NBER WORKING PAPER SERIES

HOW LONG IS A SPELL OF UNEMPLOYMENT?: ILLUSIONS AND BIASES IN THE USE OF CPS DATA

Nicholas M. Kiefer

Shelly J. Lundberg

George R. Neumann

Working Paper No. 1467

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 1984

Cornell University and Economics Research Center/NORC; University of Washington and NBER; and University of Iowa and Economics Research Center/NORC, respectively. We have benefited from comments by Dale Mortensen, Lars Muus, Richard Startz, Neils Westergard-Nielsen, and Chris Winship. The usual disclaimers apply. This research has been supported by grants from the National Science Foundation and the National Commission on Employment Policy. Research assistance by Beth Asch and Douglas MacIntosh is gratefully acknowledged. The research reported here is part of the NBER's research program in Labor Studies. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #1467
September 1984

How Long is a Spell of Unemployment?: Illusions and Biases in the Use of CPS Data

ABSTRACT

Most data used to study the durations of unemployment spells come from the Current Population Survey, which is a point-in-time survey and gives an incomplete picture of the underlying duration distribution. We introduce a new sample of completed unemployment spells obtained from panel data and apply CPS sampling and reporting techniques to replicate the type of data used by other researchers. Predicted duration distributions derived from this CPS-like data are then compared to the actual distribution. We conclude that the best inferences that can be made about unemployment durations using CPS-like data are seriously biased.

Nicholas M. Kiefer
Department of Economics
Cornell University
Ithaca, NY 14853

(607) 256-6315

Shelly J. Lundberg
Department of Economics
University of Washington
Seattle, WA 98195

(206) 543-6149

George R. Neumann
Department of Economics
College of Business Administration
University of Iowa
Iowa City, IA 52242

(319) 353-4549

I. INTRODUCTION

Despite decades of intense interest in the sources and nature of unemployment, the answer to the simple question, "How long do spells of unemployment last?" has remained elusive. This is rather surprising since economists are often tempted to interpret short spells of unemployment as a frictional or normal component of unemployment about which policy-makers need not be concerned, and long-term or chronic unemployment as a disequilibrium component embodying most of its socially wasteful aspects. Hall (1970) and Marston (1976) have pointed out that this is not a particularly useful distinction, since the high unemployment rates of many disadvantaged groups seem to be related to employment instability or frequent unemployment spells rather than lengthy individual spells. Nevertheless, the observation that the average spell of unemployment is relatively short has been used to support the position that most unemployment is essentially "voluntary." Conversely, the inability of some individuals to locate a job over a period of several months is considered to be evidence of labor market malfunctioning. As a prelude to analyzing changes in the ratio of long-term to short-term unemployment, Green *et al.* state in the *Monthly Labor Review* that

> The length of time that workers remain unemployed is an important indication of the severity of the Nation's unemployment problem.

The duration data used to study unemployment comes, almost exclusively, from the Census Bureau's monthly current Population Survey (CPS), the only representative sample obtained on a continuing basis. Despite the attractive features of CPS data that make it an invaluable tool in many research applications, the data cannot be used to answer even rather simple questions concerning unemployment durations. The CPS is a point-in-time survey and thus cannot measure important aspects of phenomena which persist over time. For example, the CPS records as duration of unemployment the average duration of unemployment spells in progress as of the survey date, rather than the duration of completed spells. Moreover, as Kaitz (1970) has noted, the point-in-time sampling of the CPS understates the frequency of short spells of unemployment since longer spells are more likely to be recorded in the survey.

In spite of these difficulties, CPS data has been "the only game in town" for analyzing unemployment durations, and thus the trend in recent work has been to glean more information from the CPS by exploiting the quasi-panel nature of the data - an individual is in the sample for four months, out for eight, and then in again for four months - and by imposing additional restrictions on the underlying distribution of spell durations. These restrictions are then used to compute descriptive statistics of the duration distribution, typically, the mean and the upper deciles. These statistics, in turn, form the basis for almost all informed discussion regarding the duration of unemployment spells.

Obviously these empirical characterizations of unemployment spells are only as good as the statistical assumptions they are based on; yet within the confines of CPS data there is no way to check their validity. Providing such a check is the major purpose of this paper.

We depart from the current trend by introducing a new sample of completed unemployment spells obtained from panel data. We apply to this data CPS sampling and reporting techniques, including the censoring of long spells, to replicate the type of data used by other researchers. We then apply conventional statistical analysis to the CPS-like data and to censored versions of the actual duration data, and use the results of that analysis to characterize the duration distribution as researchers before us have done. Comparisons of predicted distributions with the actual are then made for several statistical models, applied to five different sets of duration data. In this way we distinguish the impact of point-in-time sampling from that of censored reporting of durations, since reporting techniques can be altered more readily.

Our results suggest that point-in-time sampling, by introducing considerable error into the measurement of unemployment durations, leads to unreliable estimates of the duration distribution. When data derived from continuous monitoring is used, both applying more general parametric forms and employing more information on very long spells lead to significantly better predictions. This is not the case for CPS-like data; censored data performs nearly as well as uncensored data, and the simplest parametric forms frequently provide the best fits. The best inferences that can be made about the actual distribution using CPS-like data are seriously biased. Mean durations are underestimated by 2 to 3 weeks, and the fraction of spells lasting more than 28 weeks or more than 52 weeks are substantially understated. Based on these findings, we are skeptical about the value of CPS duration data in the study of unemployment.

II. ESTIMATING UNEMPLOYMENT DURATION FROM CPS DATA

Inferences about completed spells of unemployment depend crucially upon the type of data that are available. We begin by using "perfect" data to characterize the duration distribution. Assume that a homogeneous population has been continuously monitored. At each moment of time, the number of persons with elapsed duration in unemployment s is known. Denoting calendar time by t, let the distribution of exit times from unemployment spells started at time t be given by $F_t(s)$.

Suppose that a particular date, T_0, is chosen to analyze the unemployment records, and that we wish to estimate the distribution of completed unemployment durations. Denote the flow into unemployment at date t as p(t). The probability that an individual who entered the unemployed state at date $t < T_o$ is still unemployed at date T_o is $1 - F_t(T_o - t) = 1 - F_t(s)$. The density of elapsed durations at the interview date T_o is:

$$(1) \qquad h(s;T_o) = \frac{p(T_o - s)\ [1 - F_{T_o - s}(s)]}{\int_{-\infty}^{T_o} p(u)[1 - F_u[T_o - u]]du}$$

In general, this density will depend upon the entire previous history of the process, as the denominator of equation (1) makes plain.

