

Stuart Scott, Bureau of Labor Statistics

## 1. Introduction.

The Bureau of Labor Statistics is now completing six quarters' worth of data collection for a pilot survey of job openings in establishments. In each of four states, Florida, Massachusetts, Texas, and Utah, 1200 firms were asked to record all vacant positions in their organization which they were actively seeking to fill on the reference date (cf. Appendix for survey variables and definitions). During the first quarter, a reinterview of a one in six subsample was conducted to address issues of quality. In part, the survey data were collected again to permit measurement error estimates along the lines of the error models described by Hansen, Hurwitz, and Pritzker (1964) and others at the Bureau of the Census. The main focus of this paper is our methodology on the reinterview survey.

Government officials from Arthur Burns to William Proxmire have called for collection of job openings statistics. Unemployment statistics are regularly used to measure labor supply in terms of job seekers at a point in time. Job openings statistics are intended to provide information on labor demand in terms of jobs open to the unemployed at a point in time. At the level of economic policy, Congress has shown interest in assessing how close to full employment the country is by examining unemployment and job openings statistics together. Another potential use is examining occupational breakdowns of job openings to help determine state job training programs. This is appealing because of the large sums of money put into these programs and the frequency with which trainees are unable to land jobs they are trained for.

Intrinsically, job openings are hard to estimate since they are rare relative to total employment in a firm. Furthermore, the number of job openings in a firm is not as solid a piece of data as an hourly wage rate or number of employees. The task is to come up with a definition which is collectible, and then to check whether this definition encompasses enough job openings for intended uses of the data. Collection issues have been addressed in earlier surveys in the U.S. (Konstant and Wingard, 1968) and Canada (Ostry and Sunter, 1970).

Our reinterview survey in March, 1979, addressed collectibility in two ways:

- i) a Quality Measurement (QM) schedule to measure response error through re-collection of the data, and
- ii) a Unit Profile schedule to ascertain recruiting/hiring processes and recordkeeping practices.

The QM portion was intended to provide a quantitative measure of accuracy, given the survey definitions. The Unit Profile, in a qualitative way, explored how all-encompassing this definition is in practice. Put another way, the Unit Profile gave indications of how thoroughly a personnel officer or other respondent covered job openings in the firm. Overall, our findings with respect to reinterview methodology included the following:

- i) For all four states, respondents to the main survey also tended to respond to the reinterview survey.

- ii) The QM survey was difficult to do, but in two of the four states reinterview techniques were effective enough that measurement error statistics could be computed.

## 2. Response rates.

The initial reference date for the Job Openings Survey was March 31, 1979. The total sample size per state was 1200. Phone prenotification explained the purpose of the survey, described the forms, and elicited a contact to send the schedules to. (In one state, some firms were solicited by personal visit and by mail.) Collection was by mail, with mailings timed for arrival on or just before the reference day. Phone reminders were scheduled to begin one week later. After four weeks, each state was to conduct a field follow-up of a systematic sample of 175 units classified as nonrespondents out of the total sample of 1200. The field follow-up was a collection by personal visit, or for smaller firms, in part by phone. Eight weeks were permitted for data collection.

The reinterview survey was scheduled to begin the third week. Recognizing the perishability in quality of data, states were instructed to complete as many interviews as possible during the first four weeks. To facilitate this, priority was to be given to the reinterview units in phone reminders and in scheduling personal visits.

Table 1 shows overall response rates for the reinterview survey. The low rates for the QM, 35% for Florida and 39% for Utah, severely curtail analytic use of the QM results, but do not represent an accurate measure of cooperation. Taking into account nonresponse and "late" response on the main survey, response to both the QM and Unit Profile was positive. What is required to get a QM response? First, you must have a response on the original survey. Second, the original response must arrive in time to schedule a reinterview. Many firms were contacted by personal visit to obtain at least an original response and a Unit Profile. No firm was to receive two personal visits, so cooperation for the QM could be requested only for firms with mail responses. Thus, we adopted as a fairer measure of cooperation a conditional response rate, reinterview respondents divided by mail respondents on the original. The second half of Table 1 shows these rates. (Figures for Florida are not yet available). In Utah, 71% of eligible units responded to the QM. Of mail respondents, better than 90% responded to the Unit Profile in the three states with available data. While, on average, the QM collection took less time than the Unit Profile, respondents showed more resistance to the QM. Thus, reasonable response rates for a QM are achievable, and even higher rates for a more qualitative follow-up survey like the Unit Profile.

Calls for appointments were made for the reinterview in most cases. The requests for a follow-up or validation of the original survey were made as brief and as simple as possible. In some states, especially for smaller firms, drop-ins were conducted.

