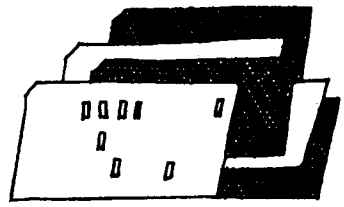
Research Summaries



Forgotten unemployment: recall bias in retrospective data

FRANCIS W. HORVATH

It is a well documented fact that the ordering and specific wording of a survey questionnaire can produce different "readings" of the same underlying event. Differences can also arise when the same general question is asked of individuals at different times.

One area in which such discrepancies have been found is in comparisons of unemployment data collected on a monthly basis with those obtained from a once-yearly survey. The source of the monthly unemployment figures is the Current Population Survey (CPS), conducted by the Census Bureau for the Bureau of Labor Statistics. Each month, representatives of approximately 60,000 households are asked questions about their labor force activity and that of other household members during a given reference week. Annual unemployment estimates are constructed by averaging monthly data for the year.

Annual unemployment data are also obtained, on a retrospective basis, from the Work Experience Supplement to the Current Population Survey in March of each year. Persons responding to the supplement questions are asked to recall events which occurred from January through December of the previous year. These respondents are asked to aggregate the year's labor force activities into summary figures. That is, they are asked questions such as: "How many weeks was (household member's name) looking for work?" and "How many weeks was (household member's name) working?" However, no attempt is made to ascertain the months in which these events took place.

The supplement data reflect the number of persons with unemployment at any time during the previous year; a person experiencing two or more spells of unemployment is counted only once. Recently, this estimate has been about 18 to 21 million persons a year. In contrast, the annual average of the unemployment data collected monthly—simply the sum of the 12 monthly estimates (unadjusted for seasonality), divided by 12—represents the mean number of persons unemployed during a "typical" week of the year. This figure was about 5 to 7 million during 1975–79. On the surface, these two figures appear unconnected. However, because the Work Experience Supplement also includes questions on the duration of unemployment in the previous year, it is possible to directly relate the retrospective supplement information to the annual averages of the monthly data.

The basic method, developed by Daniel Suits and Richard Morgenstern,[1] derives the total number of weeks of unemployment occurring in the labor force during the full year from the Work Experience Supplement. Specifically, persons are grouped according to the supplement information into "duration of unemployment" intervals. Multiplying the number of persons in each group by the midpoint (in weeks) of the duration interval yields an estimate of the total weeks of unemployment occurring to persons within that particular cell. The sum over all cells provides an estimate of the total weeks of unemployment occurring in the labor force over the year. Dividing this total by 52 gives the adjusted supplement estimate of a typical week's unemployment, roughly comparable to the results from the monthly studies.

Exhibit 1 shows this basic relationship, using hypothetical data. Over a given year, there are 155 individuals reporting some unemployment. Suppose that most of these are unemployed in 1 month only, and during all other months they are either employed or not in the labor force. However, assume that 25 of the 155 report unemployment in 2 consecutive months. In the monthly survey, then, the total reported instances of unemployment would be 180; annual average unemployment would be 180 divided by 12, or 15 individuals. With an accurate annual retrospective survey, the number of persons reporting some unemployment during the entire year would be 155—180 less the 25 who were unemployed over a 2-month period. Dividing reported weeks of total unemployment by 52 yields the "adjusted" retrospective estimate, which is exactly equal to the annual

Francis W. Horvath is an economist in the Division of Labor Force Studies, Bureau of Labor Statistics.

Exhibit 1. Derivation of annual average unemployment from the monthly CPS and from the annual Work Experience Survey

Based on monthly measurements:

- Annual average unemployment $= \frac{180 \text{ reports}}{12 \text{ months}} = 15$

Based on retrospective annual survey:

- Of 155 persons with any unemployment, 130 were unemployed for 1 month (4.33 weeks) and 25 were unemployed for 2 months (8.67 weeks).
- Annual average unemployment =

$$\frac{(130 \text{ persons} \times 4.33 \text{ weeks}) + (25 \text{ persons} \times 8.67 \text{ weeks})}{52 \text{ weeks per year}} = 15$$

average from the monthly survey.

It has previously been noted that these adjusted Work Experience Supplement estimates are usually less than the comparable annual average figures, particularly among women and youth.[2] However, the reasons offered for the overall understatement have been largely conjectural.

