Unemployment Behavior
Evidence from the CPS Work Experience Survey

Thomas S. Coleman

ABSTRACT

This paper discusses the nature and uses of data on individual unemployment experience available from the Current Population Survey (CPS). The purpose of the paper is two-fold: first, to describe the general statistical process generating such data and then to assume a specific, tractable, stochastic process by which the data could have been generated; and second, to carefully determine what, if anything, these data can tell us about the nature of unemployment. The conclusions from the empirical analysis are two: First, entry rates into unemployment and differences in entry rates across people are more important than spell exit rates for explaining unemployment during the year and levels of unemployment. Second, there appear to be some inconsistencies between inferences drawn from the experience data and those drawn from other data sets.

I. Introduction

This paper examines the CPS unemployment experience data, which gives the number of weeks of unemployment and the number of spells of unemployment experienced by a sample of people in a year. The paper is empirical and descriptive rather than theoretical. It focuses

The author is an assistant professor of economics at the State University of New York at Stony Brook, currently on leave at the Bank of Boston. *He would like to thank Yoram Weiss, Alan Harrison, and anonymous referees for suggestions, but retains responsibility for errors and omissions.*

[Submitted August 1987; accepted July 1988]

on both the entry rate or frequency (probability of entering) and the exit rate or duration (probability of exiting) of unemployment spells, for it is on both the entry and exit rates that the level and distribution of unemployment depends. The conclusion is that entry rates and differences in entry rates across people are of primary importance in explaining unemployment experience.

As an example of the differences between focusing on exit rates alone versus both entry and exit rates, consider the standard finding that the distribution of unemployment spell durations has a "fat tail" or exhibits negative duration dependence; i.e., the probability of leaving is lower for people who have been unemployed longer.[1] This paper finds the same. Focusing only on single spell exit rates, this would seem to imply a concentration of unemployment among those with long spells, and might lead to a policy prescription of reducing the number of people with long spells as the most effective way of reducing unemployment. Looking at both entry and exit rates implies the opposite might be true. Many of those with long average durations (low exit rates) appear to have low entry rates and thus do not contribute substantially to any concentration of unemployment. Those with high entry rates, however, contribute disproportionately to unemployment. Reducing the entry rate into unemployment for those with high entry rates could be more effective for reducing unemployment than shortening the mean duration (increasing the exit rate). One could argue that unemployment in the U.S. is a problem of staying out of unemployment rather than getting out of unemployment.

In addition to providing a more complete description of unemployment, focusing on the entry and exit rates provides a framework in which one can compare different data sets. The distribution of both in-progress spells (as collected monthly by the CPS) and unemployment during the year (as collected in the March CPS) can be derived as a function of the underlying estimated exit and entry rates. Different data sets can then be compared for consistency.

Modeling the CPS unemployment experience data requires that one specify a stochastic process by which the data could have been generated and then fit the process to the data. Work by Sattinger (1983, 1985; see also Ridder 1985) is used. Previous studies (in particular Clark and Summers 1979, and Akerlof and Main 1980) have analyzed the experience data, but their failure to specify carefully the statistical methodology raises some questions about their conclusions.

Observations from 1984 (the March 1985 CPS) are used to provide

1. There is substantial debate over whether these fat tails are the result of true duration dependence or population heterogeneity, but this is not important for the immediate point.

maximum likelihood estimates of entry rates and exit rates. In detail, the ML estimates lead to the following conclusions:

- There is substantial variation across the population in entry and exit rates.[2,3] The majority of the population have low unemployment entry rates during the year, a minority of the population high entry rates.
- The high entry rate minority contributes disproportionately to the level of unemployment.
- It is the cross-sectional variation in entry rates, and not variation in exit rates, that is important in accounting for the distribution of weeks unemployed during the year.
- The pattern of low entry rates for the majority, high entry rates for a minority is relatively stable across demographic groups.
- The experience data and in-progress spell data do not appear to be consistent. The exit rate estimated from the experience data is lower (mean duration higher) than from the in-progress spell data.

II. CPS Unemployment Experience Data

Every March the Current Population Survey (CPS) of the Bureau of Labor Statistics asks questions about work experience during the previous year. In particular, there is a question asking how many weeks the person spent unemployed during the year, and another asking how many spells of unemployment there were. In March 1985, for example, a respondent may answer that he was unemployed four weeks during 1984 in one spell.

The distribution of the random variable representing the answer to the CPS question is complex. For example, four weeks in one spell could be

2. The variation is assumed to be population heterogeneity. It may actually result from duration dependence. For the purpose of this paper, however, it is the presence of differences, and not their reason, that matters. This is discussed more fully below.

3. The use of heterogeneity to explain "fat-tailed" distributions is well known: see for example Blumen, Kogan, and McCarthy (1955); Spilerman (1972); Salant's (1977) application to unemployment spell data. Carlson and Horrigan (1983) have shown that heterogeneity in a Markov model can generate fat-tailed distributions to match the unemployment experience data discussed in Clark and Summers (1979) or Akerlof and Main (1980). The present paper shows, somewhat surprisingly, that it is heterogeneity in entry rates rather than in exit rates that is most important in explaining the distribution of unemployment during the year.

Table 1a
Extent of Unemployment During 1984 (Millions of Persons)[a]

	Total	1 SP	2 SP	3+ SP	NO SP
Total	21.54	12.42	3.00	3.15	2.97
<5 wks	5.29	3.89	0.46	0.26	0.69
5–10 wks	4.02	2.47	0.61	0.56	0.39
11–14 wks	2.63	1.55	0.39	0.42	0.27
15–26 wks	4.82	2.52	0.81	1.05	0.44
27–39 wks	2.25	1.20	0.42	0.51	0.12
40–52 wks	2.54	0.80	0.32	0.36	1.06

a. These are the number of people with one spell, two spells, three spells, and no spell information reported. The total population (16+) is 177.661 million. From unpublished BLS tabulations. Thanks to Shirley Smith for making the tabulations available.

the first four weeks of a 40 week spell which carries over into 1985, or it could be a single completed spell. The distribution of weeks unemployed during the year has no simple relation to the distribution of a single spell, most importantly because unemployment during the year depends on the entry rate as well as the exit rate. Weeks during the year may occur in spells which are truncated at the beginning or end, and, even worse, may represent the summation of multiple spells.

Table 1a is a tabulation of responses from the March 1985 survey, while Table 1b shows the proportions to civilian population. There are two important observations to make about Table 1b. First, there is a large proportion (33 percent) of those with unemployment who have two or more spells during the year. What does this imply about the flows of people moving in and out of unemployment? Without a description of the underlying process generating these data, it is hard to say just what this implies. The second observation is that there is a high proportion of people with more than six months of unemployment (22 percent of those with unemployment). Unemployment appears concentrated among those with many weeks of unemployment. Is this consistent with observations from single spell data, which imply that unemployment is on average of short duration? Again, without a description of the underlying process, this is difficult to say.

To highlight the potential differences between the distribution for single spells and for unemployment during the year, Figure 1 shows, for a hypothetical individual, simulated densities of time spent in a single spell of

Table 1b

Extent of Unemployment During 1984 (Percent of Civilian Population)

	1 Total	2 1 SP	3 2 SP	4 3+ SP	5 NO SP	1 ÷ (12.12) TOT/TOT[a]	(3+4) ÷ (2+3+4) 2+ SP/TOT[b]
Total unemployed	12.12	6.99	1.69	1.77	1.67	100.00	33.11
1–4 weeks	2.98	2.19	0.26	0.14	0.39	24.56	15.54
5–10 weeks	2.26	1.39	0.34	0.31	0.22	18.66	32.00
11–14 weeks	1.48	0.87	0.22	0.24	0.15	12.19	34.20
15–26 weeks	2.71	1.42	0.46	0.59	0.25	22.36	42.44
27–39 weeks	1.27	0.68	0.24	0.29	0.06	10.45	43.61
40–52 weeks	1.43	0.45	0.18	0.20	0.60	11.78	46.07

a. This is the number unemployed in each week's category as a percent of the total unemployed.

b. This is the proportion of persons who report 2+ spells, as a proportion of the total that report spell information.

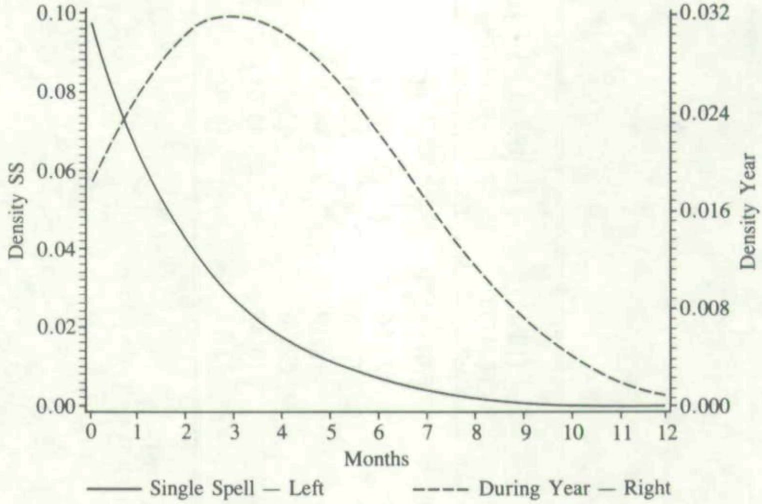


Figure 1
Density of Months Unemployed

unemployment and time spent unemployed during the year. In both cases, particular values of entry and exit rates are assumed.

The solid line (left scale) shows the density of months spent unemployed during a single spell for a newly unemployed individual.[4] The assumption is that this hypothetical individual has a mean spell length of 10 weeks (2.3 months) and an exponentially distributed leaving time (constant hazard). This means that the density of leaving times is $g(t) = \alpha e^{-\alpha t}$, with exit rate from unemployment of $\alpha = 0.4333$ and time measured in months: The solid line is simply a graph of $0.4333e^{-.4333t}$.

