

# Empirical Search Models

José J. Canals and Steven Stern  
University of Virginia.

June 18, 2001

## 1 Introduction

This chapter discusses various methods that have been used to estimate structural models of search and survival models. Most of the focus is on using available data, usually on unemployment spell lengths and accepted wage offers, to estimate the parameters of specific search models. In particular, we focus on estimating the parameters of the wage offer distribution, the reservation wage or reservation wage function, the cost of search, offer arrival rate, and discount rate. There is an added section on survival models because survival models have been used so extensively to look at unemployment spell data.

Throughout we describe actual models, estimation procedures, and issues associated with identification and estimation. There is not as much of an emphasis on results because we view the literature as still developing methodology. The frontier of the methodology for structural models involves estimating equilibrium search models and allowing for heterogeneity and interesting dynamics. We also think identification issues should receive more attention.

This chapter is divided up into sections. The next section presents the classical search model and how to estimate its parameters. We begin by presenting early "semi-structural" attempts to estimate the model. Next, there is a discussion on estimating a basic structural search model. We build on the basic model by allowing for measurement error in wages, observed and unobserved heterogeneity, right censoring, continuous time, and observed rejected offers. Next, there is a discussion of models with changing reservation wages.

The third section deals with survival analysis. We construct the basic Cox proportional hazards model with unobserved heterogeneity and discuss how to estimate the effects of observed covariates, the parameters of the baseline hazard function and the parameters of the unobserved heterogeneity density function. Then we discuss semiparametric approaches associated with estimating the baseline hazard and unobserved heterogeneity density function.

The next section discusses empirical equilibrium search models. A few theoretical models are presented. Then we discuss how to estimate their structural parameters and to deal with special problems associated with such models. Given present data, identification is an important issue. A conclusion ends the paper.

## 2 The Classical Search Model

In this section, we present a simple version of a search model in discrete time. Workers are infinitely lived, wealth maximizing agents and discount the future at a rate  $(1 + r)^{-1}$ . At each moment, a worker can be employed or unemployed. Unemployed workers search actively for job offers. Offers are distributed according to the density  $f(w)$  with finite mean. At each period, an unemployed worker incurs a cost of search  $c$  and receives a wage offer from one of many firms in the labor market. A wage offer represents a random draw from the time-invariant distribution of wages known to the worker. Once a job is accepted, employment lasts forever. The problem of the worker has the reservation property: there exists a value  $\xi$ , called the reservation wage, such that it is optimal for the worker to accept any wage offer if it is not smaller than  $\xi$  and to reject it otherwise. Analytically, the reservation wage is the unique solution to

$$\xi = \frac{1}{r} \int_{\xi}^{\infty} (w - \xi) f(w) dw - c. \quad (2.1)$$

The left-hand side of the equation (2.1) represents the benefit associated with working at the present period for a wage  $\xi$ , while the right-hand side represents the expected discounted benefit of one more period of search if wages below  $\xi$  are not acceptable to the worker. Consequently, at the reservation wage, the worker is indifferent between accepting the present wage offer or searching one more period. Notice that the left-hand side of equation (2.1) is strictly increasing in  $\xi$ , while the right-hand side is decreasing in  $\xi$ . Consequently, as long as the cost of search is not too high, there exists a unique positive solution to the last equation. A more rigorous discussion of this model is given by, for example, Lippman and McCall (1976).

### 2.1 First Attempts

The first attempts to estimate the search model (Kiefer and Neumann 1979) assumed a wage offer equation

$$w_i^o = X_i \beta + \varepsilon_i^o \quad (2.2)$$

and a reservation wage equation

$$w_i^r = Z_i \gamma + \varepsilon_i^r \quad (2.3)$$

with  $X_i$  and  $Z_i$  representing individual characteristics that affect the expected wage offer and the reservation wage, respectively,  $\varepsilon_i^o \sim N(0, \sigma_o^2)$ ,  $\varepsilon_i^r \sim N(0, \sigma_r^2)$ , and  $E\varepsilon_i^o \varepsilon_i^r = \sigma_{or}$ . Imagine we have data from a sample of workers with information about individual characteristics, the duration of unemployment, and the accepted wage if the worker is employed. Clearly, the distribution of observed wages is truncated because we observe only wages that have been accepted by

the workers. Define  $P_i$  as the probability that an individual receives an acceptable wage offer at a certain period. Then

$$\begin{aligned} P_i &= \Pr(w_i^o \geq w_i^r) \\ &= \Pr(X_i\beta + \varepsilon_i^o \geq Z_i\gamma + \varepsilon_i^r) \\ &= \Pr(X_i\beta - Z_i\gamma \geq \varepsilon_i^r - \varepsilon_i^o), \end{aligned} \tag{2.4}$$

and the probability that an individual searches exactly  $T$  periods is

$$(1 - P_i)^{T-1} P_i. \tag{2.5}$$

$P_i$  can be estimated using data on duration of unemployment for a sample of workers and maximum likelihood estimation (MLE). In particular, the log likelihood function (LLF) is

$$L = \sum_{i=1}^N \log(1 - P_i)^{T_i-1} P_i. \tag{2.6}$$

Also, we observe accepted wage data. After correcting the wage equation for selection bias as in Heckman (1979), we can estimate

$$\begin{aligned} w_i^o &= X_i\beta + E(\varepsilon_i^o \mid X_i\beta - Z_i\gamma \geq \varepsilon_i^r - \varepsilon_i^o) + \varsigma_i \\ &= X_i\beta + \frac{\phi \left[ (X_i\beta - Z_i\gamma) / \sqrt{\sigma_r^2 + \sigma_o^2 - 2\sigma_{or}} \right]}{1 - \Phi \left[ (X_i\beta - Z_i\gamma) / \sqrt{\sigma_r^2 + \sigma_o^2 - 2\sigma_{or}} \right]} + \varsigma_i. \end{aligned} \tag{2.7}$$

where  $\phi$  is the standard normal density function,  $\Phi$  is the standard normal distribution function,  $\phi / (1 - \Phi) = E(\varepsilon_i^o \mid X_i\beta - Z_i\gamma \geq \varepsilon_i^r - \varepsilon_i^o)$  is the inverse Mill's ratio, and  $\varsigma_i$  is the deviation of  $\varepsilon_i^o$  from its conditional expected value. Equations (2.6) and (2.7) identify  $\beta$ ,  $\gamma$ ,  $\sigma_o^2$ ,  $\sigma_r^2$ , and  $\sigma_{or}$  either by placing restrictions on what variables  $X_i$  enter the wage equation and what variables  $Z_i$  enter the reservation wage equation or by relying on the nonlinear functional form of the selection bias correction term.

The Kiefer-Neumann approach to estimation of the search model uses the restrictions imposed by equation (2.1) only as a guide for the econometric specification. This presents problems of consistency with the underlying theoretical framework. In particular, the theoretical model predicts that a mean preserving spread transformation of the wage offer distribution will increase the reservation wage, that a decrease in the discount rate will imply a reduction in the reservation wage, and that any anticipated change in future wage offer distributions will affect the reservation wage in the present as well as the future. It is not clear how these restrictions apply in equations (2.2) and (2.3).

## 2.2 Structural Estimation Without Heterogeneity

More recently, researchers such as Wolpin (1987), Stern (1989), and Christiansen and Kiefer (1991) have pursued a more structural approach in which the estimation is derived explicitly from the theory. This approach incorporates the

restrictions given by the condition of optimal search to the empirical work. The econometric specification is then fully consistent with the underlying theoretical framework. In addition, it becomes possible to study parameters such as the discount rate and the cost of search not otherwise present in the econometric specification. In the next paragraphs we analyze the structural estimation of the theoretical search model previously described.

Assume, for simplicity of exposition, that we have data on duration of unemployment and accepted wage offers from a population of homogeneous workers (i.e., there are no observed or unobserved characteristics affecting the wage offer distribution or the reservation wage). Let the distribution of wage offers be  $F(\bullet)$  with density  $f(\bullet)$  and let the reservation wage that solves equation (2.1) be  $\xi$ . Then the probability of observing an individual with an spell of unemployment of  $t_i$  periods and accepted wage  $w_i$  is

$$F(\xi)^{t_i} [1 - F(\xi)] \frac{f(w_i)}{1 - F(\xi)} 1(w_i \geq \xi) \quad (2.8)$$

where  $1(w_i \geq \xi)$  is equal to one if  $w_i \geq \xi$ , and it equals zero otherwise. The first term,  $F(\xi)^{t_i} [1 - F(\xi)]$ , is the probability that the first  $t_i$  wage offers are rejected. The second term,  $f(w_i) / [1 - F(\xi)]$ , is the density of a wage conditional on it being acceptable ( $w_i \geq \xi$ ). The last term,  $1(w_i \geq \xi)$ , ensures that all observed accepted wage offers are greater than the reservation wage. The LLF based on a sample of  $N$  individuals is

$$L = \sum_{i=1}^N \log \left\{ F(\xi)^{t_i} f(w_i) 1(w_i \geq \xi) \right\}. \quad (2.9)$$

For this simple specification of the search model, we need to estimate  $\xi$ ,  $r$ ,  $c$ , and the parameters associated with the distribution of wage offers  $F$ . The LLF is increasing in  $\xi$  up until it reaches the minimum observed wage. This implies that the MLE of  $\xi$  is

$$\hat{\xi} = \min_{i=1, \dots, N} w_i \quad (2.10)$$

Recall that  $\hat{\xi}$  is the first order statistic for the sample  $\{w_i\}_{i=1}^N$  of observed wage offers, and its density is equal to  $N [1 - F(w_i | w \geq \xi)]^{N-1} f(w_i | w \geq \xi)$ . Observe that the lower limit of the accepted wage offer distribution depends on the reservation wage which is an estimated parameter of the model. For this reason, the standard regularity conditions used to demonstrate the consistency and asymptotic normality of the maximum likelihood estimator do not apply to this problem. However, the parameter estimates are still consistent. In fact, under weak regularity conditions,<sup>1</sup>  $\hat{\xi}$  is a superconsistent estimator of  $\xi$ ;

<sup>1</sup>If the wage offer distribution is twice continuously differentiable, then,  $f(\xi) > 0$  is a sufficient condition.

that is  $\hat{\xi}$  converges to  $\xi$  at a rate faster than  $\sqrt{N}$ . The estimation approach in this case consists of substituting  $\hat{\xi}$  in the LLF and obtaining estimators of the remaining parameters using standard MLE applied to the resulting function. This technique, described in Christensen and Kiefer (1991), is known as pseudo-maximum likelihood estimation.

Using this approach we can obtain estimators for the reservation wage and the parameters of  $F(\bullet)$  for a large class of distributions.<sup>2</sup> Neither the cost of search  $c$  nor the discount rate  $r$  appear in the likelihood function. As a result, in order to obtain consistent estimators for these parameters, we need to make use of the restrictions implied by equation (2.1). In general, this restriction is not sufficient to independently identify both parameters. To illustrate this point, consider a case where the distribution of wages is exponential with parameter  $\lambda$ :  $F(w) = 1 - \exp\{-\lambda w\}$  for  $w \geq 0$ . For this particular specification, the reservation wage satisfies

$$\xi = \frac{1}{r} \frac{e^{-\lambda \xi}}{\lambda} - c. \quad (2.11)$$

Observe that, for fixed  $\xi$  and  $\lambda$ , there is an infinite number of pairs  $(r, c)$  that satisfy the last equation. Then, even if we can obtain consistent estimators  $\hat{\xi}$  and  $\hat{\lambda}$ , the restriction implied by equation (2.11) is not enough to independently identify  $r$  and  $c$ .