One plausible reason for the difference between the two sets of estimates is that respondents to the supplement forget events which took place further in the past. Unless very recent or relatively long, a spell of unemployment might be forgotten by the person who encountered it, and it is even more likely to be overlooked by another household member who may be answering the work experience questions for the entire household. This forgetting should be greater for unemployment which took place long ago than for that which occurred in the recent past.

Two studies on underreporting of past unemployment form the basis for this hypothesis. During 1969–71, the Census Bureau conducted a special reinterview program in which the same respondents were asked labor force questions in 2 consecutive months.[3] The second month's survey included a retrospective interview on events occurring in the first month. Analysis of the results revealed that people did not consistently place themselves in the same labor force categories they had earlier. In particular, *unemployment in the first month was understated significantly in the second interview.* Analysts noted that: "biases in labor force status due to recall are quite high and at an unacceptable level of quality." A follow-up study found the same pattern among both men and women. Apparently, recall errors are a serious problem even over as brief an interval as 1 month.

The question of "recall bias" has also been the subject of an inquiry by Statistics Canada in relation to its retrospective Annual Work Patterns Survey.[4] During January of each year since 1978, the Canadian statistical agency has conducted this supplemental survey, in which respondents are asked labor force questions pertaining to the previous calendar year. The Canadian survey differs from the U. S. Work Experience Supplement in that persons are asked to describe their labor force activities in each of the 12 preceding months. For example, the Canadian respondent actually identifies "July" as a month in which he or she worked or looked for work. When Canadian analysts compared the answers to these retrospective questions to their ongoing monthly labor force surveys, they found a "substantial, systematic recall bias." *In the early months of the year,*

Table 1. Comparison of adjusted Work Experience Survey unemployment with annual average of monthly unemployment from the CPS, 1967–79

[Numbers in thousands]

Year	Annual unemployment data from Work Experience Survey			Annual average of CPS unemployment for a typical week (4)	Percent understatement[2] (5)
	Total with unemployment during the year (1)	Average duration of unemployment (in weeks) (2)	Adjusted estimate of weekly unemployment[1] (3)		
1967	11,561	10.05	2,234	2,976	24.9
1968	11,372	9.55	2,088	2,817	25.9
1969	11,741	9.73	2,197	2,832	22.4
1970	14,565	12.33	3,453	4,088	15.5
1971	15,852	14.06	4,286	4,993	14.2
1972	15,280	13.41	3,941	4,840	18.6
1973	14,499	11.99	3,343	4,304	22.3
1974	18,535	13.04	4,647	5,076	8.5
1975	21,101	16.27	6,601	7,830	15.7
1976	20,447	15.82	6,221	7,288	14.6
1977	19,512	14.64	5,495	6,855	19.8
1978	17,758	13.73	4,682	6,047	22.6
1979	17,972	13.27	4,585	5,963	23.1
Average	16,167	12.91	4,136	5,069	19.1

[1] (Column 1 × Column 2)/52.

[2] (Column 3 − Column 4)/Column 4.

Table 2. Half-year averages of monthly unemployment from the CPS, 1967–79

[Numbers in thousands]

Year	Total		Men		Women	
	January through June	July through December	January through June	July through December	January through June	July through December
1967	3,008	2,944	1,602	1,413	1,406	1,531
1968	2,950	2,683	1,555	1,284	1,395	1,400
1969	2,798	2,865	1,419	1,387	1,379	1,478
1970	3,756	4,421	2,092	2,379	1,664	2,042
1971	5,101	4,885	2,927	2,625	2,174	2,260
1972	5,090	4,590	2,916	2,355	2,175	2,236
1973	4,476	4,134	2,438	2,043	2,037	2,091
1974	4,788	5,364	2,600	2,735	2,188	2,628
1975	8,143	7,516	4,676	4,093	3,467	3,423
1976	7,430	7,146	4,195	3,742	3,235	3,404
1977	7,281	6,430	3,993	3,183	3,288	3,247
1978	6,264	5,830	3,302	2,800	2,962	3,030
1979	6,022	5,905	3,129	2,907	2,893	2,998
Average	5,162	4,978	2,834	2,534	2,328	2,447

unemployment is seriously understated, with the bias declining and then reversing towards the end of the year. That is, some of the unemployment in January of the previous year is "forgotten," while recent unemployment in December is overstated.

Because the U. S. Work Experience Supplement relies on recall of events which took place as much as 15 months earlier, it seems logical to expect the forgetting of past unemployment to affect this survey as well. The following discussion illustrates one method with which this intuitive expectation can be tested.