Can we estimate the time-dependent distribution functions, $F_t(s)$? Clearly one cross-sectional "snap-shot" alone is not sufficient, even with p(t) known. However, with "perfect" data we can obtain estimates of $F_t(s)$ for a particular t by observing the durations of unemployment spells started at time t, which, by assumption, are observed over time.[1] In this manner it is possible to produce estimates of duration distributions that differ by time period and to compare unemployment behavior in booms and in troughs.

In practice, researchers have typically restricted the general form of (1). Two approaches are possible in view of available data — the first based on cross-section data and the second on gross flows. If only a "snapshot" of elapsed durations is available, but the p(t)'s are known, one can estimate the parameters of the distribution of exit times provided that the distribution is known up to a vector of parameters and that it is unchanging over time.[2] This is known as the synthetic cohort method. In this case we have

$$(1') \qquad h(s;T_o) = \frac{p(T_o - s)\ [1-F(s;\phi)]}{\int_{-\infty}^{T_o} p(u)\ [1-F(T_o - u;\phi)]du}$$

where by $F(s;\phi)$ we indicate specifically the requirement that F be known up to the parameter vector ϕ.

Estimates of ϕ can be used to calculate functions of ϕ such as the average duration of a completed spell, or the upper percentage points of F. The major advantage of this method is that it allows one to deal explicitly with non-stationarity. Indeed, if ϕ contains an element related to demand pressure, it is feasible to examine whether the exit distribution varies with the business cycle.

Despite the attractiveness of this specification, it rarely has been implemented in the economics literature. Unpublished papers by Luckett (1978) and Smith (1982) and the article by Bowers and Harkness (1979) are the only studies that discuss or implement this approach, to our knowledge.[3]

The more common method of using cross-section data is to assume that stochastic process generating the exit times is Markov or semi-Markov. The ergodic property of regular Markov processes implies that the inflow p(t) converges to $\bar{p}$ and thus (1') converges to

$$(1'') \qquad h(s) = \frac{\bar{p}[1 - F(s)]}{\bar{p}\int_{-\infty}^{T_o} [1 - F(T_o - u)]du} = \frac{1 - F(s)}{\bar{D}}$$

where $\bar{D} = \int_0^\infty [1-F(s)]ds$ is the mean of completed spells. This is the result of Kaitz (1970) and Salant (1977), though obtained in a different manner.[4] Typically, a parametric form of F is chosen such that $1-F(s)$ and $\bar{D}$ can be expressed as functions of a few parameters. It should be stressed that (1'') holds only in the stationary state and that even relatively minor departures from stationary inflows can have large consequences for the estimates. In particular, reporting a time series of estimated functions which are based on period-by-period stationarity assumptions is likely to be of little value.[5]

The second approach to estimating the distribution of exit times from unemployment is available when entry and possibly exit times are known for all individuals.[6] If all entry and exit times are known we are in the "perfect" data case described above. CPS data have only a quasi-panel structure; when individuals are observed for four months, only some of the labor market flows will be observed for each cohort. However, aggregate flow data for successive periods allows one to calculate the empirical hazard rate:

$$(2) \qquad \lambda_{T_o}(s) \cdot \Delta h = (U[s,T_o] - U[s + \Delta h, T_o + \Delta h])/U[s,T_o]$$

where U[s,t] is the stock of unemployed with elapsed duration s at time t and Δh is the time between samples. The fundamental relationship between hazard functions and distribution functions

$$(3) \qquad a) \qquad \lambda(s) = f(s)/(1-F(s)) \qquad \forall s$$

$$b) \qquad F(s) = 1 - \exp\left\{- \int_0^s \lambda(x)dx\right\}$$

provides the link between observed behavior and the hypothesized underlying distribution of completed spell lengths (Kalbfleisch and Prentice (1980)). In practice, $\lambda(s)$ is not observed for all values of s. Some simple curves are therefore fit to the data and the complete distribution is inferred from these estimates. In this regard, it is worth remarking that non-negativity throughout the range of s is an essential requirement for any hypothesized hazard function. Moreover, one typically requires that

$$(4) \qquad \lim_{s \to \infty} \int_0^s \lambda(x)\, dx = \infty$$

in order that the distribution function in (3) b) be non-defective. While defective distributions can be of use in some problems, e.g., mover-stayer models, our intended purpose of estimating hazard rates - identifying the distribution of durations of completed spells of unemployment - argues strongly against considering distributions with $F(\infty) < 1$. In other words, some hazard functions which appear to fit the data well may be objectionable on an a priori basis.

In principle, estimated parameters of the exit time distribution obtained from gross flow methods should be identical to estimates obtained from the point-in-time sampling methods, provided the specification is correct. Both approaches allow for non-stationary behavior in the entrance rates into unemployment, and while neither can encompass pure cohort effects on the exit distribution, each can treat fluctuation in the exit distribution as a function of aggregate economic conditions. The major advantage of gross flow data is that it allows one to ignore the relatively cumbersome weighting shown in (1').

Despite the close relationship between the sample data function — either the elapsed duration distribution in the point-in-time sample or the hazard function in the gross flow approach — and the distribution of completed spells, there are serious problems associated with using CPS data in either of its manifestations to estimate duration distributions. A primary difficulty is that the reported data are presented as grouped data in intervals of 0 - 4, 5 - 6, 7 - 10, 11 - 14, 15 - 26, or more than 27 weeks of elapsed unemployment. As these categories are not integer multiples of the sampling window (one month), it is not possible to obtain the exact weight, p(t), needed to calculate (1') in the synthetic cohort approach.[7]

Grouping of the data also causes difficulties for the gross flow approach. Since interviews are one month apart and we do not know when during the month the transition out of unemployment occurred, actual completed durations lie in the intervals 0 - 8.3, 5 - 10.3, 7 - 14.3, 11 - 18.3, 15 - 30.3 and 27^{+} weeks. As the overlapping of these intervals suggest, we know completed durations with very little precision from gross flow data.[8]

Other difficulties follow from the CPS reporting procedures. Usable data are censored at about six months. Only five of the reported intervals can be used to fit hazard functions, since the longer spells of unemployment in the open-ended final category must be excluded. Also, since many short-spells of unemployment are missed altogether — the length bias problem in sampling discussed by Kaitz (1970) — the fit of the estimated distribution depends upon behavior in both the lower and upper tails being consistent with the assumed functional form.

