Research Summaries

Response variation in the CPS: caveats for the unemployment analyst

JAMES M. POTERBA AND LAWRENCE H. SUMMERS

The Current Population Survey (CPS), conducted by the Census Bureau for the Bureau of Labor Statistics, is one of the principal sources of data on U.S. labor markets. It has been used in numerous investigations of unemployment, because it provides descriptive information about the characteristics of jobless workers and about their unemployment experience. Data on the duration of unemployment spells and on the factors affecting reported unemployment spell lengths have been subject to particularly intensive study to determine how public policies can affect the amount of time that workers spend in unemployment, and how the reason for an individual's entry into unemployment influences his or her subsequent labor market activity.

Relatively little is known about the frequency of response errors in CPS survey data and their implications for empirical research. The Census Bureau's CPS Reinterview Survey Program provides some indication of response variation by helping to determine whether respondents answer questions consistently within a particular survey month. However, the Reinterview Survey does not indicate whether individuals provide logically consistent survey responses from month to month. The recent advent of panel data sets containing information on survey participants for several consecutive months makes it particularly important to determine if individuals answer similar questions in similar ways in different survey months. If reported durations of and reasons for joblessness are logically inconsistent over time, analyses that focus on changes in individual behavior are likely to be flawed by spurious changes due to reporting error.

This article draws upon a potentially rich source of information for evaluating survey answers, a 3-month matched sample of respondents, to gauge the problem of response variability in the CPS. Our analysis is divided into four parts. The first section reviews evidence from the Reinterview Survey on individuals' reported labor market status. In the second section, we examine the consistency across time of reported unemployment durations and consider the salience of the unemployment/not-in-the-labor-force (NILF) distinction. The third section presents evidence on the consistency over time of individuals' reported reasons for unemployment. And the final section considers the implications of our results for empirical research in labor economics, using both the CPS and other data sets.

Unemployment status misreporting

Reporting errors are a substantial problem in the CPS. The incidence of errors due to response and coding mistakes is well documented by the Reinterview Surveys, during which a subsample of the households included in each month's CPS are recontacted.[1] These secondary interviews, which usually occur about a week after the original survey, ask respondents to describe their activities in the preceding week. In some cases—those included in the "nonreconciled" component of the Reinterview Survey—no attempt is made to determine which, if either, of two different responses on the original and reinterview surveys is correct. However, for the "reconciled" subgroup of the Reinterview Survey, which typically constitutes about one-third of the reinterviewed households, the second interviewer compares the responses from the first survey with the reinterview answers before leaving the household, and attempts to resolve any conflicts.[2]

The reconciled Reinterview Surveys permit analysis of employment status coding errors. For May 1976, table 1 shows the fraction of individuals in each labor market category after reconciliation, by category as reported in the initial survey. Most (99.1 percent) of the employed CPS respondents had been correctly classified in the regular CPS, as had most of those who were truly out of the labor force (99.2 percent). However, a substantial fraction of unemployed individuals had initially been reported in other categories. Ten percent of the truly unemployed had been classified as not in the labor force (NILF) and an additional 3.6 percent had been recorded as employed. There is some evidence that the mismeasurement problem was greater for

James M. Poterba is assistant professor of economics at the Massachusetts Institute of Technology. Lawrence H. Summers is professor of economics at Harvard University.The National Bureau of Economic Research (NBER) and the National Science Foundation provided financial support for this research. However, the opinions expressed in this article are those of the authors and not necessarily those of NBER.

women than for men.

The finding that some unemployed individuals are misclassified is important for studies of unemployment dynamics. If nearly 15 percent of unemployed individuals are incorrectly classified in a given month, then the effect on month-to-month transitions between labor force states must be considered. Studies of labor market behavior based on gross flows or panel data from the CPS may be adversely affected.[3]

In particular, the data in table 1 suggest that there is some confusion between the states of "unemployment" and "not in the labor force." As we will show later, many unemployed persons who drop out of the labor force at some point before again becoming unemployed report themselves as experiencing one *ongoing* spell of unemployment. According to the Reinterview Survey, only 0.25 percent of individuals initially classified as NILF are actually unemployed, because many individuals in the population are genuinely not in the labor force and are rather unlikely to be experiencing an unemployment spell. However, conditional upon an individual's having been unemployed the month before, the measurement error rates for the NILF category may be large—far larger than those in the table.[4]

