LABOR FORCE TRANSITIONS: A COMPARISON OF UNEMPLOYMENT ESTIMATES FROM TWO LONGITUDINAL SURVEYS[1]

Martha S. Hill, Survey Research Center, Univ. of Michigan
Daniel H. Hill, Survey Research Center, Univ. of Michigan

Introduction

The existence of micro-level panel data has lead to some major revisions in the way we think of labor market behaviors, bringing into question the very concept of unemployment as an unanticipated negative shock (e.g., Heckman and MaCurdy, 1980; MaCurdy, 1981; and Ashenfelter and Ham, 1979) and providing a novel perspective on the costs and even the definition of unemployment (e.g., Adams, 1985, and Abowd and Ashenfelter, 1981). The policy implications[2] are of such importance that we need to be very careful that they are not merely artifacts of the data due perhaps to measurement problems. Certain aspects of this issue have already been investigated. A number of studies of the quality of self-reported unemployment, for example, have recently appeared in the literature (e.g., Bowers and Horvath, 1984; or Poterba and Summers, 1985), and these present evidence of appreciable errors in reports of unemployment status and duration. Evidence of failure to report spells of unemployment is presented by Mathiowetz (1984), who finds that omission of reports of unemployment spells increases substantially with the length of time from the termination of the spell to the interview.

While validation studies would offer the most reliable evidence on the quality of unemployment data, comparisons across different surveys can also be informative. The Survey of Income and Program Participation (SIPP)[3] and the Panel Study of Income Dynamics (PSID)[4] are two studies with similar unemployment information collected from comparable segments of the population, but with an important design difference. The SIPP data is collected on a more frequent basis than the PSID, and Mathiowetz's findings suggest that the more frequent interviewing schedule of the SIPP should result in more complete reporting of unemployment spells, especially short ones.

The purpose of the present analysis is to compare reports of unemployment experiences for the July through December 1983 period obtained from the PSID and the SIPP. Duration of unemployment and transitions to employment are the focus since these labor force experiences are frequent enough during a short time span to allow relatively precise estimates with moderate sized surveys. The research will address two questions. First, do measures of the incidence and duration of unemployment differ between the studies?, and, second, do these differences result in different estimates of the parameters of a multivariate model of unemployment?

Survey Methodologies

In addition to the differences in frequency of contact between the PSID and the SIPP, there are also differences in question design. The PSID tends to require more of the respondent in terms of contextual detail but also provides more structure to the questions. On the other hand, the SIPP provides more precise dating of employment events. The PSID obtains information on the timing of unemployment events in two distinct question sequences. One sequence proceeds iteratively asking a series of questions about increasingly remote jobs and unemployment spells between jobs, until the entire previous calendar year is accounted for. Since this would not capture periods of temporary layoff, with returns to the same job, a second set of questions asks for total amounts and timing of work lost due to specific labor force events such as illness, unemployment, strike, or vacation during the reference year. The data from these sequences are extensively edited for completeness and consistency in the SRC's Ann Arbor Coding/Editing facility and are processed in the form of monthly dating for a variety of employment events. The SIPP procedure consists of providing the respondent a calendar of the weeks of the four-month reference period just proceeding the interview and requesting a report of which weeks were ones with a job, which were weeks with unpaid absences from a job and what the main reason was for them, and which were weeks looking for a job.

In addition to differences in frequency of interviews (and, thereby, length of recall), and the method of eliciting the information from the respondent, the two studies differ in designations of individuals (and "families") and in who is to be the informant. Since it was designed in 1966, the PSID uses the, now archaic, "head of household" definition of the designated respondent, with husbands reporting for their wives. The SIPP, on the other hand, designates each adult as a respondent and attempts to interview all such persons in its sample households. Both studies allow proxy reports when the designated respondent is not available.

Sample Restrictions and Definition of Variables

Whenever one wishes to compare or combine two datasets it is necessary to restrict them to their common content. Often this intersection of elements is only a small fraction of the total content of either study. The present analysis is no exception.

