J. R. Statist. Soc. A (1980),
143, *Part* 2, *pp.* 141–165

The Analysis of Re-Employment Probabilities for the Unemployed

By TONY LANCASTER AND STEPHEN NICKELL

University of Hull *London School of Economics*

[Read before the ROYAL STATISTICAL SOCIETY on Wednesday, December 12th, 1979, the President SIR CLAUS MOSER in the Chair]

SUMMARY

This paper is concerned with the modelling and estimation of the process covering the rate at which unemployed individuals leave unemployment. In particular it discusses the problems caused by unobservables in the estimation procedure and ways in which these problems may be attacked. Finally it comes to some rather pessimistic conclusions concerning the identification of the true process in the presence of unmeasured heterogeneity in the population.

Keywords: UNEMPLOYMENT; MAXIMUM LIKELIHOOD; IDENTIFICATION AND UNOBSERVABLES

0. INTRODUCTION

SINCE the middle of 1973, registered male unemployment in Britain has risen by some 120 per cent. Over the same period, the length of time which a male entrant on the unemployment register can expect to remain there has increased from roughly 9 weeks to 18 or 19 weeks (see *D.E. Gazette*, September 1978, p. 1056, Table 7). Thus, almost all the recent increase in unemployment may be accounted for by the rise in the length of time it takes unemployed individuals to find new jobs. The rate at which entrants flow into the register has hardly changed at all. The question then arises as to what causes these dramatic changes in the duration of unemployment spells. Is it simply that vacancies are scarcer now than they used to be because of a short-fall in aggregate demand for output? Or are there other factors at work such as the increasing generosity of unemployment compensation making the unemployed more choosy about the work they will accept and, thereby, voluntarily prolonging the duration of their unemployment spells? In order to investigate such questions as these, there are two basic approaches. The first, and perhaps the most obvious, is to try and explain shifts in aggregate unemployment or unemployment duration over time by changes in aggregate demand and any relevant supply side factors (such as benefit levels). This has been done with British data by a number of economists (see, for example, Maki and Spindler, 1975; Cubbin and Foley, 1977; Taylor (1977)) with results which are not entirely satisfactory partly because of their extreme disparity but basically because of the essential weakness of the data and the lack of any sound theoretical basis for the estimated equations.

A second alternative is to use cross-section data on individuals. If we observe "similar" individuals facing different levels of benefit or different levels of demand for their services we may hope to estimate the impact of these differences on the duration of their unemployment spells. The serious difficulty with this approach is the fact that we can only control for those differences between individuals which we observe and if there are unmeasured attributes, these can very easily bias our estimates. Nevertheless, we feel that this is the only real alternative in the investigation of unemployment duration and in a number of papers (Lancaster, 1979; Nickell, 1979a, b) the authors have already reported their results using this approach. In this paper, however, we shall be concerned with the statistical background to the generation of empirical results in this area. In particular we shall concentrate on the importance of unobservable variables in the analysis of individual unemployment durations, the problems they cause and the extent to which these problems may be overcome. In the next section we shall look

briefly at the theoretical background to the determination of unemployment duration. This is followed in Section 2 by an analysis of how the process determining unemployment duration may be modelled which leads naturally to a discussion of estimation procedures and the treatment of unobservables. After a brief look at how some of the results illustrate these theoretical problems, we proceed in Section 3 to discuss the fundamental problem of identification which arises in any study of duration data once we allow for the fact that we cannot control for all relevant individual characteristics. Our results here are somewhat pessimistic.

1. The Theory of Unemployment Duration

If we assume that unemployed individuals have incomplete information as to the location of appropriate vacancies along with their associated wages, then it is natural to suppose that they will seek out and sample the available opportunities. Suppose they sample one possibility per period. Then in each period either they will be offered a job at some wage or they will be offered nothing. Under fairly weak conditions, it can then be shown that the unemployed individual will, in each period, compare his wage offer (if any) with a predetermined reservation wage and if the latter is no larger than the former, he will accept the job but not otherwise. This reservation wage will be a function of, among other things, the distribution of wage offers, the probability of receiving a wage offer in any period, the level of unemployment compensation and the length of time the individual has already been unemployed. The (conditional) probability of leaving unemployment in any period will be equal to the product of the probability of receiving an offer and the probability of this offer exceeding the reservation wage. Clearly this conditional probability is a direct result of the behavioural mechanism at work and is therefore the natural variable for us to try and explain if we wish to gain insight into the determination of unemployment duration. This is particularly so because the way in which the conditional probability of obtaining a job changes over an unemployment spell and the way in which the impact of certain variables also changes over the spell are of considerable independent interest.

The variables which are liable to affect the probability of an individual leaving unemployment may be divided into four main categories, namely, personal characteristics, family composition, local labour demand and income variables. Personal characteristics and local labour demand can mainly be thought of as influencing the probability of receiving a job offer in any period given a particular distribution of possible wages. Family composition may also act through the job offer probability because, for example, being married may be seen as a signal of reliability. On the other hand, it comes in on the supply side since family needs may influence the period during which a head of household is prepared to remain jobless. Finally, the income variables will influence the probability of a job offer being accepted. A rise in the level of benefits received while unemployed will reduce the chances of leaving unemployment and a rise in prospective incomes in work will increase them. These two relationships are often compressed by considering only the *ratio* of incomes out of and in work, which is known as the replacement ratio, and is obviously negatively related to the conditional probability of obtaining a job.

Finally, it is worth briefly considering what we might expect to happen to an individual's conditional probability of obtaining a job as he remains unemployed for a longer and longer time. First, there is considerable evidence to suggest that individuals who have been out of work for a long time are not viewed favourably by prospective employers nor by the Employment Service. In the light of this, we might expect the conditional probability of obtaining a job to decline with the length of spell. Second, there is a lot of survey evidence which indicates that long term unemployment is particularly unpleasant for the majority of individuals (see Daniel, 1974, for example). If this is the case, then we might expect them to be fairly desperate for a job and hence to follow the strategy of accepting the first job they are offered. Under these circumstances, the current level of the replacement ratio would have no effect on the chances of such an individual leaving unemployment, at least, within some range. So it

is worth testing the hypothesis that the impact of the current replacement ratio on the conditional probability of obtaining a job diminishes over the course of an individual's unemployment spell.

2. Data and Likelihood Functions

The theoretical reasoning sketched in the last section and developed at length in the literature suggests hypotheses about the way in which the sequence of probabilities that an unemployed man will go back to work varies both between individuals and as unemployment is prolonged for the same individual. In order to test these hypotheses it is natural to attempt to estimate the conditional probability of returning to work, θ say, as a function of set of variables designated by a vector $\mathbf{x}^*$ as well as the length of time a man has already been out of work, t. The hazard function $\theta(x^*, t)$ would then be defined by

$$\theta(\mathbf{x}^*, t)\,dt = P(t<T<t+dt)/P(T>t), \tag{2.1}$$

and the corresponding distribution and density functions $G(\mathbf{x}^*, t)$ and $g(\mathbf{x}^*, t)$ are readily expressed in terms of θ by noting that

$$\theta(\mathbf{x}^*, t)\,dt = g(\mathbf{x}^*, t)\,dt/\{1-G(\mathbf{x}^*, t)\}. \tag{2.2}$$

On integration we have the survivor function

$$1-G(\mathbf{x}^*, t) = \exp\left\{-\int_0^t \theta(\mathbf{x}^*, \tau)\,d\tau\right\}, \tag{2.3}$$

and from (2.2)

$$g(\mathbf{x}^*, t) = \theta(\mathbf{x}^*, t)\exp\left\{-\int_0^t \theta(\mathbf{x}^*, \tau)\,d\tau\right\}. \tag{2.4}$$

If we have some data on individual unemployment durations and if we are prepared to write θ as a simple function of $\mathbf{x}^*$ and t with a relatively small number of parameters, then we may estimate these simply by expressing the likelihood in terms of θ and maximizing it with respect to the unknown parameters. In fact this procedure may be followed even if we do not have data on a set of completed unemployment spells but merely enough information to write down the likelihood in terms of θ. Suppose, for example, we have a random sample of individuals who entered unemployment at a given date and were re-interviewed some t_0 periods later. Suppose individuals $i = 1, \ldots, I$ had already returned to work after a spell of length t_i and individuals $j = 1, \ldots, J$ had not so far managed to obtain work. Then we may write the likelihood as

$$\mathscr{L} = \prod_{i=1}^{I} g(\mathbf{x}_i^*, t_i) \prod_{j=1}^{J} \{1-G(\mathbf{x}_j^*, t_0)\}. \tag{2.5}$$

It is worth noting that the probability of the events generating the sample is

$$\prod_{i=1}^{I} g(\mathbf{x}_i^*, t_i)\,dt_i \prod_{j=1}^{J} \{1-G(\mathbf{x}_j^*, t_0)\}$$

and that the likelihood is simply any fixed multiple of this probability. The likelihood given in (2.5) is then this probability divided by $\Pi\, dt_i$, for i from 1 to I, which explains the apparent discrepancy in units between the first and second groups of terms. Data sets of this type are not uncommon, for example the DHSS cohort study of the unemployed which is currently underway will generate data of precisely this type.

Even more limited information can be analysed along these lines. For example, suppose a random sample of unemployed individuals are selected and then interviewed some h periods later. If individuals $i = 1, \ldots, I$ have been unemployed for t_i periods when selected and have returned to work by the time of interview and individuals $j = 1, \ldots, J$ have been unemployed for t_j periods when selected and had not returned to work when interviewed, the likelihood is

$$\mathscr{L} = \prod_{i=1}^{I} \left\{ \frac{G(\mathbf{x}_i^*, t_i+h) - G(\mathbf{x}_i^*, t_i)}{1 - G(\mathbf{x}_i^*, t_i)} \right\} \prod_{j=1}^{J} \left\{ \frac{1 - G(\mathbf{x}_j^*, t_j+h)}{1 - G(\mathbf{x}_j^*, t_j)} \right\}. \tag{2.6}$$

This is the type of likelihood used in Lancaster (1979) and is in fact the likelihood for a binary-dependent variable regression model in which "success" is finding a job between selection and interview and the regressors are $\mathbf{x}^*$ and t. If the date of return to work is known and is h_i periods after the first interview, then the numerator of the first term in (2.6) may be replaced by $g(\mathbf{x}_i^*, t_i+h_i)$ and somewhat more efficient estimates may thereby be obtained. Finally it is worth pointing out that even a random sample of unemployed individuals who are merely interviewed once may be analysed. If each individual has an uncompleted spell t_i and $u_i(\tau)$ is the probability density of entry into unemployment at time τ, then the likelihood is

$$\mathscr{L} = \prod_{i=1}^{I} \left([u_i(t-t_i)\{1 - G(\mathbf{x}_i^*, t_i)\}] \Big/ \int_0^\infty u_i(t-v)\{1 - G(\mathbf{x}_i^*, v)\}\, dv \right), \tag{2.7}$$

where t is the interview date. Essentially because of the single interview date, the likelihood contains nuisance parameters, u_i, which must be eliminated. Fortunately, any individual specific effects in the u_i will cancel in the likelihood and only time variations are important. The way this problem is dealt with is discussed fully in Nickell (1979) and need detain us no further here.

