diamond and Maskin although our interpretation differs somewhat from theirs.[29]

Each potential match has two characteristics: a flow rate of output $2x$ and a rate of termination $\sigma$. Distribution $F(x)$ is assumed to be absolutely continuous with finite absolute first moments. The value of a job to a worker who receives flow rate $x$ (by virtue of the sharing assumption) is defined as $V_e(x)$.

It is determined by the following calculation:

$$V_e(x)=\frac{1}{1+r\Delta t}\{x\Delta t+(1-\sigma\Delta t)V_e(x)+\sigma\Delta t V_u\}. \tag{3.1}$$

[29]An alternative model not estimated here is one due to Telser (1980). His model generates the functional form of the equilibrium wage offer distribution but there is no search in equilibrium. The model of Jovanovic (1979) is similar in many respects to the model presented in the text.

[Terms of $o(\Delta t)$ are ignored.] The first term is the (discounted) value of the flow $x$ over the interval $\Delta t$. The second term is the (discounted) value of keeping the job times the probability of keeping the job $(1-\sigma\Delta t)$. The third term is the (discounted) value of becoming unemployed ($V_u$) times the probability of becoming unemployed. Collecting terms and passing to the limit we reach

$$V_e(x)=(x+\sigma V_u)/(r+\sigma). \tag{3.2}$$

As a consequence of $F(x)$ having finite absolute first moments $V_u$ is finite. $V_e(x)$ is increasing in $x$ and has bounded expectation. The agent is indifferent between a match that yields an $x^*$ such that $V_e(x^*)=V_u$ and unemployment. For any $x \geqq x^*$ the agent earns a rent with value $V_e(x)-V_u$. $x^*$ is the reservation wage and it is defined as that value of $x^*$ such that

$$V_e(x^*)=V_u=(x^*+\sigma V_u)/(r+\sigma),$$

so

$$x^*=rV_u.$$

The rate of arrival of encounters is $\lambda=g(L)$ assumed to be differentiable. The more participants that are searching in the market (i.e., the higher $L$), the more rapid is the rate of arrival of encounters,

$$g(0)=0, \qquad g'(L)>0 \quad \text{for all} \quad L \geqq 0.$$

The value of unemployment is

$$V_u=-\frac{c\Delta t}{1+r\Delta t}+\frac{(1-\lambda\Delta t)}{1+r\Delta t}V_u+\frac{\lambda\Delta t}{1+r\Delta t}E\max\left[V_e(x), V_u\right]. \tag{3.3}$$

[Terms of $o(\Delta t)$ are ignored.] The first term on the right is the discounted direct cost of search. The second term is the discounted value of remaining unemployed times the probability that no potential match is encountered $(1-\lambda\Delta t)$. The third term is the probability of encountering a potential match $(\lambda\Delta t)$ times the discounted expected value of the two options that confront the agent: to take the match [with value to him of $V_e(x)$] or to remain unemployed [with value $V_u$]. The expectation is computed with respect to the distribution $F(x)$.

Since the model possesses the reservation wage property, passing to the limit we may write (3.3) for $V_u>0$ as

$$rV_u+c=\lambda\int_{rV_u}^{\infty}(V_e(x)-V_u)\mathrm{d}F(x)=\frac{\lambda}{r+\sigma}\int_{rV_u}^{\infty}(x-rV_u)\mathrm{d}F(x). \tag{3.4}$$

To assure that $V_u > 0$ it is required that

$$c < \frac{\lambda}{r+\sigma} \int_0^\infty x \, \mathrm{d}F(x).^{30} \tag{3.5}$$

Note that $\lambda = (g(L))$ must be sufficiently large to support search behavior. Thus in small markets (e.g. rural areas) search may not be undertaken.

The rate of exit from unemployment is

$$h(t_u) = \lambda(1 - F(rV_u)), \tag{3.6}$$

while the rate of exit from employment is

$$h(t_e) = \sigma. \tag{3.7}$$

In steady state equilibrium, the number of individuals exiting unemployment must equal the number entering unemployment. Thus letting $P$ be the total market size and $L$ the total stock of unemployed, steady state equilibrium requires that

$$\lambda(1 - F(rV_u))L = \sigma(P - L),$$

where $P - L$ is the total stock of the employed. Thus

$$g(L)(1 - F(rV_u))L + \sigma L = \sigma P. \tag{3.8}$$

Condition (3.8) imposes further constraints on the problem. In particular for given values of the parameters there may be no value of $L$ that solves (3.8) or there may be several. The minimum size of the unemployment pool that is required to support an equilibrium with search is [from condition (3.5) using $\lambda = g(L)$ and the monotonicity of $g(L)$]

$$L = g^{-1}((r+\sigma)c/\mu_x),$$

where $\mu_x = \int_0^\infty x \, \mathrm{d}F(x)$. Thus the higher the cost of search, the rate of interest and the rate of match termination and the lower the mean of the productivity distribution, the higher the unemployment pool must be to support a search unemployment equilibrium. Determination of the uniqueness of the equilibrium hinges on the functional form of the match distribution. For example, if $F(x)$ is exponential there is at most one equilibrium and there is always one for a large enough population $P$.[31]

[30]The proof is immediate from (3.3).

[31]To prove this differentiate the left-hand side of (3.8) with respect to $L$ and notice that it is monotone increasing in $L$.

Assuming a unique equilibrium, the unemployment rate is

$$\mu = L^*/P.$$

The higher the cost of search and the lower the interest rate, the lower the equilibrium rate of unemployment. The effect of an increase in $\sigma$ is ambiguous. Positive translations of the match productivity distribution reduce the equilibrium unemployment rate.

Condition (3.8) may be used in several ways in performing empirical work. It provides a check on estimates of the model because the aggregate unemployment rate $\mu$ should equal

$$\mu = \sigma/[\sigma + \lambda(1 - F(rV_u))].$$

It may also be used to eliminate one parameter from the estimation procedure.

### 3.2. Econometrics of the two-state model

The statistical structure of the two-state model so closely resembles that of the one-state model that a lengthy discussion is unnecessary. Here we only consider the case of a homogeneous population.

The density of accepted wages is

$$f(x \mid x > rV_u) = f(x)/(1 - F(rV_u)), \qquad x \geqq rV_u. \tag{3.9}$$

The density of duration time spent in unemployment ($t_u$) is (defining $d_u = 1$ if a spell is uncensored, and zero otherwise)

$$\{\lambda(1 - F(rV_u))\}^{d_u} \exp - \{\lambda(1 - F(rV_u))t_u\}, \tag{3.10}$$

where $t_u$ is the measured length of a spell. Thus the joint density of accepted wages and duration $t_u$ is

$$m(t_u, x, d) = \{\lambda f(x)\}^{d_u} \exp - \{\lambda(1 - F(rV_u))t_u\}. \tag{3.11}$$

The density of match duration $t_e$ is (defining $d_e = 1$ if a spell is uncensored, and zero otherwise)

$$g(t_e) = (\sigma)^{d_e} \exp - \{\sigma t_e\}. \tag{3.12}$$

These equations together with eq. (3.4), inequality (3.5) and the requirement that accepted wages exceed reservation wages,

$$x \geqq rV_u, \tag{3.13}$$

define the full econometric model.

As in the one-state model, inequality (3.13) is essential in securing structural estimates of the $c$ and $\lambda$ functions. Satisfaction of inequality (3.5) is essential to make the model economically meaningful. As in the one-state model, the minimum of accepted wages (or any fixed order statistic) is a consistent estimator of $rV_u$. A recoverability condition for $f(x)$ is essential. If met we can estimate $f(x)$. From the hazard rate for leaving unemployment, $\lambda$ may be consistently estimated, and from the hazard rate for leaving employment, $\sigma$ may be consistently estimated. Provided condition (2.10) is met for $F$, we may condition on $\widehat{rV_u}$ in the concentrated likelihood to secure $\sqrt{N}$ asymptotically normally distributed and consistent estimators of the structural parameters.