Because of differences in the timing of reference period of the two studies we are forced to restrict our reference period to a rather short segment of the total era of history covered by either study. Since the period of history covered by the two studies is the beginning of what has come to be known as the Reagan Recovery, we cannot assume any stationarity in unemployment behavior and, thus, must restrict the comparisons of the unemployment data in the two studies to the same time period. Since the SIPP sample was introduced to the study on a rotating basis, we face a trade-off between the number of SIPP rotation groups included and the length of the period over which we will measure unemployment experiences. A compromise was struck with the selection of the six month period from July to December 1983, a period covered by the first two rotation groups of the SIPP. This, of course, throws out far more information from the SIPP than from the PSID, and analysts must remember in evaluating our findings that more precise estimates for both studies, and especially the SIPP, are possible if the studies were to be analyzed independently. The period is, however, long enough to capture important aspects of transitions out of unemployment, as Feldstein's (1976) work stressing the importance of temporary layoffs suggests.

A further restriction in the comparative analysis concerns the sample. In the PSID, the detailed employment questions are asked only of 'heads' for themselves and of heads about their 'wives'. This means that in order to make comparisons across the studies it is necessary to restrict the larger SIPP population of inference to individuals who would be so classified by the PSID. Thus, we restrict the SIPP sample to those individuals whose relationship to the "reference person" is either self or spouse. Furthermore, since in 1983 individuals who were out of the labor force (i.e. retired, 'housewife', permanently disabled or student) at the time of the PSID interview were not asked the detailed employment sequence,[5] it is necessary to further restrict the SIPP sample to persons who are either employed or unemployed in the second week of April 1984 (the modal 1983 PSID interview week). Finally, a natural extension of the restriction to those either employed or unemployed at the time of interview was to confine the sample further to people always in the labor force (i.e., employed or unemployed) throughout the July to December period. The net result of these restrictions are comparable samples of 5218 in the PSID and 6212 in the SIPP for what we will term the "adult

persistent labor force".

Our analysis also attempts to maintain comparability in the definition of variables. In both studies unemployment includes both time looking for work when without a job and temporary layoff. Time with a job and either working, sick, on strike, or on vacation is counted as employment time. The variables used as covariates in the multivariate analysis are limited to a set available in both studies. These variables are believed to affect either the individual's potential wage in a new job (i.e. age, education, race, and gender) or his reservation wage (i.e. family income needs level, the earnings of other family members, asset income, means tested transfer income, and whether receiving unemployment compensation). A table describing the variables is available from the authors upon request.

The bivariate analyses are weighted to correct for differences in initial probabilities of selection and for differential nonresponse. Limitations of the computer programs precluded weighting in the multivariate analysis.

Results

Table 1 presents the proportions of individuals in the two (restricted) samples who experience some unemployment during the six months for which the studies overlap. These figures are presented separately for men and women as well as both genders combined. As we would expect given the shorter recall period, respondents in the SIPP are somewhat more likely to report having some unemployment. The 11.2% average incidence estimate from the SIPP is roughly fifteen percent greater than the estimate from the PSID. While we have not yet computed complex sampling errors for either estimate, sampling errors computed under the assumption of simple random sampling would suggest that this difference is significant at the five percent level. With typical design effects from the PSID of less than 1.5, we would expect this difference to remain close to the margin of significance even with complex sampling errors. The slightly greater difference between SIPP and PSID unemployment incidence estimates for women is not, however, significant even under the assumption of simple random sampling.

Table 1
Percent of Adult Persistent Labor Force Members Unemployed at Some Time July-December 1983
[Sample Sizes in Brackets]

	Men	Women	All
SIPP	11.4% [3,666]	11.0% [2,552]	11.2% [6,218]
PSID	10.0% [2,970]	9.3% [2,242]	9.7% [5,212]

When attention is confined to the subsample reporting some unemployment, more dramatic differences between the studies appear (see Table 2). The average amount of time reportedly lost from work due to unemployment for males in the SIPP is nearly a month (4.11 weeks) longer than that in the PSID. This difference of nearly forty percent is highly significant and, combined with the fact that the average number of transitions out of unemployment reported for males in the SIPP (.53) is lower than that in the PSID (.65), suggests that a major difference between the two studies is a higher proportion of long-term unemployed in the SIPP. The corresponding differences for women are barely perceptible and are far from significant. These same patterns persist when the spell itself is used as the unit of analysis. Since we expected short spells of unemployment rather than long ones to be better reported in the SIPP because of its more frequent interviewing schedule, these results are somewhat puzzling.