Christopher Flinn and James Heckman have argued that the states of unemployment and NILF are well-defined and distinct.[5] They draw evidence from models showing clear differences between persons who are unemployed and those who are not in the labor force in the probability of becoming employed. However, this evidence is not relevant to understanding whether a large fraction of those who are unemployed drift in and out of the NILF category with little or no change in behavior. Again, the explanation of Heckman's and Flinn's finding is that there are a large number of individuals classified as NILF who are not casual entrants to the labor force. These persons—whether disabled, retired, or otherwise unable or unfit to work—are conceptually distinct from the unemployed, who are searching for work. Thus, a small fraction of all NILF respondents, but a substantial portion of those NILF respondents who were unemployed in the preceding month, may actually be searching for work and ready to accept a job in a given current month. These are the miscategorized workers on whom we focus.[6]

Table 1. Probabilities of reporting labor force status as employed, unemployed, or NILF in the regular CPS, by "true" status as determined by the Reinterview Survey, May 1976

"True" status	Status as reported in the regular CPS		
	Employed	Unemployed	NILF
Total:[1]			
Employed	.9905	.0016	.0079
Unemployed	.0356	.8602	.1041
NILF	.0053	.0025	.9923
Men:[2]			
Employed	.9922	.0013	.0065
Unemployed	.0474	.8720	.0806
NILF	.0062	.0048	.9890
Women:[3]			
Employed	.9892	.0019	.0089
Unemployed	.0194	.8442	.1363
NILF	.0049	.0015	.9936

[1]Sample size = 7,079.

[2]Sample size = 3,329.

[3]Sample size = 3,750.

SOURCE: Tables were computed from "General Labor Force Status in the CPS Reinterview by Labor Force Status in the Original Interview, Both Sexes, Total, After Reconciliation," May 1976, Bureau of the Census (unpublished).

Reported spell durations

The Current Population Survey interviews individuals in several consecutive months, and CPS "match files" contain data on all interviews with a group of survey participants. These data may be used to examine month-to-month changes in individuals' reported unemployment spell durations. Survey respondents who report that they are unemployed are asked how many weeks they have been "without a job and looking for work." If individuals who are unemployed in 2 consecutive months accurately describe their labor market experience, the reported unemployment spell duration in the second CPS monthly interview should exceed the first-month reported duration by 4 or 5 weeks.[7]

We obtained data on survey participants who were unemployed in May 1976 and were interviewed again in June 1976. These data were used to compute the *difference* between each individual's reported unemployment spell durations in May and June:

$$\text{DIFF} = \text{DUR}_{\text{June}} - \text{DUR}_{\text{May}}$$

The measurement of *DIFF* is complicated by several factors. First, some survey participants who are unemployed on both survey dates may report a much lower spell duration in the second interview because at some point between surveys they either found a job or stopped searching. Because there is no way of determining whether inconsistent reports with second-interview durations of less than 5 weeks are spurious, we report results which both include and exclude this group from the calculations. Second, some respondents may appear to make inconsistent responses because they have been unemployed for so long that the duration values for both months are coded "99." Duration is recorded in a two-digit data field, so that spells of more than 99 weeks cannot be reported. However, this problem did not appear to be substantial in our data set. Only 1.7 percent of the respondents whose spell durations did not change from month to month had reported "99" on the May survey, and a negligible fraction had had May durations of between 96 and 98 weeks.