The dashed line (right scale) shows the density of months unemployed during the year (assuming steady-state) for this hypothetical individual. The mean unemployment spell length is still 10 weeks (exit rate from unemployment constant at $\alpha = 0.4333$), and it is additionally assumed that the entry rate into unemployment is constant at $\beta = 0.2167$, giving a mean spell of nonunemployment of 20 weeks (about five months). The density of months unemployed during the year depends in a complicated way on both the entry rate into unemployment (β) and the exit rate (α). The exact formula is given in Equation (1) below and derived in Sattinger (1983) and Appendix A. The dashed line is simply a graph of points generated from Equation (1) below, using $\alpha = 0.4333$ and $\beta = 0.2167$.

4. For example, reading from the graph the probability of leaving between 2 months and $2+dt$ months is about $0.04dt$.

Figure 1 demonstrates the importance of the difference between the distribution of single spells and months unemployed during the year. The solid line is the correct density for a single spell for our hypothesized individual, while the dashed line is the correct density of months unemployed during the year for the same individual. The difference between the solid and dashed lines represents a difference in the random variables plotted, not a difference in behavior. The solid line is the exponential density commonly used as the starting point in studies of unemployment duration. The distribution of weeks during the year also assumes an exponential density of leaving times for single spells, but it is far from exponential. It has an extremely fat tail relative to the distribution of a single spell, and is not even monotonic. The point of Figure 1 is that the two distributions are generated by the same underlying behavior, but look dramatically different.[5] It is thus important to carefully specify what random variable is being studied, if correct conclusions are going to be drawn from the empirical analysis.

In a following Section I critique previous studies which have used the unemployment experience data. Figure 1, however, highlights the problem discussed there. Clark and Summers (1979) and Akerlof and Main (1980) are not careful in distinguishing between the densities for single spells (solid line) versus weeks accumulated during the year (dashed line).

III. Statistical Methodology—Unemployment During Year[6]

I will start by sketching the derivation of the distribution of weeks unemployed during the year. It will become apparent that a full treatment of the distribution is not feasible. I will then review the simplified problem, where one assumes a two-state Markov process (see Sattinger 1983, 1985, for a detailed derivation).

The density of times spent unemployed is found by following all possible paths in and out of unemployment. Start by allowing three states—

5. The parameters chosen are not necessarily representative of average behavior in the U.S. (The assumed entry rate into unemployment is quite high, i.e., the mean time spent not-unemployed is quite short.) They were chosen to highlight the potential difference between the density of time spent in a single spell versus the time spent unemployed during the year. As will become apparent below, however, high entry rates into unemployment do appear to be empirically important.

6. This section is somewhat technical. The reader can safely ignore it if he or she takes on faith the Expression (1) as the correct density of time spent unemployed during an S-month period.

unemployment, employment, and NLF (not in the labor force).[7] The distribution of leaving times from unemployment will be denoted $G_u(t)$, and the survivor function $S_u(t) = 1 - G_u(t)$. The distribution of leaving times from employment and NLF are $G_e(\cdot)$ and $G_n(\cdot)$. To start, the distributions may have arbitrary duration dependence. I will show below that one must assume constant hazards to obtain a tractable problem.

Assume time homogeneity, so that densities do not change with calendar time. Define T to be the random variable measuring the number of months a person (randomly selected at the beginning of the period) spends in unemployment during a given S month period. Assume that the process is in steady state, and that there is no occurrence dependence in the sense of Heckman and Borjas (1980), although to begin there may be duration dependence. In this case the inflows from e to u and n to u are constants (call them k_e and k_n for now, with $k = k_e + k_n$).

The following outline shows the possible ways a person can accumulate exactly t months of unemployment (the densities are shown in Appendix A):

- I. Start in unemployment (probability of this is $\int_{-\infty}^{0} kS_u(-\tau)\,d\tau$)
  - A. Have one spell of length exactly t
    - a) Completed—density = $f_{I,A,a}(t;S)$
  - B. Have two spells, summing to t weeks total
    - a) Two completed spells—density = $f_{I,B,a}(t;S)$
    - b) One completed, one uncompleted spell—density = $f_{I,B,b}(t;S)$
  - C. Etc.
- II. Start in employment or not in the labor force (NLF)
  - A. Have one spell of unemployment of length exactly t
    - a) Completed—density = $f_{II,A,a}(t;S)$
    - b) Uncompleted—density = $f_{II,A,b}(t;S)$
  - B. Etc.

The density for t weeks of unemployment during the year is found by summing, over all possible paths, the probabilities of each path:

$$f_s(t) = \int_{-\infty}^{0} kS_u(-\tau)\,d\tau[f_{I,A,a}(t;S) + f_{I,B,a}(t;S) + \ldots]$$
$$+ \left[1 - \int_{-\infty}^{0} kS_u(-\tau)\,d\tau\right][f_{II,A,a}(t;S) + f_{II,A,b}(t;S) + \ldots].$$

The fundamental problem is that even the simplest density—for one spell starting in unemployment, $f_{I,A,a}(t;S)$—cannot be evaluated. The

7. I will show below that it is necessary to simplify to two states.

density function for leaving times from unemployment is $g_u(t)$, so the probability density that $T = t$ in one spell (conditional on starting unemployed) is

$$g_u(t)P[\text{do not reenter unemployment in } S - t \text{ months}] \quad \text{for } 0 < t < S.$$

The probability represented by P[do not reenter unemployment in $S - t$ months] must be expressed in terms of G_e and G_n, the distribution functions for leaving times from employment and unemployment. Since there may be arbitrary movements between employment and NLF, this involves an infinite sum of (infinite) multiple integrals. It is thus necessary to simplify by assuming that the not-unemployed state has the distribution $G_n(t)$; i.e., by assuming a two-state rather than three-state process. Then P[do not reenter unemployment in $S - t$ months] $= 1 - G_n(S - t)$. This avoids numerical evaluation of multiple integrals.

The simplification to a two-state model is not sufficient, however. The density for two spells, with the second spell uncompleted (i.e., $f_{II,A,b}(t;S)$) involves multiple integrals over $g_n(\cdot)$ and $g_u(\cdot)$, and is thus not feasible for practical applications. The hazards for entering and leaving unemployment must be assumed constant, so that the densities are exponential. Here I will use β to denote the hazard for entering unemployment (entry rate), and α for the hazard for leaving unemployment (exit rate), so that

$$g_u(t) = \alpha e^{-\alpha t} \qquad g_n(t) = \beta e^{-\beta t}.$$

Making the two simplifications of assuming a two-state process and assuming that entry rates into and exit rates from unemployment are constant (exponentially distributed times) leads to the full density, from Appendix A, of:

$$\text{(1)} \quad f_s(t) = e^{-\alpha t - \beta(S-t)} \Bigg[(\alpha u_0 + \beta n_0) \sum_{n=0}^{\infty} [\alpha t \beta (S - t)]^n / (n!)^2 + (\alpha\beta t u_0 + \alpha\beta(S - t) n_0) \sum_{n=0}^{\infty} \cdot [\alpha t \beta (S - t)]^n / [n!(n + 1)!] \Bigg]$$

(The expression $n!$ is n-factorial, $n! = n \cdot (n - 1) \cdot (n - 2) \cdots 2 \cdot 1$.)

In the data, the random variable being measured is the time spent in one state of a multi-state process, where leaving times may have duration dependence. The statistical model I am using to approximate reality is a two-state Markov process. It is an approximation to the actual process. The simplicity of the model allows one to compare the distribution of weeks unemployed during the year, given these simplifications, with the observed distribution of spells. It should be viewed as a null against which to compare observations. (See also Ridder 1985, for recent work with more than two states.)

The statistical model I am using can be generalized by allowing various forms of unobserved heterogeneity; i.e., allowing different people to have different entry and exit rates. Such a model with heterogeneity is parameterized by the entry rates and exit rates, and the proportion of the population in each group. Below I estimate four models with heterogeneity:

1. Heterogeneity in not-unemployment only:
   Type 1's: proportion = δ, exit rate = α, entry = β_1.
   Type 2's: proportion = $1 - \delta$, exit rate = α, entry = β_2.
2. Heterogeneity in unemployment only:
   Type 1's: proportion = δ, exit rate = α_1, entry = β.
   Type 2's: proportion = $1 - \delta$, exit rate = α_2, entry = β.
3. Heterogeneity in both unemployment and not-unemployment:
   Type 1's: proportion = δ, exit rate = α_1, entry = β_1.
   Type 2's: proportion = $1 - \delta$, exit rate = α_2, entry = β_2.
4. Two types of heterogeneity in both unemployment and not-unemployment:
   Type 1's: proportion = δ_1, exit rate = α_1, entry = β_1.
   Type 2's: proportion = δ_2, exit rate = α_1, entry = β_2.
   Type 1's: proportion = δ_3, exit rate = α_2, entry = β_1.
   Type 2's: proportion = δ_4, exit rate = α_2, entry = β_2.

The density given by (1) in Appendix A is an infinite sum. It cannot be expressed analytically or evaluated easily. Nor can the distribution, or any moments of the distribution, be evaluated in closed form. Maximum likelihood estimation requires numerical evaluation of the density, and numerical integration. Because of the double factorials in the denominator of each term, however, the sum converges rapidly. All estimates presented in this paper used Romberg integration (see Gerald and Wheatley). See Appendix C for more detail on the computational aspects of the model.

IV. Review of Previous Work

A basic point in using the CPS unemployment experience data is that the random variable "weeks unemployed during the year" has a complicated relation to the duration of a spell of unemployment. Clark and Summers in their section using the unemployment experience data

(Clark and Summers 1979, 33–39) proceed as if the variable they are measuring is the time spent in a single spell of unemployment.

To clarify exactly what they are doing, a little background is necessary. The second section of Appendix A contains a review of the statistical methodology for single spell durations (see also Heckman and Singer 1984, 97–100, and the references therein). Let us start with the random variable X which measures the (random) time a new entrant into unemployment spends in unemployment.[8] The distribution function is $G(x)$, and the mean duration for a new entrant is $E(X)$. Let us also define a second random variable related to the first, X_c (time completed), which measures the total time spent unemployed conditional on being unemployed at the survey date. This random variable conditions on a person being unemployed, and so $E(X_c) > E(X)$.