As a consequence of the homogeneity assumption spells may be pooled in the log likelihood function $\mathscr{L}$. Thus we may define

$$\mathscr{L} = \sum [d_u(\ln \lambda f(x)) - \lambda(1 - F(rV_u))t_u + d_e(\ln \sigma) - \sigma t_e].$$

The concentrated likelihood can be written as

$$\tilde{\mathscr{L}} = \sum [d_u(\ln \lambda f(x)) - \lambda(1 - F(\widehat{rV_u})t_u + d_e \ln \sigma - \sigma t_e].$$

For the case of exponential matches,

$$f(x) = \theta e^{-\theta x},$$

we may write

$$\tilde{\mathscr{L}} = \sum [d_u(\ln \lambda + \ln \theta - \theta x) - \lambda e^{-\theta \widehat{rV}_u} t_u + d_e \ln \sigma - \sigma t_e]. \tag{3.14}$$

The maximum likelihood estimators of $\lambda$, $\theta$ and $\sigma$ are, respectively,

$$\hat{\lambda} = (\sum d_u) e^{\theta \widehat{rV}_u} / \sum t_u, \tag{3.15a}$$

$$\hat{\theta} = 1 / \sum d_u (x - \widehat{rV_u}), \tag{3.15b}$$

$$\hat{\sigma} = \sum d_e / \sum t_e. \tag{3.15c}$$

Since condition (2.10) is satisfied,

$$\sqrt{N} \begin{pmatrix} \hat{\lambda} - \lambda \\ \hat{\theta} - \theta \\ \hat{\sigma} - \sigma \end{pmatrix}$$

is normally distributed with mean zero and a covariance matrix that can be estimated by minus the inverse of the matrix of second partials of the log likelihood. A consistent estimator of search cost $c$ may be obtained by solving from (3.4), and its standard error can be derived from the delta method.

### 3.3. Estimates of the two-state model

This section reports maximum likelihood estimates of the parameters of two specifications of the two state equilibrium model. The first specification assumes that (firm or worker) match values are exponentially distributed; the second assumes that they are normally distributed. Both sets of estimates are derived under the assumption of homogeneity although we attempt to control for heterogeneity by estimating the model for one demographic group: young (20–24) white male high-school graduates not attending school sampled from the National Longitudinal Surveys in 1971. A more complete description of these data is given below table 1. To make the model empirically operational we define the unemployment state more broadly than is usually done to include unemployment as conventionally defined plus non-participation in the market.

The parameter estimates for the two models are presented in table 1. The estimated reservation wage is \$1.5 (stated as an hourly rate). The mean length of a job ($1/\sigma$) is 28 months. For both models viability condition (3.5) is satisfied.

Table 2 records some of the more interesting features of these estimates. The first two rows record estimates of the cost of search ($c$) secured from eq. (3.4). These estimates are obviously quite sensitive to the choice of interest rate $r$ and the assumed functional form for the distribution of match productivities. The exit rate from the non-employed ('search') state $h_u$ is 1.21 which implies a mean length of non-employment of 8.3 months. Given $\sigma$ and $h_u$ the implied equilibrium non-employment rate is 22.5% which compares closely with the actual non-employment rate of 24.5%.

As discussed in section 2, estimates of the rate of job offer or match arrivals are quite sensitive to the assumed functional form of the offer distribution. This is dramatically illustrated in sections four and five of table 2. If an exponential model is assumed the estimated mean length of time between offers ($1/\lambda$) is five months and the probability of receiving no offer in the next six months is only 0.3. The probability of accepting a match that is offered is 0.6. Thus the exponential model implies a relatively rapid rate of arrivals of offers and a relatively frequent rate of rejection of those offers.

If a normal model is assumed, the story changes. In this case the mean length of time between offers is 7.5 months and the probability of receiving no offer in the next six months is 0.45. However, the probability of accepting

Table 1

Estimates of the two-state equilibrium search model (standard errors in parentheses).[a]

| $\lambda$ (rate of arrival of job offers) | $rV_u$ (reservation wage, \$/hr) | $\theta$ (parameter of exponential match distribution) | $M$ (mean of normal match distribution) | $\sigma^2$ (variance of normal match distribution) | $\sigma$ (rate of termination of jobs) |
|---|---|---|---|---|---|
| *Exponential model* | | | | | |
| 0.201 | 1.5 | 0.339 | — | — | 0.035 |
| (0.008) | — | (0.038) | | | (0.035) |
| *Normal model* | | | | | |
| 0.1318 | 1.5 | — | 3.325 | 1.709 | 0.035 |
| (0.004) | — | | (0.34) | (0.16) | (0.035) |

The mean of the accepted wage offer distribution is $\bar{x}=3.45$ and the variance is 1.464.

[a]The data are taken from the National Longitudinal Study of Young Men, the first wave of which was collected in 1966. White, not enrolled in full-time schooling (in 1971), hourly workers (determined by their current employment status) were used. The final requirement was made to eliminate the need to compute hourly wage rates for workers not paid on an hourly basis. We use data on lengths of employment and non-employment spells occurring in 1971. We use data on accepted wages to estimate the parameters of the wage distribution. Information from 231 complete and incomplete unemployment spells and 2915 complete and incomplete non-employment spells is used. The data are available from the authors on request.

Table 2

Implications of the estimates (standard errors in parentheses).

| | | Exponential match | Normal match |
|---|---|---|---|
| *Implied cost of search, c* | $r=0.05$ | 1.139 | 0.327 |
| | | (0.616) | (0.148) |
| | $r=0.10$ | 2.692 | 1.4547 |
| | | (1.86) | (0.643) |
| Implied exit rate from 'unemployment', $h_u(=\lambda(1-F(rV_u))$ | | 1.21 | 1.21 |
| Implied equilibrium 'unemployment rate', $(\mu=L^*P)=\sigma/(\sigma+h_u)$ | | 0.225 | 0.225 |
| Actual 'unemployment rate' (1971)[a] | | 0.245 | 0.245 |
| Probability of encountering no match in the next six months | | 0.30 | 0.45 |
| Probability of encountering one match in the next six months | | 0.36 | 0.36 |
| Probability of encountering two or more matches in the next six months[b] | | 0.34 | 0.19 |
| Probability that a potential match will be accepted $(=\int_{rV_u}^{\infty} \mathrm{d}F(x))$ | | 0.60 | 0.92 |

[a]This is the CPS non-employment rate for the group taken from *Employment and Training Report of the President*, 1976.

[b]Computed from the Poisson law $P(j)=(\lambda t)^j \mathrm{e}^{-\lambda t}/j!$

a match that is offered rises to 0.92. Thus in the normal model, job arrivals are a relatively rare event and most workers accept the first job offer they receive.

In this particular case, the choice of model is clear. The accepted wage data are not exponentially distributed and an eyeball test indicates that the implied truncated normal distribution fits the data better. Before taking much comfort from this point, it is important to recall the point made in section 2. In order to estimate $\lambda$ it is necessary to assume that the match offer distribution is recoverable. This requires making an arbitrary parametric assumption about the functional form of the productivity match distribution and the choice of this functional form will critically affect estimates of certain parameters of the model. Although we can check the plausibility of any assumed functional form (assuming recoverability) by inspecting accepted wage distributions, the power of such a test is low. The recoverability condition cannot be tested.

Appendix B presents a three-state model of labor force dynamics. Two new points are made there: (A) non-anticipated shocks in non-market time are introduced, and (B) a structural economic interpretation of the competing risks model widely used in social science is provided. Point A is particularly important because it provides a good economic reason, not provided in this section, for matches between workers and firms to dissolve. A worker (or entrepreneur) may terminate a match in the model of appendix B because of an unanticipated increase in the value of non-market time. More general models of this type are presented in Coleman and Heckman (1981).