In all the different examples discussed, knowledge of the complete set of variables $\mathbf{x}^*$ and a choice of functional form which is reasonably close to the true form will enable one to produce satisfactory estimates of the parameters of θ by standard maximum likelihood techniques. Indeed in some cases (see Cox, 1972, for example) inferences can be drawn concerning the effects of the $\mathbf{x}^*$ variables on θ without making any specific assumptions about the way θ varies with t. The crucial problem with all this, however, is the fact that we are unlikely to observe all the variables $\mathbf{x}^*$. This is by no means an unusual problem but in linear regression analysis it only becomes serious if the omitted regressors are correlated with those which are included. In the present case this no longer applies. Suppose, for example, we have omitted a variable called "motivation", say, which is independent of all other regressors and is positively correlated with the chances of obtaining work. Now consider successive cohorts of longer and longer duration in the unemployed population. These are bound to have a higher and higher proportion of individuals who are deficient in "motivation" and this will show up in the data as a decline in the chances of obtaining work with increasing duration. Thus time already spent unemployed is negatively correlated with the omitted variable "motivation" which will inevitably bias the estimated impact of duration so far on the chances of obtaining work.

This problem is one which has long been recognized and crops up in numerous areas. The question of interest is generally to determine whether or not the past occupancy of some state by an individual influences the current probability of occupying that state. The difficulty arises because if individuals differ in their chances of being in the state in question, then in any heterogeneous group it will be observed that there is a correlation between previous occupancy of a state and current occupancy *even though*, for any individual, the current probability of occupancy is independent of that person's history. There is, for example, a considerable literature on the recognition of accident proneness and the question of whether an individual's accident history affects his current chances of an accident (see Bates and Neyman, 1952, for example). In our context the question is whether or not being unemployed

yesterday affects the chances of getting a job today. But this is not the only problem caused by the omitted regressors. They may also introduce a bias in other coefficients and indeed they may induce spurious duration dependence in these coefficients as well.

To analyse these effects more formally let us assume that we observe a subset x of $\mathbf{x}^*$ and that the omitted variables may be captured by a single random variable v which has a distribution function $H(v)$ (density function $h(v)$). Let us further suppose that we may write $\theta(\mathbf{x}^*, t)$ in multiplicative form as

$$\theta(\mathbf{x}^*, t) = v\mu(x, t), \tag{2.8}$$

where the scalar v is analogous to the error term in the linear model and, as in that case, we shall assume it is distributed independently of the included regressors, x, and of t. Thus, for an individual, we think of v as being randomly chosen at $t = 0$ independently of x and then remaining unchanged at least until a job is obtained.

If we let

$$I(x, t) = \int_0^t \mu(x, \tau)\, d\tau \tag{2.9}$$

then the survivor function, for individuals homogeneous with respect to $\mathbf{x}^*$, i.e. to both x and v is

$$1 - G(\mathbf{x}^*, t) = \exp\{-vI(x, t)\}. \tag{2.10}$$

However, the data of an investigator who uses only the elements of x as regressors will in fact be generated by the distribution of T conditional only on x and this is (2.10) averaged with respect to the distribution of v. Thus, in obvious notation, we have

$$1 - G(x, t) = \int_0^\infty \exp\{-vI(x, t)\}\, dH(v). \tag{2.11}$$

Differentiating minus the logarithm of (2.11) gives the conditional probability or hazard function for individuals classified by x alone, $\theta(x, t)$ say, and this may be written

$$\theta(x, t) = \mu(x, t) \int_0^\infty v w(v, x, t)\, dv \tag{2.12}$$

where

$$w(v, x, t) = \exp\{-vI(x, t)\} h(v) \Big/ \int_0^\infty \exp\{-vI(x, t)\}\, dH(v)$$

is the density function of v given $T > t$. That is, the distribution of the error term, v, over the survivors at t. Thus, equivalently,

$$\theta(x, t) = E\{\theta(x^*, t) \mid T > t\} = \mu(x, t) E(v \mid T > t) = \mu(x, t) \bar{v}(x, t) \tag{2.13}$$

Since $\bar{v}$, the average of v over the survivors at t, must fall with t as those with the larger v's obtain work earlier, it is clear that $\theta(x, t)$ must rise less rapidly or fall more rapidly than $\theta(\mathbf{x}^*, t)$. In fact, an easy calculation gives

$$\partial \bar{v} / \partial t = -\mu(x, t) \operatorname{var}(v \mid T > t) < 0.$$

Note also, however, that $\bar{v}$ depends on x as well as t and this will ensure that the ways in which changes in x affect the observed hazard $\theta(x, t)$ are different from their impact on the true hazard. To obtain simple results in this area we may specialize $\theta(x^*, t)$ to the proportional

hazard form

$$\theta(x^*, t) = v\mu(x)\psi(t) \tag{2.14}$$

which has the useful property that the proportional impact of changes in μ and hence x on θ is constant over time. Under these circumstances we may show that although

$$\frac{\partial \log \theta(x^*, t)}{\partial \log \mu} = 1$$

we have

$$\frac{\partial \log \theta(x, t)}{\partial \log \mu} = 1 - \frac{\mu(x)}{\bar{v}} \left\{ \int_0^t \psi(\tau)\, d\tau \right\} \operatorname{var}(v \mid T > t) \leqslant 1. \tag{2.15}$$

Thus the proportional impact of μ on the hazard function is diminished and also depends on t. This is an important result because it indicates that we may draw incorrect inferences concerning included regressors even though we only omit variables which are uncorrelated with them. Furthermore if, for example, we wish to test the hypothesis that the impact of a particular regressor is attenuated over time as in our hypothesis concerning the replacement ratio effect on the chances of obtaining a job discussed in Section 1, then we must proceed with caution given the spurious time dependence introduced by the omitted variables.

Before we go on to discuss how we may take account of these omitted variables in estimation, it is worth briefly presenting a specific model of the error distribution, $H(v)$, to illustrate the above results. It will be clear from (2.11) that the mathematically most convenient model is the Gamma family. Since any departure of the mean from unity can be absorbed in the constant of μ we choose the Gamma family with unit mean and parameter ω which is the reciprocal of the variance, σ^2. Thus

$$dH(v) \propto v^{\omega - 1} e^{-v\omega}\, dv.$$

Using this in (2.11) *et seq.* we readily find

$$1 - G(x, t) = \{1 + \sigma^2 I(x, t)\}^{-1/\sigma^2}; \quad \bar{v} = E(v \mid T > t) = \{1 + \sigma^2 I(x, t)\}^{-1};$$

$$\operatorname{var}(v \mid T > t) = \sigma^2 \{1 + \sigma^2 I(x, t)\}^{-2}. \tag{2.16}$$

Thus

$$\theta(x, t) = \mu(x, t)/\{1 + \sigma^2 I(x, t)\} \tag{2.17}$$

from (2.13). Note that since I increases monotonically from zero to infinity, the observed chances of obtaining work, $\theta(x, t)$, will go to zero with t whatever the properties of μ so long as it is bounded. Furthermore, in the proportional hazard case (2.14), we may differentiate (2.17) to obtain

$$\frac{\partial \log \theta(x, t)}{\partial \log \mu(x)} = \left\{ 1 + \sigma^2 \mu(x) \int_0^t \psi(\tau)\, d\tau \right\}^{-1}. \tag{2.18}$$

This implies that the derivative of the observed hazard with respect to $\mu(x)$ is always closer to zero than that of the true hazard and has the same sign. Moreover it vanishes asymptotically. However, it is worth noting that counter examples can be constructed to show that neither of these properties holds true for every form of distribution $H(v)$.

In the specific case $\mu(x) = \exp\{\mathbf{x}'\boldsymbol{\beta}\}$ we have

$$\frac{\partial \log \theta(x^*, t)}{\partial x_j} = \beta_j$$

and

$$\frac{\partial \log \theta(x, t)}{\partial x_j} = \beta_j \left\{1 + \sigma^2 \mu(x) \int_0^t \psi(\tau)\, d\tau\right\}^{-1}. \tag{2.19}$$

The effect of each regressor on the observed chances of obtaining work is thus attenuated and falls monotonically to zero in spite of its constant impact on the true hazard.

To summarize, if the data are generated by the underlying model (2.8) and we fit a model of this form which ignores the omitted regressor term v by assuming it identically equal to its expectation, we shall overestimate the decline in the individual's chances of obtaining work over a spell of unemployment. We are also likely to underestimate the impact of observed variables on his chances of getting a job and we may well observe a spurious decline in this impact as his unemployment spell lengthens. Professor D. R. Cox in private correspondence with the first author suggested an alternative way of allowing for the possible omission of systematic sources of variation, namely to allow them to enter (2.8) additively as

$$\theta(\mathbf{x}^*, t) = \mu(x, t) + v \quad \text{or} \quad \theta(\mathbf{x}^*, t) = \{\mu(x) + v\}\psi(t).$$

We might even assume the fully additive model

$$\theta(t \mid \mathbf{x}^*) = \mu(x) + v + \psi(t).$$

These all lead to rather different results from (2.17) and (2.19) and indeed in all three cases an attenuation effect like (2.19) does not arise.

In the light of these problems it is clear that we must attempt to model the omitted variables explicitly in setting up the likelihood function. One way of doing this is to suppose, as we have in the theory, that the omitted regressors may be summarized by a multiplicative random variable, v, and thus to assume that the probability of obtaining a job may be written as in (2.8), i.e.

$$\theta(\mathbf{x}^*, t) = v\mu(x, t).$$

The next problem is the choice of functional forms for both $\mu(x, t)$ *and* for the distribution of v, namely $H(v)$. Once these are selected we may write the survival probability, $1 - G$, in terms of observables and the parameters of μ and H using (2.11). Then any of the likelihoods (2.5), (2.6) or (2.7) may be written in terms of the parameters of μ and H which may then be estimated.

Two examples of this technique taken from the author's previous work illustrate the general role of omitted regressors and the kind of results which may be expected. Lancaster (1979) uses the following model,

$$\mu(x, t) = \exp(\mathbf{x}'\boldsymbol{\beta})\, \alpha t^{\alpha - 1}, \quad h(v) = v^{\theta - 1} e^{-v\theta} \quad \text{where } \theta = 1/\sigma^2. \tag{2.20}$$

In table 1 we present some results which illustrate the impact of successive inclusion of regressors on the rate at which the observed probability of obtaining work declines over the spell. These estimates were obtained assuming v constant and demonstrate clearly how the rate of decline is reduced as successive regressors are included and hence the unobserved heterogeneity is reduced.