The mean time for a new entrant will be less than the mean time completed ($E(X) < E(X_c)$) for two reasons. First, conditioning on some accumulated unemployment picks out those individuals who are unlucky enough to have long spells, and so forces the mean higher. Second, unemployment spells generally exhibit negative duration dependence[9] and so those with some accumulated unemployment tend to have longer-tailed distributions. As shown in the appendix, the density of times completed is equal to the fraction of weeks in the mean time for a new entrant:

$$f_c(x_c) = x_c g(x_c)/E(X).$$

A numerical example (taken from Clark and Summers 1979, 17–18) will clarify ideas. Each week there are 21 new entrants to unemployment. Of these, 20 will leave in one week, and one will leave in 20 weeks. In other words, the density of time spent unemployed for a new entrant is

$$g(x) = \begin{cases} 20/21 & \text{if } x = 1 \\ 1/21 & \text{if } x = 20. \end{cases}$$

In steady-state, there will be 40 people unemployed at any one time. If we sample in the middle of a week, there will be 20 people who have been unemployed 0.5 weeks. There will be another 20 who have been unemployed from 0.5 to 19.5 weeks, but all of whom will eventually complete 20 weeks. The density of completed times will be

$$f_c(x) = \begin{cases} 0.5 & \text{if } x = 1 \\ 0.5 & \text{if } x = 20. \end{cases}$$

8. Everything is assumed time homogeneous and in steady state.

9. That is, duration dependence at the population level which may be generated by either duration dependence in individuals' hazards or population heterogeneity.

This gives $E(X) = 1.905$ and $E(X_c) = 10.5$ (not 9.5 as reported in Clark and Summers). As can be seen, $f_c(1) = 1 * (20/21) \div 1.905 = 0.5$.

Clark and Summers (1979) choose to concentrate on times completed, i.e., the random variable X_c.[10] They first use CPS gross flow data, which match individuals' labor force status over consecutive months, to estimate the distribution of leaving times for new entrants; i.e., $G(x)$.[11] From this, they calculate the distribution of times completed, $F_c(\cdot)$, which appears in their Table 1 under the heading "Proportion of Unemployment (expressed as a fraction of the total weeks of unemployment)." For future reference, they find that for 1974, all groups, 49 percent of completed unemployment spells result from spells lasting three months (13 weeks) or longer. In other words, they find that for 1974 all groups,

$$\text{(2a)} \quad 1 - F_c(13) = \int_{13}^{\infty} ug(u)\,du \Big/ \int_{0}^{\infty} ug(u)\,du = 0.49.$$

When they turn to the unemployment experience data, Clark and Summers are measuring a different random variable than either times for a new entrant (X) or times completed (X_c): They are measuring the weeks spent unemployed during a twelve-month year (random variable T, with density $f_s(t)$, from the previous section). In their Table 4 they report a tabulation of the distribution function $F_s(\cdot)$ under the heading "Unemployed persons (percent of labor force)." They also report a distribution of the weeks of unemployment, under the heading "Weeks of unemployment (percent of weeks)."[12] For 1974 all groups, they find that 73.5 percent of weeks of unemployment accumulated during 1974 is accounted for by people who accumulate more than 14 weeks of unemployment. This is the ratio[13]

$$\text{(2b)} \quad \left[\int_{t}^{S} uf_s(u)\,du + Su_0p_s\right] \Big/ \left[\int_{0}^{S} uf_s(u)\,du + Su_0p_s\right].$$

10. The original random variable X seems a better choice, but that is not the issue here.

11. I call this the distribution of leaving times for new entrants, while Clark and Summers call it the distribution of completed spells. I reserve the name "time Completed" for the distribution of total time spent unemployed conditional on being unemployed at the time of the survey, which is different from the distribution of leaving times for new entrants.

12. Table 4 of Clark and Summers (1979) is retrospective data from the March CPS. The question asked in the CPS is "how many weeks were you unemployed last year?" The number of people who answer "14 weeks" is an estimate of $Nf_s(14)$, where N is the total number of people. Clark and Summers apparently calculate the total weeks unemployed over the year by adding up each person's weeks unemployed, i.e., estimating $N\int_0^S uf_s(u)\,du$. They then calculate the weeks accumulated by people who have, say, more than 14 weeks, which is an estimate of $N\int_{14}^S uf_s(u)\,du$. They then divide the latter by the former to get the "fraction of unemployment" included in spells lasting more than fourteen weeks.

13. The numerator of this expression is the integral of the number of people with unemployment who have 't' or more months of unemployment times their months of unemployment

The problem arises when Clark and Summers compare the data from Table 1 with that from Table 4:

> Compared with the spell durations of Table 1, which are estimated from the monthly CPS, a much higher fraction of unemployment and nonemployment is included in spells lasting more than fourteen weeks—73 percent of unemployment [for 1974]. [p. 35]

Clark and Summers are comparing the two statistics in (2a) and (2b). They have similar functional form, but do not measure the same random variable.

There are two points to be made about Clark and Summers's comparison of their Table 1 results (based on monthly gross flow data, and measuring single spell distributions) against their Table 4 results (based on the CPS experience data, and measuring weeks accumulated during a year). First, they are using two different sets of observations and attempting to test whether they are consistent. This is a good research strategy. Second, however, the statistics they use to compare their tables are fundamentally not comparable. The comparison quoted above is essentially one of apples and oranges, and does not imply anything about the two data sets. It should be clear that $1 - F_c(t)$, (2a), and the ratio (2b) will in general be different, even for the same underlying stochastic process.

A valid comparison between the gross flow and experience data would be to take estimates of the underlying parameters from the experience data (Clark and Summers's Table 4), calculate the implied $1 - F_c(t)$, and compare that with the results from their Table 1. Or alternatively compare the implied mean duration of a single spell for a new entrant from the experience data with that calculated from the gross flow data. This is done below, with the conclusion that there appear to be some differences between the single spell data commonly used and the experience data.

Clark and Summers conclude that "normal turnover (short spells of unemployment followed by job attainment) accounts for an insignificant proportion of measured unemployment" (p. 42). This is incomplete and not totally correct. Much of measured unemployment is the result of repeated short spells of unemployment. As pointed out above, about 30 percent of those with unemployment during the year have repeat spells. The estimates for heterogeneous Markov models presented in the next section imply that a minority of the population has high entry rates into

(up to a maximum of 'S') during a period of 'S' months (12 months or one year in this). The denominator is the integral of the number of people with any unemployment during the 'S' months (during the year) times their months of unemployment. The term u_0 is the proportion of the population unemployed at the beginning of the period, and p_s is the probability that a person unemployed at the beginning stays unemployed for all of the 'S' months.

unemployment. This generates the "concentration" of unemployment observed by Clark and Summers. The important result of this paper, relative to Clark and Summers's, is that it demonstrates that high entry rates among a minority of the population, rather than low exit rates (long duration of single spells), is important for an empirical description of unemployment during the year.

Akerlof and Main (1980) also use the CPS unemployment experience data. They concentrate on two aspects of the data: the number of persons with multiple spells, and the decrease in "spell length" for multiple spells. The number of multiple spells is critical, and will be discussed below. The length of multiple spells, however, is a necessary result of the laws of probability.

Akerlof and Main wish to classify the unemployed loosely into two types: those who repeat spells often, and those who do not (1980, p. 887). This is an important and valid classification, but the short length of multiple spells they cite is not evidence in support of such a classification. Akerlof and Main find

> an empirical regularity regarding the negative correlation between average spell lengths and the number of spells of unemployment experienced in a calendar year. [1980, p. 889]

Such an "empirical regularity" is a consequence of the conditioning in the observations. Even for a homogeneous population, conditioning on multiple spells during a fixed time period must give shorter average spell lengths than conditioning on single spells.

To be precise, what Akerlof and Main call the "average spell length" is really not the average length of a spell, since spells may be truncated at the beginning and end of a year. Indeed, those with multiple spells during a year will be more likely to have truncated spells. This will reduce the "average spell length" for those with multiple spells.

The negative correlation Akerlof and Main find can be reproduced using the distribution for a homogeneous two-state Markov process.[14] For the two-state Markov process, everyone behaves the same. Nonetheless, those with multiple spells have shorter average spells than those with single spells; they must if they are to fit those multiple spells in during a year.[15] (See Table 2.)

14. Assume that unemployment and nonunemployment follow a two state Markov process with $\alpha = 0.2407$, $\beta = 0.008731$. This corresponds to a mean duration of unemployment of 18 weeks (rather high), a mean duration of nonunemployment of 9.5 years, and a steady state unemployment of 3.5 percent of the population. (In 1974, unemployment was 3.4 percent of the population. Cf. Tables 1 and 2 of the BLS *Handbook of Labor Statistics*.)

15. The density of time spent unemployed, conditional on having only a single spell of

Table 2
Actual and Simulated Unemployment Experience and "Average Spell Length"

	Average Annual Experience of Unemployment for People with		"Average Spell Length" for People with	
	1 Spell	2 Spells	1 Spell	2 Spells
Markov (wks)	13.7	19.5	13.7	9.7
Akerlof & Main (wks) (Table 2)	11.6	15.7	11.6	7.9

The second issue Akerlof and Main raise is the number of multiple spells. This is an important issue, and is the one aspect where a homogeneous two-state Markov model is sharply at variance with the data. Akerlof and Main report (1980, p. 887) that about 34 percent of all persons with unemployment have multiple spells (average for 1965 to 1977). This turns out to be true across a variety of time periods and a variety of demographic groups. For the homogeneous two-state Markov process used in the table above, only 4 percent of those with any unemployment have multiple spells. In other words, the data on the number of multiple spells are strongly at variance with a homogeneous Markov model.