Table 2
Mean Total Unemployment and Transitions Out of It
Among Adult Persistent Labor Force Members
Unemployed Some Time July- December 1983

	Men	Women	All
Mean (Standard Deviation) Weeks Unemployed			
SIPP	15.08 (8.45)	14.00 (8.30)	14.65 (8.40)
PSID	10.97 (8.98)	14.10 (8.97)	12.17 (.9.10)
Mean (Standard Deviation) Number of Unemployment to Employment Transitions			
SIPP	0.53 (0.68)	0.56 (0.70)	0.55 (0.69)
PSID	0.65 (0.78)	0.51 (0.55)	0.59 (0.69)
Number of the Unemployed Persons in Samples			
SIPP	422	272	694
PSID	360	261	621

A Proportional Hazards Model of Transitions to Employment

In order to obtain some preliminary notion of how the study differences might affect the estimates of structural model parameters, we use the data to estimate a proportional hazards model of the transition from unemployment to employment. Following Cox (1972) we assume that the hazard rate of re-employment for individual i at time t is of the form:

$$\lambda^{i}(t) = \lambda(t)\, e^{\beta' x_i} \qquad 1)$$

where x_i is a vector of the characteristics of the individual. As was the case in selecting the time period and types of individuals to include in the analysis, we are limited in our selection of individual characteristics to those which are collected in a comparable manner in each of the studies. These consist of demographic characteristics (age, race, gender) and measures of their resources (asset income, income of other family members, welfare income, and a dummy for whether they received unemployment compensation during the period) and needs (the official poverty needs standard for the family in which they lived during the period). Estimates of the parameter vector β are obtained by maximizing the partial likelihood function:

$$L_1 = \prod_{i=1}^{n} [(e^{\beta' x_i}) \div (\sum_{h \epsilon R(t_i)} e^{\beta' x_h})] \qquad 2)$$

where n is the number of completed spells and $R(t_i)$ is the set of uncompleted spells at time t_i. This latter group can be thought of as the set of individuals still 'at risk' (unemployed) when individual i moves from unemployment to employment.

We estimate the proportional hazards model under a variety of conditions—first for the PSID and the SIPP separately and then combining the two studies to test for differences. Variants of the same comparisons suggested by differences in this analysis are then explored.

Table 3 presents the parameter estimates obtained when the proportional hazards model is estimated on the two data sets separately. Both studies tend to exhibit effects in the expected directions, and there is considerable agreement in the signs and magnitudes of their coefficients. In both studies the rate of exit from unemployment to employment is higher the younger and the better educated the individual is. Positive coefficients on the non-black indicator also appear in both studies. Being better educated and non-black would tend to raise wage offers whereas being younger would tend to increase the number of job offers since the pay-back period for training a worker would be longer. Higher wage offers and more job offers would facilitate earlier exit from unemployment to employment.

Table 3
Maximum Partial-Likelihood Estimates
of the Proportional Re-employment Hazard Model,
Separately for PSID and SIPP

	SIPP	PSID
Age	−.0135*	−.0105
	(.0069)	(.0078)
Education	.0083	.0072
	(.0189)	(.0258)
Non-Black	.1229	.4831**
	(.1585)	(.1174)
Male	−.0502	.2127+
	(.1093)	(.1212)
Needs	.0275	.0322
	(.0203)	(.0240)
Others' Earnings	.0007	.0041
	(.0053)	(.0063)
Asset Income	−.0464	.0001
	(.0356)	(.0241)
Welfare Income	−.0667+	−.1387**
	(.0347)	(.0385)
Whether Unemployment Compensation	−.2758**	−.1437
	(.1046)	(.1136)
Chi-Square	19.92	53.36
d.f.	9	9
N	797	692

+ Significant at .10 level.
* Significant at .05 level.
** Significant at .01 level.

The income and needs covariates represent factors affecting the individual's reservation wage, the wage level that a job offer would have to match or exceed in order to attract the individual to employment. The greater the availability of income other than that to be realized by the individual working, the higher the individual would tend to set his or her reservation wage, which would tend to prolong unemployment. The financial needs of the person's family would also tend to influence the reservation wage, but in the opposite direction—lowering it and thus making the individual more likely to accept a job offer. Welfare income, whether unemployment compensation received, and the needs variables produce coefficients in both studies consistent with this model.