Table 2 presents Lancaster's full estimates of the model (2.20) where, in column 1, σ^2 is constrained to be zero.

TABLE 1

Estimates of α from model (2.20)

$\hat{\alpha}$	*Included variables*
0·67	Constant
0·74	Constant, log age
0·768	Constant, log age, log unemployment rate
0·773	Constant, log age, log unemployment rate, log replacement ratio

Note: The data refer to 479 unskilled workers who are either men or single women. They are taken from the survey data on the registered unemployed described in Daniel (1974) and refer to Britain. The unemployment rate refers to that in the region in which the individual resides. The replacement ratio refers to the ratio of benefits received when interviewed to after tax income when last at work. The likelihood used was that set out in (2.6). Further details may be found in Lancaster (1979).

TABLE 2

Variable	*v constant* *Coefficient (s.e.)*	*v variable* *Coefficient (s.e.)*	*Elasticity of expected unemployment duration*
Log age	−0·57 (0·21)	−0·70 (0·30)	0·86
Log unemployment rate	−0·33 (0·18)	−0·41 (0·26)	0·44
Log replacement ratio	−0·44 (0·21)	−0·48 (0·26)	0·53
Time	$\hat{\alpha} = 0{\cdot}76$ (0·05)	$\hat{\alpha} = 0{\cdot}95$ (0·15)	
Error variance	$\sigma^2 = 0$	$\sigma^2 = 0{\cdot}15$ (0·16)	
Log likelihood	−224·33	−224·09	

Before commenting we present analogous results for the model presented in Nickell (1979) where

$$\theta(\mathbf{x}^*, t) = \{1 + \exp -(v + \mathbf{x}'\boldsymbol{\beta} + \alpha_1 t + \alpha_2 t^2)\}^{-1},$$
$$v = v_1 \text{ with probability } \phi, \quad v_2 \text{ with probability } 1-\phi. \tag{2.21}$$

Although this is not strictly a model of the multiplicative type, since θ is small, we have

$$\theta(x^*, t) \simeq \exp(v + \mathbf{x}'\boldsymbol{\beta} + \alpha_1 t + \alpha_2 t^2) = e^v \exp(\mathbf{x}'\boldsymbol{\beta}) \exp(\alpha_1 t + \alpha_2 t^2)$$

so it is approximately multiplicative and may be expected to yield results in line with our theoretical discussion. The results are presented in Table 3.

In spite of having different lists of included regressors, different observations and different types of data, the results are remarkably similar. Allowing for omitted variables changes a clearly falling hazard into one which is roughly constant (Table 2) or even increasing (Table 3) (in the somewhat unlikely event that the marginally non-significant coefficient on t^2 is, in fact, zero). Allowing for error increases the coefficient on the included regressors in modulus in every case. These results are implied by (2.15) and (2.19) and are consistent with omitted systematic factors entering multiplicatively into the model. Nevertheless, in Table 3 it appears that the time dependence of the replacement ratio effect is unaffected by allowing for omitted variables. Finally, even though allowance for omitted variables has notable effects on time and included regressor coefficients it gives quite negligible increases in the log likelihood.

What are we to make of this? A similar situation arises in ordinary multiple regression analysis. If one adds to the analysis a new regressor highly correlated with those included,

the result will be little reduction in residual variance but there may be major changes in the coefficients of the included regressors. Even though in principle the new regressor is relevant, in practice the design of the regressor matrix does not permit precise evaluation of its independent effect. In a sense the effect of time variation in the hazard and the systematic

TABLE 3

Variable	*v constant* ($v_1 = v_2$) *Coefficient* (*s.e.*)	*v variable* *Coefficient* (*s.e.*)	*Elasticity of expected unemployment duration unless otherwise stated*
Time (tenths of a week)	−0·0067 (0·0072)	0·0066 (0·0053)	
Time^2	−0·0070 (0·0073)	−0·017 (0·010)	
Replacement ratio ($t < 20$)	−1·58 (0·43)	−1·68 (0·49)	0·95
Replacement ratio ($t \geqslant 20$)	0·31 (0·55)	0·35 (0·54)	
Dependents	−0·08 (0·090)	−0·092 (0·11)	One child raises duration by 9 per cent
Marital status	0·82 (0·22)	0·88 (0·30)	Marriage lowers duration by 53 per cent
Age	−0·021 (0·0046)	−0·023 (0·0062)	0·88
Ill-health	−0·47 (0·15)	−0·50 (0·17)	Ill-health raises duration by 65 per cent
Local labour demand	−0·45 (0·22)	0·47 (0·26)	−0·05
v_1	—	−1·00 (0·34)	
v_2	—	1·13 (1·07)	
ϕ	—	0·65	
Constant	−0·98		
Log likelihood	−1005·78	−1005·43	

Notes: The data refer to the 426 unemployed males, 18 years and over, surveyed in the General Household Survey of 1972. It includes a double sample for Scotland. In this sample the unemployed are those who have had no employment for a week or more and who claim to be seeking work or who are waiting to take up a job or who would be seeking work but for temporary sickness. Most of these are registered but about 15 per cent are not. Furthermore, about 7 per cent of the registered unemployed do not fall into this category because they are not seeking work. The replacement ratio is computed in terms of family income and is allowed to vary over the spell (i.e. with t). "Dependents" refers to the number of the individual's dependents excluding his wife. Marital status and ill-health are dummy variables, the latter referring to the reason for leaving the last job. "Local labour demand" refers to the ratio vacancies/unemployment within the individual's region and three digit occupation. The replacement ratio coefficient switches at 20 weeks to capture the time dependence of its coefficient. Further details may be found in Nickell (1979a, b).

error are "collinear". This is particularly clear in Lancaster (1979) where assuming (a) no error ($\sigma^2 \equiv 0$) but possible time dependence ($\alpha \not\equiv 1$) and, (b) possible error ($\sigma^2 \not\equiv 0$) and no time dependence ($\alpha \equiv 1$) give almost identical log likelihoods. This suggests that there may be a problem of identification and this is the subject of our next section.

3. THE IDENTIFICATION PROBLEM

The fact that the data appear unable to discriminate with any precision between genuine time variation in the hazard function and unobserved sample heterogeneity suggests that it is worth pursuing the general question of identification in models such as this one. In order to investigate this problem we may start from the basic model of the previous section where the probability of obtaining a job or hazard function is given by

$$\theta(x^*, t) = v\mu(x, t)$$

with x being those independent variables which may be observed. v, representing the unobservables, has a distribution $H(v)$ and in this model the survivor function which is generated

by the data, i.e. conditional on the observables x, is given by

$$1-G(x,t)=\int_0^\infty \exp\{-vI(x,t)\}\,dH(v), \tag{3.1}$$

where I is defined in (2.9). Let us define the function $L(s)$ by

$$L(s)=\int_0^\infty \exp(-vs)\,dH(v), \tag{3.2}$$

where L is simply the Laplace transform. Combining (3.1) and (3.2) gives

$$1-G(x,t)=L\{I(x,t)\} \tag{3.3}$$

and this equation reveals precisely why there is a problem of identification. The left-hand side is the observed survivor function which may be estimated by simply looking at the rate at which various well-defined groups leave unemployment. $I(x,t)$, from which $\mu(x,t)$ may be obtained by differentiation, is the function which we wish to discover for then we would know precisely how different individuals behave when they become unemployed. Unfortunately, unless we have some prior information concerning the distribution H, L can obviously be any function of the form (3.2), that is any Laplace transform of a density function. In fact, it is possible to specify precisely the class of allowable L functions. Following a theorem due to Bernstein (see Widder, 1946), Theorem 12a, Ch. IV, p. 160 and Theorem 14a, Ch. VII, p. 310) a *necessary and sufficient* condition that a function $L(s)$ can be expressed in the form (3.2) where H is a distribution function is that L should run from unity to zero and be completely monotone in $0\leqslant s<\infty$. A completely monotone function $L(s)$ on some interval (a,b) satisfies

$$(-1)^k L^k(s)>0, \quad a<s<b, \quad \text{all } k\geqslant 0, \tag{3.4}$$

where L^k is the kth derivative of L. Thus its derivatives alternate in sign. Consequently if we observe a survivor function for individuals of type x given by $1-G(x,t)$, then the true function $I(x,t)$ can be any function of the form

$$I(x,t)=L^{-1}\{1-G(x,t)\}, \tag{3.5}$$

where L is *any* completely monotone function running from unity to zero.

In our empirical work we have attempted to get around this problem by specifying functional forms for both L and I. One glance at (3.3) reveals why the parameters of these functions may be somewhat imprecisely estimated. In a sense, however, there is a more serious problem. We may have little prior information on the correct shape of the functions L and I and we may therefore produce misleading results by mis-specifying these functional forms. The following provides a good example. For simplicity suppose there are no exogenous variables and let the true functions be

$$L^*(s)=\exp(-\bar{v}s), \quad I^*(t)=\log(at+1), \quad 1-G(t)=(at+1)^{-\bar{v}}, \quad a>0, \tag{3.6}$$

which are readily seen to satisfy (3.3). We may also look at these in terms of the hazard function, θ, for by differentiating (3.3) we obtain

$$\theta(x,t)=-\frac{L'\{I(x,t)\}}{L\{I(x,t)\}}\cdot\mu(x,t). \tag{3.7}$$

In the context of this equation, our true functions are

$$\frac{L^{*\prime}(s)}{L^*(s)}=-\bar{v}, \quad \mu^*(t)=\frac{a}{at+1}, \quad \theta(t)=\frac{a\bar{v}}{at+1}. \tag{3.8}$$

These reveal clearly that the true model has no unobserved heterogeneity and the hazard

function genuinely declines as $(at+1)^{-1}$. Suppose, however, that the investigator, not knowing L^* or I^* specifies the following functions,

$$L(s) = (\alpha s+1)^{-q}, \quad I(t) = t^{1-\beta}, \quad q>0, \quad \alpha>0.$$

Note that L has the required property of complete monotonicity and runs from unity to zero. The investigator will thus hypothesize that the data are generated by

$$L\{I(t)\} = (\alpha t^{1-\beta}+1)^{-q}$$

and he observes

$$1-G(t) = (at+1)^{-\bar{v}}.$$

He will then estimate $\beta = 0$, $\alpha = a$, $q = \bar{v}$ thereby wrongly attributing all the observed decline in the hazard to unobserved heterogeneity. The implications of this are that using functional forms to identify L and I in (3.3) as we have done may be appealing and, indeed, may be the best one can do with cross-section data. The results on the time structure of the hazard function cannot, however, be treated as particularly reliable and are unlikely to be very robust with respect to different functional form assumptions. On the other hand, problems with the measurement of the impact of included regressors are unlikely to be so serious because their estimated impact, at least on expected duration, will not depend crucially on the treatment of unobservables.