V. Analysis of CPS Experience Data

To analyze the data on extent of unemployment, I use unpublished tabulations from the Bureau of Labor Statistics shown previously in Table 1. The tabulations are similar to those that appear in Table 46 of the BLS *Handbook of Labor Statistics*, 1983, except that the distribution of weeks unemployed during the year are broken out by number of spells.[16]

unemployment, is $f_1(t) \div \int_0^S f_1(u)\,du$, where $f_1(t) = e^{-\alpha t - \beta(S-t)}[(\alpha u_0 + \beta n_0) + \alpha\beta(S - t)n_0]$, plus a point mass of $u_0 e^{-\alpha S}$ at $t = S$ and a point mass at $t = 0$. The density conditional on two spells during the year is $f_2(t) \div \int_0^S f_2(u)\,du$, $f_2(t) = e^{-\alpha t - \beta(S-t)}[(\alpha u_0 + \beta n_0)(\alpha t\beta(S - t)) + \alpha\beta(S - t)n_0(\alpha t\beta(S - t))/2 + u_0(\alpha t\beta)]$ plus a point mass at $t = 0$.

16. Thanks to Shirley Smith of the BLS for making the tabulations available.

Estimating the parameters of a homogeneous two-state Markov model by maximum likelihood is relatively simple, except that all integrations must be done numerically using a density that is itself an infinite series. The density of months (or weeks) of unemployment during the year is given by Equation (1) in the appendix. The likelihood function is described in Appendix C. For the homogeneous two-state Markov model the parameters to be estimated are α, the hazard rate for leaving unemployment (exit rate), and β, the hazard rate for entering unemployment (entry rate).

Table 3, Column (1) shows the estimation results using the observed number of people from Table 1, without accounting for the number of spells. The homogeneous two-state model estimates the total proportions rather well, but grossly misses the proportions of repeat spells. Taking the data on number of spells into account does much worse; see column two of Table 3. The overall proportions are badly misestimated, and the proportions of multiple spells are still fit poorly. These results show that a homogeneous Markov model is strongly at variance with the data; more people have multiple spells than is predicted.

The question arises whether a model with heterogeneity fits the data better. As discussed earlier, I estimate four models with heterogeneity:

1. Heterogeneity in not-unemployment only.
2. Heterogeneity in unemployment only.
3. Heterogeneity in both unemployment and not-unemployment.
4. Two types of heterogeneity in both unemployment and not-unemployment.

Models 1 and 2 are nested within Model 3. Model 4 is included because it is a generalization of Model 3, to test whether the results from Model 3 are robust.

Table 4 column one shows the results of estimating Model 1, with heterogeneity in entry rates. It is assumed that there are two types of people, those with low entry rate (type-1's, β_1), and those with high entry rate (type-2's, β_2). Everyone has the same hazard for leaving unemployment (α). The conclusion is that Model 1 (heterogeneity in entry rates only) captures most of the unemployment behavior. First, the overall distribution of weeks is not fit too badly. Second and more importantly, the number of multiple spells is fit pretty closely. Model 3 (introducing heterogeneity in both exit rates and entry rates) does somewhat but not substantially better than Model 1, and Model 4 does somewhat better than Model 3. Model 2, with heterogeneity in exit rates only, badly misestimates the proportion of multiple spells, as well as misestimating the aggregate proportions by duration category. (See the column under

Table 3
Implied Proportions and Estimated Parameters for Homogeneous Model

	Proportions (in percentages)[a]		
		Implied	
	Actual	Not Using Spell Data	Using Spell Data
Total unemployed	12.12	12.09	16.54
1–4 weeks	2.98	2.29	4.41
5–10 weeks	2.26	2.71	4.62
11–14 weeks	1.48	1.41	2.11
15–26 weeks	2.71	2.86	3.51
27–39 weeks	1.27	1.69	1.42
40–52 weeks	1.43	1.13	0.48
1 spell[b]	6.99	10.11	13.49
2 spells[b]	1.69	0.31	0.76
3 spells[b]	1.77	0.003	0.02
χ^{2c}		6.50	13.66

	Estimated Parameters	
	Not Using Spell Data	Using Spell Data
α (exit rate from unemp)	0.1818	0.2977
standard error[d]	(0.0015)	(0.0019)
implied mean duration of unemp	5.5 mths	3.4 mths
β (entry rate to unemp)	0.00741	0.01182
standard error[d]	(0.000053)	(0.000066)
implied mean duration of not-unemp	11 yrs	7.0 yrs

a. Percent of population.

b. As proportion of those reporting spell data.

c. This is a "chi-squared statistic" in the sense that $\chi^2 \equiv \Sigma_i[P_i - A_i]^2/A_i$, where P_i = predicted proportion in category i, A_i = actual proportion in category i. All categories, some not shown, are used. This is intended as a heuristic quality of fit measure, rather than a statistical measure.

d. These standard errors are only approximations to the asymptotic ML standard errors. They are based on the assumption that the sample consisted of 100,000 individuals. The actual data are unpublished BLS tabulations from the March 1985 Current Population Survey, weighted to reflect the U.S. population. The CPS consists of about 100,000 individuals.

Table 4

Implied Proportions and Estimated Parameters (Heterogeneity Models Using Data on Number of Spells)

		Implied Proportions (in percentages)[a]			
	Actual	Model 1	Model 2	Model 3	Model 4
Total unemployed	12.12	12.12	16.55	12.13	12.13
1–4 weeks	2.98	2.20	5.76	1.89	2.73
5–10 weeks	2.26	2.59	4.22	2.64	2.73
11–14 weeks	1.48	1.41	1.58	1.57	1.30
15–26 weeks	2.71	3.18	2.63	3.44	2.46
27–39 weeks	1.27	2.04	1.43	1.79	1.71
40–52 weeks	1.43	0.71	0.93	0.81	1.20
1 spell	6.99	6.80	13.49	7.13	7.01
2 spells	1.69	2.24	0.76	1.77	1.76
3 spells	1.77	1.41	0.02	1.56	1.69
χ^{2b}		3.24	12.82	2.33	1.53

	Estimated Parameters			
	Model 1	Model 2	Model 3	Model 4
α_1 (exit rate from unemp)	0.3185	0.1851	0.1748	0.1976
standard error[c]	(0.0021)	(0.0038)	(0.0047)	(0.0029)
implied mean dur of unemp	3.1 mths	5.4 mths	5.7 mths	5.1 mths
α_2 (exit rate from unemp)		0.9671	0.5131	1.209
standard error[c]		(0.035)	(0.0059)	(0.027)
implied mean dur of unemp		1.0 mths	1.9 mths	0.8 mths
β_1 (entry rate to unemp)	0.004518	0.01185	0.004396	0.004834
standard error[c]	(0.000062)	(0.000065)	(0.000050)	(0.00006)
implied mean dur of not-unemp	18 yrs	7.0 yrs	19 yrs	17 yrs
β_2 (entry rate to unemp)	0.2162		0.2145	0.2604
standard error[c]	(0.0030)		(0.0024)	(0.0038)
implied mean dur of not-unemp	4.6 mths		4.7 mths	3.8 mths
δ (proportion of α_1, β_1 types)	0.9380	0.5397	0.9465	0.6897
standard error[c]	(0.00094)	(0.016)	(0.00020)	(0.011)
δ (proportion of α_1, β_2 types)				0.0290
standard error[c]				(0.00075)
δ (proportion of α_2, β_1 types)				0.2575
standard error[c]				(0.011)

a. Percent of population.

b. This is a "chi-squared statistic" in the sense that $\chi^2 \equiv \Sigma_i[P_i - A_i]^2/A_i$, where P_i = predicted proportion in category i, A_i = actual proportion in category i. All categories, some not shown, are used. This is intended as a heuristic quality of fit measure, rather than a statistical measure.

c. These standard errors are only approximations to the asymptotic ML standard errors. They are based on the assumption that the sample consisted of 100,000 individuals. The actual data are unpublished BLS tabulations from the March 1985 Current Population Survey, weighted to reflect the U.S. population. The CPS consists of about 100,000 individuals.

Table 5
Concentration of Unemployment Implied by Estimates from Model 4 Heterogeneity

	Unempl.[a]	Pop. Share[b]	Unempl. Share[c]	Unempl ÷ Pop
Low entry, low exit (α_1, β_1)	2.39%	68.97%	43.11%	0.7
Low entry, high exit (α_2, β_1)	0.40	25.75	2.68	0.1
High entry, low exit (α_1, β_2)	56.86	2.90	43.17	14.9
High entry, high exit (α_2, β_2)	17.72	2.38	11.04	4.6

a. The steady-state unemployment-to-population ratio, $O\bar{u}_i = \beta_i/(\alpha_i + \beta_i)$.
b. The estimated population share, from Table 4.
c. The ratio of steady-state unemployment-to-population ratio for group i to total. Total is $O\bar{u} = \Sigma\delta_i O\bar{u}_i$.

"Model 2".) This supports the contention that it is entry rates of unemployment rather than exit rates of unemployment that are critical in annual unemployment experience.

Heterogeneity in exit rates does appear in Models 3 and 4, just as in single spell data. One might (incorrectly) infer from apparent heterogeneity in exit rates that there is concentration in the burden of unemployment because some people have longer average duration (lower exit rates) than others. Table 5 shows that the estimates from Model 4 heterogeneity do indeed imply substantial concentration of unemployment, but it is primarily among those with high entry rates into unemployment. The last column of Table 5 shows the ratio of the implied unemployment share to population share, and is a crude measure of concentration of unemployment. Keeping entry rates constant, there is indeed greater concentration among those with low exit rates (0.7 versus 0.1; 14.9 versus 4.6). The dramatic differences, however, are between those with low entry rates and those with high entry rates. Altogether, those with high entry rates make up only 5.3 percent of the population, but 54.2 percent of the steady-state unemployment-to-population ratio.