## 4. Summary and conclusions

This paper takes a first step toward developing rigorous structural econometric models for the empirical analysis of labor force dynamics. The analysis is presented in continuous time although most of the new points raised here apply with full force to discrete time models. Rigorous structural estimation requires the solution of new problems that have not been addressed in previous work. We identify those problems and offer some solutions. We have also stated certain unsolved problems in the belief that their clear statement will focus attention on the relevant issues that should be addressed in subsequent econometric research on structural models of labor force dynamics.

A key condition required in our analysis is a *recoverability condition* that is implicit in all econometric analyses of truncated data of which structural models for the analysis of labor force dynamics are a special case. This condition must be satisfied in order to recover an untruncated distribution from a truncated distribution with a known point of truncation. A major conclusion of our analysis is that a recoverability condition will only be

satisfied if the untruncated distribution is assumed to belong to a parametric family. It is not satisfied for all parametric families. Recoverability is not a testable proposition.

The implications of this analysis are far ranging. Economic theory is rarely so specific as to predict the functional forms of untruncated distributions.[32] Thus, at its foundation, structural estimation in all truncated data models involves an intrinsically arbitrary decomposition of reduced form parameters into structural components. Both formally, and by way of empirical examples, we have demonstrated that estimates of structural parameters are very sensitive to arbitrary choices made about functional forms required to empirically implement structural models. Put more succinctly, consistent non-parametric estimation of the parameters of most structural duration (and truncated data) models is impossible. The arbitrariness inherent in structural estimation is essential.

Granting the satisfaction of a recoverability condition (as the entire literature on the econometrics of truncated data has done) we have also demonstrated that structural estimation in a general class of dynamic labor force models requires the utilization of inequality constraints on accepted offers. Unless these constraints are used, the structural parameters of the model are not identified. Previous analysts have not noted this essential feature of these models.

Incorporation of the identifying inequality constraints into an estimation procedure raises certain non-standard statistical problems. Conventional regularity conditions are violated so that the standard theory of maximum likelihood estimation does not apply. We develop the appropriate statistical theory for certain cases likely to be of interest in applications but we have not developed the general case in this paper although we indicate how this might be done.

We also formulate a structural two-state *equilibrium* matching model of unemployment and estimate a simple version of it. Previous empirical research on labor force dynamics has been conducted within the framework of partial equilibrium search models which are not internally consistent because they imply degenerate equilibrium wage distributions. The model developed in this paper produces a non-degenerate equilibrium wage offer distribution.

Empirical results obtained from this model provide an economically meaningful decomposition of the parameters of reduced form exit rates from unemployment. However, we demonstrate that the estimated structural parameters are very sensitive to arbitrary assumptions about functional forms of estimating equations. We demonstrate that two different behavioral models are broadly consistent with the same data.

[32]A notable exception is the analysis of Telser (1980) on equilibrium price and wage distributions.

Our initial enthusiasm for structural estimation has been tempered somewhat by the analytical and computational problems reported in this paper. The econometric analysis of structural duration models is still in its infancy. We are confident that subsequent analysis will solve some of the problems left unsolved here.

Nonetheless, because much analysis of structural models remains to be done, because basic facts on labor force dynamics remain to be collated and analyzed and because of the inherent arbitrariness of the identifying assumptions used to secure structural estimates (and the sensitivity of estimates of structural parameters to such assumptions), we also feel that a reduced form approach to analyzing labor force models may be a useful complement to the structural approach presented in this paper. A companion paper [Flinn and Heckman (1982)] developes a reduced form approach to the empirical analysis of labor force dynamics.

## Appendix A

*The maximum likelihood estimator in the case of more general distributions for Z*

A complete analysis for the case of a general distribution for $Z$ is a task for the future. Here we analyze the exponential example with no censoring for a case in which $\lambda$ and $c$ are positive constants and $\theta(Z)=\exp(Z_\theta\beta_\theta)$. Analysis of this example serves as a prototype for the more general case.

The model is that of eqs. (2.19), (2.20) and (2.21) in the text with $\lambda=\mathrm{e}^{\beta_\lambda}$ and $c=\mathrm{e}^{\beta_c}$. The sampling frame is assumed to be sufficiently long that $d=1$. The log likelihood is

$$\mathscr{L}=N\ln\lambda+\sum(Z_\theta\beta_\theta-\exp(Z_\theta\beta_\theta)x)-c\sum \mathrm{e}^{Z_\theta\beta_\theta}t_u,$$

with constraints for each observation,

$$x\geqq \mathrm{e}^{-Z_\theta\beta_\theta}(\ln\lambda/c-Z_\theta\beta_\theta), \qquad \lambda, c>0. \tag{A.1}$$

A critical aspect of this problem is that $\mathscr{L}$ is monotonically increasing in $\lambda$ and decreasing in $c$ and hence monotonically increasing in $\lambda/c$. The constraints are thus always effective in securing determinate solutions for $\lambda$ and $c$. It is useful to rewrite the first constraint as

$$x\mathrm{e}^{Z_\theta\beta_\theta}+Z_\theta\beta_\theta\geqq\ln\lambda/c. \tag{A.2}$$

As a consequence of the fact that $\mathscr{L}$ is monotone increasing in $\lambda/c$ (for fixed $\beta$), the maximum likelihood estimator of $\lambda$ and $c$ is derived by choosing

$\lambda/c$ so that

$$\ln(\widehat{\lambda/c}) = \min\{x\mathrm{e}^{Z_\theta\beta_\theta} + Z_\theta\beta_\theta\}, \tag{A.3}$$

where the right-hand side term is the minimum sample value of the term inside the braces. Exponentiating both sides, and substituting for $c$ in $\mathscr{L}$ we reach a concentrated likelihood

$$\tilde{\mathscr{L}} = N\ln\lambda + \sum(Z_\theta\beta_\theta - \exp(Z_\theta\beta_\theta)x)$$
$$-\lambda\exp[-\min\{x\mathrm{e}^{Z_\theta\beta_\theta} + Z_\theta\beta_\theta\}]\sum \mathrm{e}^{Z_\theta\beta_\theta}t_u.$$

Maximum likelihood estimators of $\lambda$ and $\beta_\theta$ are achieved by maximizing this function; $c$ can be solved from (2.16).

The estimation strategy in this model is thus analogous to that presented in section 2.4 except now we require an order statistic of a nonlinear function of $x$ and $Z_\theta$ to secure estimates.

Define $M = x\mathrm{e}^{Z_\theta\beta_\theta} + Z_\theta\beta_\theta$. For $Z_\theta$ iid the density $f(\cdot)$ of $M = m$ can be shown to be[33]

$$f(m) = (\lambda/c)\,\mathrm{e}^{-m}, \qquad m \geqq \ln(\lambda/c),$$

so the density of $\min\{M\}$ is

$$f(\min\{m\}) = (\lambda/c)^N \mathrm{e}^{-N[\min\{m\}]}, \qquad \min\{m\} \geqq \ln(\lambda/c).$$

Assuming bounded parameters and bounded $Z$, $\min\{M\}$ is strongly consistent for $\ln(\lambda/c)$ and we conjecture that weak consistency of the mle estimators for $\lambda$, $c$ and $\beta_\theta$ is an almost immediate consequence (assuming the population covariance of $Z_\theta$ is of full rank) and that the estimators for $\lambda$ and $\beta_\theta$ can be shown to be asymptotically normal (with norming factor $\sqrt{N}$).