In soliciting cooperation during the reinterview, interviewers explained the purpose, outlined the interview, and, when applicable, thanked the respondent for cooperation on the main survey. Emphasis was placed on testing for data quality through examining our

Table 1

Overall Response Rates

<u>QM</u>	In scope	Respondents	Response rate
Florida	196	68	35%
Utah	191	75	39
<u>Unit Profile</u>			
Florida	196	86	44%
Massachusetts	188	157	84
Texas	195	99	51
Utah	191	123	64

Conditional Response Rates

<u>QM</u>	Mail response on original	Respondents	Response rate
Utah	105	75	71%
<u>Unit Profile</u>			
Massachusetts	141	134	95%
Texas	65	60	92
Utah	105	96	91

concepts, procedures, and forms, rather than looking for their mistakes.

Respondents seemed to accept the aims of testing our methods, testing for quality. Cooperation for the Unit Profile was easily gained, since it offered the respondent an opportunity to talk about the firm and his role in it. More determination was required to recollect the survey data for the QM. Some of the most diligent respondents were defensive about our questioning their numbers. Probably the work of examining records for specific information and trying to recall in detail the situation some weeks back were the main obstacles for respondents. Although the Unit Profile was time-consuming, it probably strengthened the overall survey. While the respondent might feel uncomfortable about discrepancies uncovered by the reinterview, and the interviewer might be disappointed when no vacancies occurred or QM collection was weak, both usually found the Unit Profile not unpleasant.

States pledged not to use these data for job placement purposes, since testing in the 1960's viewed this as an obstacle to response. While a few employers complained about previous experiences with the state employment service, they did not raise the issue of referrals or express this as a reason for nonresponse.

3. Reinterview techniques.

Let's go back to an equation for total survey error, following Hansen, Hurwitz, and Pritzker (1964). In estimating a total,

$$\begin{aligned} \text{MSE}(\text{sample estimate}) &= \text{sampling variance} \\ &+ \text{response variance} \\ &+ 2(\text{covariance between a} \\ &\quad \text{sample deviation and a} \\ &\quad \text{response deviation}) \\ &+ \text{bias}^2. \end{aligned}$$

Two different aims from a reinterview survey are estimation of response variance and estimation of bias. For Job Openings, our questions ran like this:

- Can we collect quality data by mail?
- Are definitions understood and followed?
- Will respondents make enough effort?

This translates into going after the bias. Therefore, the reinterview was planned to provide improved values, rather than a repeat of the main survey methodology. The bias is

$$B = X_G - U$$

where

- U = true total in population
- $X_G$  = expectation of sample estimate over repeated sampling under survey conditions G.

The natural estimate is the net difference

$$ND = X_O - X_R,$$

where (both computed for reinterview sample only),

- $X_O$  = estimate from the original survey data,
- $X_R$  = best estimate from the reinterview data.

To facilitate comparisons among cells, we also compute the relative net difference,

$$\text{RND} = 100\% \times (X_O - X_R) / X_R$$

Improvement on the reinterview is in the sense of smaller response error. The primary means for improvement were

1. contact by personal visit rather than mail,
2. more intensive collection instrument, and
3. independent collection, then reconciliation.

Part of the QM instrument appears below. Starting from scratch, the interviewer goes through the openings identified by the respondent, position by position. He/she checks the criteria in the definition, enters the numbers in the "QM" columns for the four basic variables (see Appendix for definition and variables), and records whether each piece of information is based on records or memory. After the QM collection is complete, an envelope with a copy of the original schedule is accessed. Data for the original is transcribed into the "O" columns. When there are no differences the interview is closed out. Otherwise, the interviewer goes over each discrepancy with the respondent to determine the best possible value, enters it under the "R" columns, and explains the reconciliation under "Comments".

In summary, the questionnaire design attempted to produce improved values through recording of the three separate values, line-by-line comparison with the original, built-in checks on the definition, and built-in explanation of discrepancies.

Were we able to collect improved values? Yes and no. We eliminated two states' data from the analysis, because they did not carry out an independent collection. The QM portion was conducted more like a confirmation. When the original was shown to the respondent, few changes were made on the QM. A necessary ingredient to improved values seems to be starting from scratch. Reasons given for not achieving the independent collection were respondent burden affecting survey response, respondent resistance, and difficulty of reconstructing accurately, given the time lag. There was some hesitancy and uncertainty about the technique. For instance, for large firms, interviewers and survey managers sometimes felt reluctant to question responses, since these firms cooperate in other BLS surveys. Prior interviewer experience was not a determining factor. Of the two states which accomplished the independent collection, one used experienced interviewers primarily, and the other had about half inexperienced interviewers. Of the other two states, again one used experienced interviewers, the other a mixture.