Summary statistics for *DIFF* are displayed in table 2. The top panel of the table shows the results of calculations which excluded all individuals for whom DUR_{June} was less than 5, while the results in the lower panel include these respondents. Only one-third (31.8 percent) of the individuals in the match sample reported spell durations which differed by 3 to 5 weeks between the two surveys. Nearly three-quarters

Table 2. Month-to-month differences in reported unemployment spell durations, May–June 1976

[In percent]

Month-to-month difference in reported spell duration	Total	Reported May duration greater than 20 weeks	Reported May duration less than 20 weeks
	Workers reporting unemployment of at least 5 weeks in June[1]		
Less than 0 weeks	14.26	25.55	7.63
0 weeks	7.41	12.34	4.52
1 to 2 weeks	9.86	7.48	11.25
3 to 5 weeks	31.78	24.67	35.96
6 to 9 weeks	15.97	11.68	18.50
10 to 15 weeks	7.74	7.71	7.76
16 to 24 weeks	4.65	3.53	5.30
25 weeks or more	8.31	7.05	9.06
	All workers unemployed in June, regardless of spell duration[2]		
Less than 0 weeks	19.62	29.29	14.60
0 weeks	9.19	11.72	7.97
1 to 2 weeks	12.09	7.11	14.60
3 to 5 weeks	27.99	23.45	30.33
6 to 9 weeks	13.55	11.09	14.80
10 to 15 weeks	6.57	7.32	6.21
16 to 24 weeks	3.94	3.76	4.24
25 weeks or more	7.00	6.28	7.25

[1]Calculations based on May 1976 CPS questionnaire participants who were classified as unemployed, who were more than 16 years of age, and who reported May unemployment durations of more than 4 weeks. The subsequent duration numbers are based on reported responses to the June 1976 survey. A total of 1,227 individuals who were recorded as unemployed in May were reinterviewed, and found to be unemployed again in June.

[2]Calculations based on May 1976 CPS questionnaire participants who were classified as unemployed, and were more than 16 years of age. A total of 1,447 such individuals were available on the May–June match.

of the respondents made inconsistent claims about their unemployment experience, and more than 20 percent reported no increase, or a decrease, in their spell durations. Thirty-seven percent of the sample reported unemployment spell durations in June which exceeded their May durations by more than 5 weeks, and many reported much longer spells; more than 10 percent of our sample reported that the length of their unemployment spells had increased by *more than 4 months*.

Workers who have experienced long spells of unemployment are particularly unreliable in reporting spell durations. We discovered this by dividing the sample into two groups. Individuals in the first group had reported being unemployed for at least 20 weeks in May, while those in the second group had been unemployed for fewer than 20 weeks. The duration-difference calculations for these subgroups are also shown in table 2. Twelve percent of the long-spell individuals reported the same duration in both months. Only 25 percent added between 3 and 5 weeks to their initial reported spell lengths, and more than one-quarter of the first respondent group claimed shorter spell durations in June than in May. These findings indicate substantial variation in the reported unemployment durations of survey participants experiencing ongoing unemployment spells.

Regression models can be used to determine those factors which are related to substantial aberrations in the reported spell durations. Table 3 reports estimates from regressions of duration differences on individuals' demographic characteristics and reasons for unemployment. Results for the model without outlier adjustment were estimated using reported duration differences as the dependent variable. Those for the model with outlier adjustments were based on data for which the outlying values of DIFF were "trimmed." Observations for which DIFF exceed 25 weeks were replaced with 25, and those for which DIFF was less than −5 were replaced with −5.

Similar results obtain for both sets of data. According to the "trimmed" regression, the average values of the duration differences (regression constant + coefficient of the independent variable) by reasons for unemployment are: job losers, 6.24 weeks; job leavers, 5.64 weeks; workers on layoff, 4.69 weeks; and reentrants and new entrants, 7.74 weeks. All of these values are larger than the 4.43 weeks which actually separated the May and June surveys. There is little evidence that demographic factors change reported

Table 3. Regression estimates of reported unemployment spell duration differences on selected demographic characteristics and reasons for unemployment, May–June 1976

[In weeks]

Independent variable[1]	Without outlier adjustment[2]	With outlier adjustment[2]
Constant	9.12 (1.66)	7.74 (.81)
Sex and age:		
Men:		
Age 16 to 19	−.18 (1.86)	−.83 (.91)
Age 20 to 24	−1.28 (1.67)	−.43 (.82)
Age 25 to 59	.68 (1.37)	.38 (.67)
Age 60 and over	1.57 (2.75)	.53 (1.34)
Women:		
Age 16 to 19	−7.29 (2.64)	−2.93 (1.28)
Age 20 to 24	−.31 (2.13)	.40 (1.04)
Age 60 and over	−2.90 (3.50)	−1.39 (1.71)
Race (nonwhite = 1)	−1.62 (1.61)	.22 (.78)
Reason for unemployment:		
Job loser	−2.18 (1.18)	−1.50 (.58)
Job leaver	−4.47 (1.62)	−2.10 (.78)
Layoff	−4.51 (1.54)	−3.05 (.75)
R^2	.022	.022
Number of observations	1,227	1,227