Despite the frailties of the CPS data for estimating duration distributions, it is currently the only sample with representative coverage of all labor market participants, and is thus potentially the most informative. It is therefore very important to know whether, despite all these problems, the duration distributions estimated from CPS data are reasonably accurate.

III. THE DENVER INCOME MAINTENANCE EXPERIMENT DATA

Our new sample of completed unemployment spells comes from surveys conducted by the Denver Income Maintenance Experiment (DIME). A principal advantage of these data is that they provide a continuous employment history for each individual in the sample for up to 48 months. Thus the completed duration of each unemployment spell is known, with the exception of a very small number of spells which overlap the beginning or end of the sample period. The periodic interviews are retrospective, so we have accurate information on the timing of each transition into and out of employment.

The DIME sample has been discussed extensively elsewhere,[9] as have the employment histories constructed from the basic Public Use Files.[10] A few characteristics of the data, however, are particularly relevant to the application at hand since they limit the comparability of our results with those of CPS-based studies. Specifically, the sample selection criteria for DIME families and the imprecision of recorded transitions between unemployment and nonparticipation in employment interviews may limit our ability to draw conclusions about the general population from this study.

1. Sample Selection

The CPS sample is representative of the U. S. population; the DIME sample is not. The experiment was designed to measure the effects of a negative income tax on labor supply. In order that the sample correspond as closely as possible to the target population of a future NIT program, the 2,657 families initially enrolled in the experiment were required to satisfy a number of eligibility requirements. These included restrictions concerning age, family structure, and ability to work but, most importantly, all families with pre-experiment "normal" earnings above a specified level were excluded. Eligible families were allocated in a non-random fashion among eleven different financial treatments and a control group containing 40% of the total sample.

The most obvious problem in using this sample to study unemployment is that the support guarantees and varying tax rates faced by financial treatment families distort labor supply decisions. This is easily avoided, however, by restricting the analysis, as we do, to the control group, which received only token payments for reporting their work histories.

The earnings truncation of the DIME sample presents a more serious difficulty. We would expect families who experience lower incomes prior to the experiment to contain members with an unusual propensity to experience long or frequent spells of unemployment. It is not obvious how restricting our analysis to individuals in low income families will affect the distribution of unemployment spell durations, but we might expect long spells to be more prevalent than in the population as a whole.

2. The Unemployment-Nonparticipation Distinction

In general, the information available in the DIME Public Use Files permits individuals to be classified as employed, unemployed, or nonparticipating according to the CPS definitions of these labor market states. However, a problem concerning the identification of transitions between unemployment and nonparticipation arose during the construction of employment histories. The data were collected via ten periodic interviews, which were administered at intervals of three to four months. At each interview all spells of employment and non-employment since the last interview were identified. After this, a series of questions relating to each spell were asked, including probes for search activity during each spell of non-employment.

This procedure had two unfortunate results. Between periodic interviews, any transitions between unemployment and nonparticipation are not identified. As a consequence, transition rates between these two states will be underestimated.[11] Secondly, all observed transitions between these states will occur during months in which periodic interviews took place, and will be placed at the end of a month in the raw data. There will thus be some error in timing of recorded transitions. These problems suggest that some caution be exercised in interpreting results from DIME data but, given the deficiencies of CPS data, do not seem to warrant abandoning the use of such special samples.[12]

3. Continuous Monitoring and CPS-like Unemployment Data

To perform the comparisons described above, we first constructed the "true" distribution of completed unemployment spells. The middle years of the sample, 1972 and 1973, were used as the sampling frame. All spells of unemployment occurring during this period were included. A small fraction of these spells were censored because of sample attrition - about 1.5% for adult men and less than 1% for women - but, in the main, the sample consists of about 700 completed spells of unemployment for men and women over twenty-one years of age.

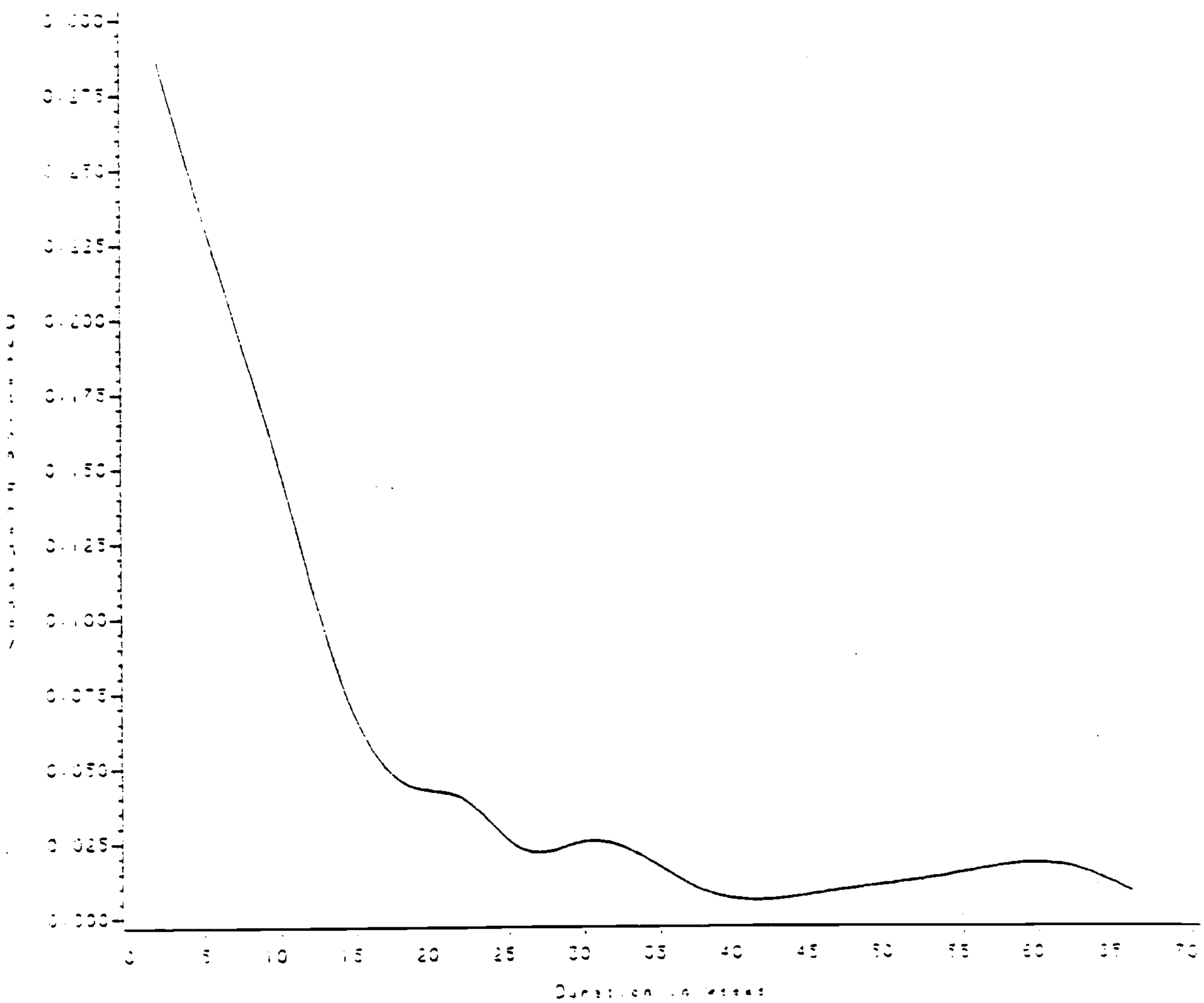
Figures 1 and 2 display the sample density functions of unemployment duration for adult men and women. These graphs are constructed by fitting a cubic spline function through the observed frequencies to obtain smoothness. Some relevant characteristics of the sample distributions are given in Table 1.