4. Concluding Remarks

We have chosen in this paper to emphasize the statistical problems which arise in attempting to analyse unemployment duration data, in particular those arising from uncontrolled heterogeneity of the data. But the economic results set out in Lancaster (1979) and Nickell (1979a, b) are interesting and we think important. The effect on unemployment durations of the relative level of unemployment benefit is consistent both with theoretical reasoning and a number of previous studies. The apparent fall in this impact as duration is prolonged revealed in the results of Nickell—which does not seem to be merely an effect of heterogeneity—is also important for purposes of public policy. We would regard the size of the effect of benefits as being now a rather firmly established parameter. The implications of this result for the interpretation of recent increases in the number of registered unemployed—the question with which we began the paper—are dealt with at length in Nickell (1979a) where it is argued that only a small fraction, perhaps 10 per cent, of the increase between 1964/65 and 1973 is attributable to the more generous level of unemployment benefit. This figure is, in fact, in close agreement with that produced by the Department of Employment Working Party (see *D.E. Gazette*, October 1976). The other regression type effects—of age in particular—also confirm previous results, though we would argue that previous studies have used less than fully efficient statistical methods.

Previous studies have analysed the variation in the completed durations of unemployment. The present work has viewed unemployment as a probabilistic process and we have tried to estimate the way in which the chance that an individual will resume work varies over time. Unmeasured systematic differences between individuals can in principle vitiate such an attempt and the evidence of our calculations suggests that this heterogeneity is important in practice. We have extended our models to allow for unmeasured heterogeneity but it seems in practice very difficult to distinguish between the effects of heterogeneity and the effect of pure time variation in the hazard function. Our discussion of the identification problem reveals precisely why this comes about and, further, how stringent the distributional assumptions have to be before identification can, in practice, be achieved. A more promising direction of research is a detailed study of a cohort of individuals with an attempt to determine how their policies and experiences change as their unemployment extends. Some information of this type is

already available and, for example, it is now rather well established that unemployed people can in practice state a least acceptable wage akin to the reservation wage of the theory of Section 2 and that such reservation wages tend to be lower the longer an individual has been out of work. Thus the policies of individuals act to create an increasing hazard function. More extensive work of this kind promises a great extension in our knowledge of the sources of variation in the unemployment statistics.

ACKNOWLEDGEMENTS

We thank Jim Heckman and John Ham for several useful discussions; also Roger Gordon, Nick Rau and the referees for helpful comments on an earlier draft of this paper.

REFERENCES

BATES, G. and NEYMAN, J. (1952). Contributions to the theory of accident proneness II: true or false contagion. *University of California Publications in Statistics*, **1**, 255–275.

COX, D. R. (1972). Regression models and life tables. *J. R. Statist. Soc.* B, **34**, 187–202.

CUBBIN, J. S. and FOLEY, K. (1977). The extent of benefit-induced unemployment in Great Britain: some new evidence. *Oxford Economic Papers*, **29**, 128–140.

DANIEL, W. W. (1974). *A National Survey of the Unemployed.* London: P.E.P. (Broadsheet No. 546.)

LANCASTER, T. (1979). Econometric methods for the duration of unemployment. *Econometrica*, **47**, 939–956.

MAKI, D. and SPINDLER, Z. A. (1975). The effect of unemployment compensation on the rate of unemployment in Great Britain. *Oxford Economic Papers*, **27**, 440–454.

NICKELL, S. (1979a). Estimating the probability of leaving unemployment. *Econometrica*, **47**, 1249–1266.

—— (1979b). The effect of unemployment and related benefits on the duration of unemployment. *Econ. J.*, **89**, 34–49.

PISSARIDES, C. A. (1976). Job search and participation. *Economica*, **43**, 33–49.

STERN, J. and SMEE, C. (1976). The unemployed in a period of high unemployment—some notes on characteristics and benefit status. Department of Health and Social Security (Mimeograph).

TAYLOR, J. (1977). A note on the comparative behaviour of male and female unemployment rates in the United Kingdom, 1951–76. University of Lancaster (Mimeograph).

WIDDER, D. V. (1946). *The Laplace Transform.* Princeton: Princeton University Press.

DISCUSSION OF THE PAPER BY PROFESSORS LANCASTER AND NICKELL

Mr A. R. THATCHER (Registrar General): I read tonight's paper with very great interest because it is not so very long since I was writing on the same subject. In fact, in Section 4 of the paper, the authors say that their results agree with the conclusion reached by a Department of Employment Working Party in 1976. It so happens that I was Chairman of that Working Party.

The conclusion we reached was that only a small part of the very large increase in unemployment since 1966 could have been due to people deliberately staying unemployed because the level of unemployment benefit had been increased. Naturally, I am very glad that tonight's authors and my Working Party both reached the same conclusion, but we reached it by very different routes.

In the President's recent address he spoke about the disturbing gap which has opened up between theoretical statistics, on the one hand, and applied statistics, on the other. It is very pleasant that tonight we can discuss a problem which can be tackled both by the highly mathematical approach adopted by tonight's authors and also by some simpler—perhaps I might say more traditional—statistical methods.

The idea that unemployment might have risen because people might be taking longer to choose their jobs was first suggested in about 1968, when it was noticed that there had been an unexpected shift in the relationship between the number of persons who were unemployed and the number of vacant jobs. People started to wonder whether this might be something to do with the fact that there had been a very large increase in benefits for some of the unemployed, because of the introduction of redundancy payments and earnings related unemployment benefits. This meant that people receiving these higher benefits were no longer compelled to take the first job that was offered, so they could stay unemployed for longer, which would increase the number of people who were unemployed at any given time. It would also increase the number of vacant jobs, and so cause a shift in the relationship between vacancies and unemployment—like the shift that had actually been observed.

In the early 1970s several academic papers fitted regressions between the level of unemployment benefit and the numbers of the unemployed. The authors of those papers suggested that a very large part of the increase in unemployment was due to the higher benefits. Of course, this theory had some highly political implications. In 1974, the Department of Employment set up a Working Party to sift the evidence. This Working Party, which included not only civil servants but also some representatives from the academic world, found that the number of people who were receiving the higher benefits at the critical dates was actually quite small. Even if all of them had doubled the time they spent on the unemployment register, this would still have increased the total number of unemployed by only 70 000 people—well under one-fifth of the increase in male unemployment at that time. Thus, it seemed to us that the people who had based their conclusions entirely on regressions must have mis-specified their models. I am afraid that this often seems to happen in economic applications and problems—but, fortunately, not tonight.

Before coming to the main substance of my comments, may I make a few brief points. First, if an unemployed person is offered a job and refuses it, that individual will certainly be unemployed for longer. It does not follow from this, however, that the total number of unemployed persons will increase, because that depends on whether somebody else takes that job instead. If he does, all that happens is that A is unemployed instead of B. The composition of the unemployed is changed, but not the total.

Next, although one can say (as the authors do in the opening sentences of their paper) that an increase in unemployment can be "accounted for" by the rise in the length of time it takes unemployed individuals to find new jobs, this is not necessarily the underlying reason. Let me explain what I mean by an example. Suppose that someone is queueing for theatre tickets and there are two clerks at the box office. After a time one of the clerks goes to lunch. The queue will get longer, and it will take longer to get to the front of the queue. Anyone standing at the back of the queue can say that the queue has got longer because people are taking longer to buy their tickets. In a way, that is true—but the real reason is that the clerk has gone to lunch. I hope to show later that this kind of explanation also affects the unemployment problem.

Next, there are two points on methodology. The authors' model assumes (in Section 1) that people will seek out and sample the available opportunities. The Department of Employment carried out some surveys of unemployed persons in 1973 and 1976, and on both occasions it was found that of all the people on the unemployment register only about one in three had been notified of even one vacancy or job for which they could apply. For the majority of the unemployed, for most of the time, there were no opportunities to sample. Of course, employment exchanges are not the only places that notify vacancies; but if they were, the idea that people are continuously sampling a large number of job opportunities and choosing between them would not be true in practice.

Next, there is the problem of the omitted variable. The authors quote the example of "motivation", which could not be included in their analysis for lack of data. In the Department of Employment surveys we made an attempt to include this. Staff in local offices were asked to make a subjective assessment for each individual in the sample of his keenness to find work and his prospects of finding work. The population was found to be heterogeneous, so the methodological problem is a genuine one. The assessments showed that motivation is very highly associated with the age of the individual and the length of time he had been unemployed. For a given age group, the individuals who had been unemployed for a long time were the least keen to find work, and they were also the ones with the worst prospects of obtaining work. Now we have to consider whether this is because a long period of unemployment saps people's keenness to find jobs, and so makes them less able to find jobs, or whether it is just that the people who were not keen in the first place are also the ones who stay unemployed for longest. This is a very important question, but we were forced to conclude (in an article in the *Department of Employment Gazette* for March 1974) that "this cannot be decided from the survey". This is really the same as the conclusion in Section 4 of tonight's paper that it is "very difficult to distinguish between the effects of heterogeneity and the effect of pure time variation in the hazard function". We were saying the same thing in different words. I agree with the authors that to settle this problem we need to have successive observations of a cohort of people, preferably assessing their keenness and prospects on each occasion.

Now at last we come to the key question: if the unemployment benefits accounted for only a small part of the rise in unemployment, what really caused that rise? Here I am afraid that I have some doubts about whether it is possible to answer this question by regarding unemployment as a completely probabilistic process. I say this because the probabilities are varying all the time, and

we really need to consider the factors which determine the probabilities. There are many such factors, but in my view there are two which dominate all the rest. These are the supply of labour and the demand for labour. If the supply of labour increases—for example, if there was a bulge in the birth rate 16 years ago so that more school leavers are coming on the market—there will be more unemployed people, and the probability that a particular individual will leave the unemployment register will decrease. If there is a fall in the demand for labour—for example, if the motor cycle industry closes down because of competition from the Japanese—again, there will be an increase in the number of the unemployed, and the probabilities of leaving the unemployment register will change.

Thus, in the last resort, in order to understand what is happening to unemployment, we have to consider what is happening to the supply and demand for labour. This is not at all easy to do, but may I try to summarize briefly the conclusions I reached after both collecting the statistics and struggling to analyse them for many years. There are a great number of complications and red herrings which I shall certainly not mention tonight, but I would like to draw the Society's attention to what I believe to be the most significant figures (Thatcher, 1979).