The central hypothesis of the structural search model is that agents make optimal choices in a simple environment. From an empirical perspective, the condition of optimal behavior imposes additional restrictions to the econometric model. A test of the search theory will imply a test of these restrictions. Unfortunately, we do not observe fundamental elements of the theory of search like the reservation wage, the cost of search, and the discount rate (to say nothing of parameters measuring heterogeneity) from the data. With this lack of information about fundamental elements of the theoretical model, it becomes difficult to construct meaningful tests of the model. Stern (1989) suggests in a model where  $r$  and  $c$  are not separately identified that one can construct the set of combinations of  $r$  and  $c$  that are consistent with the search behavior equation and then determine whether any elements of the set are "reasonable" values. This provides an informal test of the theory. Nevertheless, the difficulty in testing search models is probably the reason why search theory has been used as a restriction at hand to assist with the identification of the structural models more than as a testable theory.

---

<sup>2</sup>In order to be able to identify the distribution of wages from the distribution of accepted wage offers certain regularity conditions are needed. We will elaborate on this topic in the section about rejected offers. Flinn and Heckman(1982) present an extensive discussion of this issue.

### 2.3 Measurement Error

A major criticism of the estimation technique described in equations (2.9) and (2.10) is its sensitivity to measurement error. Observe that the estimator of the reservation wage is the first order statistic from the sample of accepted wage offers. Order statistics are known to be very sensitive to measurement error, even in large samples. All the parameters in the structural model are related to each other, via the restrictions imposed by the theory. Consequently, if measurement error is affecting the estimate of  $\xi$ , any other parameter estimated will be affected as well. Several authors have dealt with this problem by explicitly incorporating the possibility of measurement error in the estimation of the structural model. Among the papers that have followed this approach are Wolpin (1987), Stern (1989), and Christensen and Kiefer (1994). All three papers assume that the observed wages are affected by random error in measurement that is normally distributed. Stern argues that there may be two different types of measurement error in the wage-offer data. First, there may be errors in reporting wages, or in imputing tax rates or price levels. Second, the value of a job may deviate from the observed wage because of other factors such as fringe benefits and working conditions. Stern and Christensen and Kiefer present strong evidence that measurement error is an important empirical event.

Assume wages are measured with error where  $w_i$  and  $w_i^*$  represent the real and the observed wage for individual  $i$ , respectively. In particular, suppose that  $w_i^* = w_i \varepsilon_i$  where  $\varepsilon_i$  represents a random multiplicative error with distribution  $G(\bullet)$ , takes a positive value, and is independent of the wage offer. In this case, a wage offer observed without error is equivalent to a situation in which  $\varepsilon_i = 1$ . Notice that, for this specification, it is possible to observe accepted wage offers below the reservation wage without violating the underlying theoretical model.

Observe first of all that the existence of measurement error does not violate any assumptions of the theoretical model. Consequently, the structural approach to inference can rely on the restrictions implied by the conditions of optimal search. All of the papers mentioned previously have used MLE to estimate related models. Assume, as in the previous case, that we have data on duration of unemployment and accepted wage offers from a sample of homogeneous workers. The probability that a particular unemployed worker remains unemployed for one more period is still  $F(\xi)$ . On the other hand, in order to obtain the density of an observed accepted wage offer  $w_i^*$ , it is important to realize first that  $w_i = w_i^* / \varepsilon_i \geq \xi$  because the wage has been accepted. This imposes restrictions on the range of possible realizations of  $\varepsilon_i$ ,  $\varepsilon_i \in (0, w_i^* / \xi]$ . Thus, for each  $w_i^*$ , the conditional distribution of errors has the following expression,

$$G(\varepsilon | w_i^*) = \frac{G(\varepsilon)}{G(w_i^* / \xi)} \text{ for } \varepsilon_i \in (0, w_i^* / \xi]$$

and zero otherwise. Also, observe that

$$\Pr(w_i^* | \varepsilon) = f\left(\frac{w_i^*}{\varepsilon} | w_i \geq \xi\right) = f\left(\frac{w_i^*}{\varepsilon}\right) / [1 - F(\xi)].$$

Therefore, the density of observed wages is obtained after integrating the previous expression with respect to  $\varepsilon$  that is unobservable. That is

$$f^*(w_i^* | w_i \geq \xi) = \frac{\int_0^{w_i^*/\xi} f\left(\frac{w_i^*}{\varepsilon}\right) dG(\varepsilon)}{(1 - F(\xi)) G(w_i^*/\xi)}. \quad (2.12)$$

Consider a sample with  $N$  observations of the form  $(t_i, w_i^*)$  where  $t_i$  represents the length of the unemployment spell and  $w_i^*$  represents the accepted wage offer, possibly observed with measurement error, for observation  $i$ . The LLF for this sample is

$$L = \sum_{i=1}^N \log \left[ \frac{F(\xi)^{t_i}}{G(w_i^*/\xi)} \int_0^{w_i^*/\xi} f\left(\frac{w_i^*}{\varepsilon}\right) dG(\varepsilon) \right]. \quad (2.13)$$

Notice that, due to measurement error, the observed wage  $w_i^*$  is not restricted to be above the reservation wage and can take any positive value. For this reason, the first order statistic from the sample of observed wages is not an estimator of the reservation wage. Consequently, the problem of estimation in this case is reduced to standard MLE. Assuming that the distribution of wages  $F$  belongs to a certain parametric family, in most cases we will be able to obtain consistent estimators for the reservation wage  $\xi$  and the parameters associated with  $F$ . However, neither  $r$  nor  $c$  appear in the likelihood function; estimates of these parameters will have to be obtained from the restrictions imposed by the theory. At this point, the problem is not different from the previous case. As a result, these parameters cannot be independently identified. Stern (1989) presents an extensive discussion of identification in a similar context.

## 2.4 Sample Heterogeneity

Up to this point, we have considered the case of a sample of homogeneous individuals. In fact, workers with different characteristics may face different wage offer distributions and thus have different reservation wages. Also, in more complex search models in which workers can choose the search intensity, workers with different characteristics may have different costs of search and, as a result, they can optimally choose to search with different levels of intensity as is the case in Stern (1989). We divide our discussion into the effect of observed heterogeneity and the effect of unobserved heterogeneity on estimation.

### 2.4.1 Observed Heterogeneity

Within our framework it is straightforward to account for heterogeneity. Consider that  $X_i$  represents a set of characteristics associated with individual  $i$ . Assume also that these characteristics are observed by the econometrician and

affect the distribution of wage offers in a known way,  $F(\bullet | X_i)$ . In this framework, it is simple to define the LLF as an extension of equation (2.13), accounting for observed heterogeneity,

$$L = \sum_{i=1}^N \log \left\{ \frac{F(\xi_i | X_i)^{t_i}}{G(w_i^*/\xi_i)} \int_0^{w_i^*/\xi_i} f\left(\frac{w_i^*}{\varepsilon} | X_i\right) dG(\varepsilon) \right\}. \quad (2.14)$$

Where  $\xi_i$  represents the reservation wage for a worker with associated characteristics  $X_i$ . Wolpin (1987) deals with observed heterogeneity by following the spirit of equation (2.14). In order to reduce the computational burden, he transforms  $X_i$  variables, usually considered to be continuous, into dichotomous variables. Some papers divide the sample into different subsamples that contain similar workers. For example, Narendranathan and Nickell (1985) divide the sample into groups of younger and older workers, and Stern (1989) divides the sample by sex and education. This approach allows these papers to solve equation (2.14) for only a small number of individuals each representing a large group of similar individuals. This approach implies strong assumptions that are only justified for reasons of computational complexity. While the cost of computation was a reasonable concern at the time these papers were written, recent advances in computation allow for a much less restrictive treatment of observed heterogeneity.

#### 2.4.2 Unobserved Heterogeneity

Another important issue is the existence of omitted variables or unobserved heterogeneity. Even after controlling for the presence of observed characteristics  $X_i$  for individual  $i$ , there may be relevant characteristics that cannot be observed from the sample, also known as unobserved heterogeneity. Unobserved characteristics are important in explaining the behavior of the worker in the labor market. For example, workers with identical observed characteristics may face different wage offer distributions and have different reservation wages due to differences in relevant unobserved characteristics. Ignoring the problem of unobserved heterogeneity may result in biased estimators of the parameters of the model. Biased estimates can be obtained even if the omitted variables are uncorrelated with the observed ones because of the nonlinearity of the model and the restrictions imposed by the theory that incorporate implicit relations among the parameters. This problem was noticed by Nickell (1979) but has not received much attention in more recent work in this area.

One way to deal with unobserved heterogeneity is to make an explicit assumption about the distribution of the unobserved heterogeneity.<sup>3</sup> In particular, assume that agent  $i$  faces a distribution of wage offers  $F(\bullet | X_i, \nu_i)$ , with  $X_i$  and  $\nu_i$  representing the set of observed and unobserved characteristics, respectively. Even if the function  $F(\bullet | X, \nu)$  is known, this information is not enough

---

<sup>3</sup>Ignoring unobserved heterogeneity is equivalent to assuming that the distribution is degenerate. This is a much stronger assumption than needs to be made.



to construct the LLF because the value of  $\nu$  for each individual is not observed. Assume that  $\nu$  is distributed according to  $G_\nu(\bullet)$  and is independent of  $X$ . Then, the contribution to the LLF of an observation  $(t_i, w_i^*, X_i)$  after the effect of unobserved heterogeneity has been integrated out is

$$\log \int \frac{F(\xi_{i\nu} | X_i, \nu)^{t_i}}{G(w_i^*/\xi_{i\nu})} \int_0^{w_i^*/\xi_{i\nu}} f\left(\frac{w_i^*}{\varepsilon} | X_i, \nu\right) dG_\varepsilon(\varepsilon) dG_\nu(\nu) \quad (2.15)$$

where  $G_\varepsilon(\varepsilon)$  represents the distribution of the measurement error, and  $\xi_{i\nu}$  represents the reservation wage for an individual with characteristics  $(X_i, \nu)$ . This expression can be used to construct the LLF of the sample. As was the case with measurement error, the presence of unobserved heterogeneity does not violate any assumptions of the theoretical model. Consequently, the structural approach to inference can rely on the restrictions implied by the conditions of optimal search. A significant problem with this approach, at least for continuous unobserved heterogeneity distributions, is that the agent conditions on his own unobserved heterogeneity when choosing a reservation wage. Thus, the reservation wage becomes a functional. Essentially, a reservation wage must be computed for each value of  $\nu$  used to compute the integral in equation (2.15). Thus, computation costs increase significantly.