The importance of heterogeneity in unemployment entry rates holds across different years and demographic groups. Estimates for 1984 total, males, females, whites, black, and 1982 total all show a reasonable fit under Models 1 (heterogeneity in not-unemployment) and 3 (heterogeneity in unemployment and not-unemployment), but a very poor fit under Model 2 (heterogeneity in unemployment only). What is most surprising, however, is the stability of the pattern of the estimated parameters across demographic groups. For all demographic groups, the same pattern emerges: Most of the population is stable type-1's (low unemployment entry rates), with a minority of unstable type-2's (high unemployment entry rates). Table 6 shows the parameter estimates (together with implied mean durations) under Model 3 heterogeneity for 1984 males, females, whites, blacks, and 1982 total. The pattern is the same across all the groups. The results across demographic groups imply that the shape of the distribution of weeks unemployed during the year can be described by the same pattern of heterogeneity across demographic groups. The results of Table 6 (together with untabulated results by age for 1982) supports the contention that the estimated heterogeneity for the total population is not the result of aggregating different demographic groups together.

Although the pattern of heterogeneity is similar across demographic groups, i.e., the shape of the distribution of weeks unemployed during the year is similar, other aspects of unemployment are not. Table 6 shows the percent of the population with any unemployment during the year, both predicted by the model and observed. The percent with unemployment shows substantial variation across time and demographic groups, from 10.0 percent for females to 17.3 percent for blacks. Even though the pattern of heterogeneity is similar, it does not imply similar unemployment experience or unemployment rates. On the other hand, the percent of those unemployed who have two or more spells of unemployment is surprisingly similar across demographic groups. In other words even though the level of unemployment varies considerably, the pattern of unemployment (shown in Table 6 by the percent with multiple spells) is stable.

The estimates above use unpublished tabulations from the Bureau of Labor Statistics. One can also work directly with the CPS tapes. Appendix D shows that working with the CPS tapes gives qualitatively similar results. In addition, Appendix D shows that the results do not change when one uses alternative category definitions.

The estimates displayed above raise some questions. First, could the estimated heterogeneity in entry rates result from observed differences across demographic groups, or an incorrect assumption of a two-state as opposed to three-state model? I argue that this does not appear to be the full explanation for the estimated heterogeneity. Second, is the high entry rate for the minority a result of population heterogeneity, or the result of declining individual hazards? The answer to this question must remain for future research.

Collapsing the two states of unemployment and NLF into one (i.e., ignoring observed differences) is clearly a drastic simplification. Nonetheless, separating the two states of employment and NLF (while retaining the assumption of a homogeneous Markov model), still misses the propor-

Table 6
Comparison of Parameters and Implied Mean Durations, Heterogeneity Model 3, 1984 and 1982, Various Demographic Groups

	1984 Total	1984 Male	1984 Female	1984 White	1984 Black	1982 Total
Type 1's						
Proportion	94.65%	93.37	95.11%	94.96%	92.73%	93.10%
α	0.1748	0.1523	0.1907	0.1961	0.1165	0.1309
Mean U Dur	5.7 months	6.6 months	5.2 months	5.1 months	8.6 months	7.6 months
β	0.0044	0.0049	0.0036	0.0043	0.0060	0.0047
Mean ~U Dur	19.0 years	17.1 years	23.0 years	19.6 years	13.8 years	17.6 years
Type 2's						
Proportion	5.35%	6.63%	4.89%	5.04%	7.27%	6.90%
α	0.5131	0.4871	0.5849	0.5341	0.4156	0.4748
Mean U Dur	1.9 months	2.1 months	1.7 months	1.9 months	2.4 months	2.1 months
β	0.2145	0.2412	0.1568	0.2123	0.2461	0.2133
Mean ~U Dur	4.7 months	4.1 months	6.4 months	4.7 months	4.1 months	4.7 months
Percent of Pop with U (in percentages)						
Predicted	12.1	14.5	10.0	11.4	17.8	14.7
Observed	12.1	14.5	10.0	11.4	17.3	14.7
Percent of U with 2+ Spells[a] (in percentages)						
Predicted	31.8	35.6	27.3	31.7	31.8	33.3
Observed	33.1	36.9	27.7	32.7	35.3	35.5
Unemployment-to-Population Ratio (in percentages)						
Predicted	3.9	5.1	2.8	3.5	7.3	5.4

a. Those reporting 2+ spells as a percent of all those reporting number of spell information.

tion of repeat spells. It is possible to calculate the proportion of repeat spells for a three-state process, even though calculating the full density of times spent unemployed is not feasible.[17] When one does this and substitutes reasonable values for the transition rates (taken from Marston 1976), the proportion of repeaters is still too low.[18]

Ignoring differences across demographic groups would also seem, a priori, to be important. The estimates displayed in Table 6, however, show that the pattern of estimated parameters is remarkably stable across demographic groups. One can also see that differences across demographic groups is not of dominating importance by examining the proportion of those unemployed with multiple spells among males of different ages: in 1982 it is 37.0 percent for 16–19, 36.5 percent for 20–24, and 36.2 percent for 25–44. If differences between demographic groups accounted for the estimated heterogeneity (and thus the proportion of repeat spells), then the proportion of repeat spells would be significantly higher for young males, who tend to have shorter job spells and more frequent unemployment spells. Clearly, both the distinction between employment and NLF, and differences across demographic groups, will account for some of what I estimate as unobserved heterogeneity. It simply does not seem to account for all of it.

With respect to unobserved heterogeneity versus true duration dependence, the answer must be left to future research. To account for the large proportion of repeat spells, the entry rate for a person who has just left unemployment must be (on average) higher than the average over the whole population. It may be higher for one of two reasons. First, people might be intrinsically different in their entry rates. Those with high entry rate are more likely than average to be unemployed, and so more likely than average to be among the pool leaving unemployment. This would lead to the average unemployment leaver having a higher entry rate than over the whole population. Second, people might be all the same, but have a quickly falling hazard for entering unemployment. All unemployment leavers would have high entry rates, because they had just left unemployment. Those lucky enough to remain out of unemployment for a time would be less likely to reenter unemployment.

It would be possible to distinguish between heterogeneity and duration dependence using panel data on unemployment experience, such as that available from the National Longitudinal Survey or the Michigan Panel Study of Income Dynamics. Pure duration dependence implies that only

17. But see Ridder (1985) for recent work in this area.

18. See Appendix B for the calculations.

time since leaving unemployment matters. Thus, the only thing that predicts whether a person will have high entry rate is the time of last unemployment, not the number of previous spells (i.e., it excludes occurrence dependence in the sense of Heckman and Borjas 1980). Heterogeneity implies that the number of spells helps predict whether an individual has high entry rate. Further work clearly needs to be done in this area.

One final question to address is how well the unemployment experience data and the in-progress spell data match. A direct comparison between raw unemployment experience versus in-progress spell data cannot be made because the data measure different random variables. Nonetheless, two comparisons are possible. First, one can estimate the mean duration of a new spell from both data sets and compare the estimates. The estimate from unemployment experience data using Model 3 for 1984 is that mean duration is 3.2 months.[19] The estimate from the in-progress spell data in Table 3 is that mean duration is 1.8 months.[20] These are rather different.

The second possible comparison is to calculate the distribution of in-progress spells implied by the parameters estimated from the experience data. Appendix A shows the formula for the distribution of in-progress spells as a function of the distribution of new spells. Table 7 shows the actual and implied distributions of in-progress spells for 1984.

The second column of Table 7 shows the distribution for in-progress spells implied by estimates from the unemployment experience data. This does not match the in-progress spell data very closely. (For measuring the quality of fit, a "chi-squared" statistic is included.) In particular, it underpredicts the number of short spells. It also, however, underpredicts the number of very long (over 52 week) spells.

From these two comparisons, it appears that the unemployment experience and in-progress spell data are at variance over the distribution of single spells of unemployment. Without further analysis one cannot decide which is more correct. Nonetheless, the in-progress spell data provide absolutely no information on repeat spells or entry rate, so that one must rely on the experience data for information on this important aspect of unemployment behavior.

---

19. Mean duration $= p/\alpha_1 + (1-p)/\alpha_2$, where p = proportion of type-1's entering unemployment (the proportion in the population $= \delta$). $p = \delta[\alpha_1\beta_1/(\alpha_1+\beta_1)]/[\delta[\alpha_1\beta_1/(\alpha_1+\beta_1)] + (1-\delta)\,[\alpha_2\beta_2/(\alpha_2+\beta_2)]]$.

20. This is from a two-point heterogeneity model with constant hazards. A simple alternative is a gamma heterogeneity model with constant hazards, originally used in Salant (1977). The estimate from such a model is that mean duration is 1.6 months.

Table 7
Distribution of In-Progress Spells Actual and Implied from Various Estimates

		Implied by Estimates from		
	Actual[a]	Experience[b]	In-prog.[c]	In-prog.[d]
Total unemployed	100.0	100.0	100.0	100.0
1–4 weeks	39.2	24.1	36.5	37.2
5–10 weeks	20.6	23.7	25.0	23.7
11–14 weeks	8.1	10.6	7.9	8.2
15–26 weeks	12.9	18.8	10.3	12.4
27–51 weeks	7.6	15.0	8.5	8.9
52+ weeks	12.3	7.7	11.7	8.3
χ^2		0.187	0.0179	0.0211
Mean duration (mths)		3.2	1.8	1.6

a. From BLS *Employment and Earnings* January 1985.
b. Distribution of in-progress spells implied by the estimates of Model 3 heterogeneity using 1984 unemployment experience data.
c. Estimates of a two-point heterogeneity model with constant hazards using the in-progress spell data. $\alpha_1 = 0.08858$, $\alpha_2 = 0.7688$, $p = 5.44\%$.
d. Estimates of a gamma heterogeneity model with constant hazards (as in Salant 1977) using the in-progress spell data. $a = 8.482$, $r = 2.204$.

VI. Conclusion

It is a truism to say that the level and dynamics of unemployment depend on the entry rate as well as the exit rate of unemployment spells. This paper has examined the distribution of unemployment during the year, and found that it depends critically on the entry rate into unemployment. More precisely, differences in entry rates across people are more important than differences in exit rates in accounting for the distribution. The empirical results imply that a relatively small proportion of the population have high unemployment entry rate, and consequently repeated spells during the year. This group accounts for a disproportionate share of unemployment. It is not possible to determine whether the differences in entry rates result from true population heterogeneity or falling hazard for the entry into unemployment.