[1]The dependent variable in the equation was $DUR_{June} - DUR_{May}$. As indicated in text footnote 8, the specification of the equation also included control variables for the respondents' rotation group in the CPS. These variables never proved statistically significant,and are not reported here.

[2]Estimates with outlier adjustments are based on "trimmed" data; that is, observations for which reported differences exceeded 25 weeks were replaced with "25," and those for which differences were less than −5 were replaced with "−5."

NOTE: Standard error of the estimate indicated in parentheses.

duration differences, the one exception being teenage women, who appear to systematically underreport their duration increment. The reason for unemployment is a strong predictor of duration differences. Workers who were on layoff reported differences which were up to 2 weeks less than those for other unemployed individuals, while reentrants and new entrants have the greatest tendency to overstate duration differences.[8]

Beyond being interested in the average bias in reporting increments to the unemployment duration, we might be concerned about the absolute size of reporting errors. To address this issue, table 4 reports the results of four regression specifications explaining the absolute value of (DUR_{June} − (DUR_{May} + 4)). We analyze the absolute value of (DIFF-4) to prevent positive and negative errors in the duration increment from cancelling each other, as they would if we studied only the *average* duration increment.

The reported cause of unemployment affects the error in reported durations in a significant and important way. Job losers are about 2.5 weeks more accurate than the "control" group of reentrants and new entrants. Job leavers are 2 weeks more accurate than the controls, on average, and persons on layoff have still smaller response errors. For individuals on layoff, errors are on average between 3 and 6 weeks less than the control, and as many as 3 weeks less than those of either losers or leavers. The salary that the individual earned at his last job also has a statistically significant but economically small impact. A $10-per-week rise in wages reduces an individual's predicted inconsistency by about one-tenth of a week.[9]

The most important finding is that the duration of the unemployment spell affects the consistency of the individual's responses. An additional month of unemployment increases the absolute value of the difference between the reported duration difference and "truth" (4.43 weeks) by about 5 days. However, the effect of duration is more complicated than this simple model suggests. We included three linear segments in specification III to capture the possibly different duration effects of short and long spells. These linear segments are designed to allow the marginal effect of longer duration to differ as duration changes. The three variables we used, and their values for some representative initial durations, are shown below:

Variable	Value of variable if		
	$DUR_{May}=6$	$DUR_{May}=16$	$DUR_{May}=26$
DUR_{May}	6	16	26
$DUR_{May} - 12$ if $DUR_{May} > 12$	0	4	14
$DUR_{May} - 24$ if $DUR_{May} > 24$	0	0	2

To compute the effect of spell duration on absolute error, using the regression coefficients reported in column III of table 4, we evaluated each of these duration variables and multiplied them by their respective coefficients. For an individual who had been unemployed for 30 weeks in May, the calculation yields an absolute error contribution of:

$$.17(30) - .18(30-12) + .18(30-24) = 2.94 \text{ weeks}$$

This value, and the duration-related "errors" for other spell lengths, are presented below.

Table 4. Regression estimates of the magnitude of spell duration reporting error on selected characteristics, May–June 1976
[In weeks]

