JAMES P. LUCKETT
Brookings Institution and Cornell University

A Communication: Estimating Unemployment Duration

THE ESTIMATOR for the duration of completed unemployment spells used by Clark and Summers is upwardly biased by at least half a month.[1] They implicitly assume unemployment spells ending during a month continue until the last day of that month when a more reasonable assumption would be that they end in the middle of the month. In a world in which cohort escape rates were constant, this would result in a bias of about half a month. In the real world, where escape rates decline with duration, the bias is larger. Correcting the bias in the completed spell distribution would lower the estimates of the proportion of unemployment due to spells greater than or equal to a given length.

Errors of this general type are fairly common. It is difficult to estimate unemployment flows accurately using monthly data because many spells are short enough to begin and end between consecutive surveys.[2] If all spells were at least one month long, it would be possible to construct an unbiased estimator along the lines followed by the authors. Under present circumstances, however, there does not appear to be an easy way to fix their estimator.

For the sake of simplicity, assume for a moment that every unemployed

1. Kim B. Clark and Lawrence H. Summers, "Labor Market Dynamics and Unemployment: A Reconsideration," *BPEA, 1:1979,* pp. 13–60.

2. It is a common mistake, for example, to estimate the mean duration of a completed spell as the ratio of the unemployed stock to the monthly outflow as measured in the gross-flow data of the U.S. Bureau of Labor Statistics. (See, for example, Stephen T. Marston, "Employment Instability and High Unemployment Rates," *BPEA, 1:1976,* pp. 169–203.) This estimator is upwardly biased because the gross-flow data do not capture the unemployment outflows from spells that begin and end during a single survey month.

0007-2303/79/0002-0477$00.25/0

person has the same probability of leaving unemployment and that these probabilities are constant over time. A Clark-Summers estimator for mean duration, D, in this simplified world would be equivalent to

$$D = \sum_{t=1}^{\infty} (1 - p)^{t-1} pt, \tag{1}$$

where t is measured in months and p is the proportion of those unemployed at one point in time who are not unemployed one month later. The p term can be interpreted as the probability that a spell will end during a one-month period, given it has not ended before the period began. Because spells can end at any time during a month, they can be expected to end, on average, at about the middle of the month.[3] Thus p can also be interpreted as the probability that a spell has a completed length of about half a month. But in equation 1, p is used as if it were the probability that a spell has a completed length of one month. The probability of a completed spell of half a month is treated as if it were zero. If time were measured in one-month increments beginning at $t = 0.5$ instead of 1.0 (which would also require raising $(1 - p)$ to the $(t - 0.5)$ power) this source of bias would be avoided, reducing the estimate of mean length of a completed spell by half a month.

Now let me move closer to the real world and allow escape rates to decline with duration of unemployment. A Clark-Summers estimator for mean duration in this more realistic world is equivalent to

$$D = \sum_{t=1}^{\infty} \prod_{s=1}^{t-1} (1 - p_s) p_t t. \tag{2}$$

They estimate p_t using unpublished data on the proportion of in-process spells of given length that end in the succeeding month. The data are grouped into several ranges of spell duration. If all in-process spells in a particular range are assumed to have a single length equal to the midpoint of that range, and all spells that end in a particular month end in the middle of that month, then the probabilities estimated by the authors may be interpreted as follows:

p_1 = probability that the completed length of a spell is one month, given that the completed length is at least half a month;

3. Actually, the average would be somewhat before the middle of the month, even with p constant, because I am concerned here only with the group that was unemployed at the beginning of the month. As the month progresses, fewer members of this group remain unemployed so the absolute size of the unemployment outflow from the group declines with time.

p_2 = probability that the completed length of a spell is two months, given that the completed length is at least one and a half months; and so on.

When the first source of bias is corrected by moving the time index down half a month, a second source of bias is revealed. Because p declines with duration, each p_t is now too small. What is needed is $p_{0.5}$, $p_{1.5}$, and so on, where p_t is the probability the completed length of a spell is less than or equal to $t + 0.5$ given it is greater than $t - 0.5$. But this estimator is not available: there is no observation on $p_{0.5}$ because there is no way of identifying in the data those who have just begun unemployment. Using p_1 and p_2, and so on, in place of $p_{0.5}$ and $p_{1.5}$, and so on, would shift the weight in the probability distribution toward longer spells by making each probability of ending a spell too small and each probability of continuing unemployment, $(1 + p_t)$, too large. Fortunately, there are maximum likelihood estimators based on weekly models that avoid these problems.[4]

4. An advantage of the procedure used by Clark and Summers is that, unlike most weekly models, it does not require an assumption of labor market equilibrium. However, disequilibrium can be handled in a weekly model. See James P. Luckett, "The Estimation of Completed Unemployment Spell Durations from Published Data under Non-Steady-State Conditions," paper prepared for the 1978 annual meeting of the Econometrics Society.