Both male and female distributions exhibit skewness of the sort commonly thought to be characteristic of unemployment: most spells are quite short but the distribution has a long skew to the right.[13]. For example, in this sample the mean is 78% higher than the median for men, and 45% higher for women. For men the sample density function looks rather like an exponential, at least up to about the 20^{th} or 25^{th} week, while the density of completed spell lengths for women does not resemble any textbook form of which we are aware.

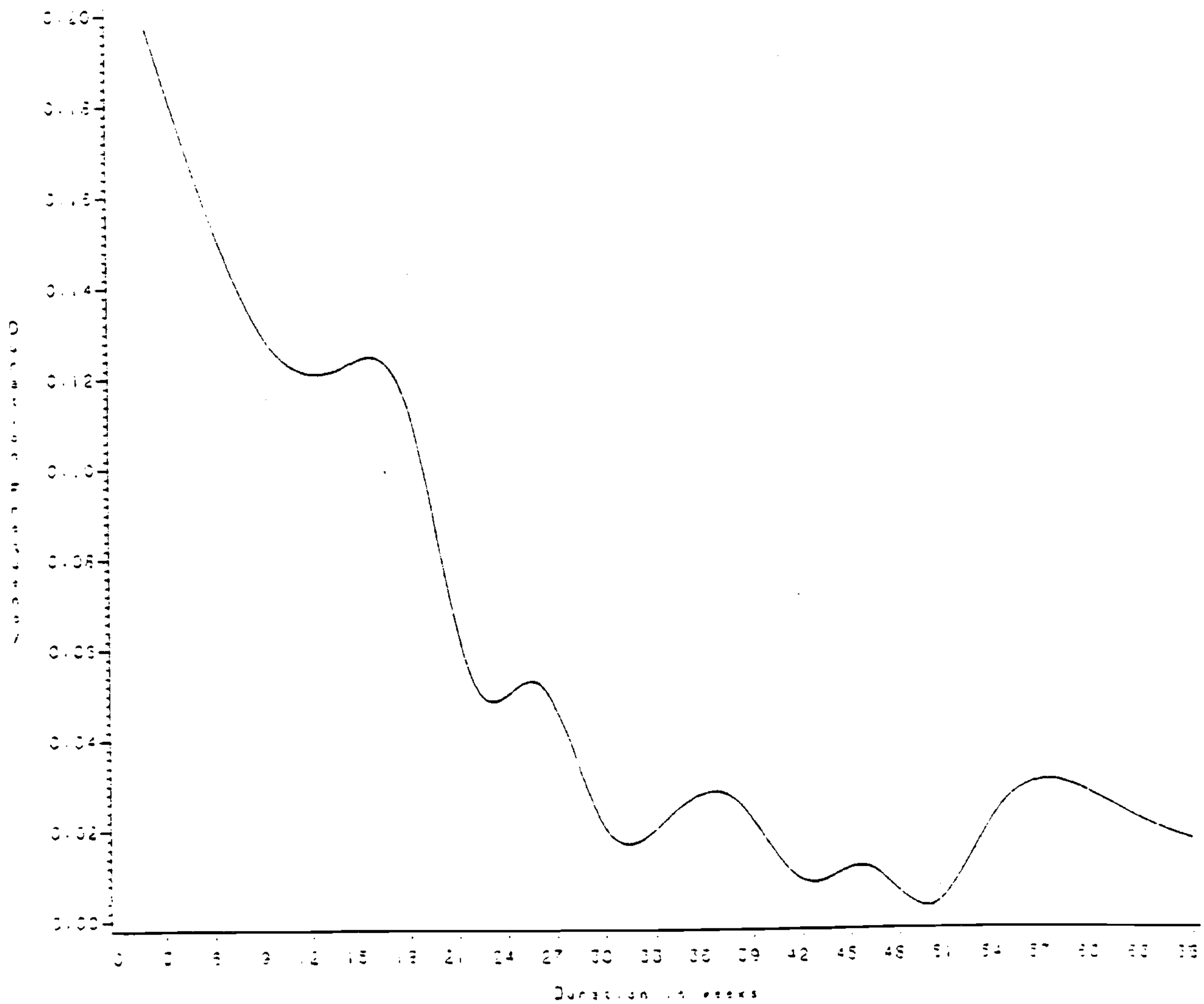
Table 1

Characteristics of Unemployment
Duration Data from DIME

	Adult Males	Adult Females
Mean (weeks)	14.2	18.4
Median (weeks)	8.0	12.7
Proportion of spells		
< 8 weeks	51.0	35.3
< 28 weeks	84.5	81.9
< 52 weeks	90.7	91.3
Number of observations	355	343



Actual Completed Spell Unemployment
Duration Frequencies, Adult Males



Actual Completed Spell Unemployment
Duration Frequencies, Adult Females

From this sample of all spells of unemployment we constructed hazard functions after grouping the data, separately by sex, into four-week intervals. These are reported in Appendix A. Hazard functions could be estimated for smaller intervals from these data, but we wished to keep the analysis close to the practical limits of what could be obtained using CPS data. Also, we have avoided using intervals of unequal length, as the CPS uses, in order to eliminate inessential complications. The effects of censoring duration data at 28 or 52 weeks can be determined by employing only the first seven or thirteen hazard rates.