Independent variable	Without outlier adjustment[2]				With outlier adjustment[2]			
	Model specification				Model specification			
	I	II	III	IV	I	II	III	IV
Constant	10.84	7.81	7.77	10.61	6.80	6.17	5.10	6.42
	(1.43)	(1.47)	(1.71)	(2.03)	(.62)	(.64)	(.75)	(.91)
Race (nonwhite = 1)	2.57	2.44	2.42	2.80	.32	.29	.31	.56
	(1.38)	(1.35)	(1.35)	(1.47)	(.59)	(.59)	(.59)	(.66)
Reason for unemployment:								
Job loser	−2.63	−3.27	−3.10	−3.35	−1.52	−1.66	−1.80	−1.93
	(1.02)	(1.00)	(1.01)	(1.08)	(.44)	(.44)	(.45)	(.48)
Job leaver	−3.20	−3.19	−3.17	−3.68	−2.17	−2.17	−2.18	−2.39
	(1.38)	(1.36)	(1.39)	(1.41)	(.60)	(.60)	(.60)	(.62)
Layoff	−6.71	−6.57	−6.35	−6.72	−3.46	−3.43	−3.46	−3.46
	(1.32)	(1.30)	(1.31)	(1.41)	(.57)	(.57)	(.58)	(.62)
Spell duration reported in May (DUR_{May})	—	.13	.17	.19	—	.027	.19	.19
		(.02)	(.15)	(.15)		(.008)	(.06)	(.07)
DUR_{MAY}-12, if $DUR_{MAY} > 12$	—	—	−.18	−.22	—	—	−.21	−21
			(.25)	(.25)			(.11)	(.11)
DUR_{MAY}-24, if $DUR_{MAY} > 24$	—	—	.18	.19	—	—	.03	.14
			(.14)	(.14)			(.06)	(.06)
Hourly earnings	—	—	—	−.13	—	—	—	−.007
				(.004)				(.002)
R^2	.028	.065	.066	.071	.035	.043	.049	.060
Number of observations	1,227	1,227	1,227	1,098	1,227	1,227	1,227	1,098

[1] The dependent variable in the equations was the absolute value of ($DUR_{June} - DUR_{May} - 4$). All equations also included demographic variables and rotation group dummies, as in table 3.

[2] See footnote 2, table 3.

NOTE: Standard error of the estimate indicated in parentheses.

Duration (DUR_{May})	Contribution of DUR_{May} to $\lvert DIFF\text{-}4 \rvert$
0 weeks	0
6 weeks	1.02
12 weeks	2.04
20 weeks	1.96
30 weeks	2.94
50 weeks	6.34

Additional weeks of unemployment spell duration are particularly poorly reflected in responses of individuals who have been unemployed for very long periods. For spells which had lasted more than a year, the predicted absolute value of the response error was over 6 weeks.

Further evidence on the reported spell durations of "new entrants" to unemployment can be obtained by studying the individuals who were categorized as employed or NILF in May and who became unemployed in June. Of those experiencing employment-to-unemployment transitions, 76 percent reported June spell durations of not more than 4 weeks. About 8 percent of this newly unemployed group, however, reported durations of more than 25 weeks after not more than 4 weeks of unemployment. Findings for the NILF-to-unemployment transitors were similar. Seventy-one percent reported spells of less than 5 weeks, but 7 percent reported very long spells (more than 25 weeks). This latter category may include individuals who were misclassified as NILF in May.

Distinguishing unemployment from NILF

A third, but closely related, problem of response error concerns the reported unemployment spell durations of individuals making labor market transitions. Forty-four percent of unemployment spells end when jobseekers choose to leave the labor force.[10] However, there are frequent transitions between the states of unemployment (U) and not in the labor force. Of the individuals who were unemployed in May 1976 and for whom three consecutive CPS questionnaires were available, 3 percent were reported as NILF in June and unemployed again in July. By comparison, 21 percent of the May unemployed sample were reported as unemployed for 3 consecutive months.

An individual who leaves the labor force is technically considered to have completed his spell of unemployment. If, at some later date, he chooses to reenter the pool of the unemployed to search for work, he begins a second unemployment spell. If survey respondents adhered to this convention, individuals who were out of the labor force in June would not report July spell durations which exceeded 4 weeks. As the lower panel of table 5 demonstrates, however, only 26 percent of the U-NILF-U survey respondents considered themselves to have begun new spells. One-third of the U-NILF-U group reported lower spell durations in the second survey, but this is not appreciably different from the fraction of shorter spells discovered in the 1-month match reported in table 2. However, it would also be incorrect to characterize the data as suggesting that time out of the labor force is treated by respondents as the equivalent of time spent unemployed. Fewer than 30 percent of the group added a full 8 weeks to their reported May unemployment spell duration. And among those individuals who did not report spells of less than 5 weeks in July, the share of responses for which $DUR_{July} - DUR_{May}$ is between 7 and 9 weeks is only 12 percent.