The problem of unobserved heterogeneity has been ignored in most structural models of search. The approach implied by equation (2.15) has been followed in some of the recent papers that consider estimation of structural search models like Eckstein and Wolpin (1990), Van den Berg and Ridder (1993) or Engberg (1994).<sup>4</sup> Since the treatment of this issue here will mimic the approach taken in a later section on equilibrium models, we postpone the detailed description of this technique. However, it is appropriate at this point to notice that, in most cases, it is impossible to obtain information about  $G_\nu(\bullet)$ . In particular, the search theory does not offer any guidance about  $G_\nu(\bullet)$ . It is probably for this reason that most papers that take this problem into consideration, including Nickell (1979), Eckstein and Wolpin (1990), Van den Berg and Ridder (1993), or Engberg (1994), have adopted a nonparametric approach similar to Heckman and Singer (1984b) which is discussed in Section 3.

## 2.5 Censoring

A right censored observation is one where the end of the spell has not occurred by the end of the time of observation. For example, consider a sample of newly unemployed people observed between dates  $\tau_1$  and  $\tau_2$ . Any unemployment spells that have not ended by  $\tau_2$  are right censored. For those spells, we do not know when, if ever, they would have ended; we know only that the spell had not ended by  $\tau_2$ . An alternative cause of censoring would occur in unemployment insurance administrative records where the individual is observed

---

<sup>4</sup>Eckstein and Wolpin and Engberg evaluate the reservation wage for each value of unobserved heterogeneity. Van den Berg and Ridder have closed form solutions to equation (2.15) that preclude evaluating the reservation wage conditional on  $\nu$ .

until his UI benefits run out. When they run out, the researcher does not know whether the unemployment spell ended in a job or if just the UI coverage ended.<sup>5</sup> The methods for estimation of structural search models described in the previous section can be adapted to handle right censored observations. Consider data of the form  $(t_i, c_i, w_i, X_i)$ , where  $t_i$  is the observed length of the spell,  $c_i$  is an indicator of whether spell  $i$  is right censored ( $c_i = 1$  iff spell  $i$  is right censored),  $w_i$  is the observed accepted wage offer for (uncensored) spell  $i$ , and  $X_i$  is a set of observed characteristics of worker  $i$ . As long as the censoring process is independent from the process generating the data, the LLF can be modified straightforwardly to account for right censoring. Consider, for example, the structural model with measurement error. The probability of observing  $(t_i, c_i = 1, X_i)$  is  $F(\xi_i | X_i)^{t_i}$  (note that no wage offer is observed), while the probability of observing  $(t_i, c_i = 0, w_i, X_i)$  is similar to the term in equation (2.14). This information can be incorporated to the LLF as

$$L = \sum_{i=1}^N (1 - c_i) \log \left\{ \frac{F(\xi_i | X_i)^{t_i}}{G(w_i^*/\xi_i)} \int_0^{w_i^*/\xi_i} f\left(\frac{w_i^*}{\varepsilon_i} | X_i\right) dG(\varepsilon) \right\} + c_i \log \left\{ F(\xi_i | X_i)^{t_i} \right\}. \quad (2.16)$$

Equation (2.16) handles right censoring, but many data sets also have left censored observations. In our context, left censoring appears, for example, if we do not observe the starting time of unemployment spell. Several authors have tried to avoid this problem by selecting data sets with special characteristics. An early treatment of the left censoring problem can be found in Nickell (1979). Lancaster (1990) provides a good discussion as well.

In principle, left censoring should not be a problem for the simple search model described in this section because this model is stationary and the probability of leaving the unemployment state is independent of the time spent in this state. However, there are examples in the literature of models in which this is not the case. Kiefer and Neumann (1979, 1981) consider a specification that accounts for the possibility of decreasing reservation wages. Wolpin (1987) presents a model in which agents search for a certain period of time after which they accept the first wage offered. This feature of the model implies decreasing reservation wages. In addition, if we take into account the possibility of unobserved heterogeneity (Nickell 1979), left censoring becomes an important problem, even in a stationary environment. Since unobserved characteristics of the agents affect the probability of escape from the unemployment state, a sample with problems of left censoring cannot be assumed random. Consequently, left censoring can produce inconsistent estimators even in stationary models.

---

<sup>5</sup>This problem could be mitigated by linking UI administrative data with wage data. However, this is not done typically.

## 2.6 Continuous Time Models With Offer Arrival Rates

In the discrete time search model described in equation (2.1), the theory does not provide any guidance for the choice of the time period length. In this sense, this approach is somewhat arbitrary since it is time unit dependent. In order to overcome this shortcoming some researchers have developed continuous-time versions of the previous model. In this section we present a simple version of a search model in continuous time. The only difference between this model and the discrete time version is in the process of wage offer arrivals. Other than that, this model parallels the structure of the search model in discrete time.

Workers are infinite-lived wealth maximizing agents, and discount the future at a rate  $r$ . In order to receive wage offers at a certain time interval  $\Delta t$ , the worker incurs a cost  $c\Delta t$ . Job offers arrive from a Poisson process with parameter  $\alpha$ . The probability of receiving a wage offer in time interval  $\Delta t$  is  $\alpha\Delta t + o(\Delta t)$ ; the probability of receiving two or more offers in time interval  $\Delta t$  is negligible. Offers are distributed according to the time-invariant density  $f(w)$  with finite mean. Once a job is accepted, employment lasts forever. The worker's search problem has the reservation wage property. Analytically, the reservation wage  $\xi$  is the unique solution to

$$\xi = \frac{\alpha}{r} \int_{\xi}^{\infty} wf(w) dw - c. \quad (2.17)$$

As in the discrete time model, this implicit equation represents the restriction implied by the condition of optimal search.

In order to obtain estimators for the parameters of the structural model in continuous time, it is convenient to start by defining the hazard function associated with the unemployment state. This function represents the instantaneous probability of receiving an acceptable wage offer or equivalently the instantaneous probability of leaving the unemployment state conditional on not having left yet. The hazard in this case is the product of the arrival rate  $\alpha$  and the probability that the wage offer is acceptable to the worker  $1 - F(\xi)$ . The hazard function for leaving unemployment is

$$\lambda = \alpha [1 - F(\xi)]. \quad (2.18)$$

Observe that the hazard function is time independent which is a result of the stationarity of the model. Given the hazard function  $\lambda$ , the probability of an unemployment spell of length  $t$  or less is

$$1 - \exp\{-\lambda t\}. \quad (2.19)$$

This is an exponential distribution with parameter  $\lambda$ , where  $\lambda^{-1}$  is the average length of an unemployment spell. The associated density function will have the

form  $\lambda \exp\{-\lambda t\}$ . Also, the probability of observing an accepted wage offer  $w$  (with no measurement error) is

$$f(w) [1 - F(\xi)]^{-1} 1(w \geq \xi) . \quad (2.20)$$

Consider a sample with  $N$  observations of the form  $(t_i, w_i)$ , where  $t_i$  represents the length of unemployment spell  $i$  and  $w_i$  represents the accepted wage offer. The LLF for this sample is

$$L = \sum_{i=1}^N \log \left\{ \lambda \exp\{-\lambda t_i\} f(w_i) [1 - F(\xi)]^{-1} \right\} . \quad (2.21)$$

As in the discrete case, the estimators have to be obtained using Pseudo Maximum Likelihood techniques because no measurement error is assumed.

This expression can be generalized to allow for right censoring and measurement error. Consider a sample of the form  $\{t_i, c_i, w_i^*, X_i\}_{i=1}^N$  and measurement error as in the discrete case. In this case, the probability of observing  $(t_i, c_i = 1, X_i)$  is  $1 - (1 - \exp\{-\lambda(X_i) t_i\}) = \exp\{-\lambda(X_i) t_i\}$  with  $\lambda(X_i) = \alpha [1 - F(\xi_i | X_i)]$  and  $\xi_i$  incorporates the restriction imposed by equation (2.17) given the distribution of wage offers  $F(\bullet | X_i)$ . Meanwhile, the probability of  $(t_i, c_i = 0, w_i^*, X_i)$  is

$$\lambda(X_i) \exp\{-\lambda(X_i) t\} \frac{\int_0^{w_i^*/\xi_i} f\left(\frac{w^*}{\varepsilon} | X_i\right) dG(\varepsilon)}{(1 - F(\xi_i | X_i)) G(w_i^*/\xi_i)} \quad (2.22)$$

This information can be incorporated to the LLF to obtain an expression similar to equation (2.14). As in the discrete case, estimation is standard MLE.

## 2.7 Rejected Offers

With information about unemployment spells and accepted wage offers, it is possible to identify nonparametrically the distribution of accepted wage offers and the reservation wage. The distribution of observed wage offers is equal to the truncated distribution of offers with truncation at point  $\xi$ . Since we are primarily interested in the underlying distribution of wages, it is important to know if the distribution of wages can be recovered from the observed distribution of accepted wage offers; this property is known as the recoverability condition. Flinn and Heckman (1982) present a comprehensive discussion of this condition. For example, they show that the Normal and Lognormal distributions, commonly used to describe wage distributions, satisfy the recoverability condition. An example of distribution that does not satisfy this condition is the Pareto distribution.

Flinn and Heckman argue that "most econometric models for the analysis of truncated data are non-parametrically underidentified, and some are parametrically underidentified as well." Obviously, it is possible to choose a parametric

family for the distribution of wages such that the recoverability condition is satisfied. A simple example will suffice to explain why this choice is, to a certain extent, arbitrary. Let  $f(w)$  be a density of wages, and define

$$f_\pi(w) = (1 - \pi)f(w | w \geq \xi)1(w \geq \xi) + g(w)1(w < \xi) \quad (2.23)$$

for any  $\pi \in (0, 1)$ , and for any density  $g(w)$  satisfying  $G(\xi) = \pi$ . Observe that  $F_\pi(w)$  is a valid distribution function. In addition  $F_\pi(w | w \geq \xi) = F(w | w \geq \xi)$ , and  $F_\pi(\xi) = \pi$ . This example shows that we cannot identify nonparametrically the underlying distribution of wages from the observed distribution of accepted wage offers in that there is a continuum of other wage offer distributions,  $F_\pi(w)$  for all  $\pi \in (0, 1)$ , with the same accepted wage offer distribution. In other words, given the available data we are unable to distinguish between a labor market with a large availability of job offers and a high probability of rejection from a labor market with a small availability of job offers and a high probability of acceptance. In particular, because we do not know  $\alpha$ , the unemployment duration data identifies only  $\lambda = \alpha[1 - F(\xi)]$  but not  $F(\xi)$ .<sup>6</sup> Some authors have overcome this identification problem by assuming that the distribution of wages belongs to a specified parametric family satisfying the recoverability property. Wolpin (1987) for example, assumes that the distribution of wages is either normal or lognormal. Other authors have fixed the number of per period offers received by the unemployed worker; once the number of offers per period is fixed,  $\pi$  is identified.

Other authors have considered data sets containing more complete information about the search behavior of the unemployed workers. Lancaster and Chesher (1983) and Van den Berg (1992) for example consider a sample containing a subjective measure of the reservation wage. Jensen et al. (1987) consider a set of data that contains information about the number of applications and number of offers received. In most cases, subjective information or information that requires a perfect recall of the past by the agents is critiqued for being not reliable.