Construction of CPS-type data required an entirely different procedure. The CPS interviews individuals at one point of each month, determines the current labor market state, then matches observations in consecutive months to give the gross monthly flows among states.[14] We have produced what we will call CPS-like data, or, when there is no possibility of confusion, simply CPS data, by matching observations on the individual's state on the first day of adjacent months, indexed by elapsed duration. The resulting gross flow data are used to produce exit probabilities grouped into duration categories at 4 week intervals. This grouping corresponds closely, but not exactly, to the one used by the CPS in reporting both gross flow and elapsed duration distributions. In what follows we will refer to the data obtained using the full panel nature of the data as continuous monitoring data, whether censored or uncensored. CPS data will be censored at 28 or 52 weeks.

In principle these data could be used to estimate the distribution of exit times by either the synthetic cohort or the gross flow approach. As a practical matter, the number of spells of unemployment recorded during the two year period was about 350 each for men and women; consequently, the number of persons unemployed at any one time is relatively small, and the grouped duration distribution will have sparse cells at any point in time. Hence, it does not appear feasible to apply the synthetic cohort approach to these data. However, because labor market conditions were stable during the 1972-73 period, as measured by both the Denver SMSA unemployment rate and the unemployment rates of DIME controls,[15] it seems reasonable to treat the completed spells of unemployment as observations from a single exit time distribution, which allows us to use the gross flow approach. The empirical work that follows focuses solely on this approach.

IV. Fitting Parametric Forms to Continuous Monitoring and to CPS-Like Data

In this section we report the results of fitting several specific distribution functions to the five sets of data described above. Continuous monitoring data will be referred to as CM28, CM52, or CMall to identify the censoring limit, and CPS-like data will be denoted CPS28 or CPS52. We report results based on three functional forms for the hazard function, described in Table 2. Other functional forms were also considered, but the estimates failed to satisfy restrictions on the hazard presented in Section II.[16]

Users of actual CPS gross flows are presented with censored data on unemployment spells, from which they wish to infer moments, or functions of moments, of the full distribution. Excluding the open-ended final duration interval, censoring occurs at about 26 weeks, leaving only five observations on grouped data. Our CPS28 data, which is grouped into regular four-week intervals, has six usable observations. This limits the choice of distribution function substantially, and we have restricted our set of hazard functions to those which are estimable within the confines of actual CPS data.

TABLE 2

PARAMETRIC MODELS OF THE HAZARD RATES

	FORM	SPECIFICATION	RESTRICTION
1)	Exponential	$\ell n\lambda(d) = B_o$	$-\infty < B_o < \infty$
2)	Weibull	$\ell n\lambda(d) = B_o + B_1 \ell n(d)$	$-1 < B_1 < \infty$
3)	Gompertz	$\ell n\lambda(d) = B_o + B_1 d$	$0 \leq B_1 < \infty$

Our primary interest is in how well the fitted distributions enable us to extrapolate to durations beyond the censoring limit.

We have estimated the functional forms described in Table 2 by maximum likelihood, using the data reported in Appendices A and B.[17] Table 3 contains these estimates for the three candidate distributions. These results suggest that models which allow non-constant hazard rates are superior to the constant-hazard exponential for both men and women. The Weibull appears to fit most versions of the data better than the exponential, according to a likelihood ratio test, as do several forms of the Gompertz with admissible parameter values. It should be noted, however, that the exit rate derived from CPS data increases with elapsed duration, while the continuous monitoring exit rates usually decline over time. Since the Weibull and Gompertz we cannot choose between them are not nested models, on the basis of a likelihood ratio test.

More important than the fit of each model to the data used to estimate it, is the ability to each parametric model to predict the actual (unobserved) distribution of unemployment spells. In the case of uncensored continuous monitoring data, these considerations coincide, since the entire distribution is used in estimation. We can compare this "best possible" fit of simple parametric forms to those achieved when data is simply censored and, in turn, to those which result when CPS-type monitoring and reporting procedures are used.

We focus on three aspects of the distribution of unemployment spells: the duration of a typical spell as measured by the mean and median, the frequency of very short spells, and the frequency of very long spells. Predicted values of these magnitudes from all acceptable parametric models are

Table 3A
Parameter Estimates for Alternative
Hazard Function Models:
Maximum Likelihood

Adult Men

DATA	Continuous Monitoring			CPS-Like	
Censoring	28 weeks (CM28)	52 weeks (CM52)	None (CMall)	28 weeks (CPS28)	52 weeks (CPS52)
			Exponential		
B_o	-2.2008	-2.3833	-2.6959	-2.5017	-2.5562
	(.0584)	(.0563)	(.0539)	(.0681)	(.0659)
			Weibull		
B_o	-2.3185	-2.2343	-2.2380	--2.9961	-2.7207
	(.0188)	(.0190)	(.0180)	(.0189)	(.0189)
B_1	-0.1242	-0.2182	-0.2138	0.2572	0.0806
	(.0096)	(.0088)	(.0075)	(.0086)	(.0082)
			Gompertz		
B_o	-2.6564	-*	-*	-2.6805	-*
	(.0426)			(.0218)	
B_1	0.0368	-	-	0.0187	-
	(.0016)			(.0018)	

* Converged at inadmissable parameter values ($B_1 < 0$).

Table 3B

Adult Women

DATA	Continuous Monitoring			CPS-Like	
Censoring	28 weeks (CM28)	52 weeks (CM52)	None (CMall)	28 weeks (CPS28)	52 weeks (CPS52)
			Exponential		
B_o	-2.5100 (.0600)	-2.6883 (.0567)	-2.9217 (.0541)	-2.7593 (.0663)	-2.7658 (.0631)
			Weibull		
B_o	-2.9531 (.0173)	-2.7994 (.0168)	-2.8200 (.0160)	-3.8733 (.0180)	-3.5229 (.0174)
B_1	0.0632 (.0078)	-0.0559 (.0069)	-0.0417 (.0061)	0.5229 (.0074)	0.3363 (.0068)
			Gompertz		
B_o	-3.0776 (.0336)	-3.2618 (.0802)	-*	-3.2260 (.0204)	-2.9768 (.0215)
B_1	0.0768 (.0021)	0.0439 (.0028)	-	0.0448 (.0014)	0.0155 (.0011)

presented in Table 4, together with the true values from the continuous monitoring sample.[18]

Means and Medians

When uncensored data are used, the estimated mean and median unemployment durations are very close to their true values for both men and women, regardless of the parametric model used.[19] If the continuous monitoring data are censored, however, the choice of a model becomes very important. CM52 performs nearly as well as CMall with the Weibull form, but usually results in a significant underestimate of both mean and median with either of the other forms. CM28 does even worse, but the undershooting is kept to a minimum by the Weibull distribution.