Some researchers, after examining the unemployment experience data, have claimed that the implied mean durations are inconsistent with estimates from other data sources, and this appears to be correct. The additional claim is sometimes made that long mean durations imply both substantial concentration of unemployment and a nondynamic unemployment process—contrary to the "new view" of unemployment as transitory within a lifetime of employment. These conclusions are not correct. First, most concentration of unemployment appears to result from high entry rates—i.e., repeated spells rather than long spells. Second, all that a study like the current one can do is carefully delineate the observations within which economists must operate, not decide whether unemployment is dynamic. The present study estimates that differences in entry rates are important, and that mean single spell durations are about three months. Neither 1.6 versus 3.2 month mean durations, nor differences in entry rates, however, imply that unemployment is dynamic or sluggish. They simply imply that economic theories must be consistent with such differences in entry rates and with durations on the order of two to four months if they are to be consistent with observation.

Appendix A

Distribution of Unemployment During the Year

I will now proceed to derive the distribution for the simplified, two-state Markov process. (See Sattinger 1983, 1985, for a detailed exposition.) The exit rate from unemployment is $-\alpha$, while the exit rate from not-unemployed is $-\beta$ (in units of months^{-1}). Employment and not in the labor force (NSF) are lumped into one category: not-unemployed. Call the distribution of leaving times from unemployment $G_u(t)$, and the distribution of leaving times from not-unemployment $G_n(t)$. Then G_u and G_n are exponential:

$$G_u(t) = 1 - e^{-\alpha t} \qquad G_n(t) = 1 - e^{-\beta t}$$

To match the outline in the text, the possible ways a person can accumulate exactly t ($\leq S$) months and the possible ways a person can accumulate T ($\leq S$) months are:

I. **Start in Unemployment** (all densities below are conditional on starting in unemployment, which has probability u_0)
   A. Have one spell of length exactly t
      a) This spell must end in not-unemployment.

$$f_{I,A}(t;S) = g_u(t)[1 - G_n(S - t)] = \alpha e^{-\alpha t} e^{-\beta(S-t)} \quad \text{for } t < S$$

with point mass

$$1 - G_u(S) = e^{-\alpha S} \quad \text{for } t = S$$

B. Two spells of unemployment summing to t months.

a) End in not-unemployed: Unemployed for t_1, not-unemployed for t_2, unemployed for $T - t_1$, and then not-unemployed for the remaining $S - T - t_2$ months. The probability density for $T = t$ is:

$$f_{I,B,a}(t;S) = \int_0^t \int_0^{S-t} g_u(t_1) g_n(t_2) g_u(t - t_1)[1 - G_n(S - t_2 - t)]\, dt_1 dt_2 = \alpha^2 \beta t(S - t) e^{-\alpha t} e^{-\beta(S-t)}$$

b) Have two spells of unemployment and end in unemployment: Unemployed for t_1, not-unemployed for $S - T$, unemployed for $T - t_1$. The probability density for $T = t$ is:

$$f_{I,B,b}(t;S) = \int_0^t g_u(t_1) g_n(S - t)[1 - G_u(t - t_1)]\, dt_1 = \alpha\beta t e^{-\alpha t} e^{-\beta(S-t)}$$

• The general term for the probability density, with $n + 1$ spells of unemployment, is

$$\alpha e^{-\alpha t} e^{-\beta(S-t)}[[\alpha t\beta(S - t)]^n/(n!)^2 + \beta t[\alpha t\beta(S - t)]^{n-1}/[n!(n - 1)!]]$$

II. **Start in not-Unemployed** (all densities are conditional on starting in not-unemployed, which has probability n_0)

A. Have one spell of unemployment of length t

a) End in not-unemployment: not-unemployed for t_1 months, unemployed for T months, then not-unemployed for the remaining $S - t_1 - T$ months. The probability density for $T = t$ is:

$$f_{II,A,a}(t;S) = \int_0^{S-t} g_n(t_1) g_u(t)[1 - G_n(S - t_1 - t)]\, dt_1 = \alpha\beta(S - t) e^{-\alpha t} e^{-\beta(S-t)}$$

b) End in unemployment: not-unemployed for $S - T$ months, unemployed for T months. The probability density for $T = t$ is:

$$f_{II,A,b}(t;S) = g_n(S - t)[1 - G_u(t)] = \beta e^{-\alpha t} e^{-\beta(T-t)}$$

B. Have two spells of unemployment

a) End in not-unemployment: not-unemployed for t_1, unemployed for t_2, not-unemployed for t_3, unemployed for $T - t_2$, not-unemployed for the remaining $S - T - t_1 - t_3$. The probability that $T = t$ is:

$$f_{II,C,a}(t;S) = .5\beta^2(S - t)^2 e^{-\beta(S-t)}\alpha^2 t e^{-\alpha t}$$

b) End in unemployment: not-unemployed for t_1, unemployed for t_2, not-unemployed for $S - t_1 - T$, and unemployed for $T - t_2$. The probability density for $T = t$ is:

$$f_{II,C,b}(t;S) = \beta^2(S - t)e^{-\beta(S-t)}\alpha t e^{-\alpha t}$$

- The general term for n spells of unemployment is

$$\beta e^{-\alpha t}e^{-\beta(S-t)}[\alpha(S - t)[(\alpha t)^n/n!][(\beta(S - t))^n/(n + 1)!] + [(\alpha t)^{n-1}/(n - 1)!][(\beta(S - t))^{n-1}/(n - 1)!]]$$

Combining all these terms (and weighting the conditional densities by the probabilities of the respective conditioning events) gives the density function of the random variable T (for $0 < T < S$) as

$$\begin{aligned}
(1)\quad f_s(t) &= e^{-\alpha t-\beta(S-t)}\Big\{\alpha u_0\Big[\sum_0^\infty [\alpha t\beta(S - t)]^n/(n!)^2 \\
&\quad + \beta t \sum_0^\infty [\alpha t\beta(S - t)]^n/[n!(n + 1)!]\Big] \\
&\quad + \beta n_0\Big[\alpha(S - t)\sum_0^\infty [\alpha t\beta(s - t)]^n/[n!(n + 1)!] \\
&\quad + \sum_0^\infty [\alpha t\beta(S - t)]^n/(n!)^2\Big]\Big\} \\
&= e^{-\alpha t-\beta(S-t)}\Big[(\alpha u_0 + \beta n_0)\sum_0^\infty [\alpha t\beta(S - t)]^n/(n!)^2 \\
&\quad + (\alpha\beta t u_0 + \alpha\beta(S - t)n_0)\sum_0^\infty \\
&\quad \cdot [\alpha t\beta(S - t)]^n/[n!(n + 1)!]\Big]
\end{aligned}$$

with a point mass of $u_0 e^{-\alpha S}$ at $t = S$ and a point mass of $n_0 e^{-\beta S}$ at $t = 0$. Under the assumption of steady state, $u_0 = \beta/(\alpha + \beta)$, $n_0 = \alpha/(\alpha + \beta)$. (Sattinger 1983, derives this distribution. He derives the densities by a more elegant argument, expresses the density in terms of the modified Bessel functions of the first kind, and derives further properties of the density and distribution function.) This density clearly cannot be evaluated analytically. Because of the double factorial in the denominator, however, the sums converge quickly. Numerical techniques must be used to calculate the distribution function, mean, etc. Romberg integration, also called extrapolation to the limit (see Gerald and Wheatley 1984) is used throughout.

Distribution of Single and In-Progress Spells

For completeness, a review of the statistical methodology for single spells is necessary. (See Heckman and Singer 1984, pp. 97–100.) Assume

a) Time homogeneity (so that densities of duration times and intake rates do not change with calendar time).
b) X is a nonnegative scaler random variable representing the time in a *single spell* of unemployment.
c) $g(x)$ is the density of X.
d) The intake rate (entry rate into unemployment) is k.
e) Current (calendar) time is $\tau = 0$.

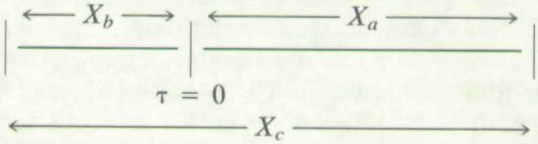
Then

$S(x) = 1 - G(x) = 1 - \int_0^x g(u)\,du =$ survivor function
$\quad =$ proportion of a cohort that survives to have duration greater than or equal x.
$kS(x) =$ number of people (at $\tau = 0$) who have uncompleted duration exactly x
$f(x) = kS(x)/\int_0^\infty kS(u)\,du = S(x)/E(X)$. (Integrating by parts, $\int_0^\infty S(u)\,du = E(X)$.)
$\quad =$ "density" of uncompleted durations observed at $\tau = 0$ (assuming $E(X) = m < \infty$).
$\quad =$ reported BLS "distribution of unemployment durations"
$\quad =$ "unemployment of x weeks as fraction of total weeks"
$h(x) = g(x) \div [1 - G(x)]$
$\quad =$ hazard, or conditional exit rate.

It is useful to define three further random variables that depend on the sampling scheme:

$X_a =$ time to completion for spells currently underway, i.e., sampled at $\tau = 0$ (time After).
$X_b =$ time so far spent in unemployment for spells currently underway (at $\tau = 0$) (time Before).
$X_c =$ total time spent in unemployment for spells currently underway (time Completed). Note that $X_c = X_a + X_b$.