The argument can be extended to permit $\lambda$ and $c$ to depend on exogenous

[33]For a given value of $\beta_\theta$ and a density for $Z_\theta$, $Z_\theta\beta_\theta$ has a density $h(Z_\theta\beta_\theta)$. The density of $x$ conditional on $Z_\theta\beta_\theta$ and $x$ satisfying constraint (A.2) is

$$f(x \mid Z_\theta\beta_\theta) = (\lambda/c)\exp(-\mathrm{e}^{Z_\theta\beta_\theta}x), \qquad x \geqq \mathrm{e}^{-Z_\theta\beta_\theta}(\ln(\lambda/c) - Z_\theta\beta_\theta).$$

The joint density of $x$ and $Z_\theta\beta_\theta$ is

$$g(x, Z_\theta\beta_\theta) = (\lambda/c)[\exp-(\mathrm{e}^{Z_\theta\beta_\theta}x)]h(Z_\theta\beta_\theta), \qquad x \geqq \mathrm{e}^{-Z_\theta\beta_\theta}(\ln(\lambda/c) - Z_\theta\beta_0).$$

Use the definition of $m$ given in the text, and applying standard change of variable methods,

$$f(m) = (\lambda/c)\int \mathrm{e}^{-m}h(Z_\theta\beta_\theta)\mathrm{d}(Z_\theta\beta_\theta)$$
$$= (\lambda/c)\,\mathrm{e}^{-m}, \qquad m \geqq \ln(\lambda/c).$$

The fact that $f(m)$ does not depend on the parameters of the $Z_\theta$ distribution and on $\beta_\theta$ is an artifact of the exponential model.

variables. For specificity we write $\lambda$, $c$, $\theta$ as in (2.19). $K_\lambda$ is the number of parameters in $\beta_\lambda$ and $K_c$ is the number of parameters in $\beta_c$.

Inequality (A.1) which must be satisfied for each observation may be written as

$$x e^{Z_\theta \beta_\theta} + Z_\theta \beta_\theta \geqq Z_\lambda \beta_\lambda - Z_c \beta_c. \tag{A.4}$$

For any set of $K_c$ values on the left-hand side there are associated $Z_c$ and $Z_\lambda$ vectors. We define the units of $Z_c$ so that any submatrix of $K_c$ vectors of $Z_c$ is positive.

Array the $N$ values of the left-hand side of (A.4) into a $N \times 1$ vector $S$. Define $[Z_c]$ and $[Z_\lambda]$ as the data matrices for the associated $N$ values of $Z_c$ and $Z_\lambda$. The vector of inequalities (A.4) may be written as

$$S \geqq [Z_\lambda]\beta_\lambda - [Z_c]\beta_c. \tag{A.5}$$

For the special but plausible case in which $Z_c = Z_\lambda$, this vector inequality can be written as

$$S \geqq [Z_c](\beta_\lambda - \beta_c). \tag{A.6}$$

From the assumed positivity of $Z_c$, the likelihood can be increased by sending each element in $\beta_c$ to minus infinity and each element of $\beta_\lambda$ to plus infinity. Constraint (A.6) defines a region which is the intersection of the hyperplanes defined by each inequality in the vector of inequalities (A.6). To find the boundary of the intersection region solve the following linear programming problem.

Define $P$ a positive vector conformable with $\beta_\lambda - \beta_c$. Let $\phi = (\beta_\lambda - \beta_c)$. Then find $\phi$ that solves the programming problem

$$\max \phi' P \quad \text{subject to} \quad S \geqq [Z_c]\phi.$$

The values of $\phi$ expressed as a function of $P$ trace out a $\phi$ frontier. The maximum likelihood estimator selects a point on this frontier.[34] There are as many edges on the frontier as elements in $\beta_\lambda - \beta_c$ (i.e., $K_c = K_\lambda$). For any given value of $\beta_\theta$ and for given values of $Z_\theta$, $Z_\lambda$ and $Z_c$, the simplex algorithm may be used to solve for the frontier.

In an iteration cycle, for fixed $\beta_\theta$ and $\beta_\lambda$, it is thus possible, in general, to solve out for $\beta_c$ as a function of $\beta_\lambda$, $Z$ and $\beta_\theta$. The likelihood function is then maximized with respect to $\beta_\lambda$ and $\beta_\theta$ and new values of the frontier are solved out, etc.

[34]This is so because the function is maximized by choosing the largest possible value of $\phi$ given any selected set of ratios in the $P$ vector.

Estimation in the more general case (A.5) proceeds along similar lines replacing $S$ in the preceding paragraph with $S-[Z_\lambda]\beta_\lambda$ and $\phi$ with $-\beta_c$. Establishing the statistical properties of these estimators is a task left for the future.

## Appendix B

*A multistate model of labor force dynamics*[35]

This appendix presents a prototypical model of labor force dynamics. With the exception of pioneering work by Toikka (1976), there are few examples of multistate *structural* models of labor force dynamics. The multistate models estimated by Marston (1976) were not given a structural interpretation.

Agents have linear utility functions and they are assumed to live forever. There is stochastic variation in the (monetary) value of non-market time in addition to stochastic variation in the wage offers received by agents. Variation in non-market time can arise because of variation in the demand for home time caused by children, illness, and the like. Considering such sources of unanticipated *stochastic* fluctuation in the value of non-market time (not known to the agent) enables us to produce a tractable economic model of labor force participation and job turnover that is especially helpful in explaining the labor force movements of women. Variation in the value of non-market time provides one motivation, not provided in section 3, for termination of matches between workers and firms.

We permit wage offers to arrive while individuals are employed. Wage growth occurs over the life cycle as individuals take repeated drawings from common wage distributions. Burdett's (1978) model of on the job search ignores non-market participation, and so his theory produces a life cycle theory of the growth of wages with age. Our model produces a theory of the growth of wages with work experience. Thus an individual who withdraws from the work force because of a high value of household time will forfeit his wage growth. That wages grow with work experience and not, primarily, with age, is the key finding of Mincer (1974) so that our model is more in accord with the facts than Burdett's and is especially suited to explain the wage growth of women. By permitting employed workers to receive job offers we also produce a simple model of job turnover.

Our model also departs from previous work by permitting individuals who are not in the labor force to receive wage offers. Our model thus predicts that it is possible to observe individuals who transit from being out of the labor force to the employed state who never report themselves as being unemployed, as in fact is observed in the data. The labor market state

[35]This material is based on Coleman and Heckman (1981).

'unemployed' is *defined* as that state in which the rate of arrival of job offers is higher than in the other labor force states. Individuals may choose to become unemployed if the rate of job offer arrivals in the state is sufficiently high. The cost of being unemployed is a direct money cost as well as the value of household time forgone. In order to make the unemployed state an economically viable option, the rate of arrival of job offers while unemployed must exceed the rate in the out of the labor force state.[36]

There are many other ways to model unemployment which we do not pursue here. Unemployment may simply be a non-market state in which individuals collect unemployment payments and consume non-market time. Given the fixed duration of most unemployment benefit programs, such a model would be inherently non-stationary although analytically tractable (i.e., it delivers qualitative results).

Our model is similar to one developed by Toikka (1976). Both models consider search in an environment in which the value of household time changes randomly. Our model is more general in that (a) wage offers and changes in the value of non-market time are not constrained to arrive one per period as in his discrete time model, and (b) we do not assume the arbitrary sequential ordering of decisions that he imposes in his model. The rate of time preference is $\rho$.

The distribution of wage offers $F(x)$ is assumed to be absolutely continuous with finite absolute moments. Successive wage offers are statistically independent. The value of non-market time is a random variable with distribution function $G(n)$ with finite absolute moments. Successive values of non-market time are statistically independent. In all states the arrivals of new values of $n$ are governed by a Poisson process with parameter $\mu$. The probability of receiving a new value of $n$ in the interval $\Delta t$ is $\mu\Delta t + o(\Delta t)$.

In the employed state, wage offers arrive by a Poisson process with parameter $\lambda_e$. In the non-market labor state, wage offers arrive by a Poisson process with parameter $\lambda_n$. In the unemployed state wage offers arrive by a Poisson process with parameter $\lambda_u$, and to satisfy viability $\lambda_u > \lambda_n$ and $\lambda_u > \lambda_e$.