With CPS data, the degree of censoring has almost no effect on the predicted means and medians. In sharp contrast to the CM data, use of the Weibull or Gompertz does not generally improve the predictions. When compared to Weibull estimates using CM data with similar censoring, CPS28 and CPS52 lead to more serious underestimates of mean durations, and slightly worse estimates of medians.

Frequency of Short Spells

One well-known feature of the CPS sampling method is that short spells of unemployment are undercounted. For example, ignoring the overall sampling rate, as interviews are conducted at four week intervals, a spell with a complete duration of four weeks or more is certain to be caught, a two-week spell has a .5 probabilty of being detected, and so forth. Since shorter spells of unemployment are less likely to be included in the CPS sample, the distribution of completed spell lengths may be biased.

How important are these missing spells? Table 4 shows that, in this sample, 51% of men's unemployment spells are less than 8 weeks long, as are 35% of women's. Using the Weibull distribution results in good estimates with CM data, regardless of censoring. With the same distribution, CPS-type data do appear to lead to length bias, as the proportion of short spells is significantly underestimated for both men and women. This impression does not persist, however, if we turn to the exponential where, for women, the bias is in the opposite direction. Although CPS sampling results in censoring at the lower tail of the distribution as well as at the upper tail, the resulting bias in the estimated distribution depends solely on how well the assumed parametric model fits in the lower tail.

Frequency of Long Spells

Table 4 contains the actual and predicted proportions of unemployment spells which end in less than 28 weeks, or less than 52 weeks. Even with uncensored data, the models we use proved too simple to fit these tail probabilities very consistently. For men, prediction is good up to 28 weeks, but the number of extremely long spells is underestimated. For women, prediction is reasonably good for 52 weeks, but underestimates the proportion of spells less than 28 weeks. The shape of the actual distributions in the upper tails clearly does not conform very well to these standard distributions. Note that this has some rather important considerations: under ideal data conditions, the best estimate suggests that 5% of men's unemployment spells are longer than 52 weeks, while the true proportion is over 9%.

As might be expected, censored versions of the data do even worse. None of the models pick up the fat upper tail of the actual distribution for men; the proportion of long spells is consistently underpredicted. For women,

Table 4A

Summary Statistics from Alternative
Hazard Function Models

Adult Men

	Mean	Median	Proportion of Spells with Duration less than 8 weeks	28 weeks	52 weeks
			Exponential		
CM28	9.03	6.26	58.8	95.5	99.7
CM52	10.84	7.51	52.2	92.4	99.2
CMall	14.82	10.27	41.7	84.9	97.0
CPS28	12.20	8.46	48.1	89.9	98.6
CPS52	12.89	8.93	46.2	88.6	98.2
			Weibull		
CM28	12.97	7.98	50.1	87.5	97.2
CM52	14.65	7.96	50.1	84.3	95.1
CMall	14.56	7.96	50.1	84.5	95.2
CPS28	12.09	9.72	41.9	92.7	99.7
CPS52	12.92	9.49	43.8	89.3	98.7
			Gompertz		
CM28	10.23	8.42	48.0	96.8	99.9
CPS28	12.09	9.26	44.6	92.0	99.8
Actual	14.2	8.0	51.0	84.5	90.7

Table 4B

Summary Statistics from Alternative
Hazard Function Models

Adult Women

	Mean	Median	Proportion of Spells with Duration less than 8 weeks	28 weeks	52 weeks
			Exponential		
CM28	12.30	8.53	47.8	89.7	98.5
CM52	14.71	10.19	42.0	85.1	97.1
CMall	18.57	12.87	35.0	77.9	93.9
CPS28	15.79	10.94	39.8	83.0	96.3
CPS52	5.89	11.02	39.6	82.8	96.2
			Weibull		
CM28	16.63	12.07	36.1	81.7	96.2
CM52	18.73	12.38	36.8	77.6	93.2
CMall	18.50	12.38	36.6	78.0	93.6
CPS28	15.11	13.18	27.7	88.7	99.6
CPS52	15.92	13.18	29.9	85.0	98.7
			Gompertz		
CM28	10.76	10.00	39.9	98.9	100.0
CPS28	14.87	13.31	30.7	87.9	99.9
CMS28	14.42	12.89	31.8	89.2	99.9
CPS52	15.92	12.34	35.2	83.2	98.3
Actual	18.4	12.7	35.3	81.9	91.3

there is a clear tradeoff between predicting the proportion of spells less than 28 weeks, and the proportion less than 52 weeks. The Weibull fits well at 28 weeks with CM28, and well at 52 weeks with CM52.

Predictions using CPS data are remarkably insensitive to changes in functional form or censoring and remarkably poor. For men, long spells are underpredicted by a very large amount (8 points for over 52 weeks). For women, the exponential predicts rather well at 28 weeks, but all models do very badly at 52 weeks.

General Goodness-of-Fit

Table 5 presents a chi-square test for the fit of each model for the first 17 four-week intervals, and the maximum distance between the predicted and actual distribution functions.[20] The results have already been suggested by the discussion above. The best fit is achieved by the Weibull applied to CM data. CM52 is nearly as good as CMall; CM28 is only a little worse. CPS data does best with the exponential model, but is always clearly inferior to the best of the CM models. Censoring has no effect on the fit of the CPS-exponential.

Several conclusions are suggested by the preceding results:

1. A very simple and commonly-used functional form - the Weibull - fits the actual distributions of unemployment durations very well. This is true for both men and women, with the only major problem being an inability to fit the upper tail of the male distribution in this sample.

2. With continuous monitoring data, censoring at 52 weeks does not lead to a marked deterioration of the fit, and censoring at 28 weeks leads to only moderate problems when the Weibull form is used.

3. CPS-like data censored at 52 weeks fits the actual distributions very

Table 5

Goodness of Fit Measures
over First 68 Weeks
of Duration Distribution

	ADULT MEN		ADULT WOMEN	
	χ^2	$\mathrm{Max}\lvert\hat{F}-F\rvert$	χ^2	$\mathrm{Max}\lvert\hat{F}-F\rvert$
		Exponential		
CM28	52.4	.11	62.4	.14
CM52	26.4	.08	20.6	.08
CMall	31.3	.10	4.3	.05
CPS28	19.0	.08	10.0	.05
CPS52	19.0	.08	9.3	.05
		Weibull		
CM28	11.1	.07	4.6	.05
CM52	7.4	.04	4.3	.05
CMall	7.5	.05	3.7	.05
CPS28	51.1	.09	57.8	.10
CPS52	28.9	.08	28.5	.07
		Gompertz		
CM28	45.2	.12	88.7	.19
CM52	-	-	31.1	.10
CPS28	34.3	.09	31.3	.11
CPS28	-	-	10.2	.07

badly, and the use of functional forms more general than the exponential generally make the predictions worse. Censoring at 28 rather than 52 weeks makes almost no difference.