Differences between the two studies do appear, however. When the two samples are combined, and a study indicator is included both additively and multiplicatively with other predictors, we are able to reject the hypothesis that the two studies yield equivalent measures of employment transitions and their determinants.[6] Significantly different effects appear for the non-black indicator and welfare income. The PSID estimates larger absolute effects of these two factors than does the SIPP. While part of the difference in the effects of welfare income could be due to differences in the components included (the SIPP includes WIC and energy assistance, whereas the PSID does not), it is unlikely that definitional differences are the underlying factor in the differential effects of race.

The results of Table 3, with the two studies examined separately, indicate that, in terms of the overall goodness-of-fit, the PSID measures have a significantly stronger systematic component than do the SIPP data. While the χ-square of 19.92 with 9 degrees of freedom for the SIPP does indicate a significant overall relationship between the hazard rate for re-employment and its explanatory variables, it is only marginally significant, and is much less significant than the χ-square of 53.36 obtained with the same variables for the PSID. This difference is due primarily to the much stronger effects of the race and welfare income on the probability of exiting unemployment to employment in the PSID. We should note, of course, that a superior fit of our re-employment model does not, in itself, indicate that the PSID data are better. It indicates only that the types of transitions observed with the PSID data are more strongly related to the predictor variables. It may well be that SIPP is better at detecting short spells of unemployment but that these episodes are less predictable than are the more salient spells reported in the PSID.

Both the SIPP and the PSID samples contain spells with observed start dates and 'left-censored' spells (ones with start dates predating the onset of the observation period). Left-censored spells present problems for hazard analysis, thus the sensitivity of our results to their inclusion is an issue. To investigate this, we stratified the samples in both the SIPP and the PSID into left-censored spells and non-left-censored spells and calculated distinct survival probabilities for each type of spell in each study.[7] The results are presented in Figures 1 and 2, for the PSID and the SIPP, respectively.

Focusing first on the non-left-censored curves, we find substantial agreement between the two studies, but some notable differences. For both the SIPP and PSID the survival curve becomes flatter with time, as we would expect.[8] In addition, during the four weeks following the onset of unemployment and over the long-duration range of unemployment (16–26 weeks) the curves are similar in shape. In the intermediate range, though, they differ; there the SIPP shows a more gradual re-employment process than does the PSID. Thus, again, the SIPP indicates more unemployment. But, apparently, the additional unemployment is from more intermediate-length spells rather than more spells of short duration.

Deviations of left-censored curves from expectations provide additional insight into the SIPP-PSID difference. A left-censored curve can be expected to follow a particular pattern in terms of its shape and placement relative to its corresponding non-left-censored curve. Since the non-left-censored survival curve flattens with time, we would expect the left-censored curve to start at a higher level, be even flatter, and so end at an even higher level than the non-left-censored one[9].

With the PSID (Figure 1) we find roughly this pattern, although the left-censored curve flattens less rapidly than expected. Measurement error from using monthly data to measure weekly dating could be causing this divergence from expectations.

A more dramatic divergence from expectations arises in the SIPP data (Figure 2). For the left-censored cases the likelihood of exit to employment is much greater in the 13–17 week range than anywhere else. This distorts the pattern of