Changes in 1971–76 *in Great Britain* (*thousands*)

	Males	*Females*
Total supply of labour	−123	+768
Employees in production industries	−565	−248
Employees in other industries	+238	+975
Registered unemployed	+383	+208
Unregistered unemployed	−34	−54

The table shows changes between 1971 and 1976. The top row gives the change in the total supply of labour, estimated as best one can. It shows a slight drop for males, but a very large increase for females—the reason being that an increasing number of married women were taking jobs or looking for work after having their families.

The second row shows a fall in the demand for labour in the industries which are included in the Index of Production. There were many reasons for the fall, one of which was that we did not do so well as Japan, Germany and France in the competition for exports. The demand fell more for males than for females because the majority of the workers in these particular industries are men.

The third row shows what happened in the other industries. There was a large increase in the demand for labour in education, health and social services, where the employees are mostly women.

These figures show the available estimates of what happened to the supply and demand for labour during the period 1971 to 1976. If we now work down the columns, taking the supply minus the demand, we find that these changes on their own could be expected to have led to an increase in unemployment of rather over 200 000 men and 40 000 women. Registered unemployment rose more than this, but partly because there was a fall in the numbers who said in sample surveys that they were looking for work but did not register as unemployed—these are the so-called "unregistered unemployed", shown on the last row of the table. Thus there was a shift in the proportion of the unemployed who registered. But this still leaves an unexplained margin of about 140 000 men and 110 000 women, part of which will be due to the effects of unemployment benefits discussed tonight, and part to other causes.

Of course, I would not for a moment suggest that these calculations are precise. There are many difficulties and uncertainties. But in conjunction with the corresponding figures for 1966–71 given in the reference they at least give tentative orders of magnitude which may perhaps help to put into perspective the various factors which have led to the level of unemployment today.

I must apologize to tonight's authors for spending so much of my allotted time in giving my own interpretation of the reasons for the increase in unemployment, rather than in commenting on theirs. I have done this because, as I said at the beginning, it is not often that there is an opportunity nowadays to compare the mathematical approach with what may be called the traditional approach. May I say that I feel extremely grateful to the authors, first, for confirming by their more rigorous methods the conclusions which we reached by more primitive means and, secondly, for giving us

an opportunity to discuss tonight a problem which is of major national importance.

I have very much pleasure in proposing the vote of thanks.

Mr J. P. BURMAN (Bank of England): Most of us are familiar with the extensive time-series work on modelling the level of unemployment—the Phillips curve and so on. I think we should warmly welcome the use by Professors Lancaster and Nickell of cross-section methods on the flow of people from the unemployment register.

This paper is rather unusual in that it reports separate empirical work by the two authors; so it is inevitable that they should be compared. Speaking as an outsider on cross-section techniques, I find Professor Lancaster's model—Cox's proportional hazard model—the more intuitively appealing. But Professor Nickell's variables are more complete.

En passant, I must say that the terminology carried over from medical research sounds a little odd when applied to unemployment: how would an unemployed man react to being told there was a measurable hazard of his finding a job?

Lancaster's data comprise a sample from the unemployment register in 1973, who were interviewed five weeks later. I am a bit puzzled about the footnote to Table 1, which suggests that people were interviewed twice. Nickell's data form a single sample from the General Household Survey for 1972. The authors seem to have squeezed a great deal of information out of quite heavily censored data—for instance only 18 per cent of Lancaster's sample of 479 actually found jobs during the five weeks between registration and interview.

The authors suggest four types of variable which would affect the probability of leaving the unemployment register—the hazard function:

1. Those characteristics of the individual which do not vary in the short-term—age, health, marital status, number of dependants.
2. A demand variable—either the local unemployment rate or the ratio of vacancies to unemployment.
3. Characteristics of the individual which may change with the duration of his unemployment—notably unemployment benefit and other family income, and the "reservation wage", the lowest wage he is prepared to accept. Both authors use as a regressor the ratio of these, the "replacement ratio".
4. Duration of unemployment itself.

Both Lancaster and Nickell introduce duration explicitly in their models, and, as the present paper explains fully, there is a major problem in distinguishing between time variation in the individual's replacement ratio and global duration effects, such as the probability of job offers by employers varying with the duration of the individual's unemployment. Professor Lancaster assumes the reservation wage to be a constant for the individual, his last wage in employment, apart from an allowance for inflation. But Professor Nickell also allows for variation of income with duration of unemployment, and bases the estimated reservation wage on the average earnings for the appropriate occupational group. In neither survey was a direct question asked about the reservation wage, to test the hypothesis that it tends to fall with duration.

It has also been shown in this paper that omitted variables will tend to exaggerate the decline in the hazard function with duration and to bias downwards the estimated elasticities of the known regressors. The very full discussion of this problem—including the use of the Laplace Transform to show that it is *virtually* insoluble with the given data—leaves one wondering whether the authors are now in a state of total scepticism about their results! It seems that the parameter estimates cannot be treated as reliable until one is sure that *all* important variables have been measured. I am led to ask why information was not collected for Professor Lancaster's sample on the additional variables used by Professor Nickell (health, marital status, number of dependants)? What other variables do the authors think are important? What are those included in the DHSS cohort study?

Professor Nickell's model estimates a replacement ratio which is constant for 20 weeks and then virtually zero. Do the authors now think that, because of omitted variables, this inference about the interaction of duration and replacement ratio is suspect?

On a notational point: the formula for the likelihood in equations (2.1) to (2.5) is developed in continuous time, necessitating an explanation of the differing dimensions in equation (2.5). It would be easier for presentation if discrete time were used throughout, and integrals replaced by sums, e.g. $g(x_i^*, t)$ = probability of ith person taking job in week t, especially since people are rather imprecise about how long has elapsed since they took a job (see Definition 2 in Lancaster,

1979). [It did not appear possible to work out, for those who had found a job, precisely when, in the interval t, $t+5$ weeks, they had found it.] As I understand it, Cox's model requires some modification of the formula on account of multiple departures in a single week.

I was delighted with the definition of "family needs" in Nickell (1979): $1+0{\cdot}6\times$ number of wives, etc.

I have very much pleasure in seconding the vote of thanks on this paper.

Mr G. HYMAN (Centre for Environmental Studies): Empirical investigations of labour mobility and migration indicate that individuals who have not moved for a long period of time are less likely to move than others (cf. Ginsberg, 1971, 1973; Hyman, 1974; Hyman and Gleave, 1977). Two rival hypotheses have been advanced to explain these findings:

(i) the hypothesis of cumulative inertia: that individual propensities to move decline with increasing duration in the same state (Myers, McGinnis and Masnick, 1967);

(ii) the hypothesis of heterogeneity: that some individuals are inherently more mobile than others (McFarland, 1970).

It is a substantive problem to choose between these two hypotheses from the results of a statistical experiment that yields an unambiguous result. I believe that such an experiment can only be conducted using life history data for individuals. Any unobserved heterogeneity could be detected in a one way analysis of variance with individuals as the rows and completed times between moves as the columns. The hypothesis of cumulative inertia could be investigated by a determination of these distributions for homogenous classes of individuals. The absence of any duration dependency of individual moving propensities would yield a null hypothesis consistent with exponential forms for these distributions. Such an experiment may not fully resolve questions concerning the causes of recent increases in male unemployment as this may be the cause of increased unemployment durations and not a consequence of them. Investigations are being conducted (Hyman and Palmer, 1978, 1979), using multivariate time series methods applied to unemployment and vacancy flow data, to produce models for predicting regional unemployment which represent the comparative influences of changes in labour supply and labour demand. Such investigations do not yield unique interpretations in terms of individual behaviour and are usefully complemented by studies based on individual life histories.

A suitable framework for such studies might be the non-stationary cross-sectional model recently proposed by Goodrich and Caines (1979). We could, for example, consider estimating an equation of the form

$$Z_i(t) = \sum_{n=1}^{p} \sum_j a_{ij}(n)\, X_{ij}(t-n) + r_i + u(t) + \varepsilon_i(t).$$

In this equation $Z_i(t)$ could represent the duration of unemployment of an individual i who is unemployed at time t whilst the $X_{ij}(t-n)$ are observed characteristics j of the individual and the individual's environment, which could include durations of previous spells of unemployment. The r_i are time invariant zero mean unobserved random processes and the $u(t)$ are time dependent processes affecting all individuals equally. The $\varepsilon_i(t)$ are zero mean normal random variants which are independent across time periods and individuals. The terms r_i and $u(t)$ capture unmeasured heterogeneity and pure time variation respectively. Whilst cumulative inertia leads to a non exponential distribution of times between moves, within a renewal framework, a dependency on past durations, reflected in the a_{ik} parameters, brings us outside the purview of renewal models.

I would like to thank Tony Lancaster and Stephen Nickell for their stimulating presentation which has caused me to think further about the interesting methodological problems that arise in the study of labour mobility.

Professor D. J. BARTHOLOMEW (London School of Economics): As the authors remark, the near impossibility of disentangling the effects of time and heterogeneity is not confined to the duration of unemployment. The same problem arises with the duration of time that people spend in a job or living in a particular place. In this connection Tuma (1976) has introduced covariates into the hazard function in a manner very similar to the authors in this paper. The conclusions reached will therefore be of interest to sociologists and manpower planners as well as economists.

The authors have chosen to study re-employment probabilities via the hazard function. There are some advantages in looking instead at the length of time before a person, who is already out of work, finds a job. The expected time to find a job is, in general, a function of time already

unemployed and is equivalent to the hazard function as a way of describing the unemployment process. This function has an obvious intuitive meaning and a simple mathematical form for the authors model. It also arises naturally when data are collected by sampling the unemployed at a point in time and then following the sample members until they find work. McGregor (1978) used such data and tried to distinguish the effects of time and personal attributes by regressing remaining time unemployed on past time and various personal attributes. However, he failed to recognize that the apparent time dependence which he found could have been due to omitted variables.

If we suppose, with the authors, that

$$\theta(x^*, t) = v \exp(\beta' \mathbf{x})$$

and that v has a gamma distribution with parameter θ then

$$E(T \mid \tau) = (\tau + \exp(-\beta' \mathbf{x}))/(\theta - 1),$$

where τ is the time a person had been out of work at the time of sampling and T is the time before he finds a job. Morrison (1978) has shown that the gamma mixing distribution is the only one to produce a linear function of τ by this means. To a first approximation this result could be used to justify fitting a linear regression function of τ and x as in McGregor's case but the assumption of constant residual variance is violated because

$$\text{var}(T \mid \tau) = \left(\frac{\theta}{\theta - 2}\right)^2 \{E(T \mid \tau)\}^2.$$

The formula for $E(T \mid \tau)$ does, however, provide a convenient way of expressing the results of fitting the model. It also offers a simple graphical method of judging the suitability of the model and estimating θ.