Table 5. Unemployment spell durations reported by transitors from unemployment to not in the labor force and back to unemployment, May–July 1976

Item	All U-NILF-U transitors[1]	
	Number of respondents	Percent of total
Difference in reported spell durations, May–June		
Total	81	100.0
Less than 0 weeks	28	34.6
0 weeks	10	12.3
1 to 6 weeks	20	24.7
7 to 9 weeks	7	8.6
10 to 15 weeks	9	11.1
More than 15 weeks	7	8.6
Reported duration in July		
Total	81	100.0
1 to 4 weeks	21	25.9
5 to 12 weeks	31	38.3
13 to 24 weeks	12	14.8
25 to 48 weeks	10	12.3
49 weeks or more	7	8.6
	Transitors reporting durations of at least 5 weeks in July	
Difference in reported spell durations, May–June		
Total	60	100.0
Less than 0 weeks	15	25.0
0 weeks	6	10.0
1 to 6 weeks	16	26.7
7 to 9 weeks	7	11.7
10 to 15 weeks	9	15.0
More than 15 weeks	7	11.7

[1]All calculations based on the May–June–July 1976 CPS match file. A total of 81 individuals were classified as unemployed (U) in May, not in the labor force (NILF) in June, and were "unemployed" again on the July questionnaire. The reported statistics are based on these individuals' responses in May and July to questions about the length of their present unemployment spell.

The fact that about two-thirds of the unemployed individuals who are classified as experiencing U-NILF-U transitions appear to view themselves as in the midst of an ongoing unemployment spell implies that there is a substantial amount of "hidden unemployment" in the U.S. economy and that, for many U-NILF-U transitors, the state of "not in the labor force" is functionally equivalent to unemployment. This emphasizes the ambiguity of current measures of labor market status, and helps to explain the strongly procyclical behavior of labor force participation.

Reasons for unemployment

The CPS match files also afford an opportunity to make intermonth comparisons of respondents' stated reasons for entering unemployment. Using the May-June 1976 match

Table 6. Reason for unemployment reported in June by reason reported in May, 1976

Reason reported in May	Reason reported in June[1] (percent of May respondents)[2]				
	Job loser	Job leaver	Layoff	New entrant	Reentrant
Job loser	82.1	5.2	6.0	0.7	6.3
Job leaver	25.1	56.6	1.7	1.7	14.8
Layoff	30.6	1.9	63.6	0.0	4.3
New entrant	.6	1.8	.6	79.9	17.2
Reentrant	17.5	9.5	.9	6.2	66.0

[1]Reported unemployment in June by reason was: job loser, 44.5 percent; job leaver, 11.3 percent; layoff, 11.8 percent; new entrant, 10.9 percent; and reentrant, 21.5 percent.

[2]Calculations were performed using the 1,497 records on the May–June 1976 CPS match tape for which the respondent was unemployed in both May and June. The calculations show, for example, that the percentage of May job losers who also reported themselves as job losers in June was 82.1 percent.

file, we cross-tabulated respondents' May "reasons" with their June "reasons."[11] Table 6 shows that only about 70 percent of the respondents cited the same reason for unemployment in both May and June. The correlation between the two responses is lowest for those originally reported as job leavers; only 56 percent of the May job leavers reported themselves as leavers again in June. Of those who changed classification, 58 percent moved to the category of job loser and 34 percent became reentrants. The groups with the highest intermonth correlations were job losers and new entrants; roughly 80 percent of the May respondents in these groups provided similar responses in the June survey. The largest intercategory movement was from layoff to job loser: Thirty percent of those reported to be on temporary or permanent layoff in May reported themselves as job losers in June. There also appears to be a surprisingly large amount of movement between the categories of reentrant and job loser.

The large incidence of reported changes from the layoff to the job loser category is of particular significance. Although the economic importance of temporary layoff unemployment has been proclaimed by several analysts, the evidence here suggests that its significance may well have been overstated. A natural interpretation of the frequent changes in the responses of persons initially on layoff is that, at some point, these individuals realize that they cannot return to their original employers. If this interpretation is correct, it implies that the reported amount of unemployment attributable to layoffs in May substantially overstated the proportion of the unemployed who would ultimately be able to return to their original employers.