Figure 1

Unemployment Survival PSID

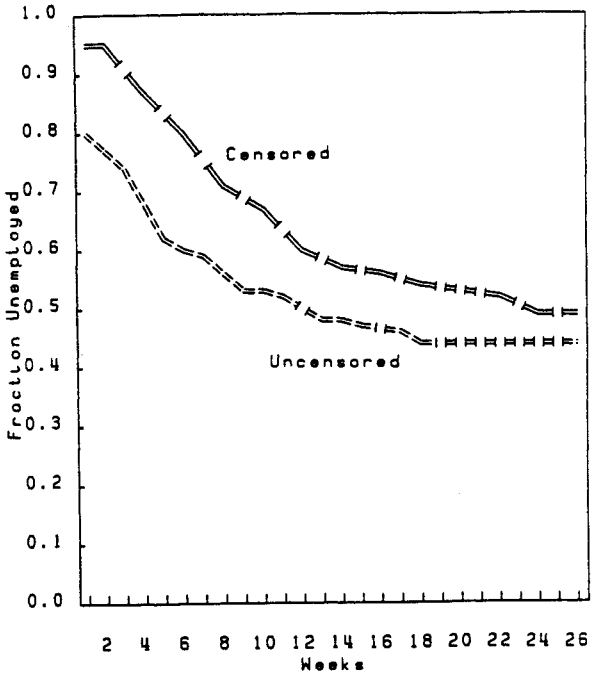
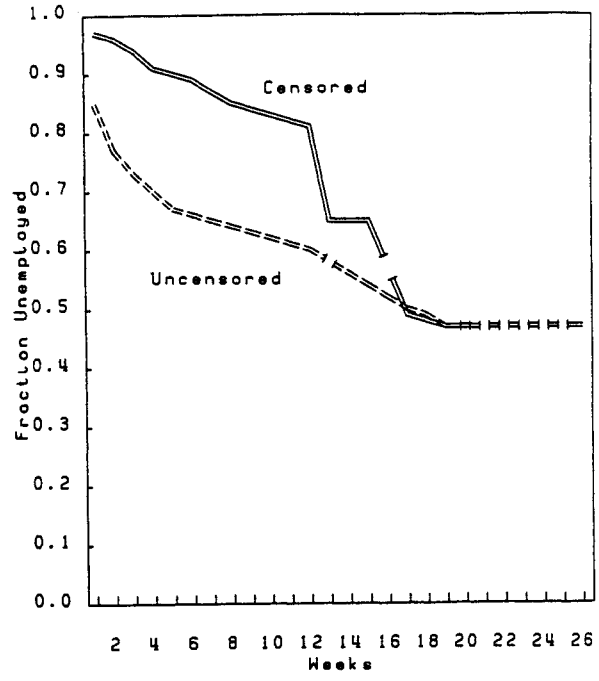


Figure 2

Unemployment Survival Functions-SIPP



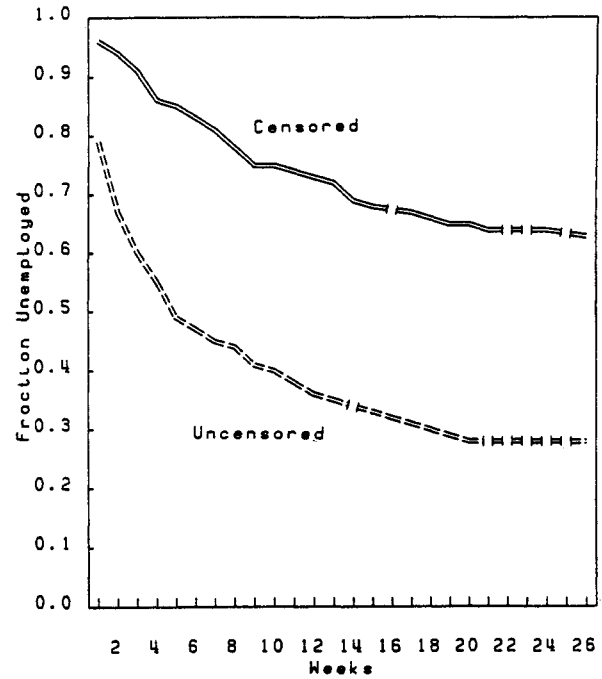
an ever-flatter slope, and, in fact, causes the left-censored curve to merge with the non-left-censored curve beginning at the 17-week point.