The implications for our current knowledge of the unemployment duration distribution, which is derived from the CPS, are almost entirely negative. Although the results reported here are specific to our data set, it seems likely that predictions from parametric models applied to CPS data are no better than those from our very similar CPS28. Even more disturbing, these predictions can probably not be improved by using more general parametric forms, or by extending the censoring limit to make additional data available. With existng sampling techniques, we cannot do much better than leave the censoring at 26 weeks and estimate an exponential hazard model. The major problem appears to be the imprecision with which unemployment durations are measured by the CPS. Measurement error obscures in CPS data the hazard rate patterns which are so evident in the continuous monitoring data.

How serious is the poor fit which results? The underprediction of very long spells and the bias in predicted mean duration might be reduced in magnitude by a more representative data set with fewer long spells. The apparent unreliability of estimates derived from CPS data is likely to be a more persistent problem. When using actual CPS data, investigators can choose among functional forms for the hazard only by examining how well they fit the available data. On this basis, we would prefer parametric forms other than the exponential. Table 5 shows, however, that such "preferred" estimates from CPS-like data may predict the underlying distribution of completed spells very poorly.

V. CONCLUSIONS

The findings reported in this study are specific to the data set used, and the other data sets could in principle produce different results. Yet, to our knowledge, the DIME data provide the only means of assessing the value of using data derived from point-in-time sampling, such as that reported in the CPS, to estimate durations of completed spells of unemployment. Our results indicate that the bias in mean durations is about 2-3 weeks, and that the frequency of long spells is underpredicted. More importantly, we find that the currently popular method of estimating hazard functions on censored and aggregated data in order to provide estimates of the complete distribution and of functions of the complete distribution, are quite unreliable. This arises in part because durations are censored and some parametrizations yield defective distribution functions, but primarily because CPS sampling and reporting techniques introduce measurement error into observed unemployment durations. Functional forms which fit the true distribution very well produce misleading estimates when applied to CPS Data. Consequently, we conclude that there is little to be coaxed out of CPS unemployment data other than what is present in non-parametric estimates of the distribution function available from the censored data.

FOOTNOTES

1 We ignore here, and elsewhere throughout this article, any distinction between a "true" escape rate distribution and that produced by the mixture of a true distribution with a distribution generated by unobserved population heterogeneity. For the purpose of estimating the distribution of completed unemployment spells, this distinction is irrelevant.

2 By unchanging over time we mean that the form of the distribution function and its dependence upon variables such as aggregate demand, are known. This dependence upon external covariates is not completely general, since it cannot encompass pure cohort effects.

3 Bowers and Harkness (1979) do not directly estimate the parameters of the exit distribution based on likelihood construction. Instead they use the empirical survivor function $L(s) = U(s, T_o)/p(T_o - s)$, where $U(s, T_o)$ is the number of spells of length S in process at time T_o, and fit smooth curves to these data. Since $L'(s)$ is not guaranteed to be non-positive, smoothing by regressing L(s) on s is necessary to produce a well behaved survivor function. In principle, this smoothing amounts to assuming a specific form for F(s)).

4 See also the papers by Akerloff and Main (1980), Bjorklund (1981), Cripps and Tarling (1974), Frank (1978), and Marston (1976).

5 European authors following Cripps and Tarling (1974), are sensitive to the effect of non-stationary state stocks on estimates of the exit distribution, and tend to report moments of the exit distribution only when a "stationary position (of the register) has been reached." (Cripps and Tarling, p. 301).

6 See Clark and Summers (1979) for an application of this approach.

7 This problem is not insurmountable. The existing gross flow data are available monthly and one can form weighted averages of the monthly flows assuming that the entry rate is uniformly distributed within a month.

8 There are further problems with the gross flow data that hinder their usefulness for the analysis of unemployment, for example, the reported flows do not add up to the current stocks. See Abowd and Zellner (1982) for a discussion of these issues.

9 See Keeley <u>et al</u> (1978), and Tuma and Robins (1980).

10 See Lundberg (1980), and Kiefer and Neumann (1982).

11 This is documented in Kiefer and Neumann (1982).

12 The appropriate way to assess these difficulties, of course, is to compare the characteristics of DIME data with those of the alternative, CPS data. It must be noted, then, that the exact date on which a transition between unemployment and nonparticipation takes place is not available from CPS gross flow data either. In fact, for <u>any</u> recorded

transition, we know only that a change in labor force status has occurred at some point during the month. Thus, the error in the reported duration of unemployment spells beginning or ending in nonparticipation will be larger in DIME data than in CPS data, while for spells beginning or ending in employment it will be smaller. In addition, Woltman (1980) suggests that simple response variability results in inflated CPS gross flow data, and that this problem is particularly serious for movements into and out of the labor force. This implies that part of the discrepancy between the unemployment-nonparticipation transition rates in DIME and CPS may be the existence of spurious transitions in the latter.

13 See Burdett *et al* (1982), Kiefer and Neumann (1982) and Salant (1977) for other evidence of unemployment spell length and its distribution.

14 We are ignoring the complications brought about by movement into and out of CPS due to sample rotation. Users of the CPS have generally assumed that there is no rotation effect, and we do not attempt to reproduce the problem in our sample.

15 The Denver SMSA unemployment rate was 3.6% in 1972 and 3.4% in 1973. For DIME controls, the equivalent unemployment rates were 7.6% in each year. (Kiefer and Neumann, 1982, p.334).

16 We considered generalizations of the Weibull and Gompertz hazards obtained by adding squared terms in log duration and duration respectively. In all cases, the maximum-likelihood estimates violated condition (4) on the hazard. A linear-log model, $\lambda(d) = B_0 + B_1 \ln(d)$ routinely fit best with $B_1 < 0$, violating the nonnegativity condition on the hazard. These results are not reported here. Note that Clark and Summers (1979) do not impose these restrictions on the hazard, and end up with an estimated c.d.f. having a regular maximum.