Employed workers may leave their current job for two reasons:[37] they may leave if a new value of household time $n$ is drawn that is sufficiently high, but they also may leave if they receive a wage offer elsewhere that is sufficiently high. To simplify the presentation we ignore any transaction costs in movement among these states although the framework developed can readily be extended to accommodate such costs.

The value of a job that pays instantaneous wage $x$ when a household

[36]This insight is the basis of a test of search theory; see below.

[37]Coleman and Heckman (1981) consider a more general model with exogenous terminations. In addition, they permit wage offers and non-market time offers to be correlated contemporaneously and over time.

opportunity $n$ is available is $V(x)$. It is determined by computing the expected value of all the outcomes that can occur to an individual who holds such a job.

There are three possibilities. First, in a small interval ot time, no wage offer may be received and there may be no change in the value of non-market time. If it is optimal to hold a job at the beginning of the interval it is also optimal to hold it at the end. Second, the agent may receive a new non-market value of time $n$. (This could be due to birth of children or the like.) In this event he compares the value of staying on the job with the value of the non-market state defined here as $S(n)$. The expected value of the maximum of these two options is computed at the beginning of the interval $\Delta t$ and is a component of $V(x)$. Third, the agent may receive a new market wage offer $x'$ in time interval $\Delta t$. If this event occurs, he compares the value of the current job to the value of the offered job. Because of the Poisson assumption, the probability of receiving both a new wage and a new non-market value in interval $\Delta t$ is negligibly small.[38]

Thus $V(x)$ is defined as

$$V(x)=\frac{x\Delta t}{1+\rho\Delta t}+\frac{(1-(\lambda_e+\mu)\Delta t)}{1+\rho\Delta t}V(x)+\frac{(\mu\Delta t)}{1+\rho\Delta t}\underset{n'}{E}\max\{V(x),S(n')\}$$

$$+\frac{(\lambda_e\Delta t)}{1+\rho\Delta t}\underset{x'}{E}\max\{V(x),V(x')\}+\mathrm{o}(\Delta t). \tag{B.1}$$

The first term is the present value of wages $x$ received over the next small interval of time $\Delta t$. The second term is the present value of the value of a job ($V(x)$) times the probability that no new wage offer is received $(1-\lambda_e\Delta t)$ times the probability that no new market offers arrive $(1-\mu\Delta t)$. Neglecting terms of order smaller than $\mathrm{o}(\Delta t)$, $(1-(\lambda_e+\mu)\Delta t)$ is the probability of no new non-market offers, and no job termination. If this event occurs, the optimal policy is to continue working at wage $x$ because no new information has arrived to change the agent's optimal decision, and because the optimal policy is stationary (because of the infinite life assumption).

The third term is the probability of receiving a new non-market draw $(\mu\Delta t)$ times the discounted expected value of the maximum of two options that exist if he receives such a draw: to continue to work at wage $x$ ($V(x)$) or to switch to non-market activity with value $S(n)$. The expectation is computed with respect to the distribution of the value of non-market time $G(n)$.

[38]Note that if we permit exogenous job terminations, the value of non-market time is *not* irrelevant in valuing a job. In the more general case considered in Coleman and Heckman (1981), when an individual is terminated, he may be forced to accept an option that he rejected, including previously foresaken non-market opportunities.

The fourth term is the probability of receiving a new wage offer ($\lambda_e \Delta t$) times the discounted expected value of the maximum of two options that exist if he receives a wage offer: to continue to work at wage $x$ (V (x)) or to switch to a new job which pays wage $x'$ with value $V(x')$. The expectation is computed with respect to the distribution of wage offers $F(x)$.

$S(n)$ is the value of household activity given non-market draw $n$. Its derivation is analogous to that given for $V(x)$. Three things can happen to a person engaged in household activity at the beginning of interval $\Delta t$. (1) No new non-market or market draws arrive. This event occurs with probability $(1-(\lambda_n+\mu)\Delta t)$. In this case, the individual continues to stay out of the workforce. (2) A new non-market draw arrives. This event occurs with probability $\mu\Delta t$. If it does occur, the individual chooses the labor force option that gives the maximum of the new $S(n)$ or the value of search $Q$. If he chooses to search, the individual is unemployed. (3) A market wage offer may arrive. This event occurs with probability $\lambda_n \Delta t$. If it does occur, the individual chooses the option with the highest value: he either continues on in the non-market state or takes the job offer. As before, the simultaneous arrival of a wage offer and fresh non-market draw is of negligible probability.

Defining $Q$ as the value of search, we may write $S(n)$ as

$$S(n)=\frac{n\Delta t}{1+\rho\Delta t}+\frac{(1-(\lambda_n+\mu)\Delta t)}{1+\rho\Delta t}S(n)+\frac{(\mu\Delta t)}{1+\rho\Delta t}\underset{n'}{E}\max[S(n'),Q]$$

$$+\frac{\lambda_n\Delta t}{1+\rho\Delta t}\underset{x'}{E}\max[V(x'),S(n)]+\mathrm{o}(\Delta t). \tag{B.2}$$

$Q$ is derived by an analogous argument. A searching individual is assumed to consume no non-market time (an assumption that can easily be relaxed) and to incur instantaneous costs of search $c$. Thus we write

$$Q=\frac{-c\Delta t}{1+\rho\Delta t}+\frac{(1-(\lambda_u+\mu)\Delta t)}{1+\rho\Delta t}Q+\frac{\lambda_n\Delta t}{1+\rho\Delta t}\underset{x'}{E}\max[Q,V(x')]$$

$$+\frac{\mu\Delta t}{1+\rho\Delta t}\underset{n'}{E}\max[Q,S(n')]+\mathrm{o}(\Delta t). \tag{B.3}$$

An unemployed individual may transit to employment or to the non-market participation state.

Formal properties of this model (and more general models) are established in Coleman and Heckman (1981). A standard contraction mapping argument establishes the existence and uniqueness of an optimal policy under the stated conditions. The model has the reservation wage property. (For a proof, see Coleman and Heckman.)

Define $R_{en}$ as the reservation wage value of $n$ that makes an employed individual who has just received a new non-market value of $n$ indifferent between going to the non-market state and continuing to work at wage $x$. $R_{ee}$ is the reservation value of $x'$ that makes an employed individual indifferent between his current job with wage $x$ and an offered job. In the absence of transactions costs, $R_{ee}=x$. $R_{nu}$ is the reservation value of $n$ that makes a non-market participant with a new value of $n$ indifferent between staying on in the non-market state and switching to unemployment. It is implicitly defined by $S(R_{nu})=Q$. $R_{ne}$ is the reservation market wage that makes a non-particpating individual with non-market value $n$ indifferent between continuing in the non-market state and accepting a job. It is implicitly defined by $V(R_{ne})=S(n)$. $R_{ue}$ is the reservation wage for an unemployed person and it is solved from $V(R_{ue})=Q$. $R_{un}$ is the reservation non-market draw that makes an unemployed individual indifferent between unemployment and non-market participation. It is implicitly defined by solving $S(R_{un})=Q$.