Dr H. P. WYNN (Imperial College, London): I just want to underline some of the comments made by Mr Hyman relating to job history analysis. For some years I have thought that the key to describing the heterogeneity of the job market is the analysis of job histories. I am sorry that this paper did not make a contribution to this although it has made a valuable one to cross-section analysis. From the economic standpoint the heterogeneity is the basis for clarifying the continuing debate about whether unemployment is due to purely economic supply/demand aspects of the labour market or the structural nature of the labour force. In other words it is important to understand the *mismatch* between the supply of and demand for labour and try to identify the characteristics of the hard-to-employ people. This can teach us what good training policies are and hence make a genuine contribution towards the economic progress of the country.

G. J. A. STERN (ICL): The authors' theory is interesting and potentially very useful in analysing unemployment and other data, but so far they have applied far too little data to test it realistically. Thus, in 2.21 they consider a model with about ten coefficients, and in Table 3 they use one data set relating to 1972 to estimate the coefficients. The sizes of the standard errors make it clear that neither the form of the model nor the values of the coefficients have been established in a way which permits comments about actual data based on the model. This remark would be valid, I believe, even if they commented on 1972 data, but I am still more surprised that they feel able to say that the model shows that the level of unemployment benefits hardly affects levels of unemployment now, eight years later. Since 1972 the world has changed out of recognition in ways which are very relevant to this study, and make it difficult to take 1972-based theory to 1980. Obviously unemployment has more than doubled over previous levels, and estimates of coefficients are hardly reliable over such a range. Less obviously, we have seen a huge expansion in the employed non-working economy: state corporations which produce little that anyone wants even with considerable subsidy; a greatly expanded bureaucracy which in wide areas cannot contribute more than zero to the economy, although in fact its contribution is usually heavily negative through interference, complex taxes, etc. State interference and greatly increased public spending, and hence taxation, regardless of the nominal political complexion of governments, has led to, or at any rate coincided with, a hugely expanded "black economy" unknown to officialdom and still more unknown to statistics used by the authors. By its nature the size of the black economy is unknowable, but most of us are aware of a great expansion in unrecorded cash payments, barter, do-it-yourself, mutual repairs of each other's

watches, cars, roofs. I recall a large restaurant chain run almost entirely by illegal Turkish workers: few facts about them will be reflected in official figures. I would suggest that in 1972 the black economy may have been 10–25 per cent, where now it is 30–50 per cent of the economy, certainly if housewives' work is taken into account.

More data for more years and more countries and regions of countries would go some way to solving these problems. Problems of compatibility exist when moving from one country to another, but less so when moving from one region to another. At one time New York had some of the most generous welfare benefits and had enormous numbers out of work and immense social misery generally: would the various states of the USA be fruitful ground to apply the authors' theory? The question they raise of the impact of the level of unemployment benefits on unemployment levels is a very important one and needs more extensive research. My own view would be that I would welcome some of the measures promised by the Conservative Party when out of office, but not implemented when they get office, which would enable one to observe the results of lower levels of benefit. Is it visionary to suggest that to help determine this important matter, if, as seems to be the case, the present Government has not got the courage to do what it seemed to want to do, then they might at least implement such policies in one region? Controversies raised by matters such as the level of unemployment and other welfare benefits will never, one suspects, be settled by academic arguments, and indeed, as suggested above, analysis is possibly intrinsically not powerful enough to do so with present data. A sort of free-port low-welfare area in Great Britain, or better, several such at various levels and with various combinations of policies would do more to settle these questions than any amount of analysis on insufficient data.

May I conclude by saying that I hope that even if the authors regard my last suggestion as impractical, they will at least try and analyse more data. May I add too that what I have said represents my own views only and must not be taken as reflecting the views of my employers.

Mrs M. VENNING (Engineering Industry Training Board): I feel rather out of place in this gathering because I am merely a collector of data. However, I particularly welcome this paper and the various contributions that have been made in the Discussion because it has revealed to me the possibilities there are of other people using the data which we already have. I speak as from the Engineering Industry Training Board, which has conducted a series of manpower studies over the last 12 or 13 years and has career histories for craftsmen, professional engineers and for managers.

Our interest was primarily in trying to understand how people moved around between jobs as a guide to how we might intervene with training policies. But certainly in the craftsmen survey, we have information on periods during which respondents were unemployed. We would be delighted to make these data available to those people who have suggested that they might be further analysed.

Mr A. G. CARRUTHERS (Department of Employment): The interest to me in this paper lies in two directions. First, the methodology pioneering that it contains and, secondly, the degree to which it may help in the understanding of levels and trends in unemployment.

On methodology, the authors are to be congratulated on the pioneering effort which their work represents. It has broken new ground in using cross-section data from surveys, and data relating to individual people, rather than types of person to which work in this area has often had to have resort. Further, if I may say so, it is all presented thoughtfully, in modest and tentative fashion. Clearly, work of this kind cannot hope to conquer all at the first attempt. There are some problems, most of which are to be expected. One, obviously enough, is that of the missing variables because it means that it has not been possible to take into account the whole of the picture and all the factors at work. If some of these are important or influential, this limits the weight that can be attached to the conclusions. Another obvious and important factor which it has not been possible to take into account is redundancy payments. These seem very likely to be important since it is likely that they may encourage people to take their time in looking for an acceptable job. Another likely factor not covered is relative mobility which may have been changing in recent years. As has already been mentioned, longitudinal aspects may also have a bearing in analysing the problem.

Secondly, turning to the recent trends in unemployment and the light the work may shed on all this, it has been noted that the findings in the paper are consistent with work done by the Department of Employment, so there is no need to expand on this.

Finally, may I comment on the background to the study, with slightly different emphasis from that contained in the first paragraph of the paper. In that paragraph it is suggested that the rise in

male unemployment since the middle of 1973 can be accounted for almost wholly by the rise in duration of unemployment, with the flow having changed hardly at all. This leaves the implication that if it is possible to explain the increase in duration, it would then be possible to explain the rise in unemployment. In fact, there was a very marked rise in the inflow on to the register between the end of 1973 and the second half of 1975. That is clearly shown on a graph in the September 1976 issue of the *Department of Employment Gazette* in an article on flows. Over that period male unemployment rose by nearly 500 000. Duration of unemployment also increased, but it was not the sole contributor. Of course, this is to be expected on the down phase in the economic cycle when numbers joining the register increase and, at the same time, it takes longer to find a job. The increased inflow and the longer duration march side-by-side, and causal relationships must be regarded with care.

It might be noted that another observation of the down phase is probably beginning at the present time. The inflow figures in recent months show a marked increase compared with earlier in 1979.

Dr FRANK HANSFORD-MILLER (Inner London Education Authority): I would like, first, to congratulate tonight's authors on being brave men. Obviously, research that deals with human motivation is one of the most difficult areas. It reminds me of what S. W. Wooldridge (1956) once said about geography:

> "I have no doubt that it is in the field of human phenomena that it has its greatest opportunity, even if it there faces its greatest difficulties."

I believe this also applies today to Statistics.

Listening this evening, my mind goes back to the time of the General Election, when I was a candidate, and canvassing in Combe Martin, in the North Devon constituency. Many people there were very distressed at the large numbers of young people who they alleged had come down from Merseyside and Scotland and were spending their time on the beach in the sunshine, describing themselves to the employment exchange as "oil-rig fitters"—there being no oil rigs in that area—or of other similar unlikely occupations for that area. In other parts of North Devon I was told of men who had given up work because it paid them to do so. There is considerable concern about these matters among the people at large. In this connection the size of the family, with the increased benefits it can bring, is relevant, and in Table 3 we see, under "Dependents" that "one child raises the duration [of unemployment] by 9 per cent". It would be interesting to know the corresponding changes for two children, three, four, five and so on.

The fact must be faced that in the present situation unemployment can be a way of life—and is undoubtedly a way of life for many people. After all, this was to be an age of leisure, and what is unemployment but leisure if one's bodily needs and so on are satisfied, if one has sustenance, if one can look at television all day or go fishing and such like and if, as has been suggested, one has a little bit of the black economy? It should be realized that there is a psychological factor involved, and it is possible the replacement ratio should include a factor for the increase in personal benefit resulting from being unemployed—and for most individuals today there is no longer the old feeling of stigma at being unemployed. I think this is an important difference between today and earlier times.

There are two other points I should like to make. First, how far is lack of mobility a factor in the high figures of unemployment? There is a large number of council houses in the country—almost 50 per cent of the total houses—and people in them are more or less tied to that area. Particularly when there are large lay-offs, as in British Steel recently, no doubt many of those workers are in council housing. They cannot really move around as in previous times. In my boyhood, between the wars, there were so many more rented homes, and people could move around freely and rent a house somewhere else. In other parts of the world, for example in Australia, there is still a great deal of rented accommodation, which means that anyone can rent a house and go in search of work. Mobility is a most relevant factor. There is also the present difficulty in changing one's home if one owns it, because of several people having to be involved and all signing agreements at the same time—as one person moves out, another moves in, and so on. These kinds of difficulties must considerably increase the numbers of unemployed.

Secondly, it seems to me there is hardly any consideration in the paper of regional differences. One economic policy that has been consistently pursued since the war has been of development areas, assisted areas, and dispersion. Vast sums of money, billions of pounds, have been put into

the areas of so-called high unemployment. If those figures are studied today it will be seen that those billions of pounds have not brought down the unemployment. The whole country has just become a land of economic decline. Today we are in an economic war situation and I recall my studies during the last war. I believe the classic principles of war as enumerated by Von Clausewitz (1968) are highly relevant to us today. One of his major principles is to reinforce success. What we have done economically in the UK since the war has been the exact opposite. We have reinforced failure, and in so doing impoverished the successful areas. We have moved people and business away from the economically successful areas, and thus deliberately made those areas unsuccessful, in the hope that by so moving people and giving capital to the unsuccessful areas those areas will become successful. It has been a disastrous policy and no-one has benefited. I am on record many times as putting the alternative policy, which is to follow Von Clausewitz, and by reinforcing success make all the limbs prosperous as well. Even the outlying limbs will gather strength from the heart if that is in good health.

In this connection I think the difference in levels of wages in different parts of the country is highly relevant to unemployment. I discovered this in North Devon, where wages are low, but where there is the same rate of unemployment benefits. This means that what is called in the paper the "replacement ratio" is considerably higher in those low-wage areas than it is, say, in London or other major urban areas. This is because the surrounding level of wages is considerably lower, and it is of greater benefit, or it brings less hardship, for someone to be out of work in those areas.

Finally, on a point of enquiry, regarding the data in Table 3, no explanation is given why the data should include a double sample for Scotland.

The following contributions were received in writing after the meeting.