Conclusions

Our findings call into question some of the individual responses to fundamental parts of the monthly CPS questionnaire. They buttress the evidence from Reinterview Surveys which suggests that misreporting or misrecording takes place. While information of the type presented here cannot be used to evaluate the bias in CPS responses, it does imply that measures of behavioral change may be overstated because of response error.

Our analysis also sheds light more generally on the problem of response error in survey research. For a number of reasons, the CPS is likely to generate more accurate and consistent responses than other sample surveys. For example, the CPS questions ask only about recent behavior, rather than behavior over the course of a year or a longer interval. More safeguards are used to ensure reliability than in most other studies of labor market behavior. And, to a greater extent, CPS questions probe objective behavior rather than subjective intent. Our focus on the CPS was motivated solely by its widespread use by researchers and policymakers, and by the availability of data necessary for consistency checks.

We believe that our findings suggesting the need for caution in performing statistical analysis of these data are applicable to other surveys of labor market behavior, although more research on this question would be valuable. Especially when investigations focus on period-to-period changes, errors in variables problems are likely to be serious. Unfortunately, most of the methods currently used to examine aspects of dynamic labor supply behavior are not at all robust with respect to errors in variables. Future research should examine more thoroughly the causes of misreporting and alternative techniques for developing consistent data. In the meantime, statistical techniques for adjusting data, and for constructing estimates in the presence of errors in variables, should be improved. □

——*FOOTNOTES*——

ACKNOWLEDGMENT: The authors wish to thank Francis Horvath of the Bureau of Labor Statistics for comments on an earlier draft of this article.

[1] See Dorcas W. Graham, "Estimation, Interpretation, and Use of Response Error Measurements" (Washington, U.S. Department of Commerce, 1974); Henry Woltman and Irv Schreiner, "Possible Effects of Response Variance on the Gross Changes Data," Memo, Bureau of the Census, May 11, 1979; and *The Current Population Survey Reinterview Program: January 1961 through December 1966*, Technical Paper 19 (Washington, U.S. Bureau of the Census, 1968).

[2] This procedure fails to detect those individuals who report consistent, but incorrect, responses in both months.

[3] See J. M. Poterba and L. H. Summers, "Spurious Transitions and the Gross Flows Data," mimeo, 1983, for a discussion of methods for adjusting BLS gross flows data based on estimated response error probabilities.

[4] See J. M. Poterba and L. H. Summers, "A Multinomial Logit Model with Errors in Classification," mimeo, 1983, for a description of analytical procedures for studying labor market transitions when some responses are measured with error.

[5] Christopher J. Flinn and James J. Heckman, *Are Unemployment and Out-of-the-Labor Force Behaviorally Distinct States?* Working Paper 979 (Cambridge, Mass., National Bureau of Economic Research, 1982).

[6] After completing this paper, we became aware of closely related research by Norman Bowers and Francis Horvath. See "Keeping Time: An Analysis of Errors in the Measurement of Unemployment Duration," unpublished.

[7] Between the May and June Surveys which are the focus of our work, 4.43 weeks elapsed.

[8] We also experimented by adding the individuals' reported May du-

ration to the regression models. This had a substantial negative effect on the reported duration difference. However, it is difficult to determine whether this is genuinely the result of the longer-duration unemployed responding with smaller differences. An alternative explanation is that the finding is purely a statistical artifact. Conditional on a high reported May duration, the difference between the June and May durations is likely to be less than if the value of DUR_{May} is low. This means that in a regression model for $DIFF$, DUR_{May} will have a negative coefficient. This hypothesis also predicts that, by similar reasoning, DUR_{June} should have a *positive* coefficient. Some support for this view was provided when we substituted DUR_{June} for DUR_{May} and observed a significant positive coefficient. Therefore, because the results appear spurious, we have not reported equations which include duration variables.

[9]Our equations also include control variables for the respondents' rotation groups in the CPS. Rotation Group I indicates individuals who participated in the survey in May, June, July, and August; Rotation Group II denotes those who participated only in May, June, and July. The omitted dummy variable is for those who participated only in the May and June surveys. These variables, not reported in the tables, never proved statistically significant.