The following figure shows these graphically:



These are actually three different random variables, each different from the original variable, X, the time spent in a randomly chosen new spell. The probability that a spell is in progress is proportional to the length of the spell; i.e., long spells are over-represented and so $E(X_c) \geq E(X)$. It has been pointed out many times with reference to unemployment that the distribution of these three random variables are not the same, and are different from the distribution of leaving times, X. (See, e.g., Heckman and Singer 1984a, Akerlof and Main 1981, Salant 1977, Carlson and Horrigan 1983.) The distributions of these variables are (cf. Heckman and Singer 1984a):

X_b: the density of an interrupted spell is the density of uncompleted spells, derived above. Thus

$$f_b(x_b) = kS(x_b) \Big/ \int_0^\infty kS(u)\,du = S(x_b)/E(X)$$

X_c: in the population, the density of x_c conditional on $0 < x_b < x_c$ is

$$g(x_c|x_b) = g(x_c)/[1 - G(x_b)]$$

Using the expression above for $f_b(x)$,

$$f_c(x_c) = \int_0^x g(x_c|x_b) f_b(x)\,dx_b = \int_0^{x_c} (g(x_c)/E(X))\,dx_b = x_c g(x_c)/E(X).$$

X_a: the density of the forward time x_a can be derived using $x_c = x_a + x_b$:

$$f_a(x_a) = \int_0^\infty g(x_a + x_b|x_b) f_b(x)\,dx_b = \int_0^\infty [g(x_a + x_b)/E(X)]\,dx_b$$

$$= [1/E(X)] \int_{x_a}^\infty g(z)\,dz = [1 - G(x_a)]/E(X) = S(x_a)/E(X).$$

In a time homogeneous environment X_a and X_b have the same distributions. Thus (in a time homogeneous environment) $E(X_c) = 2E(X_a) = 2E(X_b)$.

It is easy to show that $E(X_c) > E(X)$ unless $\mathrm{Var}(X) = 0$ (cf. Heckman and Singer, p. 100). The density $f_c(x)$ of the random variable X_c is the measure suggested by Clark and Summers, and $E(X_c)$ is Akerlof and Main's (1981) "experience weighted" mean duration, S_{ew}. The usefulness of this as a measure of amount or concentration of unemployment, however, is not so clear.[21] The random variable X measures the duration of unemployment for a new entrant. The variable X_c weights less heavily those who have the luck to have short spells. If there are many of these, they may be quite important in unemployment. In fact, it is the distribution function, $G(x)$, of the random variable X which tells the complete story, and from which we can derive the densities and means of new-entrant spells ($g(x)$ and $E(X)$) or completed spells ($f_c(x)$ and $E(X_c)$).

Appendix B

Single Spells for Three-State Model

The distribution of time spent in not-unemployed under the assumption of a three-state Markov process is a mixture of Markovs. Define the instantaneous transition matrix to be:

$$Q = \begin{bmatrix} -q_{ee} & q_{eu} & q_{en} \\ q_{ue} & -q_{uu} & q_{un} \\ q_{ne} & q_{nu} & -q_{nn} \end{bmatrix}$$

Then the density function for leaving from not-unemployed to unemployed is:

$$g(t) = e^{bt}(\alpha b + \beta)/(b - a) - e^{at}(\alpha a + \beta)/(b - a)$$

$$a,b = [-(q_{ee} + q_{nn}) \pm [(q_{ee} + q_{nn})^2 - 4(q_{ee}q_{nn} - q_{ne}q_{en})]^{1/2}]/2$$

$$\alpha = q_{eu}e_0/(e_0 + n_0) + q_{nu}n_0/(e_0 + n_0)$$

$$\beta = (q_{nn}q_{eu} + q_{nu}q_{en})e_0/(e_0 + n_0) + (q_{ee}q_{nu} + q_{eu}q_{ne})n_0/(e_0 + n_0)$$

where e_0 and n_0 are the proportions of the population in employment and NLF at the start of the spell. (See Cox and Miller, section 4.6.) For a

21. Some discussions have not made the issue any clearer. For example, Akerlof and Main (1983) make an analogy between the means of the random variables X and X_c on the one hand, and the mean, median, and mode of a normal variate on the other. Their point is that the mean, median, and mode are three measures of central tendency which happen to coincide for a normal variate, but that like with the mean, median, and mode, one should not assume $E(X)$ and $E(X_c)$ will be the same in all cases. X and X_c are actually different random variables, not just different measures of the same variable. Carefully distinguishing what random variable one is actually measuring is important.

person randomly picked from a steady-state population, these are the steady-state population proportions, e^*, n^*. For a person newly arriving from unemployment, these are

$$e_0/(e_0 + n_0) = q_{ue}/q_{uu}$$
$$n_0/(e_0 + n_0) = q_{un}/q_{uu}.$$

Call the density function for leaving from not-unemployed to unemployed using the steady-state values for employment and NLF $g_s(t)$, and the density for someone just arriving from unemployment $g_n(t)$.

The proportion of the population who have no unemployment in a period of length T (assuming the population is in steady state) is

$$(e^* + n^*)[1 - G_s(T)].$$

The probability of a single spell in a period of length T when in the steady state (letting $q_{uu} = \lambda$) is

$$\begin{aligned}\int_0^T &[P[1 \text{ spell of length } t \mid \text{start in } u] \\ &+ P[1 \text{ spell of length } t, \text{ ending in } e \mid \text{start in } e] \\ &+ P[1 \text{ spell of length } t, \text{ ending in } u \mid \text{start in } e]]\, dt \\ &+ u^* e^{-\lambda T}\end{aligned}$$

where

$$P[1 \text{ spell of length } t \mid \text{start in } u] = \lambda e^{-\lambda t}[1 - G_n(t)]$$

$$P[t, \text{ ending in } e \mid \text{start in } e] = \int_0^{T-t} g_s(t_1)\lambda e^{-\lambda t}[1 - G_n(T - t - t_1)]\, dt_1$$

$$P[t, \text{ ending in } u \mid \text{start in } e] = g_s(T - t)e^{-\lambda t}.$$

This can be analytically integrated and then evaluated.

For white males 25–59, the instantaneous transition matrix is

$$Q = \begin{bmatrix} -0.0148 & 0.0115 & 0.0033 \\ 0.4726 & -0.6186 & 0.1460 \\ 0.0744 & 0.0543 & -0.1287 \end{bmatrix}$$

This is derived from Marston's (1976) average monthly probability matrix, $P(1)$, by

$$Q = \ln P(1) = \sum_0^\infty [P(1) - I]^i/i$$

where I is the identity matrix.[22]

Using these transition rates, 16.5 percent of the population have some unemployment during the year, while 15.5 percent of the population have a single spell of unemployment. Thus only 6.1 percent of those with unemployment have repeat spells. This compares with observed proportions between 29.9 percent and 32.7 percent for men 25–54 for the years 1973–78.

Appendix C

Estimation of Homogeneous Two-State Model

First for the case where no spell information is used. The probability of observing an individual with between t_1 and t_2 months of unemployment (not taking any account of how many spells this occurs in) is:

$$p_{t_1,t_2} = \int_{t_1}^{t_2} f_s(t)\, dt,$$

where $f_s(t)$ is given by Equation (1) above. The contribution to the likelihood function for an individual with between t_1 and t_2 is simply p_{t_1,t_2}, so the contribution to the log likelihood of N people is $N\ln(p_{t_1,t_2})$.

The data used to estimate the model are from Table 1. There are 5.290 million people with between 0 and four weeks of unemployment, 4.019 with between four and 10 weeks, etc. Thus the total log likelihood (ignoring data on number of spells) is

(3a) $\pounds = 5.290^*\ln(p_{0,4}) + 4.019^*\ln(p_{4,10}) + \ \ldots$

When spell information is used, the probability of observing an individual with between t_1 and t_2 months of unemployment, in exactly i spells, is:

$$p^i_{t_1,t_2} = \int_{t_1}^{t_2} f_{i,s}(t)\, dt,$$

where

(4) $f_{i+1,s}(t) = \Big[(\alpha u_0 + \beta n_0) \sum_{0}^{\infty} [\alpha t \beta(S - t)]^i/(i!^*i!)$

22. This expression calculates the instantaneous Markov transition matrix from the observed monthly probability matrix. This is not always possible; i.e., the resulting matrix Q may not be a valid transition matrix. (See J. Coleman 1964, and Singer and Spilerman 1976a, 1976b). It turns out that the formula above does work for the gross flow data in Marston (1976); see Coleman (1984, pp. 26–28 and 130–31).

$$+ \alpha\beta(S - t)n_0 \sum_{0}^{\infty} [\alpha t\beta(S - t)]^i/(i!^*(i + 1)!)$$

$$\left. + [u_0/(S - t)] \sum_{0}^{\infty} [\alpha t\beta(S - t)]^i/(i!^*(i - 1)!) \right] e^{-\alpha t} e^{-\beta(s-t)}$$

The probability of observing someone with no unemployment is then

$$p^0 = 1 - \sum_{1}^{\infty} p^i_{0,t_1} - \ldots - \sum_{1}^{\infty} p^i_{t_{n-1},t_n}.$$

Again, taking the data from Table 1, the log likelihood function (using data on number of spells) is

$$\text{(3b)} \quad £ = 3.885^*\ln(p^1_{0,4}) + 2.467^*\ln(p^1_{4,10}) + \ldots + 156.125^*\ln(p^0)$$
$$+ .690^*\ln(p_{0,4}) + .391^*\ln(p_{4,10}) + \ldots$$

The last terms in the log-likelihood function represent observations on people for which there is no spell information.

The likelihood is a function of α and β. The likelihood is maximized with respect to these two parameters. The problem is that the likelihood function involves integrating over infinite sums. Let us focus on the likelihood which does not use spell information (Equation 3a). The likelihood involves the integral $\int_{t_1}^{t_2} f_s(t)\,dt$. The density $f_s(t)$ is a function of α and β, and is given by Equation (1), which involves infinite sums. Sattinger (1983) has expressions for the distribution function, but these also involve infinite sums. The strategy I have taken is to approximate the density $f_s(t;\alpha,\beta)$ by a truncated sum. Because of the double factorials in the denominator (see Equation 1), a very good approximation can be obtained with a relatively few terms. Once this is done the integral $\int_{t_1}^{t_2} f_s(t)\,dt$ can be calculated by numerical integration.

I use Romberg integration (also known as extrapolation to the limit), using a FORTRAN program from Gerald and Wheatley (1984, pp. 281–83). This allows control of the error in the numerical integration. (The numerical integration is done to an accuracy of about 10^{-6}, and the density $f_s(t)$ to about 10^{-7}.) The IMSL subroutine ZXMIN, which uses the Davidson-Fletcher-Powell algorithm with numerical first derivatives, is used to actually minimize the negative of the log likelihood function.