The reservation wages are thus defined by the following equations:

$$V(x)=S(R_{en}(x)), \tag{B.4a}$$

$$V(x)=V(R_{ee}) \quad \text{so} \quad R_{ee}=x, \tag{B.4b}$$

$$S(R_{nu})=Q, \tag{B.4c}$$

$$V(R_{ne}(n))=S(n), \tag{B.4d}$$

$$V(R_{ue})=Q, \tag{B.4e}$$

$$S(R_{un})=Q. \tag{B.4f}$$

Collecting terms and passing to the limit ($\Delta t\to 0$), eqs. (B.1)–(B.3) become

$$V(x)=\frac{x}{\rho}+\frac{\mu}{\rho}\int_{R_{en}}^{\infty}(S(n')-V(x))\,\mathrm{d}G(n')+\frac{\lambda_e}{\rho}\int_{R_{ee}}^{\infty}(V(x')-V(x))\,\mathrm{d}F(x), \tag{B.5}$$

$$S(n)=\frac{n}{\rho+\mu}+\frac{\mu}{\rho+\mu}\int_{R_{nu}}^{\infty}(S(n')-Q))\,\mathrm{d}G(n')$$

$$+\frac{\lambda_n}{\rho+\mu}\int_{R_{ne}}^{\infty}(V(x')-S(n))\,\mathrm{d}F(x')+\frac{Q}{\rho+\mu}, \tag{B.6}$$

$$Q=\frac{-c}{\rho}+\frac{\lambda_u}{\rho}\int_{R_{ue}}^{\infty}(V(x')-Q))\,\mathrm{d}F(x')+\frac{\mu}{\rho}\int_{R_{un}}^{\infty}(S(n')-Q))\,\mathrm{d}G(n'). \tag{B.7}$$

This three-equation system can be solved for the functionals $V(x)$, $S(n)$ and $Q$ in terms of $c$, $\mu$, $\lambda_u$, $\lambda_e$, $\lambda_n$ and the parameters of $\mathrm{d}G(n)$ and $\mathrm{d}F(x)$. Using the reservation wage functions defined above, these three equations produce restrictions across the parameters of the transition matrix governing transitions among the three states. [For more details, see Coleman and Heckman (1981).]

The densities of duration times in each labor force state have a simple representation in terms of the reservation wages and arrival rates of offers. The econometric model produced by the theory is a competing risks model of the sort widely used in the duration analysis literature [see Kalbfleisch and Prentice (1980)]. We establish the analogy by briefly reviewing that model using terminology from the biostatistics literature.

A patient can die from many causes that are assumed to operate independently. Suppose that there are two diseases, each of which is fatal. The time to death from disease one is $t_1$, if it is the only disease afflicting the patient, while the time to death from disease two is $t_2$, if disease two is the only one afflicting the patient. In the presence of both diseases, the time to death is $\min(t_1, t_2)$. $t_1$ and $t_2$ are clearly only hypothetical random variables each of which would characterize time to death if the other disease is not at work.

Suppose that each disease considered in isolation produces an exponential time to death. Thus

$$f(t_1) = h_1 \mathrm{e}^{-h_1 t_1}, \qquad f(t_2) = h_2 \mathrm{e}^{-h_2 t_2}.$$

Direct calculation reveals that the time to death $[t = \min(t_1, t_2)]$ has the density

$$f(t) = (h_1 + h_2)\mathrm{e}^{-(h_1+h_2)t}.$$

The probability that the patient dies from disease one is $P_1$ defined as

$$P_1 = h_1/(h_1 + h_2) = \int_0^\infty \int_{t_1}^\infty h_1 \mathrm{e}^{-h_1 t_1} h_2 \mathrm{e}^{-h_2 t_2} \mathrm{d}t_2 \mathrm{d}t_1.$$

The competing risks model can be applied to the three-state model of labor force dynamics. For example, an employed individual is subject to two risks: (a) getting a higher wage on another job and (b) getting a non-market draw that causes him to terminate his current job. Whichever risk occurs first causes him to exit his current job. Define $\tilde{t}_{ee}$ as the *hypothetical* duration of employment in the current job if the only possible exit is to another job. $\tilde{t}_{en}$ is the hypothetical duration of employment if the only possible exit is to the non-market state. Each random variable has an exponential distribution as a

consequence of the time stationary Markovian process produced by the theory. The hazard rate for the density of $\tilde{t}_{ee}$ is

$$h_{ee}=\lambda_e(1-F(x)), \tag{B.8}$$

where $x$ is the current wage in the employed state. The hazard for the density of $\tilde{t}_{en}$ is

$$h_{en}=\mu(1-G(R_{en}(x))).^{39} \tag{B.9}$$

By the preceding argument, the density of the observed employment spell $t_e$ is the density of $\min(\tilde{t}_{ee}, \tilde{t}_{en})$,

$$f(t_e)=(h_{ee}+h_{en})\exp-(h_{ee}+h_{en})t_e. \tag{B.10}$$

Using methods exposited in Flinn and Heckman (1982) or Kalbfleisch and Prentice (1980), it is possible to use duration data to consistently estimate $h_{ee}$ and $h_{en}$ for each $x$. The probability that an individual leaves his current job to take another job, $P_{ee}$, is

$$P_{ee}=h_{ee}/(h_{ee}+h_{en}). \tag{B.11}$$

For exponential market and non-market wages with parameters $\theta_1$ and $\theta_2$ this model is a logit in reservation wages and the current wage $x$,

$$P_{ee}=1/(1+\mathrm{e}^{a_e+\theta_1 x-\theta_2 R_{en}(x)}),$$

where $a_e=\ln(\mu/\lambda_e)$. Note that direct estimation of transition probabilities enables us to estimate $\theta_1$ [provided we solve for $R_{en}(x)$ from eqs. (B.4)–(B.7)].

The density of accepted wage offers for individuals who quit one job with wage $x$ to get another with wage $x'$ is

$$f(x')/(1-F(x)), \qquad x'\geqq x.$$

Given knowledge of $x$ and $x'$, and an assumed functional form for $F$, it is possible to estimate consistently the parameters of $F$ as the number of observed $e$ to $e$ transitions becomes large. Consistent estimates of this density are *not* possible without imposing a recoverability condition. Thus the model is non-parametrically underidentified but parametrically identified. Since it is

[39]The analogy to competing risks model is somewhat strained because the reservation wage rules for exit to one state would change if the other state were eliminated from this model. We ignore this point in constructing our analogy.

possible to use duration data to consistently estimate $h_{ee}$, assuming recoverability is satisfied, $\lambda_e$ can be consistently estimated [see (B.8)].

A parallel argument may be developed for transitions from other states. Thus the duration of time spent in an unemployment spell has density

$$f(t_u) = (h_{ue} + h_{un}) \exp - (h_{ue} + h_{un}) t_u, \tag{B.12}$$

where

$$h_{ue} = \lambda_u (1 - F(R_{ue})), \tag{B.13}$$

$$h_{un} = \mu (1 - G(R_{un})). \tag{B.14}$$

From duration by destination data it is possible to estimate consistently $h_{ue}$ and $h_{un}$. [See Flinn and Heckman (1982).]

The probabilities of exiting to $e$ and $n$ from $u$ are

$$P_{ue} = h_{ue}/(h_{un} + h_{ue}), \qquad P_{un} = 1 - P_{ue}.$$

For exponential market and non-market wages,

$$P_{ue} = 1/(1 + \mathrm{e}^{a_u + \theta_1 R_{ue} - \theta_2 R_{un}}),$$

where

$$a_u = \ln (\mu/\lambda_n).$$

The density of accepted wage offers for individuals who exit from unemployment to employment is

$$f(x)/(1 - F(R_{ue})), \qquad x \geqq R_{ue}.$$

From a sample of accepted wages for unemployed individuals it is possible to use the minimum accepted wage to estimate consistently $R_{ue}$. Hence it is possible to estimate consistently $\lambda_u$ [from eq. (B.13)], since it is possible to use the duration data to estimate consistently $h_{ue}$.

The density of time spent in a non-market spell is

$$f(t_n) = (h_{ne} + h_{nu}) \exp - (h_{ne} + h_{nu}) t_n, \tag{B.15}$$

where

$$h_{ne} = \lambda_n (1 - F(R_{ne}(n))), \tag{B.16}$$

$$h_{nu} = \mu G(R_{nu}). \tag{B.17}$$

As before, it is possible to use the duration by destination data to consistently estimate $h_{ne}$ and $h_{nu}$. To simplify the exposition we assume that $n$ is known.[40] The probabilities of exit to $e$ and $u$ from $n$ are

$$P_{ne} = h_{ne}/(h_{ne} + h_{nu}), \qquad P_{nu} = 1 - P_{ne}.$$

For exponential market and non-market wages,

$$P_{ne} = 1/(1 + e^{a_n + \theta_1 R_{ne} - \theta_2 R_{un}}),$$

where

$$a_n = \ln(\mu/\lambda_n).$$

From data on accepted wages in the $n$ to $e$ transitions it is possible to consistently estimate the reservation wage $R_{ne}$ as the number of such transitions becomes large. Hence [from (B.16)] it is possible to estimate consistently $\lambda_n$, since it is possible to use the duration data to estimate consistently $h_{ne}$. Using (B.14) and (B.17) one can estimate $\mu$ (simply sum the two hazards). If $n$ is observed, $R_{ne}$ can be consistently estimated by the smallest $n$ observed in the non-market state. Assuming recoverability, $G(n)$ can be consistently estimated.