This length for an unemployment spell coincides very closely with the spell length associated with a transition from unemployment throughout Wave 1 of the SIPP to employment at the very beginning of Wave 2.[10] Since the reported unemployment-to-employment transitions in the SIPP are, indeed, more likely to happen at the wave-to-wave seam than at other times[11], it is likely that transitions reported then are more prone to misreporting. Erroneous reports cannot be individually identified, but likely candidates include those obtained for persons with movements from unemployment to employment, or the reverse, at the wave-to-wave seam and no other time. While eliminating the seam-only transition cases is not a viable solution to the reporting error problem, it is a useful technique for obtaining an approximate idea of its magnitude.[12]

With the SIPP sample modified in this way, the pattern of the left-censored survival curve relative to the non-left-censored one is exactly as expected—initially higher and flattening out more quickly over time so that the final level difference is greater than the initial one.[13] (See Figure 3.) Further, reestimating the combined PSID-SIPP model with the SIPP portion modified as described above substantially reduces the differences in the model estimates for the two studies; the global χ-squares of the models with and without the study-specific interactions drops from about 30 (d.f. 10) to 17, a difference which is significant at only the .10 rather than .01 level. Erroneous seam transitions, thus, appear to account for a substantial part of the difference in the multivariate results of the SIPP and the PSID.

Figure 3

Unemployment Survival SIPP w/o Seams



Conclusions and Directions for Future Research

While the PSID and SIPP employment event history sequences are intended to measure the same labor market behaviors, differences in the two study designs do have significant effects on the measures obtained. While both studies undoubtedly miss some episodes of unemployment, the more frequent interviewing schedule of the SIPP does seem to result in a more complete accounting of unemployment than does the PSID. For comparable periods of history and populations of inference the SIPP obtains estimates of unemployment incidence which are roughly fifteen percent higher than those obtained in the PSID. Since it is less likely that individuals will report unemployment when they have had none than it is that casual unemployment will be forgotten, this result, along with the larger sample size of the SIPP, would argue in favor of analysis of SIPP data over PSID data for studies of unemployment incidence. However, the SIPP is not without

some questionable aspects.

For males with some unemployment the estimated amounts of unemployment are dramatically higher in the SIPP than in the PSID. In this case, however, it is less clear that more is better. This result could, for instance, be a reflection of a greater tendency, say, with a short reference period to report being unemployed the entire reference period, even when there were, in reality, some periods of employment. Proxy reports may, however, be the major problem. SIPP and PSID results concerning levels of unemployment among women are quite similar, and for that subgroup the frequency of proxy reports is of comparable size in the two studies. The PSID, however, has a much lower frequency of proxy reports for men than does the SIPP.

It is unclear which study is superior for the purpose of estimating the parameters of multivariate behavioral models of unemployment durations or transitions. Although the signs of the parameters of our proportional hazards model were found to be in general agreement across the studies, there were marginally significant differences in the magnitudes of a number of parameters and strongly significant differences in the overall goodness of fit. Overall the PSID data tended to yield stronger associations between the probability of re-employment and the exogenous variables included in the analysis. The effects of race and welfare income on the probability of becoming re-employed are much stronger in the PSID. The larger sample and better coverage of short spells of unemployment would seem to argue for use of the SIPP data in studies focusing on only one of the two studies, but the problem of inordinately large numbers of reported transitions occurring at the seams of the waves is sufficiently serious and difficult to model as to make the PSID an attractive alternative.

The study differences we have detected in the present analysis suggest a number a lines of potential future research. First, whether the SIPP does in fact obtain better reports of short spells needs to be investigated with longer observation periods which will allow the inclusion of all four SIPP rotation groups. Second, the predominance of transitions at the seams needs to be further investigated to see to what extent it is a reflection of the higher proportion of proxy interviews in the SIPP overall, and thereby more switches between self- and proxy reports. Imputations may also be a source of concern. Finally, the 'seam problem' should be investigated in the PSID as soon as the second wave of detailed employment event histories are merged to the first. Since in the PSID the reference periods are designed to overlap by six months, it should provide considerable methodological leverage in the analysis of these within/between wave inconsistencies.

REFERENCES

Abowd, J.M. and Ashenfelter, O.C. (1981) "Unemployment and Compensating Wage Differentials" (Report No. 7923) (Chicago: University of Chicago, Center for Mathematical Studies in Business and Economics).

Adams, J.D. (1985) "Permanent Differences in Unemployment and Permanent Wage Differentials" *Quarterly Journal of Economics* 26–56.

Ashenfelter, O.C. and Ham, J. (1979) "Education, Unemployment, and Earnings" *Journal of Political Economy*, 87:99–116.