## 2.8 Structural Dynamic Models of Search

The search models presented so far have assumed a stationary environment. In such cases, we have observed that the behavior of the workers in the labor market can be described by a constant stopping rule. The stationary search model has serious limitations. For example, Meyer (1990) finds evidence for duration dependence in unemployment spells even after controlling for unobserved heterogeneity (see Section 3); this can not be consistent with a stationary model.

Not many models of search in a nonstationary environment have been estimated. The model we describe, Wolpin (1987), will suffice to illustrate the additional level of complexity added by eliminating the assumption of stationarity from the search model. Wolpin(1987) presents a simple model in discrete

---

<sup>6</sup>Note that even if we assume  $\alpha = 1$ , there is too much freedom in choice of  $G(\bullet)$  to nonparametrically identify  $F(\bullet)$ .

time. The two differences of this model relative to a standard stationary model are that a) an unemployed worker is allowed to search for a finite number of periods  $T$ , after which, if the worker remains unemployed, he is assumed to accept the first job available and b) offer probabilities change over time; we ignore the second source of nonstationarity. Let  $V_t$  denote the value of search in period  $t$  conditional on entering period  $t$  without a job:

$$V_t = E \max \left[ w_t, -c + (1+r)^{-1} V_{t+1} \right] \text{ if } t = 1, \dots, T-1 \quad (2.24)$$

and  $V_T = w_T$ . From equation (2.24), we observe that an unemployed worker at time  $t < T$  will accept a job offer if it is at least  $\xi_t = \left\{ -c + (1+r)^{-1} V_{t+1} \right\}$ , the reservation wage, and the employed worker at time  $T$  will accept any wage offer. Intuitively, the reservation wage after  $t-1$  periods of search  $\xi_t$  is the expected profit obtained from search the next period. If  $T$  is finite, the sequence  $\{\xi_t\}_{t=1}^T$  is decreasing. Observe that

$$\xi_{T-1} = -c + (1+r)^{-1} E[w] \quad (2.25)$$

and  $E[w | w \geq \xi_{T-1}] \geq E[w | w \geq \xi_T] = E[w]$ . This implies that

$$\xi_{T-2} = -c + (1+r)^{-1} E[w | w \geq \xi_{T-1}] \geq \xi_{T-1}. \quad (2.26)$$

In general, assuming  $\xi_{T-i} \geq \xi_{T-i+1}$  for  $i = 2, \dots, k-1$ , it follows that

$$E[w | w \geq \xi_{T-k+1}] \geq E[w | w \geq \xi_{T-k+2}] \quad (2.27)$$

which implies that

$$\xi_{T-k} = -c + (1+r)^{-1} E[w | w \geq \xi_{T-k+1}] \quad (2.28)$$

$$\geq -c + (1+r)^{-1} E[w | w \geq \xi_{T-k+2}] = \xi_{T-k+1}.$$

This proves by induction that the sequence  $\{\xi_t\}_{t=1}^T$  is decreasing. In this model, the reservation wage changes through time in a systematic and recursive manner described by equations (2.25) through (2.28).

In order to estimate the model, Wolpin assumes that wage data is observed with measurement error. Assume  $\ln w_t^* = \ln w_t + \nu_t$  where  $w_t^*$  represents the observed wage,  $w_t$  represents the true wage, and  $\nu_t$  represents measurement error. In order to be consistent with our previous notation, we can consider that  $\nu_t = \ln \varepsilon_t$ . By assumption, the distribution of wage offers for an individual  $i$ ,

$F(\bullet | X_i)$  is lognormal,<sup>7</sup> and its median is a function of specific observed characteristics of individual  $i$ ,  $X_i$ ; in addition, the distribution of the measurement error  $G(\bullet)$  is normal and independent of the distribution of wages. In this case, the contribution to the LLF of an observation  $(t_i, w_i^*, X_i)$  from the sample is

$$\sum_{\tau=1}^{t_i-1} \log \frac{F(\xi_{i\tau} | X_i)}{G(\xi_{i\tau})} + \log \int_0^{w_i^*/\xi_{it_i}} f\left(\frac{w_i^*}{\varepsilon} | X_i\right) dG(\varepsilon). \quad (2.29)$$

For each guess of the parameters in equation (2.29) and for each observation, the sequence of reservation wages  $\{\xi_t\}_{t=1}^T$  is obtained from equations (2.25) and (2.28). The value  $T$  is also chosen to optimize the value of the LLF of the sample.

Miller (1984) and Pakes (1986) are both dynamic stopping rule models. Miller uses a special structure to solve an occupational choice problem, and Pakes uses a special case with simulation methods to simulate when firms reject the option to renew a patent. While not the same economic problem as discussed in this chapter, they have a very similar generic structure to dynamic search models.

### 3 Survival Analysis.

The analysis of duration data, also known as survival analysis, was developed to describe the timing of events such as when a person dies (thus the name) or when a machine breaks. These statistical techniques have become a subject of increasing interest in economics, especially labor economics. Numerous empirical papers have addressed such issues as unemployment duration (Lancaster 1979 and Nickell 1979), the effects of unemployment benefits on the spells of unemployment (Moffit 1985, Solon 1985, Meyer 1990, and McCall 1996), turnover (Burdett et al. 1985), occupational matching (McCall 1990), retirement (Diamond and Hausman 1984), strike length (Kennan 1985 and Gunderson and Melino 1990), and job search (Jovanovic 1984). Sometimes this type of technique is regarded as a reduced form for behavioral economic theories like the theory of job search or the theory of job matching. More appropriately, it is a flexible approximation of behavior or an informative method of describing the data.

Consider a random variable  $T$  which takes positive values and describes the length of time until an event of interest occurs. Assume that the distribution of duration  $T$  can be specified by a distribution function  $F(t)$ , with associated density function  $f(t)$ . Other functions of interest associated with the duration process are the survivor function  $S(t) = 1 - F(t)$  which represents the probability that a spell will last a time period  $t$  or longer, and the hazard function

---

<sup>7</sup>Observe that, in principle, the assumption that agents accept any offer after  $T$  implies that the distribution of wage offers can be identified nonparametrically by observing the accepted wage offers and spell lengths after this time.

$\lambda(t)$  which represents the instantaneous probability of the spell ending at  $t$  conditional on it not having ended prior to  $t$ . The hazard function can be written as

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T \leq t + \Delta t \mid T \geq t)}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{\Delta F(t)}{\Delta t} S(t)^{-1} = \frac{f(t)}{S(t)}. \quad (3.1)$$

From the previous definition, it is evident that, given a certain density, we can determine the associated hazard function; as we will show later on, the opposite is also true. The derivative of the hazard function  $\partial\lambda(t)/\partial t$  is called the duration dependence. If  $\partial\lambda(t)/\partial t > 0$ , there is positive duration dependence, and if  $\partial\lambda(t)/\partial t < 0$ , there is negative duration dependence. Given the hazard rate for a certain duration process, it is easy to determine the associated survival function. Observe that,  $\lambda(t) = -d \ln S(t)/dt$ . Then, solving a simple differential equation, we obtain

$$\begin{aligned} S(t) &= \exp \left\{ - \int_0^t \lambda(z) dz \right\} = \exp \{ -\Lambda(t) \}, \\ f(t) &= \lambda(t)S(t) = \lambda(t) \exp \{ -\Lambda(t) \} \end{aligned} \quad (3.2)$$

where  $\Lambda(t) = \int_0^t \lambda(z) dz$  is the integrated hazard function.

Up to this point we have considered a simple duration process that does not depend on additional covariates. From an economic perspective, the main concern is usually to study the impact of key exogenous variables on the distribution of  $T$ . Consider a data set of the form  $\{(t_i, X_i)\}_{i=1}^N$  from a population of individuals where  $t_i$  represents the spell of time until an event of interest occurs for observation  $i$  and  $X_i$  represent a vector of characteristics for  $i$ . The distribution of duration for agent  $i$  can be specified as  $F(t \mid X_i)$  with associated density function  $f(t \mid X_i)$ . Similarly, we can define the survivor function  $S(t \mid X_i) = 1 - F(t \mid X_i)$  and the hazard function  $\lambda(t \mid X_i)$ . Furthermore, in most data sets we should account for the possibility of right censoring. As we have already mentioned in the previous section, this does not represent a problem as long as the censoring process is independent of the data generating process. Consider a sample of the form  $\{t_i, c_i, X_i\}_{i=1}^N$ , where  $c_i$  is an indicator of whether spell  $i$  is censored ( $c_i = 1$  iff spell  $i$  is censored). In this case, the probability of an observation  $(t_i, c_i, X_i)$  will be  $f(t_i \mid X_i) = \lambda(t_i \mid X_i) S(t_i \mid X_i)$  for  $c_i = 0$  and  $S(t_i \mid X_i)$  for  $c_i = 1$ . Thus, the log likelihood function is

$$\begin{aligned} LLF &= \sum_{i=1}^N (1 - c_i) \log f(t_i \mid X_i) + c_i \log S(t_i \mid X_i) \\ &= \sum_{i=1}^N (1 - c_i) \log \lambda(t_i \mid X_i) - \Lambda(t_i \mid X_i). \end{aligned} \quad (3.3)$$

In empirical applications, it is common to start with the specification of the hazard function. Equations (3.2) and (3.3) show how to use  $\lambda(t \mid X)$  to specify the log likelihood function. Consider the following hazard function:



$$\lambda(t | X) = \lambda_0(t) \lambda_1(X, \beta) \quad (3.4)$$

where  $\lambda_0(t)$  is called the baseline hazard. If there is enough variation in the  $X$  variables and all of them are observable,  $\lambda_0(\bullet)$  and  $\lambda_1(\bullet)$  are identified up to a constant. This specification is known as the proportional hazard model. It includes most of the parametric models considered in empirical applications such as Lancaster (1979), Solon (1985), and Narendranathan, Nickell, and Stern (1985). In this case, the function  $\Lambda(\bullet)$  can be written as

$$\Lambda(t | X) = \lambda_1(X, \beta) \int_0^t \lambda_0(s) ds. \quad (3.5)$$

Usually  $\lambda_1(X, \beta)$  is modelled as  $\exp\{X\beta\}$ .<sup>8</sup> In this case we obtain

$$\lambda(t | X) = \lambda_0(t) \exp\{X\beta\}. \quad (3.6)$$

If  $\lambda_0(t) = 1$ , equation (3.6) becomes the hazard function associated with an exponential distribution with parameter  $\exp\{X\beta\}$ ;  $F(t) = 1 - \exp\{-\lambda_1 t\}$ . If  $\lambda_0(t) = \alpha t^{\alpha-1}$ , equation (3.6) becomes the hazard of a Weibull distribution;  $F(t) = 1 - \exp\{-\lambda_1 t^\alpha\}$ . The proportional hazard model does not arise from any theoretical economic model. Its popularity is probably due to the fact that the estimated parameters provide a straightforward interpretation: the estimated coefficients are the derivative of the log hazard with respect to the associated  $X$  variable.