Thus under the assumptions made, it is possible to consistently estimate $\lambda_e$, $\lambda_n$, $\lambda_u$, $F(x)$, $\mu$, $G(n)$, $R_{ue}$, $R_{en}$, $R_{nu} = R_{un}$, $R_{ne}$ and $R_{ue}$. The asymptotic distribution theory for these estimators is produced by an analysis similar to that given in section 2. For more detail, see Coleman and Heckman (1981).

The distinguishing feature of this model is that the theory imposes *cross-transition* restrictions. These restrictions have not been noted or imposed in previous empirical work on labor force turnover [see e.g. Tuma and Robbins (1980)]. These restrictions aid in model identification and provide tests of the theory. We have already noted that the model predicts growth as a function of experience. By parameterizing non-market distributions we are able to produce a rich class of turnover models which enable us to introduce economic analysis from household economics into turnover models. Children, non-market income, health status and the like determine the value of non-market time.

Tests of hypothesis $\lambda_u > \lambda_e$ and $\lambda_u > \lambda_n$ provide indirect evidence on the empirical validity of search models. If, for example, $\lambda_u < \lambda_n$, considerable doubt would be cast on the standard search model. This hypothesis can be tested assuming that a recoverability condition is satisfied. For further analysis of this model and more general models, see Coleman and Heckman (1981).

[40]This is an enormous simplification. Some data sets ask respondents questions which can be used to infer $n$. An analysis of the more general case in which $n$ is unknown to the econometrician is presented in Coleman and Heckman (1981).

## Appendix C

*The Kiefer–Neumann procedure*

This appendix discusses the relationship between the analysis in section 2 of the text and the Kiefer–Neumann (1979, 1981) analysis of search unemployment. We demonstrate that the Kiefer–Neumann analysis is critically dependent on the assumption that agents receive one offer per period (however the period is defined). We also note that even under their assumptions, their estimator is less efficient than the one presented in the text because they overlook a critical piece of information which we exploit. We further demonstrate that their inefficient estimation approach does not generalize to a continuous time setting and so cannot be used as a substitute for the analysis of the more general model presented in the text.

The Kiefer and Neumann approach is a direct application of statistical results developed in the sample selection bias literature [see Heckman (1976b, 1979) and Heckman and MaCurdy (1981)] and in the literature on female labor supply [see Smith (1979)]. $F(x)$ is a log normal distribution so $\ln X$ is normally distributed with mean $z\beta$ and variance $\sigma^2$. Wage offers arrive one per period. Following Lippman and McCall (1976a) the reservation wage $rV$ is defined as the solution to

$$c+rV=(1/r)\int_{rV}^{\infty}(x-rV)\mathrm{d}F(x) \quad \text{for} \quad rV\geqq 0.^{41} \tag{C.1}$$

We may solve for $R^*=rV$ (for $r$ assumed known),

$$R^*=R^*(c,z\beta,r). \tag{C.2}$$

Properties of $R^*$ have already been established in the text (fix $\lambda=1$ in the analysis of section 2.1). Given their log normality assumption, it is convenient to work with

$$R=\ln R^*=R(c,z\beta,r). \tag{C.3}$$

Conditional on $z$, the probability that an unemployment spell terminates at period $j$ is (see section 2.2)

$$P(j)=(F(R^*))^{j-1}(1-F(R^*)). \tag{C.4}$$

Invoking log normality for $X$,

[41] Kiefer and Neumann ignore this inequality restriction in deriving their estimates.

$$P(j)=\left[\Phi\left(\frac{R(c,z\beta,r)-z\beta}{\sigma}\right)\right]^{j-1}\Phi\left(\frac{z\beta-R(c,z\beta,r)}{\sigma}\right), \tag{C.5}$$

where $\Phi$ is the unit normal cdf. The density of accepted wages is

$$f(x\mid X>R)=f(x)/(1-F(R)). \tag{C.6}$$

For the log normal case,

$$f(\ln x\mid \ln X>R)$$
$$=\left((2\pi\sigma^2)^{-\frac{1}{2}}\exp-\frac{1}{2}\frac{(\ln x-z\beta)^2}{\sigma^2}\right)\Big/\Phi\left(\frac{z\beta-R(c,z\beta,r)}{\sigma}\right). \tag{C.7}$$

Using standard sample selection bias formulae [e.g. Heckman (1976)],

$$E(\ln X\mid \ln X>R)=z\beta+\sigma\lambda\left(\frac{z\beta-R(c,z\beta,r)}{\sigma}\right), \tag{C.8}$$

where

$$\lambda(n)=((2\pi)^{-\frac{1}{2}}\exp-\tfrac{1}{2}n^2)/\Phi(-n). \tag{C.9}$$

The joint density of date to termination and accepted wages is

$$P(j)f(x\mid X>R)=[F(R^*)]^{j-1}f(x), \tag{C.10}$$

$$x\geqq R^*. \tag{C.11}$$

Exactly as in the analysis of sections 2.3–2.5, the likelihood function formed from the joint density is monotonic in $c$ and constraint (C.11) is always satisfied as an equality for at least one observation.

*Kiefer and Neumann do not maximize the likelihood for their model.* Instead, they use an indirect two-stage approach patterned exactly after previous models estimated in the female labor supply literature. They first use data on unemployment duration in a likelihood formed from (C.5) to estimate $(R(c,z\beta,r)-z\beta)/\sigma$, as a reduced form index. (This stage corresponds to estimating an equation determining whether or not a woman works.) Using this estimated index, they then follow previous econometric work in female labor supply and estimate $\lambda$ [see eqs. (C.9) and (C.10)]. From a regression of accepted wages on $z$ and estimated $\lambda$ it is possible [under conditions set forth in Heckman (1979)] to consistently estimate $\beta$ and $\sigma$.

From the estimated $\beta$ and $\sigma$ it is possible to solve from the estimated index $(R(c,z\beta,r)-z\beta)/\sigma$ for each observation, to estimate $R(c,z\beta,r)$ for each

observation. Using a nonlinear procedure, it is possible to consistently estimate $c$ using the restriction implied in (C.1). Kiefer and Neumann (1981) use an iterated variant of this procedure.

The estimates derived from this procedure are not efficient because they do not exploit all the information in the sample (i.e., they are not derived from the likelihood function).[42] Moreover, the two-stage procedure is critically dependent on the assumption that agents receive one offer per 'period'. In that case, the per-period probability of accepting a job $(1-F(R))$ is also the denominator of the conditional density of wage offers [see (C.6)]. From the duration probit, it is possible to estimate the selection index, $(z\beta - R(c, z\beta, r))/\sigma$, and so it is possible to correct for truncation bias in the estimated accepted wage offer distribution. In the more general case in which the probability of receiving an offer is less than one, the per-period probability of accepting a job is *not* the denominator of (C.6). [For a continuous time example, see eqs. (2.4) and (2.5) in the text; the point applies to either a continuous time or discrete time model.] Thus in the more general case, if the rate of arrival of job offers is not known, it is not possible to estimate the selectivity variable [$\lambda$ in eq. (C.8)] from the duration data alone; thus the two-stage procedure outlined above breaks down. Structural estimation requires utilization of the inequality restrictions exploited in section 2.