Bowers, N. and Horvath, F.W. (1984) "Keeping Time: An Analysis of Errors in the Measurement of Unemployment Duration", *Journal of Business and Economic Statistics* 2:140–149.

Cox, D.R. (1972) "Regression Models and Life Tables." *Journal of the Royal Statistical Society* ser. B, 34:187–220.

Cox, D.R. (1975) "Partial Likelihood", *Biometrika* 62:269–844.

Feldstein, M. (1976) "Temporary Layoffs in the Theory of Unemployment", *Journal of Political Economy* 937–957.

Heckman, J.J. and MaCurdy, T. (1980) "A Life Cycle Model of Female Labor Supply" *Review of Economic Studies* 47: 47–74.

MaCurdy, T. (1981) "An Empirical Model of Labor Supply in a Life Cycle Setting" *Journal of Political Economy* 89: 1059–1085.

Mathiowetz, N.A. (1984) "The Problems of Omissions and Telescoping Error: New Evidence from a Study of Unemployment" *Proceedings of the American Statistical Society, Session on Survey Methods.*

Poterba, J.M. and Summers, L.H. (1985) "Adjusting the Gross Changes Data: Implications for Labor Market Dynamics", *Proceedings of the Conference on Gross Flows in Labor Force Statistics* (Washington: U.S. Departments of Commerce and Labor) 81–95.

FOOTNOTES

[1]The authors gratefully acknowledge the helpful comments of Greg Duncan, Nancy Mathiowetz, and Willard Rodgers.

[2]See Feldstein (1976) for discussion of the theoretical and policy implications of temporary layoffs.

[3]The SIPP is a survey of households conducted by the Bureau of the Census with data for a panel collected every four months for a two and one-half year period. It began collecting information in October 1983.

[4]The PSID is an annual survey of households conducted by the Survey Research Center of the University of Michigan. It has been collecting income and labor market information, among other things, from the same sample of individuals, and their descendents, each year since 1968. Since 1983, detailed information on the timing of work history events has also been collected.

[5]Beginning in 1984 the detailed sequences were asked of all PSID primary adults regardless of labor-force status at the time of the interview.

[6]The global χ-square obtained when all the study indicator interactions plus a study indicator dummy are included is 82.82 (d.f. 19), which is significantly higher than the 52.94 (d.f. 9) when they are not. A table with the full detail of this analysis is available from the authors upon request.

[7]Ideally one would like to know the beginning date of all spells, but when this is not possible a second-best solution is to restrict the sample to non-left-censored spells. For our samples, though, the left-censored spells constitute a very sizable portion of all unemployment spells: in the SIPP 424 of the 797 unemployment spells sampled are left-censored, and in the PSID 339 out of 692 spells are left-censored. Thus eliminating left-censored spells would also present problems. The stratification approach was the third best alternative.

[8]The flattening with time is expected since, over time, the people remaining at risk will increasingly become ones with circumstances not amenable to becoming employed.

[9]This follows because spells classified as left-censored are known to have begun on or before what we treat as their beginning, and thus most are at a further stage of unemployment than a non-left-censored case tracked for the same length of time.

[10]The length of time between the end of our observation period and the 'seam' between the waves is thirteen weeks for rotation group 1 (as is the length of time from the beginning of our reference period and the seam for rotation group 2), and seventeen weeks is the length of a full wave.

[11]When we partition the SIPP into pairs of weeks for each person we obtain 17,350 person-week pairs. The mean fraction of these pairs which involve a transition from unemployment to employment is .0218. Looking only at the 694 pairs of weeks at the seam, however, the comparable figure is .1211, and for those 16,656 person-week pairs not at the seam it is .0177. The chances of a re-employment event being reported between waves is, therefore, slightly less than seven times as great as within a wave.

[12]We do not, of course, recommend this as a solution to the seam problem in estimating behavioral models from the SIPP. We drop the 'seam' cases merely as a means of seeing if the study differences could be caused by them.

[13]The level of the modified-SIPP-sample curves should not be taken as representative of the unemployment process since removing those individuals from the SIPP who reported transitions from one homogeneous state to the opposite at the wave seam can be expected to eliminate a disproportionate number of long spells of unemployment. It is not surprising, then, that the modified SIPP sample yields lower survival probabilities than either the PSID or the full SIPP sample.