Professor G. CHAMBERLAIN (University of Wisconsin—Madison): The authors have stressed the important point that the distribution of a single spell does not allow us to distinguish heterogeneity from duration (or time) dependence. Any mixture can always be regarded as an alternative functional form for a homogeneous population. The authors respond to this point by cautiously interpreting particular functional forms. An alternative response is to determine what is identified, in the absence of particular functional forms. For example, simple search models predict an exponential distribution for duration. Can we determine if a distribution is consistent with some mixture of exponentials?

Say that we form intervals of equal length and consider the probability $G(k)$ that the spell of unemployment extends beyond the kth interval, $k = 1, \ldots, K$. We do not know the distribution of completed spell durations for those spells extending beyond the Kth interval. If we are observing a mixture of exponential distributions, then $G(k)$ is the kth moment of a distribution concentrated on the interval from zero to one. A necessary and sufficient condition for the $G(k)$ to be the first K moments of such a distribution can be obtained as follows: if K is odd, form the Hankel matrices A with i,j element equal to $G(i+j-1)$, $i,j = 1, \ldots, (K+1)/2$, and B with i,j element equal to $G(i+j-2)-G(i+j-1)$, $i,j = 1, \ldots, (K+1)/2$, where $G(0) = 1$; if K is even, form the Hankel matrices A with i,j element equal to $G(i+j-2)$, $i,j = 1, \ldots, (K+2)/2$, and B with i,j element equal to $G(i+j-1)-G(i+j)$, $i,j = 1, \ldots, K/2$. Then the necessary and sufficient condition is that A and B be positive definite or positive semi-definite. A necessary condition, which becomes sufficient as K tends to infinity, is that the successive differences of the $G(k)$ be non-negative. (Hausdorff; see the monograph by Karlin and Shapely, 1953.) These conditions can be the basis for a statistical test.

Professor J. CREEDY (University of Durham): This is a rather curious paper. In the Introduction it is stated that there are "two basic approaches" to the analysis of unemployment duration. The first, aggregate time series analysis, is rapidly dismissed in favour of cross-section analysis of individuals. This seems rather dogmatic, especially as the criticisms of the first approach apply also to the cross-section studies. However, as the authors themselves suggest in the conclusion, the analysis of longitudinal data is really the most "obvious" and potentially most fruitful approach.

The authors later refer to data obtained by sampling individuals, who enter unemployment at a given date, at discrete points in time, and suggest that such data are "not uncommon". This is certainly an exaggeration of the true position, especially in this country, and indeed the only example given is of a small DHSS survey which has not yet been completed. The problems of using discrete time periods are also well known.

The Introduction also promises, "an analysis of how the process determining unemployment duration may be modelled", whereas Section 2 simply begins with a general statement that the conditional probability of returning to work can be written as a function of "a set of variables" and the duration of the current incompleted spell. This is no substitute for specification of the process.

However, Section 1 of the paper does contain a potted survey of some considerations which may be relevant in specifying a model of the process of "leaving unemployment". Surprisingly the authors seem quite content to juxtapose conventional naive search theory with heterogeneous populations. It is also worth noting here that a sizeable proportion of individuals who leave the unemployment register do not actually return to work, but experience sickness or even leave the labour force. There is, however, no reference to this in the context of the use of cross-section data.

The discussion also contains no reference to the fact that unemployment benefit is not received unconditionally (contributions are required and benefits can be exhausted) and that benefits may be cut if the individual is thought to be placing "undue" restriction on the nature, hours, remuneration and location of the type of work he is willing to accept. Indeed, one of the main features of unemployment in recent years is that less than half of the unemployed are in receipt of unemployment benefit. A further important point is that unemployment is highly concentrated among a relatively small proportion of the population who experience repeated spells. It is very surprising to read in the conclusion that the authors, "regard the size of the effect of benefits as being now a rather firmly established parameter". This can only be interpreted as a statement of faith; the authors' "pessimism" does not seem to extend to their own results.

The identification problem discussed in Section 3 is of course well known, but it is perhaps worth repeating. The authors state that "Previous studies have analysed the variation in the completed durations of unemployment", and contrast this with their approach of considering unemployment, "as a probabilistic process". This considerable distortion of the literature is unfortunate, especially as very few references are given to the vast amount of useful work on unemployment flows which has been carried out in recent years, especially in the United States.

Dr ARNOLD KATZ (University of Pittsburgh): Lancaster and Nickell make a significant contribution to the methods of studying distributions of unemployment. Their approach is an artful attack on the problems of giving the dynamic sequential character of unemployment events a tractable empirical characterization. The negative conclusions which they come to about the prospects of separating state-dependence from unobserved sample heterogeneity will inevitably cause other investigators to exercise even greater caution in drawing policy inferences from cross-section data.

It may be worth noting that Lancaster model (2.20) imposes what could be potentially misleading restrictions on the relationships between re-employment probabilities and the exogeneous observed x characteristics of unemployed persons. The nature of the difficulty may be briefly indicated by considering the implications if the unemployed behave as recent job search theories predict. Such theories imply that re-employment probabilities depend on the differences between market wage offers and individual asking wages. Letting s stand for the conditional mean value of the differences between market offers and the asking wages for persons of a given set of x attributes, it may be shown that the β coefficients that Lancaster estimates are not constant, as (2.20) implies. Their predicted value $[f(s)/F(s)]\,s'$ depends instead on the probability density f and distribution function F of s and the s' change in s with respect to a particular x. Since f, in turn, depends on market conditions, Lancaster's coefficients seem likely to vary with the business cycle and other factors. Because Nickell's model (2.21) is a close approximation, it is subject to the same difficulty. The variability in the coefficients may be particularly important for the replacement ratio influences, where changes over the cycle matter for policy purposes.

This difficulty is a characteristic of the specific models and not the general methodology. The methods Lancaster and Nickell propose invite comparisons with the recent and very similar work of Kiefer and Neumann (1). The Kiefer–Neumann study is more comprehensive because it has been adapted to samples with observations of re-employment wages. On the other hand, they treat periods of unemployment as a discrete variable, losing some of the flexibility and perhaps reliability of the continuous models that Lancaster and Nickell propose. Taken together, however, the work of both sets of authors promises to be highly influential in future research on unemployment issues.

Professor N. M. KIEFER (University of Chicago): The authors survey succinctly the statistical issues involved in the analysis of unemployment duration data and conclude pessimistically that the situation is bleak. I would like to argue with the conclusion, for two reasons: One is that the identification problem is not new or unfamiliar to economists; the role of economic theory is to provide guidance in making identifying assumptions. The second is that different data configurations provide information on these and even more complex questions.

A theoretical underpinning for the analysis of unemployment duration is provided by the theory of job search, as treated in classic papers by McCall (1970) and Mortensen (1970). The model has been elaborated variously without changing the nature of the basic insights. The important aspect of this work from an empirical point of view is the provision of a *behavioural* model of unemployment duration, as opposed to the purely descriptive models described by Lancaster and Nickell. The behavioural model allows consideration of issues which cannot be dealt with in the descriptive model. In particular, behavioural coefficients have direct economic interpretations, so alternative identifying assumptions can be compared in terms of their economic implications. The situation is directly analogous to the problem of identifying a structure in the familiar linear simultaneous equations model. Typically a "structure" $By+\Gamma x = u$ is specified but only the "reduced form" coefficients $\Pi = B^{-1}\Gamma$ are identified (e.g. from regressing y on x). The decomposition of Π into B and Γ is made on the basis of identifying restrictions which usually fix certain elements of B and Γ. Economically implausible values for other elements of B and Γ cast doubt on the validity of the identifying restrictions. Sorting duration dependence in aggregate hazard functions into heterogeneity and duration dependence in individual hazard functions is a similar process. Structural or behavioural models of unemployment duration are discussed and estimated in Kiefer and Neumann (1979a, b). The latter paper presents a model in which heterogeneity and duration dependence are distinguished, under certain identifying assumptions.

Lancaster and Nickell and I agree about what the problems are. It is true that a knowledge of a density which is a mixture of other densities does not generally identify those densities without further assumptions. Nonetheless, assumptions should be made, and their consequences explored. This is typically somewhat easier to do in a behavioural model than in a descriptive model, which fact may explain the authors' pessimism.

Finally, data on repeated spells of unemployment are becoming available for some samples in the US; these data permit the use of far more robust methods for sorting out heterogeneity from duration dependence.

The AUTHORS replied later, in writing, as follows.

The basic question which we all wish to elucidate concerns the extent to which the recent rise in unemployment is a supply as opposed to a demand side phenomenon. When we say that the rise can be "accounted for" by the increase in unemployment duration this, of course, says nothing about this basic question. In the terminology of Mr Thatcher's example, the time to buy tickets can increase either because of a reduction in the number of ticket clerks or because of a general increase in the desire of individuals to lounge about in queues for tickets. Before going on to discuss the nub of Mr Thatcher's contribution it is worth noting one important point. Contrary to his argument, an increase in the supply of labour should not lead to any permanent increase in unemployment if the macro-economy is functioning properly. Thus, when National Service was done away with, there must have been a considerable increase in the supply of labour but little or no impact on unemployment in that period.

The core of Mr Thatcher's analysis is the discussion surrounding his table of changes in labour supply, employment and unemployment. The procedure is as follows. The supply of labour is equal to the labour force. The demand for labour is equated to employment. But the difference between these is equal to unemployment since if one is in the labour force and not in employment, one must presumably be unemployed. (There may, of course, be other categories such as self-employed but their existence does not affect the structure of the argument.) So if L_1, U_1, N_1 are labour supply, unemployment and employment in 1971 and L_6, U_6, N_6 are the same in 1976 we have, by definition,

$$L_1 = U_1+N_1, \quad L_6 = U_6+N_6$$

and hence

$$U_6-U_1 = (L_6-L_1)-(N_6-N_1).$$

Thus the change in unemployment must, by definition, be entirely explained by the change in supply less the change in "demand". There are now two points to be made. First, the equating of the demand for labour with employment is clearly invalid since this would imply that employers always have precisely the number of workers they want. Fluctuations in vacancies and the CBI skilled labour shortage indicators could hardly exist under such circumstances. In other words, equating employment with demand rules out entirely any possibility of a supply side explanation of unemployment. For example, even if all the labour force voluntarily withdrew their labour and became unemployed, employment and hence "demand" would drop to zero and the consequent 100 per cent unemployment would be entirely explained by the "demand" side collapse.

Second, however, the numbers in Mr Thatcher's table do not seem to add up as they should do if the labour force were only divided into employed or unemployed. This unexplained number clearly, therefore, implies that there must be some other missing category whose change is not presented in the table. The fact that the increase in unemployment is larger than the change in labour force less the change in employment tells us nothing about benefit or any other supply side effects on unemployment. It merely tells us either that the numbers are incorrect or that the excess has emerged from some other category not recorded in the table. It is not, therefore, clear to us how such information can possibly answer the basic question with which we started out.