Estimation of Heterogeneous Two-State Model

For Model 3, the density of leaving times is given by $f^1_s(t;\alpha_1,\beta_1)$ for type-1's and $f^2_s(t;\alpha_2,\beta_2)$ for type-2's. One does not know, however, whether a particular individual is a type-1 or type-2. The "type" of an individual is binomially distributed, with probability δ of being type-1. The density for an individual with t months of unemployment during the year is

$$f_s(t,\delta;\alpha_1,\alpha_2,\beta_1,\beta_2) = \delta f_s^1(t;\alpha_1,\beta_1) + (1-\delta) f_s^2(t;\alpha_2,\beta_2).$$

The probability of observing an individual with between t_1 and t_2 months of unemployment, is

$$p_{t^1,t^2} = \int_{t^1}^{t_2} f_s(t)\, dt.$$

The probability of observing an individual with between t_1 and t_2 months of unemployment, in exactly i spells, is:

$$p^i_{t^1,t^2} = \int_{t^1}^{t_2} f_{i,s}(t;\alpha_1,\alpha_2,\beta_1,\beta_2)\, dt,$$

where

$$f_{i,s}(t,\delta;\alpha_1,\alpha_2,\beta_1,\beta_2) = \delta f^1_{i,s}(t;\alpha_1,\beta_1) + (1-\delta) f^2_{s,i}(t;\alpha_1,\beta_2).$$

and

$$\begin{aligned} f^1_{i+1,s}(t) = \Bigg[(\alpha_1 u_0 + \beta_1 n_0) \sum_0^\infty [\alpha_1 t \beta_1 (S-t)]^i/(i!*i!) \\ + \alpha_1\beta_1(S-t) n_0 \sum_0^\infty [\alpha_1 t \beta_1 (S-t)]^i/(i!*(i+1)!) \\ + [u_0/(S-t)] \sum_0^\infty [\alpha_1 t \beta_1 (S-t)]^i/(i!*(i-1)!)\Bigg]\, e^{-\alpha_1 t - \beta_1 (S-t)} \end{aligned}$$

The probability of observing someone with no unemployment is then

$$p^0 = 1 - \sum_1^\infty p^i_{0,t_1} - \ldots - \sum_1^\infty p^i_{t_{n-1},t_n}.$$

The log likelihood function (using data on number of spells) is thus

$$£ = 3886*\ln(p^1_{1,4}) + 2781*\ln(p^1_{5,10}) + \ldots + 93751*\ln(p^0)$$

For Model 1, the likelihood is the same except that there is only a single α, while for Model 2 there is only one β.

Appendix D

Unemployment During 1984—From March 1985 CPS Tapes

One problem that quickly becomes apparent upon examination of the CPS tapes is that reported weeks cluster around months and half-years. For

Table 8
Weeks Unemployed During 1984—Those with Spell Data from March 1985 CPS

Weeks	Frequency	Percent	Weeks	Frequency	Percent
1	256	2.3	27	92	0.8
2	417	3.8	28	151	1.4
3	506	4.6	29	39	0.4
4	1,103	10.1	30	172	1.6
5	235	2.1	31	47	0.4
6	453	4.1	32	223	2.0
7	235	2.1	33	48	0.4
8	704	6.4	34	87	0.8
9	257	2.3	35	115	1.0
10	392	3.6	36	136	1.2
11	96	0.9	37	57	0.5
12	864	7.9	38	53	0.5
13	405	3.7	39	109	1.0
14	146	1.3	40	225	2.1
15	154	1.4	41	22	0.2
16	304	2.8	42	98	0.9
17	266	2.4	43	50	0.5
18	166	1.5	44	117	1.1
19	65	0.6	45	21	0.2
20	372	3.4	46	61	0.6
21	54	0.5	47	38	0.3
22	393	3.6	48	83	0.8
23	38	0.3	49	56	0.5
24	156	1.4	50	41	0.4
25	71	0.6	51	32	0.3
26	693	6.3			

example, there is a sharp spike at four weeks (one month) and 26 weeks (six months) of unemployment last year. This is shown in Table 8 which has tabulations from the March 1985 CPS tape. What is most likely going on is that people do not recollect exactly how many weeks they were unemployed last year, but rather round to the nearest "natural" break point—a month or half-year. If that is the case, then the apparent detail in the distribution shown in Table 6 is spurious.

One way to circumvent such a problem is to use weeks reported in

Table 9
Estimated Parameters and Implied Proportions for 1984 Using March 1985 Tape Data

	Standard Categories (in percentages)[a]		
	Estimated	Actual	
Total unemployed	10.95	10.94	α_1 = .1657 (.0033)
1–4 weeks	1.66	2.34	mean = 6.0 months
5–10 weeks	2.36	2.16	α_2 = .5134 (.0074)
11–14 weeks	1.42	1.44	mean = 1.9 months
15–26 weeks	3.12	2.57	β_1 = .003711 (.000077)
27–39 weeks	1.62	1.20	mean = 22 years
40–52 weeks	0.77	1.23	β_2 = .2104 (.0042)
1 spell[b]	6.32	6.99	mean = 4.8 months
2 spells[b]	1.66	1.69	Proportion of type 1 =
3 spells[b]	1.45	1.77	0.9489 (.0012)
	Non-Standard Categories (in percentages)[a]		
	Estimated	Actual	
Total unemployed	10.94	10.94	α_1 = .1700 (.0030)
1–2 weeks	0.83	0.78	mean = 5.9 months
3–6 weeks	1.61	2.23	α_2 = .5007 (.0074)
7–10 weeks	1.53	1.49	mean = 2.0 months
11–14 weeks	1.40	1.44	β_1 = .003810 (.000070)
15–20 weeks	1.79	1.23	mean = 22 years
21–32 weeks	2.28	1.99	β_2 = .2185 (.0041)
33–42 weeks	0.86	0.89	mean = 4.6 months
43–52 weeks	0.64	0.89	Proportion of type 1 =
1 spell[b]	6.34	6.99	0.9504 (.0011)
2 spells[b]	1.62	1.69	
3 spells[b]	1.47	1.77	

a. Percent of population.
b. As proportion of those reporting spell data.

categories, as in the tabulations from the BLS reported in Table 1.[23] The question then arises, however, whether the estimates are sensitive to the categories chosen. Table 9 shows estimates using the March 1985 CPS tapes using two alternative category definitions. The first is the standard BLS categories used above. The second are chosen so that months fall in the middle of categories, in an attempt to correctly categorize people who round to the nearest month in reporting. The important observation from Table 9 is that the estimates are quite similar using the two different category definitions.

A formal likelihood ratio test fails to reject the hypothesis that the estimates are different, even at a 10 percent confidence level. The likelihood ratio test statistics is calculated as $\lambda = 2[\pounds(\text{unrestricted}) - \pounds(\text{restricted})]$. In this case, there are two test statistics which can be calculated. The first uses the data in the standard categories, as in the top panel, so that £(unrestricted) = maximum of likelihood function from top panel, and £(restricted) = likelihood function evaluated at parameters from bottom panel. The value of this statistic is 7.8. The second test statistic uses £(unrestricted) = maximum of likelihood function from bottom panel, and £(restricted) = likelihood function evaluated at parameters from top panel. The value of this statistic is 9.2. Both are below the 5 percent critical value of a χ^2 variable with five degrees of freedom, which is 11.1.

References

Akerlof, G. A., and B. G. M. Main. 1980. "Unemployment Spells and Unemployment Experience." *American Economic Review* 70(5):885–93.

———. 1981. "An Experience-Weighted Measure of Employment and Unemployment Durations." *American Economic Review* 71(5):1003–1011.

———. 1983. "Measures of Unemployment Duration as Guides to Research and Policy: Reply." *American Economic Review* 73(5):1151–52.

Blumen, I., M. Kogan, and P. J. McCarthy. 1955. *The Industrial Mobility of Labor as a Probability Process*. Cornell Studies of Industrial and Labor Relations, vol. 6. Ithaca, N.Y.: Cornell University.

Carlson, J. A., and M. W. Horrigan. 1983. "Measures of Unemployment Duration as Guides to Research and Policy: Comment." *American Economic Review* 73(5):1143–50.

Clark, K. B., and L. H. Summers. 1979. "Labor Market Dynamics and Unemployment: A Reconsideration." *Brookings Papers on Economic Activity* 1979(1):13–72.

23. The likelihood of observing 0 to four weeks is simply the integral of the density from 0 to four weeks: $\int_0^4 f_s(t)\,dt$.

Coleman, J. S. 1964. *Introduction to Mathematical Sociology*. New York: Free Press.

Coleman, T. S. 1984. "Essays on Aggregate Labor Market Business Cycle Fluctuations." Doctoral dissertation, University of Chicago, December 1984.

Gerald, C. F., and P. O. Wheatley. 1984. *Applied Numerical Analysis*. Reading, Mass.: Addison-Wesley.

Heckman, J. J., and G. J. Borjas. 1980. "Does Unemployment Cause Future Unemployment? Definitions, Questions, and Answers from a Continuous Time Model of Heterogeneity and State Dependence." *Economica* 47:247–83.

Heckman, J. J., and B. Singer. 1984. "Econometric Duration Analysis." *Journal of Econometrics* 24:63–132.

Marston, S. T. 1976. "Employment Instability and High Unemployment." *Brookings Papers on Economic Activity* 1976(1).

Ridder, G. 1985. "The Number of Transitions and the Occupation Time in Continuous Time Markov Chains: Distributions and Algorithms." Unpublished. Report AE 4/85, University of Amsterdam.

Salant, S. W. 1977. "Search Theory and Duration Data: A Theory of Sorts." *Quarterly Journal of Economics* 91:39–57.

Sattinger, M. 1983. "Distribution of the Time Spent in One State of a Two-State Continuous Time Markov Process." Memo 1983–88, Institute of Economics, University of Aarhus, Denmark.

———. 1985. *Unemployment, Choice, and Inequality*. Heidelberg: Springer-Verlag.

Spilerman, S. 1972. "Extensions of the Mover-Stayer Model." *American Journal of Sociology* 78:599–627.