A new look at retrospective bias

Table 1 lists annual averages from the monthly surveys and the adjusted supplement estimates for the years 1967–79. As previously indicated, the retrospective unemployment figures from the supplement are consistently less than the annual averages from monthly data. For the labor force as a whole, the degree of understatement ranged from about 9 to 25 percent and averaged 19.1 percent over the period.[5] The understatement appeared to be smaller during periods of increasing unemployment, such as 1974–75.

The annual averages in table 1 treat each month equally, in the sense that each month constitutes one-twelfth of the average. If the Work Experience survey provided completely parallel estimates of both the incidence and total duration of unemployment, figures from the supplement and from the monthly surveys should be exactly equal. Because the annual averages weight each month equally, it is perfectly acceptable to combine the monthly data into two sets—January through June, and July through December—and treat them each as contributing one-half the annual average. While this treatment is nothing more than a mathematical identity, it provides a convenient tool which may be used to examine the retrospective bias of the March supplement. That is, each half year average of monthly information can be examined for correspondence with the Work Experience data.

Table 2 shows the 6-month averages of monthly unemployment figures for the total, and by sex, for the 1967–79 period. In general, total unemployment has been greater during the 6 six months of the year. Note, however, that this does not hold for women, who report higher levels of unemployment during the last 6 months.

Because the corresponding adjusted supplement figures (not shown) all fall below these half-year averages, a simple comparison of the data would make it appear that the 6-month period with lower observed unemployment is "closer" to the Work Experience estimates. For men, the second half of the year is "more similar" to the supplement figures, while the first-half averages are closer for women.

Obviously, each of the 6-month averages contributes one-half to the annual average. If the adjusted supplement count is exactly analogous to the annual average of monthly data, then it too should be a composite of the two periods. That is, given:

(1) Annual average unemployment $\equiv$ (0.5) January-June 6-month average + (0.5) July-December 6-month average

and

(2) Annual average unemployment = Adjusted Work Experience unemployment

then

(3) Adjusted Work Experience unemployment = (0.5) January-June 6-month average + (0.5) July-December 6-month average

A simple test is proposed to examine this relationship. The adjusted Work Experience figure, W_e, is re-

gressed on the two 6-month averages as:

$$W_e = a + b_1 (JJAV) + b_2 (JDAV) + u$$

where a is a constant term, JJAV is the 6-month average of the unemployment counts for January through June, JDAV is the average for July through December, and u is the disturbance term, assumed to be independently distributed, with a zero mean and constant variance.

Because both halves of the year should figure equally into the cumulative total, we should expect $b_1 = b_2$. More specifically, if the adjusted supplement unemployment were an exact measure of the incidence and the length of unemployment, we should find $a = 0$, and $b_1 = b_2 = 0.5$. That is, there would be no understatement, and each half of the year would count equally towards the cumulative total. This should hold regardless of any size differences in unemployment between the two halves of the year.

Table 3 shows the coefficients obtained when the relationship was tested on total, male, and female unemployment. The differences between expectations and the results of this simple test are stunning.

In each case, the second half of the year appears to have much more weight in the adjusted supplement estimates than the first half. The coefficients for the second half of the year are all greater than 0.5, and are significantly related to the degree of unemployment reported in the Work Experience Supplement. On the other hand, the coefficients for the first half of the year indicate little relationship at all; no estimate of b_1 was significantly different from zero at a 90-percent confidence level. All of the second-half coefficients are at least four times as great as those of the first half.

These results seem to support the contention that unemployment in the first half of the year is "forgotten" in the Work Experience Supplement relative to the more recent second half, which is closer to the week in March when supplement questions are asked. The Work Experience data still unexplainably understate the unemployment of certain demographic groups by differential amounts. However, the foregoing analysis indicates that recall bias does play an important part in the general shortfall, which appears strongly in evidence even when the data are disaggregated by sex. □

Table 3. Regression of adjusted Work Experience unemployment on half-year averages of monthly unemployment from the CPS, 1967–79

[t-statistics in parentheses]

Adjusted Work Experience unemployment	Estimates of—			R^2
	a	January–June coefficient b_1	July–December coefficient b_3	
Total	−416.30 (−2.20)	0.07 (0.53)	0.84 (5.34)	0.98 —
Men	−107.28 (−1.17)	.18 (1.64)	.81 (6.26)	.98 —
Women	−222.93 (−1.78)	.06 (0.25)	.73 (2.95)	.96 —

FOOTNOTES

[1] Daniel B. Suits and Richard D. Morgenstern, "Duration as a Dimension of Unemployment," Paper presented at the Econometric Society Meetings, Washington, December 1967.