Another point not noted by Kiefer and Neumann is that under the assumption that agents receive one wage offer per period, it is *always* possible to recover the untruncated wage offer distribution (above the truncation point) from the truncated distribution precisely because the denominator of the truncated distribution is known (see the discussion in section 2.3 and footnote 11). Thus, provided that the reservation wage is estimated in the manner we suggest in the text, the untruncated distribution of wage offers is always recoverable from the sample truncated distribution if agents receive one wage offer per period. Thus it is not necessary to impose a functional form for accepted wages onto the data. Under their assumptions all the structural parameters of the model are non-parametrically identified.

Adding unobserved heterogeneity to the model in the fashion of Heckman and Willis (1977) or Kiefer and Neumann (1981), does not alter our conclusions in this appendix in any fundamental way. For the sake of brevity, we refrain from extending our analysis to include this case.

[42]One offsetting advantage is that their estimators are less sensitive to measurement error because they do not use extreme value observations while the estimators in the text are based on extreme observations.

## References


Aigner, D. and S. Chu, 1968, On estimating the industry production function, American Economic Review 58, Sept., 826–839.

Amemiya, T., 1973, Regression analysis when the dependent variable is truncated normal, Econometrica 41, no. 5, 997–1016.

Billingsley, P., 1968, Convergence of probability measures (Wiley, New York).

Bishop, Y., S. Feinberg and P. Holland, 1975, Discrete multivariate analysis (MIT Press, Cambridge, MA).

Blumen, I., M. Kogan and P. McCarthy, 1955, The industrial mobility of labor as a probability process, in: Cornell studies of industrial and labor relations, Vol. 6 (Cornell University Press, Ithaca, NY).

Burdett, K., 1978, Employee search and quits, American Economic Review 68, April, 212–220.

Burdett, K. and D. Mortensen, 1978, Labor supply under uncertainty, in: R. Ehrenberg, ed., Research in labor economics, Vol. 2 (JAI Press, London, CT) 109–157.

Burdett, K. and D. Mortensen, 1980, Search, layoff and labor market equilibrium, Journal of Political Economy 88, no. 4, Aug., 652–672.

Coleman, T. and J. Heckman, 1981, Dynamic models of labor supply, Unpublished manuscript (University of Chicago, Chicago, IL).

Cox, D.R. and D.V. Hinkley, 1974, Theoretical statistics (Chapman and Hall, London).

DeGroot, M., 1970, Optimal statistical decisions (McGraw-Hill, New York).

Diamond, P., 1971, A model of price adjustment, Journal of Economic Theory 3, June, 156–168.

Diamond, P. and E. Maskin, 1979, An equilibrium analysis of search and breach of contract, I: Steady states, Bell Journal of Economics 10, no. 1, Spring, 282–318.

Durbin, J., 1973, Distribution theory for tests based on the sample distribution function, in: SIAM regional conference series in applied mathematics, No. 9 (Arrowsmith, Bristol).

Flinn, C. and J. Heckman, 1982, Models for the analysis of labor force dynamics, in: G. Rhodes and R. Basmenn, eds., Advances in econometrics (JAI Press, London, CT).

Forsund, F., C. Lovell and P. Schmidt, 1980, A survey of frontier production functions and their relationship to efficiency measurement, Journal of Econometrics 13, 5–25.

Galambos, J., 1978, The asymptotic theory of extreme order statistics (Wiley, New York).

Greene, W., 1980, Maximum likelihood estimation of econometric frontier functions, Journal of Econometrics 13, no. 1, May, 27–56.

Heckman, J., 1976a, Sample selection bias as a specification error: Report to the Rand Corporation (Rand, Santa Monica, CA).

Heckman, J., 1976b, The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models, Annals of Economic and Social Measurement 5, no. 4, Fall, 475–492.

Heckman, J., 1979, Sample selection bias as a specification error, Econometrica 47, no. 1, Jan., 153–162.

Heckman, J., 1981, Heterogeneity and state dependence, in: S. Rosen, ed., Conference on labor markets (University of Chicago Press, Chicago, IL).

Heckman, J. and T. MaCurdy, 1980, A life cycle model of female labor supply, Review of Economic Studies 47, no. 146, Jan., 47–74.

Heckman, J. and T. MaCurdy, 1981, Recent theoretical studies of labor supply, Part II, in: R. Ehrenberg, ed., Research in labor economics, Vol. 5 (JAI Press, London, CT).

Heckman, J. and B. Singer, 1981, A method for minimizing the impact of distributional assumptions in econometric models for duration data, Econometrica, forthcoming.

Heckman, J. and B. Singer, 1982, The identification problem in econometric models for duration data, in: W. Hildenbrand, ed., Advances in econometrics, Proceedings of world meetings of the Econometric Society, 1980 (Cambridge University Press, Cambridge).

Heckman, J. and R. Willis, 1977, A beta logistic model for analysis of sequential labor force participation by married women, Journal of Political Economy 85, 27–58.

Jennrich, R., 1969, Asymptotic properties of nonlinear least squares estimators, Annals of Mathematical Statistics 40, 633–643.

Jovanovic, B., 1979, Firm specific capital and turnover, Journal of Political Economy 87, no. 6, Dec., 1246–1260.

Kalbfleish, J. and R. Prentice, 1980, The statistical analysis of failure time data (Wiley, New York).
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, no. 1, Feb., 89–108.
Kiefer, N. and G. Neumann, 1981, Individual effects in a nonlinear model, Econometrica 49, no. 4, July, 965–980.
Lewis, H.G., 1974, Comments on selectivity biases in wage comparisons, Journal of Political Economy 82, no. 6, Nov., 1145–1156.
Lippman, S. and J. McCall, 1976a, The economics of job search: A survey, Part I, Economic Inquiry 14, June, 155–189.
Lippman, S. and J. McCall, 1976b, Job search in a dynamic economy, Journal of Economic Theory 12, no. 3, June, 365–390.
Lippman, S. and J. McCall, 1981, The economics of belated information, International Economic Review 22, no. 1, Feb., 135–146.
Lucas, R.E. and E. Prescott, 1974, Equilibrium search and unemployment, Journal of Economic Theory 7, Feb., 188–209.
Marston, S., 1976, Employment instability and high unemployment rates, Brookings Papers on Economic Activity 7, 169–203.
Mincer, J., 1974, Schooling, experience and earnings (Columbia University, New York).
Mortensen, D., 1977, Unemployment insurance and job search decisions, Industrial and Labor Relations Review 30, no. 4, 505–517.
Nickell, S., 1979, Estimating the probability of leaving unemployment, Econometrica 47, no. 5, Sept., 1249–1266.
Pearson, K., 1903, Mathematical contributions to the theory of evolution — XI, Royal Society Philosphical Transactions A 200, 1–66.
Robbins, H., 1970, Optimal stopping, American Mathematical Monthly 77, 333–343.
Ross, S.M., 1970, Applied probability models with optimization applications (Holden-Day, San Francisco, CA).
Rothschild, M., 1973, Models of market organization with imperfect information: A survey, Journal of Political Economy 81, no. 6, 1283–1308.
Silcock, H., 1954, The phenomenon of labor turnover, Journal of the Royal Statistical Society A 117, 429–440.
Smith, J., 1980, Female labor supply theory and estimation (Princeton, NJ).
Takayama, A., 1974, Mathematical economics (Dryden Press, Hinsdale, IL).
Telser, L., 1980, Economic theory and the core (University of Chicago Press, Chicago, IL).
Toikka, R.S., 1976, A Markovian model of labor market decisions by workers, American Economic Review 66, Dec., 821–834.
Tuma, N.B. and P.K. Robins, 1980, A dynamic model of employment behavior: An application to the Seattle and Denver income maintenance experiments, Econometrica 48, no. 4, May, 1031–1052.
Wilson, C., 1980, Equilibrium wage distributions, Unpublished manuscript (University of Wisconsin, Madison, WI).