Up to this point, we have assumed that all the variables that matter for the duration process are included in  $X$  and are observed by the econometrician. Even in early work (Lancaster 1979 and Lancaster and Nickell 1980), researchers were aware that ignoring the possibility of omitted variables in a duration model can heavily bias the included parameter estimates and lead to misleading conclusions. To illustrate this point, suppose that the proportional hazard model is the correct specification for the study of unemployment spells. In addition, assume that the econometrician does not observe all the variables affecting the duration of unemployment; i.e.,  $\lambda(t | X) = \exp\{X\beta + \varepsilon\}$  where  $\varepsilon$  measures the effect of omitted variables. For the sake of concreteness, assume the  $\varepsilon$  takes on two values,  $\varepsilon^+ > \varepsilon^-$ , and that  $\Pr[\varepsilon = \varepsilon^+] = \Pr[\varepsilon = \varepsilon^-] = 1/2$  at time zero. As the process evolves, people with high values of  $\varepsilon$  ( $= \varepsilon^+$ ) will leave the sample faster than people with low values  $\varepsilon$  ( $= \varepsilon^-$ ). This will change the relative proportions of people (with respect to  $\varepsilon$ ) toward people with low values of  $\varepsilon$ . As time goes by, since the average value of  $\varepsilon$  is falling, the average hazard rate will fall. Thus, we will observe decreasing hazards, even after controlling for observed heterogeneity (the  $X$ 's) even though the hazard for each particular worker is constant over time. Consequently, ignoring the problem of

---

<sup>8</sup>In particular, this specification is equivalent to the one introduced by Cox (1972).

omitted variables will lead to negative duration dependence bias. Because of the nonlinearity of the model, it also leads to inconsistent estimates of all of the model's parameters (Flinn and Heckman 1982).

In order to be able to identify the effects of unobservables and observables, it is necessary to make explicit assumptions about the way in which they interact.<sup>9</sup> Assume then that the true specification of the hazard, for a certain agent  $i$  in the sample, is as follows:

$$\lambda(t | X_i, \nu_i) = \lambda(t | X_i) \nu_i = \lambda_0(t) \lambda_1(X_i, \beta) \nu_i \quad (3.7)$$

where  $X_i$  represents the set of observed variables, and  $\nu_i = \exp\{\varepsilon_i\}$  summarizes the effect of any other variable that affects the duration process and is not observed (unobserved heterogeneity). Then the survival function is

$$\begin{aligned} S(t | X_i, \nu_i) &= \exp\left\{-\int_0^t \lambda(z | X_i, \nu_i) dz\right\} = \exp\{-\Lambda(t | X_i, \nu_i)\} \\ &= \exp\{-\Lambda(t | X_i) \nu_i\}, \\ f(t_i | X_i, \nu_i) &= \lambda(t | X_i, \nu_i) S(t | X_i, \nu_i). \end{aligned} \quad (3.8)$$

In this case, the LLF for the sample  $\{t_i, c_i, X_i, \nu_i\}_{i=1}^N$  is

$$\sum_{i=1}^N (1 - c_i) \log f(t_i | X_i, \nu_i) + c_i \log S(t_i | X_i, \nu_i). \quad (3.9)$$

In principle, this expression cannot be used to estimate the model because the values of  $\{\nu_i\}_{i=1}^N$  are not observed. This problem can be overcome if the distribution generating the unobserved heterogeneity is known. Then we can define a LLF based on the marginal probabilities once the unobserved heterogeneity has been integrated out. More precisely, assume that  $\nu_i$  represents a particular realization from a distribution  $G(\bullet)$  independent of  $X_i$ . It is usual in this case to choose the normalization  $E(\nu) = 1$ .<sup>10</sup> We can define

$$S(t | X) = \int S(t | X, \nu) dG(\nu) = \int \exp\{-\Lambda(t | X) \nu\} dG(\nu) = M_\nu[-\Lambda(t | X)] \quad (3.10)$$

and

$$\begin{aligned} f(t | X) &= \frac{-\partial S(t | X)}{\partial t} \\ &= \lambda(t | X) \int \nu \exp\{-\Lambda(t | X) \nu\} dG(\nu) = \lambda(t | X) M_\nu^{(1)}[-\Lambda(t | X)] \end{aligned} \quad (3.11)$$

<sup>9</sup>See Heckman (1991) for an illuminating review of this issue.

<sup>10</sup>Observe that this is not a restriction because a constant is included in  $X_i\beta$ , and any  $E(\nu) \neq 1$  would be captured in the constant.

where  $M_\nu[\bullet]$  and  $M_\nu^{(1)}[\bullet]$  represent the moment generating function of  $\nu$  and its derivative, respectively.<sup>11</sup> We can now specify the LLF for the sample  $\{t_i, c_i, X_i\}_{i=1}^N$  using the integrated, or marginal probabilities:

$$LLF = \sum_{i=1}^N (1 - c_i) \log f(t_i | X_i) + c_i \log S(t_i | X_i) \quad (3.12)$$

$$= \sum_{i=1}^N (1 - c_i) \left\{ \log \lambda(t | X) + \log M_\nu^{(1)}[-\Lambda(t | X)] \right\} + c_i \log M_\nu[-\Lambda(t | X)].$$

Lancaster (1979) lets

$$\lambda(t | X_i, \nu_i) = \alpha t^{\alpha-1} \mu_i \nu_i \quad (3.13)$$

where  $\mu_i = \exp\{X_i \beta\}$  and the unobserved heterogeneity has a density  $g(\nu) \propto \nu^{\sigma-1} \exp\{-\nu\sigma\}$ . In this case,  $M_\nu[z] = [1 - (z/\sigma)]^{-\sigma}$  and we obtain:

$$S(t | X) = M_\nu[-\Lambda(t | X)] = \left( 1 + \frac{\mu I(t)}{\sigma} \right)^{-\sigma}, \quad (3.14)$$

where  $\mu_i = \exp\{X_i \beta\}$  and  $I(t) = \int_0^t \alpha t^{\alpha-1} dt$ . From this expression, it is straightforward to obtain  $f(t | X)$  and to construct the LLF of the sample. Using MLE, we can obtain consistent estimators of the parameters associated with the duration process  $(\alpha, \beta, \sigma)$ .

It is important to recognize the significance of controlling for unobserved heterogeneity. The true hazard for an individual with unobserved heterogeneity  $\nu$  is

$$\lambda(t | X, \nu) = \alpha t^{\alpha-1} \mu \nu. \quad (3.15)$$

By controlling for unobserved heterogeneity, we consistently estimate the expected hazard

$$\lambda(t | X) = E_\nu \{ \lambda(t | X, \nu) \} = \alpha t^{\alpha-1} \mu. \quad (3.16)$$

On the other hand, if we control just for the observed characteristics  $X$ , the survival function of the duration process is given in equation (3.14) with associated hazard

---

<sup>11</sup>For some distributions  $G$ , the moment generating function does not exist. More generally  $M_\nu[-\Lambda(t | X)]$  is the Laplace transform of  $G$ .

$$\lambda^*(t | X) = \frac{\partial \log S(t | X)}{\partial t} = \alpha t^{\alpha-1} \mu S(t | X)^{1/\sigma} = \lambda(t | X) S(t | X)^{1/\sigma}. \quad (3.17)$$

If we do not control for unobserved heterogeneity, we obtain a consistent estimator for  $\lambda^*(\bullet)$  in equation (3.17). As  $t$  increases, the survival probability goes to zero and the hazard  $\lambda^*(t | X)$  decreases over time with respect to the true hazard  $\lambda(t | X)$ ; this confirms our previous intuition. This argument can be generalized easily by observing that

$$\lambda^*(t | X) = E(\lambda(t | X, \nu) | T > t) = \lambda(t | X) E(\nu | T > t) \quad (3.18)$$

where  $T$  represents the time of employment and  $E(\nu | T > t)$  represents the expected value of the unobserved heterogeneity among the remaining unemployed workers at time  $t$ . Under the assumption that  $\nu$  is independent of the observed heterogeneity, this expectation should be decreasing over time.

We have learned that it is important to take into account the possibility of unobserved heterogeneity. This is especially true in many applications of duration analysis to labor economics in which the concern has been to study the effect of unemployment benefits or the behavior of the unemployed workers over the spell of unemployment. In this sense, we have shown that by ignoring the existence of unobserved heterogeneity in the sample, we are potentially estimating more negative duration dependence than actually exists.

In the previous example, given data on spell duration and observed characteristics of the workers, it was possible to estimate the distribution of unobserved heterogeneity, the baseline hazard, and the effect of observed characteristics on the probability of leaving the unemployment state. In order to achieve that goal, it was necessary to make strong parametric assumptions about the form of the hazard function and the distribution of unobserved heterogeneity. In most cases, we may have little prior information about the correct distribution of unobserved heterogeneity, and we therefore may produce misleading results by misspecifying this distribution. It is then pertinent to ask to what extent the available data can provide a nonparametric identification of each one of the relevant functions separately. Lancaster and Nickell (1979) find that "it seems in practice very difficult to distinguish between the effects of heterogeneity and the effect of pure time variation in the hazard function." Elbers and Ridder (1982) show that, at least for the proportional hazard specification, it is possible to identify nonparametrically the distribution of unobserved heterogeneity independently of the other components of the duration process as long as there is enough variation in the observed characteristics. The intuition for this result is that changes in the duration variable and the covariates allow us to trace out the different components of the hazard. This result depends crucially on the form of the proportional hazard model, in particular, the separability of the hazard into one function of the duration, another of the covariates and the

independence of the unobserved characteristics respect to the observed ones.<sup>12</sup> Gurmu, Rilstone and Stern (1996) present a similar result for the proportional hazard model allowing for the possibility of interactions between  $t$  and some of the covariates. The identification result of Elbers and Ridder (1982) and Heckman and Singer (1984a) opens the possibility of nonparametric estimation of the distribution of the unobserved heterogeneity.

Authors like Lancaster (1979) and Heckman and Singer (1984b) have shown that ignoring unobserved heterogeneity can lead to biased estimates of the parameters of the hazard function. In addition, Heckman and Singer (1984b) have shown that different specifications for the distribution of the unobserved heterogeneity can lead to very different estimates. This finding led them to propose a flexible nonparametric method to control for unobserved heterogeneity. In this paper we pursue a heuristic description of this method. Readers interested in a more technical description should refer to the original paper. From equation (3.12), we obtain

$$LLF = \sum_{i=1}^N (1 - c_i) \left\{ \log \lambda(t | X) + \log M_\nu^{(1)}[-\Lambda(t | X)] \right\} + c_i \log M_\nu[-\Lambda(t | X)]. \quad (3.19)$$

where

$$\begin{aligned} M_\nu[-\Lambda(t | X)] &= \int \exp\{-\Lambda(t | X)\nu\} dG(\nu) \\ M_\nu^{(1)}[-\Lambda(t | X)] &= \int \nu \exp\{-\Lambda(t | X)\nu\} dG(\nu). \end{aligned} \quad (3.20)$$

The Heckman and Singer approach consists of approximating the unknown distribution of unobserved heterogeneity  $G(\nu)$  with a discrete distribution with positive probability mass at a finite number of points. Consider then the set of points  $\{\nu_1, \nu_2, \dots, \nu_k\}$  with probability mass  $\pi_j > 0$  associated to  $\nu_j$ ,  $j = 1, \dots, k$  and  $\sum_{j=1}^k \pi_j = 1$ . Substituting the discrete approximation of  $G(\nu)$  in equation (3.20), we obtain

$$\begin{aligned} \hat{M}_\nu[-\Lambda(t | X)] &= \sum_{j=1}^k \pi_j \exp\{-\Lambda(t | X)\nu_j\} \\ \hat{M}_\nu^{(1)}[-\Lambda(t | X)] &= \sum_{j=1}^k \pi_j \nu_j \exp\{-\Lambda(t | X)\nu_j\} \end{aligned} \quad (3.21)$$

where  $\hat{M}_\nu[\bullet]$  and  $\hat{M}_\nu^{(1)}[\bullet]$  represent the approximations to the true functions,  $M_\nu[\bullet]$  and  $M_\nu^{(1)}[\bullet]$  respectively. Substituting of the approximations,  $\hat{M}_\nu[\bullet]$