[2] Richard D. Morgenstern and Nancy S. Barrett, "The Retrospective Bias in Unemployment Reporting by Sex, Race and Age," *Journal of the American Statistical Association*, June 1974, pp. 355–57; and Wayne Vroman, "Measuring Annual Unemployment," Unpublished paper (Washington, The Urban Institute, 1979).

[3] See Louis E. Williams, "Methods Tests Phase III, First Report on the Accuracy of Retrospective Interviewing and Effects of Nonself Response on Labor Force Status," Memo, Bureau of the Census, June 24, 1969; Charles Jones and Robert Aquilino, "Methods Tests Phase III, Second Report . . .," Memo, Bureau of the Census, Jan. 29, 1970; and Robert Aquilino, "Methods Tests Phase III, Third Report . . .," Memo, Bureau of the Census, Apr. 2 1971. These special studies should not be confused with the regular CPS reinterview program, in which a percentage of households are reinterviewed as a quality check on the data. The special program allowed separate study of proxy response versus self-response, as well as "recall" bias.

[4] For a description of the Canadian Annual Work Patterns Surveys and selected data from them, see *Patterns of Full- and Part-Year Employment and Unemployment: Results From the Annual Work Patterns Surveys for 1977 to 1980* (Statistics Canada, Catalog No. 71–531, forthcoming).

[5] These figures are slightly higher than estimates provided by Barrett and Morgenstern, and Vroman. One reason for the difference is that this study used unpublished Work Experience data which shows finer detail for duration-of-employment intervals. The midpoints of the most detailed interval size available were used for all years. For the open-ended interval, "40 weeks or more" of unemployment, 44 weeks was used as the midpoint, rather than 46. The procedure indicated was used by Morgenstern and Barrett in the original article on unemployment underreporting. Unpublished information on persons unemployed 27–39 weeks was also used.

Easing the hardship of plant shutdowns

The Bureau of Labor Statistics' 1980–81 study of major collective bargaining agreements found that companies are providing more protection for covered employees affected by plant shutdowns and movements, greater opportunities for interplant transfers, and more frequent relocation allowances. The study reviewed 1,593 agreements, each covering 1,000 workers or more in private industry, excluding railroads and airlines.

Among a sample of 522 of the contracts, 36 percent placed some restriction on management's right to close or relocate plants—up from 22 percent in a 1966–67 study. Although the majority of the sample agreements were in the nonmanufacturing sector, more than two-thirds of the contracts imposing restrictions were in manufacturing.

Interplant transfer provisions were included in 552, or 35 percent, of the 1,593 contracts surveyed. Such clauses appeared in 37 percent of the manufacturing agreements, and 33 percent of those in nonmanufacturing. When a similar study was conducted 13 years earlier, 32 percent of all agreements dealt with interplant transfers.

In 1980–81, provisions that required a company to pay all or part of an employee's relocation expenses were found in 41 percent of the agreements having interplant transfer clauses, compared with 34 percent in 1966–67. In manufacturing, 36 percent of the contracts specified these payments, compared with 46 percent in nonmanufacturing.

The detailed report on the study, *Major collective bargaining agreements: plant movement, interplant transfer, and relocation allowances,* BLS Bulletin 1425–20, is available from the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. Price $4.75. □

Bilateral protection

The mutual accommodations and adjustments to the hard issues of collective bargaining that the parties have displayed in regard to wages, employee benefits, and institutional issues is no less in evidence when one inspects the current status of the administrative issues in our labor-relations system. Management has increasingly recognized the job-protection and working-condition problems of the industrial employee and has made important concessions in these areas. At the same time, however, there has been reciprocal recognition on the part of unions that the protection of the employee cannot be at the expense of the destruction of the business firm. The axiom that employees cannot receive any protection from a business that has ceased to exist appears to have been fully appreciated by all but the extreme recalcitrants of the labor movement, and workable compromises have been possible with respect to the areas of seniority, discipline, and the various other dimensions discussed no less than in the case of previous topics.

—ARTHUR A. SLOANE AND FRED WITNEY
Labor Relations, 4th ed.
(Englewood Cliffs, N.J., Prentice-Hall, Inc., 1981), p. 442.