17 For the continuous monitoring data, elapsed duration is set at the midpoint of each 4-week interval. For CPS data, elapsed duration to date takes on the values of 4.15, 8.65, 12.65, 16.65, 20.65, and 24.65; the midpoint of the prior month duration category plus the sampling window (4.3 weeks).

18 Closed form solutions exist for the exponential and Weibull mean value functions and these were used in Table 4. The Gompertz mean values were obtained via numerical integration. Median values were obtained from the fitted survivor functions.

19 The only exception is that the exponential overestimates the median for men only by about two weeks.

20 The chi squared statistic is formed as $\chi^2 = N \sum_{t=1}^{17} (\hat{F}(\varepsilon) - F(t))^2/\hat{F}(\varepsilon)$, where N is the number of spells of unemployment in the sample, $\hat{F}$ is the predicted distribution, and F is the empirical distribution. The summation runs over the seventeen duration intervals reported. Table 5 also reports the maximum deviation between the predicted and actual distribution for each specification, which is useful for any distance tests such as the Kolmogorov-Smirnov test.

APPENDIX A

Hazard Rates
from Continuous Monitoring Data

Weeks of Unemployment	Adult Males	Adult Females
0 - 4	.2873	.1983
5 - 8	.3581	.1927
9 - 12	.3046	.1937
13 - 16	.2314	.2346
17 - 20	.1720	.2847
21 - 24	.1818	.1837
25 - 28	.1270	.2250
29 - 32	.1636	.1129
33 - 36	.1522	.1455
37 - 40	.0764	.1915
41 - 44	.0417	.0789
45 - 48	.0417	.1143
49 - 52	.0758	.1452
53 - 56	.0758	.1452
57 - 60	.1071	.1818
61 - 64	.1071	.1818
64 - 68	.1428	.4286

APPENDIX B

Hazard Rates
from CPS-Like Data

	Hazard Rates over Succeeding Month	
Weeks of Unemployment at Initial Interview	Adult Males	Adult Females
0 - 4	.3760	.2317
5 - 8	.3025	.2550
9 - 12	.3396	.2697
13 - 16	.2466	.2632
17 - 20	.3137	.2632
21 - 24	.2571	.2728
25 - 28	.1600	.2000
29 - 32	.2222	.2903
33 - 36	.2222	.3462
37 - 40	.1292	.1250
41 - 44	.2054	.1765
45 - 48	.1052	.0751

REFERENCES

ABOWD, JOHN, and ARNOLD ZELLNER, "Estimation of Gross Labor Force Flows From The Current Population Survey," H. G. B. Alexander Mimeo, Chicago, April 1982.

AKERLOF, GEORGE A., and BRIAN G. M. MAIN, "Unemployment Spells and Unemployment Experience," American Economic Review, vol. 70, December 1980, pp. 885-893.

BJORKLUND, A., "On the Duration of Unemployment in Sweden, 1965-76," The Scandinavian Journal of Economics, vol. 83, 1981.

BOWERS, J. K., and D. HARKNESS, "Duration of Unemployment by Age and Sex," Economica, vol. 46, no. 183, 1979.

BURDETT, KENNETH E., et al, "Earnings, Unemployment, and the Allocation of Time Over Time," Review of Economic Studies, (forthcoming).

CLARK, KIM B. and LAURENCE H. SUMMERS, "Labor Market Dynamics and Unemployment: A Reconsideration," Brookings Papers on Economic Activity, 1: 1979, pp. 13-60.

CRIPPS, T. F. and R. J. TARLING, "An Analysis of the Duration of Male Unemployment In Great Britain," The Economic Journal, June 1974, pp. 289-316.

FRANK, ROBERT H., "How Long Is A Spell of Unemployment," Econometrica, vol. 46, March 1978, pp. 285-301.

GREEN, MICHAEL, et al, "Employment and Unemployment - Trends During 1977," Monthly Labor Review, vol. 101, February 1978.

HALL, ROBERT E., "Turnover In The Labor Force," Brookings Papers on Economic Activity 3, 1972, pp. 709-56.

KAITZ, HYMAN, "Analyzing the Length of Spells of Unemployment," Monthly Labor Review, vol. 93, November, 1970, pp. 11-20.

KALBFLEISCH, JOHN D. and ROSS L. PRENTICE, The Statistical Analysis of Failure Time Data, (New York: John Wiley and Sons), 1980.

KEELEY, MICHAEL C., et al., "The Estimation of Labor Supply Models Using Experimental Data," American Economic Review, vol. 68, December 1978, pp. 873-887.

KIEFER, NICHOLAS M., and GEORGE R. NEUMANN, "Wages and the Structure of Unemployment Rates," in Martin N. Baily, ed., Workers, Jobs and Inflation, (Washington: The Brookings Institution), 1982, pp. 325-351.

LUCKETT, JAMES P., "The Estimation of Completed Spell Durations from Published Data Under Non-Steady State Conditions," Paper presented at the 1978 Winter Econometric Society meeting.

LUNDBERG, SHELLY, J., Unemployment and Household Labor Supply", unpublished dissertation, Northwestern University, August, 1980.

MARSTON, STEPHEN T., "Employment Instability and High Unemployment Rates," Brookings Papers On Economic Activity, 1: 1976, pp. 169-203.

SALANT, STEVE W., "Search Theory and Duration Data: A Theory of Sorts," Quarterly Journal of Economics, vol. 91, February 1977 pp. 39-57.

SILCOCK, H., "The Phenomenon of Labour Turnover," Journal of the Royal Statistical Society, Series A, vol. 107, 1954, pp. 429-440.

SMITH, NINA, "Varighed of Ledighedsperioder og Indstromning tie Arbejdsloshed: En Analyse Pa Danske Data 1965-1978", ("Duration of Unemployment and the Flow into Unemployment: An Analysis of Danish Data 1965-1978"), Working paper 82-5, Aarhus University, November 1982.

TRIVEDI, P. K., and G.M. BAKER, "Estimation of the Average Length of An Unemployment Spell from Labour Force Experience Data", Australian National University mimeo, February, 1982.

TUMA, NANCY B., and PHILLIP K. ROBINS, "A Dynamic Model of Employment Behavior: An Application to the Seattle and Denver Income Maintenance Experiments," Econometrica, vol. 48, May 1980, pp. 1031-1052.

WOLTMAN, HENRY, "Possible Effects of Response Variance on the Estimate of Gross Changes from Month-to-Month in the CPS", in Using the Current Population Survey as a Longitudinal Data Base, BLS Report no. 608, August 1980, pp. 11-13.