---

<sup>12</sup>A similar result in the context of risk models can be found in Heckman and Honore( 1989).

and  $\hat{M}_\nu^{(1)}[\bullet]$ , for the true functions,  $M_\nu[\bullet]$  and  $M_\nu^{(1)}(\bullet)$ , in equation (3.19), we obtain an approximation for the *LLF*,

$$LLF_k = \sum_{i=1}^N (1 - c_i) \left\{ \log \lambda(t | X) + \log \hat{M}_\nu^{(1)}[-\Lambda(t | X)] \right\} + c_i \log \hat{M}_\nu[-\Lambda(t | X)]. \quad (3.22)$$

The method proceeds by obtaining estimators of  $\{\nu_j, \pi_j\}_{j=1}^k$  and estimators of the parameters of the hazard function. These estimators will be the values that maximize equation (3.22). Observe that this problem is not a standard MLE problem because the number of parameters necessary to estimate the approximate distribution of unobserved heterogeneity,  $2k$ , can be in principle infinite and the asymptotics rely on  $k \rightarrow \infty$ . In practice, the estimation strategy consists of choosing the  $k$  for which  $LLF_k$  stops growing in  $k$ . Sometimes researchers use a criterion such as the Aikeke number; most of the time researchers increase  $k$  until the LLF stops growing (Card and Sullivan 1988, Gunderson and Melino 1990, and Gritz 1993) or fix  $k$  ahead of time (Heckman and Walker 1990, Behrmann, Sickles, and Taubman 1990, and Johnson and Ondrich 1990). One major criticism of this approach is the lack of asymptotic distribution theory for the parameters estimated. In practice, the results of the estimation will depend of the parameters of the unobserved heterogeneity.

Some authors, as for example Han and Hausman(1990) and Meyer (1990), have shown that a nonflexible specification of the baseline function  $\lambda_0(t)$  can bias the estimates of the other parameters. They suggest to estimate  $\lambda_0(t)$  semi-parametrically assuming that it can be represented as a step function. Furthermore, these authors have argued that the biases in the proportional hazard model may be larger for misspecification of the baseline hazard than for misspecification of heterogeneity distribution. For a similar model, Sueyoshi (1992) presents Monte Carlo evidence indicating that estimates are sensitive to misspecification of the unobserved heterogeneity distribution. In addition, this type of misspecification yields biased estimates of the baseline function  $\lambda_0(t)$ . In conclusion, the literature on this subject shows that both types of misspecification are important and should be taken into account when estimating a proportional hazard model.<sup>13</sup>

## 4 Empirical Equilibrium Search Models

The theoretical search models described in section two assume that unemployed workers search sequentially for a job. The worker accepts the first offer above his reservation wage and remains in that job forever. These models are formally inconsistent because each (homogeneous) firm acting optimally will offer a wage which depends upon only the distribution of reservation wages. If all firms face

---

<sup>13</sup>Recent papers that use this approach are Holt, Merwin, and Stern (1996) and McCall (1996).

the same distribution of reservation wages, each will offer the same wage, and there will be a degenerate distribution of wage offers. Also, there is no search because either the worker's reservation wage is above the common wage offer and the first wage offer received is always accepted or the worker's reservation wage is below the common wage offer and he does not search. These implications are at odds with the phenomena that these models try to explain, the search behavior of the agents as well as the variation in wages. This inconsistency was first pointed out by Diamond (1971).

In order to respond to Diamond's critique, several authors have developed theoretical models that depict wage (or price) dispersion and search as an equilibrium outcome (Reinganum 1979, Albrecht and Axell 1984, Burdett and Judd 1983, and Mortensen 1990). These equilibrium models are variations of the more simple models presented in Section 2. An important difference is the incorporation of the optimal behavior of firms into the models. Several options have been suggested in order to obtain a nondegenerate distribution of wages at equilibrium. For example, some models assume that workers may face several job offers simultaneously. In this case, workers with more than one job offer are able to bargain for a wage above the reservation wage (Burdett and Judd 1983). Others depict a framework with heterogeneous agents. For example, Albrecht and Axell (1984) assume that the firms in the market differ in their productivity and the agents, searching sequentially for a job, differ in their valuation of leisure. In this scenario, it may be optimal for firms that differ in their productivity to have different wage policies.<sup>14</sup> Mortensen (1990) models agents searching sequentially for a job when unemployed and when employed, moving from a job to a new one offering a higher wage. In this framework, firms face a tradeoff between short run and long run profits. A firm offering high wages keeps workers for a longer period of time than a firm offering lower wages but extracts a smaller surplus per period from each worker. Intuitively, different wage policies can be optimal for the firms. As a result, the equilibrium in this model is characterized by a nondegenerate continuous distribution of wage offers where each wage in the support of this distribution represents an optimal wage policy for the firm.

Equilibrium models of search have not been incorporated to the empirical literature until recently. Eckstein and Wolpin (1990) estimate a generalization of the Albrecht and Axell model. Eckstein and Wolpin note that "estimating equilibrium labor market models does not require data on both workers and firms, although it does require assumptions about the structure of the distributions of preferences and technology." More recently, several other authors have estimated the Mortensen model with homogeneous agents, or different generalizations of this model with heterogeneous agents.

We begin this section with a description of the Eckstein and Wolpin model. Next, we present several attempts to estimate the Mortensen model. We conclude the section pointing out some of the shortcomings of the existing literature

---

<sup>14</sup>Burdett and Judd show that the model in Albrecht and Axell works only if there is a finite mass of people with zero search costs. One might consider this to be an unreasonable assumption.

and some possible avenues for future research.

## 4.1 Eckstein and Wolpin

Eckstein and Wolpin (1990) represents the first attempt to estimate an equilibrium model. Individuals work in a market where firms differ in their productivity. All workers have the same productivity at the same firms, i.e. they are homogeneous with respect to their market skills, but are heterogeneous with respect to their nonmarket productivity, or preferences for leisure. Individual preferences and firm productivity are private information, although the distribution of preferences over individuals and the distribution of wage offers are known to all agents. There are  $n + 1$  types of individuals with many people of the same type and many firms with heterogeneous productivities. The problem of the worker is as described in Section 2, and each firm maximizes intertemporal profits.

A Nash equilibrium wage offer distribution is shown to exist. Wage dispersion arises due to heterogeneity in worker tastes for leisure and differences in productivity across firms. Different worker types have different reservation wages. In addition values between two adjacent reservation wages are not optimal wage offers for the firms. Consequently, the distribution of wages in equilibrium has a discrete support that corresponds to the reservation wages of the different worker types. More productive firms offer high wages, attracting high reservation wage types in addition to the low reservation wage types. The model also predicts that an average high reservation wage worker spends more time looking for a job and eventually obtains a higher wage; this contradicts other theoretical models that rely on unobserved heterogeneity to generate a negative correlation between unemployment spell length and accepted wage offer.

In order to be able to estimate the model, the authors introduce several parametric assumptions about the distribution of productivity among firms and the distribution of workers types. The model is estimated first just using information on duration of unemployment and then using information on duration of unemployment and accepted wage offers. In order to be able to fit the wage data to a discrete distribution with  $n + 1$  points of support, the authors assume that wages are measured with errors. A novel feature of this paper is that the restrictions imposed by the equilibrium are incorporated to the estimation procedure. However, any heterogeneity in productivity across workers, whether observed or unobserved, would cause the distribution of wage offers to have more than  $n + 1$  points of support. Thus, their implementation of the model relies upon an unreasonable assumption about individuals. Both estimated equilibrium models perform poorly. In fact, measurement error accounts for almost all of the dispersion in observed wages. Nevertheless, the model estimated with duration data is used to simulate the effects of alternative levels of the minimum wage on unemployment and wages.



## 4.2 Mortensen model

In this section, we present a simple version of the equilibrium model described in Mortensen (1990). The economy consists of a continuum of homogeneous workers and homogeneous firms. Firms set wages and unemployed and employed workers search for wage offers among firms.

Workers are infinite-lived wealth maximizing agents. Unemployment income net of search cost is given by  $b$ . Job offers arrive from a Poisson process with parameter  $\alpha_u$  if the worker is unemployed and  $\alpha_e$  if employed. When employed, a worker is laid off at rate  $\delta$ . Assume that each worker faces a known, stationary, nondegenerate distribution of wages  $F(w)$  with a finite mean. Mortensen and Neumann (1988) show that the problem of the worker can be characterized by a reservation wage  $\xi$ . Analytically, the reservation wage is the unique solution to

$$\xi = b + (\alpha_u - \alpha_e) \int_{\xi}^{\infty} \frac{1 - F(w)}{\delta + \alpha_e [1 - F(w)]} dw. \quad (4.1)$$

An unemployed worker accepts any wage offer above the reservation wage, and an employed worker accepts any wage offer that exceeds his current wage.

In this framework, there is a flow of workers between different states, employment and unemployment, and among different jobs. Let  $m$  denote the measure of workers in the economy and  $u$  denote the measure of them that are unemployed. The instantaneous flow of workers from unemployment to employment is  $\alpha_u [1 - F(\xi^-)] u$ ,<sup>15</sup> the product of the offer arrival rate, the acceptance probability, and the measure of unemployed workers. Similarly, the instantaneous flow of workers from employment to unemployment is equal to  $\delta(m - u)$ . If the economy is at its steady state, these two flows should be equal implying

$$u = \frac{m}{\delta + \alpha_u [1 - F(\xi^-)]}. \quad (4.2)$$

Let  $G(w)$  represent the fraction of all unemployed workers who earn wage  $w$  or less. An expression for  $G(w)$  can be obtained at the steady state. Observe that the measure of workers with wage less or equal to  $w$  is  $G(w)(m - u)$  and should be constant in steady state. On the one hand, the flow of workers into this group, that is those unemployed workers that accept a wage less than or equal to  $w$ , is  $\alpha_u [F(w) - F(\xi^-)] u$ . On the other hand, the flow of workers out of this group is the sum of those in the group who become unemployed,  $\delta G(w)(m - u)$  and those who receive a job offer that exceeds  $w$ ,  $\alpha_e [1 - F(w)] G(w)(m - u)$ . Since the inflow and outflow should be equal in steady state,

$$G(w) = \frac{F(w) - F(\xi^-)}{[1 - F(\xi^-)]} \frac{\delta}{\delta + \alpha_e [1 - F(w)]}. \quad (4.3)$$

<sup>15</sup>The wage offer distribution  $F(\bullet)$  might have a discontinuity at  $\xi$ . Thus, we use the notation  $\xi^-$  to indicate a wage infinitesimally less than  $\xi$ .