The footnote to which Mr Burman refers was in error—the individuals were selected from the unemployed and then interviewed once later. Daniel's original analysis (1974) also exploited the fact that in the time between selection and interview some people found jobs. The "replacement ratio" is the ratio of income out of work to income in work, Lancaster assumed this constant though the reservation wage was not assumed constant. The Daniel data does contain some information on asking wages and an analysis of this is under way.

Our scepticism is far from total but refers only to the possibility of distinguishing genuine from spurious effects in estimating the *time variation* in the probability of leaving unemployment due to the apparent arbitrariness of assumptions about the distribution of the individual effect, v in the notation of (2.8). We are, in contrast, reasonably confident about the estimates of the effects of $\mathbf{x}$ variables on the *mean* duration of unemployment since these are likely to be nearly independent of assumptions about the distribution of v as long as v is distributed independently of $\mathbf{x}$. Indeed in the model of Lancaster such estimates are exactly independent of the distribution of v. To see this note that $E(T|\mathbf{x}) = E\{E(T|v,\mathbf{x})\}$ where the outer expectation is with respect to v. But with the model of (2.20) $E(T|v,\mathbf{x}) \propto v^{-1/\alpha}$ so that $\delta \log E(T|\mathbf{x})/\delta x_i$ does not depend upon the distribution of v. More generally, assumptions about the precise way in which the hazard varies over time seem to be in effect assumptions about the higher moment of the distribution of unemployment duration and point estimates of regression effects, i.e. of shifts in the mean, should be rather insensitive to such details.

We agree with Mr Carruthers that there are significant variations in the rate of entry to the register of unemployed. Our comments on the *secular* stability of the inflow are simply intended to draw attention to the fact that roughly the same numbers flow onto the register now as did in 1967 whereas unemployment has more than doubled. This indicates that the vast *secular* changes in the level of unemployment correspond to equally vast *secular* changes in duration accompanied by no *secular* change in the inflow.

With regard to redundancy payments, while they are obviously important to those who receive them, the proportion of entrants into unemployment who do actually receive them is less than 15 per cent and hence they are not very significant in the aggregate picture.

Regarding the contribution by Mrs Venning and the Engineering Industry Training Board's generous offer of data—we accept.

The models referred to by Mr Hyman really require that we know more of the individual's labour market history than his current unemployment spell. Whatever the data, it seems to us sensible to model directly the sequence of transition probabilities as we indicated in Section 1. The recent work by mathematical sociologists referred to by Mr Hyman and Professor Bartholomew is clearly relevant here.

We agree with Dr Wynn and several other speakers that repeated observations on the same individuals will be much more informative than cross-sectional data.

Professor Creedy wonders about other samples of entrants to unemployment. The data of McKay and Reid's (1972) SSRC-financed study and of Daniel's earlier study (1972) come to mind, though the former data are unhappily not available to other investigators. The DHSS study

involves 2300 individuals which is scarcely small. Regarding the brevity of our discussion of the economics of the process we made it clear that the present paper was intended to emphasize the statistical problems. It is true that many individuals who leave the register do not return to work. Our samples consisted of unemployed individuals who were actively seeking work and doubtless some of these would eventually leave the register for reasons other than return to employment. However, our data are uninformative about these events and, moreover, it is always possible to interpret our duration distributions and transition probabilities as being conditional on movement when it occurs, being to a new job. Professor Creedy's qualification is well worth making though. We have discussed those effects we think reasonably firmly established and those which are not in our reply to Mr Burman. With regard to Professor Creedy's final points we regret spending much time on a "well-known" problem though a reference or two would have made his remark somewhat more conclusive. Finally we are well aware of the other literature on unemployment flows—we were rather hoping to improve upon it.

Mr Stern's comments are wide ranging, to say the least. One or two points are worth making. His suggested figures for the increase in the percentage of the "black" economy indicates that true GNP in Britain has been growing at over 4 per cent throughout the 1970's which should reassure those of us who are worried about our declining rate of productivity growth. Work of the kind he suggests has already been applied to cross section data from the various states in the USA with very similar results to our own. Finally, the kind of social experiments he suggests have been carried out in a limited form in the USA in the form of the various negative income tax experiments. The results, in so far as they can be said to be conclusive, are of the same order of magnitude as our own.

Dr Hansford-Miller's point on the difficulties of moving from one region to another experienced by council house tenants is well taken. In some other work by the second author, he finds that controlling for income, socio-economic status and other relevant factors, a council house tenant is more likely to be unemployed than similar individuals who live in other kinds of accommodation. This need not necessarily reflect immobility, however, since it could be the case that council estates are concentrated in areas of low labour demand.

We agree with Professor Katz that the wage offer distribution can only be thought of as fixed for relatively short time periods.

We were glad to see a contribution from Professor Kiefer referring to his own interesting recent work with Professor Neumann.

We are agreed on the necessity of using all the restrictions implied by the economic theory. The difficulty is that restrictions on the form of the distribution of individual effects are not implied by any economic theory we know of. Data on repeated spells of unemployment and employment currently being analysed both in the US and in Britain will clearly be much more informative about the determinants of labour market transitions, as will data arising out of repeated interviews of individuals within a spell of unemployment. Both types of data pose new and fascinating problems of econometric method. We are a little unhappy about the categorization of our models as "descriptive". It is not clear that reduced form models such as ours deserve such an epithet for they, like those of Kiefer, are founded on optimal job search theory. Perhaps the difference arises out of our major concern with the way in which re-employment probabilities vary over time. Behavioural restrictions on the dynamics of the job search process are harder to come by than they are in stationary, infinite horizon, models.

Professor Chamberlain's remarks are interesting and we look forward to seeing them developed. Presumably they do not extend to mixtures of Wiebull distributions homogeneous with respect to some unknown α in the notation of (2.20), for non-integer moments then arise.

Finally, we thank all contributors for their comments.

References in the Discussion

DANIEL, W. W. (1972). Whatever happened to the workers in Woolwich? A survey of redundancy in S.E. London. London: P.E.P. (Broadsheet no. 537).

—— (1974). A national survey of the unemployed. London: P.E.P. (Broadsheet no. 546).

DEPARTMENT OF EMPLOYMENT GAZETTE (1976). Unemployment and notified vacancies—flow statistics. September 1976, 976–979.

GINSBERG, R. B. (1971). Semi-Markov processes and mobility. *J. Math.*, Social I, 233–263.

—— (1973). Stochastic models of residential and geographic mobility for a heterogeneous population. *Environment and Planning*, **5**, 113–124.

GOODRICH, R. L. and CAINES, P. E. (1979). Linear system identification from nonstationary cross-sectional data. *IEE Trans. Autom. Control*, **AC-24**, No. 3, 403–411.

HYMAN, G. (1974). Cumulative inertia and the problem of heterogeneity in the analysis of geographic mobility. Centre for Environmental Studies, Report CES RP 11.

HYMAN, G. and GLEAVE, D. (1977). A reasonable theory of migration. *Trans. Inst. Brit. Geogr.*, 3(2), 179–201.

HYMAN, G. and PALMER, D. (1978). A regional analysis of the dependency between registered unemployment and the stock and flows of notified vacancies. Centre for Environmental Studies, Report CES WN 471. Also in *Environment and Planning* **10**, 853–866.

—— (1979). A causal analysis of unemployment and vacancies in British regions 1969–74). Centre for Environmental Studies, Report CES WN 535. Also *Regional Studies* (to appear).

KARLIN, S. and SHAPELY, L. S. (1953). *Geometry of Moment Spaces.* American Mathematical Society, Monograph No. 12, pp. 50, 55.

KIEFER, N. M. and NEUMANN, G. R. (1979a). An empirical job search model with a test of the constant reservation wage hypothesis. *J. Polit. Econ.*, **87**(1), 89–107.

—— (1979b). Individual effects in a non-linear model: explicit treatment of heterogeneity in the empirical job-search model. University of Chicago.

MCCALL, J. (1970). Economics of information and job search. *Quart. J. Econ.*, **84**, 113–126.

MCFARLAND, D. D. (1970). Intergenerational social mobility as a Markov process (including a time stationary Markovian model that explains observed declines in mobility rates over time). *Ann. Sociol. Rev.*, **33**, 712–722.

MCGREGOR, A. (1978). Unemployment duration and re-employment probability. *Econ. J.*, **88**, 693–706.

MCKAY, D. I. and REID, G. L. (1972). Redundancy, unemployment and manpower policy. *Econ. J.*, **82**, 1256–1272.

MORRISON, D. G. (1978). On linearly increasing mean residual life-times. *J. Appl. Prob.*, **15**, 617–620.

MORTENSEN, D. T. (1970). Job search, the duration of unemployment and the Phillips curve. *Amer. Econ. Rev.*, **60**, 847–862.

MYERS, G. C., MCGINNIS, R. and MASNICK, G. (1967). The duration of residence approach to a dynamic stochastic model of internal migration: a test of the axiom of cumulative inertia. *Eugenics Quart.*, **14**, 121–126.

THATCHER, A. R. (1979). Labour supply and employment trends. In *De-industrialisation* (F. Blackaby, ed.). London: Heinemann.

TUMA, N. (1976). Rewards, resources and the rate of mobility: a non-stationary stochastic model. *Amer. Soc. Rev.*, **41**, 338–360.

VON CLAUSEWITZ, K. (1968). *On War* (A. Rapaport, ed.; J. Graham, translator). London: Penguin.

WOOLDRIDGE, S. W. (1956). The geographer as scientist, p. 18. Inaugural lecture, Birkbeck College, University of London.

As a result of the ballot held during the meeting the following were elected Fellows of the Society.

ABOABA, Fadlul O.
ADENA, Michael A.
ARNOLD, Jonathan
BATES, Douglas M.
BIRTWHISTLE, Andrew
BRAAK, Carl J. F.
BROWN, Morton B.
BUDGETT, David A.
DAFFIN, Christopher
DE SILVA, Niroma H.
DONNELLY, Kevin P.
EAMES, Margaret A.
FRANKS, Caroline R.
GOODALL, Colin R.
GULEZIAN, Ronald C.
HAWKINS, Michael M.
IRVINE, John M.
JONES, Timothy C.
LEE, Martin L.
MARLES, Vivien A.
MARQUES, Victor A.
MEHTA, Mrs Amarjeet K.
OKELL, Eric G.
OSMOND, Clive
PARSONS, Celia M.
PESMATZOGLOU, Michael
RICHARDSON, Sylvia T.
RUTHERFORD, Gilbert S. W.
SIMPSON, Christine
THIVVIANESAN, Chelliah
UNDERWOOD, David D.