[10]This was calculated as:

$$\frac{\textit{Prob(transition from unemployment to NILF)}}{\textit{Prob(transition from unemployment to employment or NILF)}}$$

For further discussion of labor market dynamics in this framework, see Kim B. Clark and Lawrence H. Summers, "Labor Market Dynamics and Unemployment: A Reconsideration," *Brookings Papers on Economic Activity*, Vol. I, 1979, pp. 13–60.

[11]Job losers and leavers were categorized on the basis of the "why did . . . start looking for work?" question. Workers who explained that they were on permanent or temporary layoff in response to the question "why was . . . absent from work last week?" were classified as on layoff. New entrants were those nonleavers and nonlosers who claimed either that (i) they had never worked at all, or (ii) they had never worked full time for more than 2 consecutive weeks. Any workers who did not fall into any of these four categories were classified as reentrants.

BLS' 1982 survey of work-related deaths

JANET MACON

The number of work-related deaths in private sector establishments with 11 employees or more was 4,090 in 1982, compared with 4,370 in 1981.[1] The corresponding fatality rate was 7.4 deaths per 100,000 full-time workers in 1982, and 7.6 in 1981. (See table 1.)

Employers participating in the Bureau of Labor Statistics' Annual Survey of Occupational Injuries and Illnesses were asked to supply specific information about deaths caused by hazards in the work environment, that is, the object or event most closely associated with the circumstances of the fatality. Estimates of the percentage of fatalities by cause represent the average for the 1981 and 1982 surveys. Percentages were calculated for the 2 years combined because large sampling errors at the industry division level preclude precise comparisons based on year-to-year changes.

The 4,090 fatalities in 1982 represent all reported deaths resulting from a job-related injury or illness in 1982, regardless of the time between the injury or onset of illness and death. About 340 of these fatalities were related to illness.

Among industry divisions, fatality rates ranged from 44.3 per 100,000 full-time workers in mining industries to 2.5 in finance, insurance, and real estate industries. Between 1981 and 1982, rates decreased in 5 of the 8 industry divisions, and increased by more than 15 percent in agriculture, forestry, and fishing; transportation and public utilities; and services.

Transportation and public utilities industries reported the largest number of fatalities. The percentage of total fatalities increased in three of the industry divisions, decreased in three, and remained unchanged in two. Although the number of fatalities decreased in construction and mining, the percentage of the total remained unchanged.

Analysis by cause

More than half of all fatalities were caused by over-the-road motor vehicles, falls, heart attacks, or industrial vehicles or equipment. (See table 2.) About 1 of every 4 fatalities involved over-the-road motor vehicles. Falls, heart attacks, and industrial vehicles combined contributed 32 percent of total fatalities; falls, 12 percent; heart attacks, 10 percent; and industrial vehicles or equipment, 10 percent.

Over-the-road motor vehicles were the major cause of death in 5 of the 8 industry divisions. About 1 of every 3 of these fatalities occurred in transportation and public utilities industries, which had only 7 percent of total employment. (See table 3.)

Twelve percent of all fatalities involved falls. The construction and manufacturing industries together accounted for about 2 of every 3 falls.

About 10 percent of all fatalities were due to heart attacks. Heart attacks occurred at a slightly higher frequency in construction and transportation and public utilities, based on employment percentages.

Industrial vehicles or equipment were involved in 10 percent of all fatalities. More than half of these cases occurred in construction and manufacturing industries. Another 14 percent occurred in oil and gas extraction, which accounts for only 1 percent of total employment.

The "all other" category accounted for 3 percent of total fatalities. This category includes, for example, contact with radiation or toxic substances, drowning, train accidents, and death from various occupational illnesses.

Analysis by industry

Agriculture, forestry, and fishing. Industrial vehicles or equipment were involved in 27 percent of the fatalities, while over-the-road motor vehicles contributed 18 percent of the cases. Electrocution accounted for 16 percent and falls, 12 percent.

Janet Macon is a statistician in the Office of Occupational Safety and Health Statistics, Bureau of Labor Statistics.