The distribution in equation (4.3) is useful to determine the supply of labor to firms. Observe that  $F(w) - F(w - \varepsilon)$  is a measure of the percentage of wage offers within the range  $(w - \varepsilon, w]$ . On the other hand,  $[G(w) - G(w - \varepsilon)](m - u)$  is a measure of the number of workers with wages within the range  $(w - \varepsilon, w]$ . Consequently, the steady state number of workers per firm offering a wage  $w$  can be defined as

$$l(w) = \lim_{\varepsilon \rightarrow 0} \frac{[G(w) - G(w - \varepsilon)](m - u)}{F(w) - F(w - \varepsilon)}. \quad (4.4)$$

The steady-state profit flow earned by a firm who offers a wage  $w$  is

$$\pi(w; p) = (p - w)l(w), \quad (4.5)$$

the product of the per-worker profit times the steady state number of workers. Firms maximize the steady state level of benefits.

In this framework, a steady-state market equilibrium is a reservation wage  $\xi$  for the workers and a market wage offer distribution  $F(\bullet)$ . The equilibrium conditions are that the reservation wage maximizes the expected wealth for each worker, i.e. it satisfies equation (4.1) given  $F(\bullet)$ , and every wage offer in the support of  $F(\bullet)$  maximizes the steady state level of profits of the firm, i.e. it satisfies equation (4.5).

Observe that firms are allowed to make profits. All firms should earn the same amount of profits in equilibrium. Consequently  $\pi(w; p)$  should be constant for any wage in the support of  $F(\bullet)$ , the distribution of wage offers in equilibrium. A nice feature of this model is that a closed form solution for  $F(\bullet)$  is obtained (see Mortensen 1990):

$$F(w) = \left[ \frac{1 + k_e}{k_e} \right] \left[ 1 - \left[ \frac{p - w}{p - \xi} \right]^{1/2} \right] \quad \forall w \in [\xi, h], \quad (4.6)$$

with  $h = p - \left[ \frac{1}{(1 + k_e)} \right]^2 (p - \xi)$ , and  $k_e = \frac{\alpha_i}{\delta}$ .

$F(\bullet)$  is continuous with connected support.<sup>16</sup> It can be shown that the distribution of wage offers  $F(\bullet)$  and the distribution of earnings  $G(\bullet)$  have increasing densities  $f(\bullet)$  and  $g(\bullet)$ , respectively. Also, a firm never offers wages below the reservation wage; consequently all wage offers are accepted. In addition, transitions from one job to another should result in an increase in the wage of the worker, and wage growth on the job is not allowed.<sup>17</sup>

<sup>16</sup>Notice that, if there are a mass of employers offering the same wage in equilibrium, then it is optimal for each of these employers to increase its wage offer by a small amount in order to attract workers from other employers.

<sup>17</sup>The model does not consider the possibility of entry of new firms into the market. If we allow for this possibility, entry will occur until the point in which profits are not possible for new entrants. In this case, the most one can do is to choose a cost of entry consistent with the model, that is  $c = \frac{\pi^*}{r}$ , where  $\pi^*$  represents the flow of profits at the steady state equilibrium.

### 4.2.1 Fit of the Model

Several authors have estimated Mortensen's equilibrium model or generalizations of it. A rigorous analysis of the estimation and identification of the model with homogeneous agents has been conducted by Christensen and Kiefer (1994). They show that the model can be estimated, using MLE techniques, with a data set containing unemployment duration, reemployment wage, and job duration from a random sample of workers.

The probability of a spell of unemployment of duration  $t_i^u$  is equal to  $\alpha_u \exp\{-\alpha_u t_i^u\}$ . Since firms never post offers below the reservation wage, offers from any firm to an unemployed worker are always accepted. In addition, the probability of a wage offer  $w_i$  is equal to  $f(w_i)$ . The probability of an spell of employment of duration  $t_i^e$ , given that the wage at the present job is  $w_i$ , can be defined as,

$$(\delta + \alpha_e [1 - F(w_i)]) \exp\{-(\delta + \alpha_e [1 - F(w_i)]) t_i^e\}. \quad (4.7)$$

The contribution to the likelihood function of an observation  $(t_i^u, w_i, t_i^e)$  is the product of these three probabilities:

$$L_i(\theta) = \alpha_u \exp\{-\alpha_u t_i^u\} \cdot f(w_i) \cdot (\delta + \alpha_e [1 - F(w_i)]) \exp\{-(\delta + \alpha_e [1 - F(w_i)]) t_i^e\} \quad (4.8)$$

where  $\theta = (\alpha_u, \alpha_e, \delta, p, b)$  represents the parameters of the structural model to be estimated.

The restrictions imposed by equilibrium are implicitly stated in equations (4.1) and (4.5) and give rise to the closed form solution for the distribution of wages described in equation (4.6). Observe that equation (4.6) implies that any observed wage  $w_i$  belongs to the interval  $[\xi, h]$ . Using the same arguments presented in Section 2.2, we can consider the estimators  $\hat{\xi} = w_{\min}$  and  $\hat{h} = w_{\max}$  for  $\xi$  and  $h$ . Where  $w_{\min}$  and  $w_{\max}$  represent the minimum and maximum wage observed in the sample, respectively. Assuming wages are observed without measurement error, these estimators are consistent and converge at a rate faster than  $\sqrt{n}$ . As explained in Section 2.2, the estimators,  $\hat{\xi}$  and  $\hat{h}$ , can be substituted directly in the likelihood function. After this, we obtain an expression that depends only on  $\alpha_u$ ,  $\alpha_e$ , and  $\delta$ . These remaining parameters can be estimated using MLE. Christensen and Kiefer (1994) present a more exhaustive description of these issues.

A different way to approach this problem is to assume that wages are measured with error. This approach is essentially the same that we summarized in Section 2.3 with the main distinction being that the distribution of wages is obtained here endogenously. This technique has been applied to this framework by Van den Berg and Ridder (1993).

In order to improve the fit of the model to the data, it is necessary to introduce changes in the basic model allowing for the possibility of heterogeneous

agents. Bowlus, Kiefer and Neumann (1995 a, b) consider the case of homogeneous workers and a finite number of firm types differing in productivity. Koning, Ridder and Van den Berg (1995) present a model in which the labor market is segmented and consists of a large number of separate submarkets within which workers and firms are homogeneous.

Bowlus, Kiefer and Neumann (1995 a, b) assume that there are  $Q$  types of firms with productivity  $p_1 < p_2 < \dots < p_Q$ , and let  $\pi_j$ ,  $j = 1, \dots, Q$ , represent the fraction of firms having productivity  $p_j$  or less. This generalization of the homogeneous model is described in Mortensen (1990).<sup>18</sup> The equilibrium in this case is defined as in the homogeneous case with the distinction that here firms of the same productivity should have the same profits but firms of different productivities may have different profits. Mortensen shows that the equilibrium wage distribution is

$$F(w) = F_j(w) \quad \forall w \in [w_{lj}, w_{hj}] \quad (4.9)$$

where

$$F_j(w) = \frac{1 + k_e}{k_e} \left[ 1 - \frac{1 + k_e [1 - \pi_{j-1}]}{1 + k_e} \right] \left[ \frac{p_j - w}{p_j - w_{hj-1}} \right], \quad \forall w \in [w_{lj}, w_{hj}], \quad (4.10)$$

and the equilibrium implies that  $w_{l1} = \xi$ , and  $w_{hj-1} = w_{lj} \forall j$ ; that is, firms with higher productivity offer higher wages. The set of pairs  $\{p_j, \pi_j\}_{j=1}^Q$  have to be estimated with the rest of parameters of the model. The wage policies of different firm types create discontinuities in the wage offer distribution; for this reason, non standard techniques have to be used to estimate the model.

Koning, Ridder and Van den Berg (1995) address the problem of unobserved heterogeneity from a different perspective. They assume that there is a large number of separate markets in which workers and firms meet. Within each market workers and firms are homogeneous, and an equilibrium wage distribution of the form presented in equation (4.6) exists. The productivity of firms among different segmented markets follows a continuous distribution. They effectively have a continuum of submarkets which differ in the value of productivity of workers.

The econometrician observes a mixture of the information generated in different markets. In particular, the value of productivity  $p$  at a certain market is not observed by the econometrician and varies among different markets. In this case, assuming that  $p \sim H(\bullet)$ , the contribution to the likelihood function of an observation  $(t_i^u, w_i, t_i^e)$  is

$$\int L_i(\theta) dH(p) \quad (4.11)$$

---

<sup>18</sup>The informational structure is important in order to compute the equilibrium. Here workers and firms know the distribution of wage offers but they have no information about each other's type.

with  $L_i(\theta)$  defined as in equation (4.8). The authors assume that  $H(p)$  follows a lognormal distribution. The parameters associated with this distribution have to be estimated jointly with the rest parameters of the model. Van den Berg and Ridder (1993) follow a similar approach.

### 4.3 Policy Implications

The most common policy application of the equilibrium search models has been the study of the labor market effects of an increase in the minimum wage. Eckstein and Wolpin (1990) simulate the effect of different values of the minimum wage finding that an increase in the minimum wage increases the unemployment rate as well as the expected duration of unemployment. In particular, an increase in the minimum wage reduces the number of employers by expelling from the market the less productive ones. This produces a reduction in the rate of job offer arrival per worker and consequently increases the expected duration of unemployment per-worker. Intuitively, the same result should hold in the model presented by Bowlus, Kiefer and Neumann (1995 a, b). Taking into account the poor fit of the model to the data, Eckstein and Wolpin recommend caution when interpreting their quantitative results.

Koning, Ridder and Van den Berg (1995) use data from the Netherlands to estimate their model with only between market heterogeneity. Intuitively, this can be interpreted as a situation in which workers with the same productivity interact in the same market and do not have the option of moving to markets with higher productivity. Thus, an increase in the minimum wage will result in the closure of the markets with the lower productivity. Consequently, workers in these markets will become unemployed without any chance of getting a job in the future. Koning, Ridder and Van den Berg use the estimated model to examine the effect of changes in the mandatory minimum wage on the magnitude of structural unemployment. They find that "a 10% increase in the minimum wage increases structural unemployment from 5.2% to 10.1%."

While in Eckstein and Wolpin (1990) and Bowlus, Kiefer and Neumann (1995 a, b) an increase in the minimum wage produces an increase in the average length of unemployment spells, in Koning, Ridder and Van den Berg (1995) or Van den Berg and Ridder (1993) it produces an increase in the number of workers that are permanently unemployed. This difference is due to differences in the theoretical models; thus the data provides no information about which is correct.<sup>19</sup> These different implications about the effect of an increase in the minimum wage are relevant for policy analysis.

The efforts of several authors in estimating equilibrium models of search represent the most recent contributions to the empirical search literature. The equilibrium models with homogeneous agents produce a bad fit of the data. Once the hypothesis about homogeneous productivity on the model of Mortensen is removed, there are several generalizations that, in principle, are not distinguishable and that can fit the data similarly. We see this as an unpleasant feature of

---

<sup>19</sup>A nonnested testing procedure would be necessary to test these models against each other.

the existing literature; this is an issue to be addressed in future research. One possible way to deal with this problem is to exploit richer data sets, possibly including also information about firm behavior. Another possible avenue of research consists of investigating the nonparametric identification of equilibrium models.

## 5 Conclusions

The field of empirical search models is an important and growing field. There has been significant progress made in developing models to deal with measurement error, observed and unobserved heterogeneity, dynamics, and equilibrium conditions. While we have much hope for this field, we view there to be two significant problems to overcome in working with these models. The first is the need for powerful computers. This need is being met over time although no faster than models become more demanding of computer time. In particular, handling observed and unobserved heterogeneity and many nonparametric and semiparametric methods require very large computer resources.

The other problem is limited data. Presently, we are using essentially spell length data and accepted wage data to identify all of the parameters of a search model. In even the simplest model, such data can not separately identify a search cost and discount rate. Allowing for important other functions implied by unobserved heterogeneity or equilibrium may be asking too much of the limited data available. We feel it would be useful to identify other potential (possibly new) sources of data with direct information on possibly search cost, rejected offers, or firm behavior.

Nevertheless, we enthusiastically encourage others to pursue this field. In particular, we think there is significant work to be done on identification, testing, nonparametrics, data acquisition. We are also excited about empirical implementation of equilibrium models though we see significant identification issues not yet addressed.

## 6 References

- Albrecht, J. W. and B. Axell (1984). "An Equilibrium Model of Search Unemployment." *Journal of Political Economy*. 92(5): 824-840.
- Behrmann, J. R. Sickles, and P. Taubman (1990). "Age-Specific Death Rates with Tobacco Smoking and Occupational Activity: Sensitivity to Sample Length, Functional Form, and Unobserved Frailty." *Demography*. 27(2): 267-284.
- Bowlus, A. J., N. M. Kiefer, G. R. Neumann (1995a). "Estimation of Equilibrium Wage Distributions with Heterogeneity." *Journal of Applied Econometrics*. 10: S119-S131.
- Bowlus, A. J., N. M. Kiefer, G. R. Neumann (1995b). "Fitting Equilibrium Search Models to Labor Market Data." Working Paper Series. 95-17. University of Iowa.
- Burdett, K. and K. Judd (1983). "Equilibrium Price Dispersion." *Econometrica*. 5(4): 955-970.
- Burdett, K., N. Kiefer, D. Mortensen, and G. Neumann (1984). "Earnings, Unemployment and the Allocation of Time over Time." *Review of Economic Studies*. 51: 559-578.
- Card, D. and D. Sullivan (1988). "Measuring the Effect of Subsidized Training Programs on Movements In and Out of Employment." *Econometrica*. 56(3): 497-531.
- Christensen, B. J. and N.M. Kiefer (1991). "The Exact Likelihood Function for an Empirical Job Search Model." *Econometric Theory*. 7: 464-486.
- Christensen, B. J. and N. M. Kiefer (1994). "Measurement Error in the Prototypical Job-Search Model." *Journal of Labor Economics*. 12(4): 618-639.
- Cox, D. R. (1972). "Regression Models and Life Tables." *Journal of the Royal Statistical Society. Series B*. 34: 187-220.
- Diamond, P. (1971). "A Model of Price Adjustment." *Journal of Economic Theory*. 3: 156-168.
- Diamond and Hausman (1984). "Individual Retirement and Savings Behavior." *Journal of Public Economics*. 23: 81-114.
- Eckstein, Z. and K. I. Wolpin (1990). "Estimating a Market Equilibrium Search Model from Panel Data on Individuals." *Econometrica*. 58(4): 783-808.
- Elbers, C. and G. Ridder (1982). "True and Spurious Duration Dependence: The Identifiability of the Proportional Hazard Model." *Review of Economic Studies*. 49. 403-410.
- Engberg, J. B. (1994). "Search Duration, Accepted Wages and Selection Bias." Heinz School, Carnegie Mellon U.. Working paper.
- Flinn, C. and J. J. Heckman (1982). "New Methods for Analyzing Structural Models of Labor Force Dynamics." *Journal of Econometrics*. 18(1): 115-168.
- Gritz, R. (1993). "The Impact of Training on the Frequency and Duration of Employment." *Journal of Econometrics*. 57 (1): 21-51.

- Gunderson, M. and A. Melino (1990). "The Effects of Public Policy on Strike Duration." *Journal of Labor Economics*. 8 (3): 295-317.
- Gurmu, S., P. Rilstone and S. Stern (1996). "Nonparametric Hazard Rate Estimation." Unpublished manuscript. University of Virginia.
- Han, A. and J. A. Hausman (1990). "Flexible Parametric Estimation of Duration and competing Risk Models." *Journal of Applied Econometrics*. 5: 1-28.
- Heckman, J. J. (1979). "Sample Selection Bias as a Specification Error." *Econometrica*. 47(1): 153-161.
- Heckman, J. J. (1991). "Identifying the Hand of Past: Distinguishing State Dependence from Heterogeneity." *American Economic Review*. Papers and Proceedings. 106(3): 75-79.
- Heckman, J. J. and B. Honore (1991). "The Identifiability of the Competing Risks Model." *Biometrika*. 76(2): 325-330.
- Heckman, J. J. and B. Singer (1984a). "The Identifiability of the Proportional Hazard Model." *Review of Economic Studies*. 51: 231-241.
- Heckman, J. J. and B. Singer (1984b). "A Method of Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica*. 52(2): 271-320.
- Heckman, J. J. and J. Walker (1990). "The Relationship Between Wages and Income and the Timing and Spacing of Births: Evidence from Swedish Longitudinal Data." *Econometrica*. 58(6):1411-1443.
- Holt, F., E. Merwin, and S. Stern (1996). "The Length of Psychiatric Hospital Stays and Community Stays." University of Virginia, unpublished paper.
- Jensen, P. and N. C. Westergaard-Nielsen (1987). "A Search Model Applied to the Transition from Education to Work." *Review of Economic Studies*. 54: 461-472.
- Johnson, W. and J. Ondrich (1990). "The Duration of Post-Injury Absences from Work." *The Review of Economics and Statistics*. 72(3): 578-586.
- Jovanovic, B. (1984). "Wages and Turnover: a Parameterization of the Job Matching Model." *Studies in Labor Market Dynamics*. (eds.) G.R. Neumann and N.C. Westergaard-Nielsen. Springer-Verlag, Berlin.
- Kennan J. F. (1985). "The Duration of Contract Strikes in U.S. Manufacturing." *Journal of Econometrics*. 28: 5-28.
- Kiefer, N. M. and G. R. Neumann (1979). "An Empirical Job Search Model with a Test of the Constant Reservation Wage Hypothesis." *Journal of Political Economy*. 87(1): 89-107.
- Kiefer, N. M. and G. R. Neumann (1979). "Individual Effects in a Nonlinear Model: Explicit Treatment of Heterogeneity in the Empirical Job Search Model." *Econometrica*. 49(4): 965-979.
- Kiefer, N. M. and G. R. Neumann (1994). "Wage Dispersion with Homogeneity: The Empirical Equilibrium Search Model." (eds.) N. C. Westergaard-Nielsen and P. Jensen. *Proceedings of the Third Symposium on Panel Data and Labour Market Dynamics*. North-Holland, New York. 57-74.
- Koning, P., G. Ridder and G. J. Van den Berg (1995). "Structural and Frictional Unemployment in an Equilibrium Search Model with Heterogeneous



- Agents," *Journal of Applied Econometrics*. 10: S133-S151.
- Lancaster, T. (1979). "Econometric Methods for the Duration of Unemployment." *Econometrica*. 47(4): 939-956.
- Lancaster, T. (1990). "The Econometric Analysis of Transition Data." *Econometric Society Monographs* (17). Cambridge University Press.
- Lancaster, T. and A. Chesher (1983). "An Econometric Analysis of Reservation Wages." *Econometrica*. 51(6): 1661-1776.
- Lancaster, T. and S. Nickell (1980). "The Analysis of Reemployment Probabilities for the Unemployed." *Journal of the Royal Statistical Society Series A*. 135: 257-271 and 143, pt. 2: 141-165.
- Lippman, S. A. and J. J. McCall (1976). "The Economics of Job Search: A Survey." Parts I and II. *Economic Inquiry*. 14: 155-189 and 347-368.
- McCall, B. (1990). "Occupational Matching: A Test of Sorts." *Journal of Political Economy*. 98(1): 45-69.
- McCall, B. (1996). "Unemployment Insurance Rules, Joblessness, and Part-Time Work." *Econometrica*. 64(3): 647-682.
- Meyer, B. D. (1990). "Unemployment Insurance and Unemployment Spells." *Econometrica*. 58(4): 757-782.
- Miller, R. (1984). "Job Matching and Occupational Choice." *Journal of Political Economy*. 92(6): 1086-1120.
- Moffit, R. (1985). "Unemployment Insurance and the Distribution of Unemployment Spells." *Journal of Econometrics*. 28(1): 85-101.
- Mortensen, D. T. (1990). "Equilibrium Wage Distributions: A Synthesis." (eds.) J. Hartog, G. Ridder, and J. Theeuwes. *Panel Data and Labor Market Studies*. North-Holland, New York. 279-296.
- Mortensen, D. T. and G. R. Neumann (1988). "Estimating Structural Models of Unemployment and Job Duration in Dynamic Econometric Modeling." *Dynamic Econometric Modeling. Proceedings of the Third International Symposium in Economic Theory and Econometrics*. Cambridge University Press. 335-357.
- Narendranathan, W. and S. Nickell (1985). "Modeling the Process of Job Search." *Journal of Econometrics*. 28(1): 29-49.
- Narendranathan, W. and S. Nickell and J. Stern (1985). "Unemployment Benefits Revisited." *Economic Journal*. 95: 307-329.
- Nickell, S. (1979). "Estimating the Probability of Leaving Unemployment." *Econometrica*. 47(5): 1249-1266.
- Pakes, A. (1986). "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks." *Econometrica*. 54(4): 755-784.
- Reinganum, J. F. (1979). "A Simple Model of Equilibrium Price Dispersion." *Journal of Political Economy*. 87(4): 851-858.
- Solon, S. (1985). "Work Incentive Effects of Taxing Unemployment Benefits." *Econometrica*. 53(2): 295-306.
- Stern, S. (1989). "Estimating a Simultaneous Search Model." *Journal of Labor Economics*. 7(3): 348-369.
- Sueyoshi, G. (1992). "Semi-parametric Proportional Hazards Estimation of

Competing Risks Models with Time-Varying Regressors." *Journal of Econometrics*. 51: 25-58.

Van den Bert, G. J. (1992). "A Structural Dynamic Analysis of Job Turnover and the Cost Associated with Moving to Another Job." *The Economic Journal*. 102: 1116-1133.

Van den Berg, G. J. and G. Ridder (1993). "On the Estimation of Equilibrium Search Models from Panel Data." (eds.) J. C. Van Ours, G. A. Pfann, and G. Ridder. *Labor Demand and Equilibrium Wage Formation*. North-Holland. New York. 227-245.

Van den Berg, G. J. and Geert Ridder (1993). "An Empirical Equilibrium Search Model of the Labour Market." Unpublished manuscript. Free University of Amsterdam.

Wolpin, K. (1987). "Estimating a Structural Search Model: The Transition from School to Work." *Econometrica*. 55(4): 801-818.