The QM process had as advantages the presence of an interviewer knowledgeable in the survey concepts, the expression of interest associated with a personal visit, and some prodding to use the best sources of information available. Useful types of records included logs of job openings, job requisitions systems, recruiting materials, and job descriptions. The Unit Profile itself contained some useful preparation for the QM. It included sketching an organization chart,

going over the recruiting/hiring procedures, and discussing types of records available. In the interviewers' minds, all these advantages were, on balance, less than the disadvantages of time lag and dependence on memory. QM collection mostly came 4-8 weeks after the reference date. Reconstruction was especially difficult for the substantial number of medium-sized firms without pertinent records. Here are a couple of the difficulties mentioned:

- i) Once an offer to hire is made, there is no opening. In checking whether a particular opening existed on the reference date, often the records contained the starting date, but not the date of the offer.
- ii) Firms often kept track of openings by means of a file of position descriptions; once the job was filled, the position description went into the personnel file of the new employee, and was lost for any statistical purposes.

The reconciliation process worked. Sometimes, it followed QM collection directly; sometimes, the respondent was offered a break, while the interviewer prepared for reconciliation. Interviewers reported some awkwardness in the procedure and some guesswork when evidence was weak, but they were able to match entries from the two schedules and to obtain a reconciled value from the respondent. Even in the small fraction of cases with many vacancies, this was accomplished. When there were discrepancies, the reconciled value could be the QM value, the original value, or a third value, and all three cases occurred. Thus, in the states with independent collection, we feel the reconciled values did represent improved values.

Even if the respondent is asked to start from scratch, the QM results are likely to be affected when the interviewer knows the original results. This issue was addressed in a recent experiment in the CPS reinterview program (Schreiner, 1980). There is a natural inclination for the interviewer to look at the original schedule beforehand to avoid "going in cold." Our instructions, however, called for waiting until QM collection was complete, and, after some early resistance, the instructions were followed. Enforcing this approach by not making the original available makes necessary a third contact with the respondent to accomplish reconciliation. The burden, cost, and time lapse of a third contact usually override such a safeguard.

#### 4. Reinterview data.

With small sample sizes and low overall response rates, no reliable estimates of quality can be made. The available data have been examined for indications of quality, but conclusions must remain tentative. In 57 of 70 units in Florida, or 81%, there was agreement on the most important number, total current job

Extract of QM Schedule

Company Job Title (a)	Unit (b)	Recruiting Code (c)	Avail. Outside Firm Source (d)	Cur- rently Avail. Source (e)	No. of Current Job Openings			QM Occupation	
					(f) QM (20-21)	(g) O (22-23)	(h) R (24-25)	Source (i) (26-27)	Title (j) (28-41)
		(15)	(16-17)	(18-19)					

openings in the firm. In Utah, the corresponding numbers were 56 out of 75, or 75%. This provides evidence that the definition was manageable and that most firms made enough effort to report this number with accuracy. A closer look at the data, however, suggests that questions of data quality remain.

Table 2 gives counts of agreements and disagreements by state and rough size class. " $X_O$ " and " $X_R$ " stand for original number and reconciled number of current job openings, respectively. Firms are partitioned into "both values zero", "values equal and greater than zero", "original greater than reconciled", and, "reconciled greater than original." Notice that over half of all respondents reported no openings. This is a reminder that job openings, compared to establishment employment, are small in number. Overall, survey estimates of the vacancy rate were between 2% and 3%. This rarity of job openings means that small absolute errors may be sizable relative errors.

Second, naturally enough, disagreements increase with size since the proportion of firms with openings increases. Looking at firms with disagreements by size class, we count 9 out of 24, or 38% of the large firms, 250 and over, in Florida and 7 out of 14, one-half, in Utah. Utah found more disagreements in the medium sized firms. Since exact agreement is not a very good measure, especially in large firms, let's look at relative net differences,  $(X_O - X_R) / X_R \cdot 100\%$ . (Table 2). In Florida, they are small and in opposite directions. In the large firms, underreporting and overreporting tended to balance out. In Utah, for both these size classes, underreporting was observed. As indicators, these numbers are large enough to suggest the possibility of significant undercounting. This is further reflected in the comments during reconciliation in Utah. There, "insufficient effort" by the respondent was cited as the cause of error in several instances.

A final cautionary note on data quality: Weaknesses of the QM as a yardstick lead to overstating the amount of agreement. In order to get a difference between " $X_O$ " and " $X_R$ ", the interviewer had to get a QM value different from the original. In both Utah and Florida, approximately one-third of the error comments were "original better than QM." In other cases, due to time lag and dependence on

memory, the QM may have failed to pick up actual errors. In addition, we might well expect a bias component from nonresponse since number of job openings in a firm is related to effort required for response.

## 5. Conclusions.

Planners and interviewers both were pleased with the willingness of employers to participate in a survey addressing survey quality. Tying in more qualitative approaches to quality and collection issues, fruitful in their own right, can enhance recollection of data. Recollecting survey data can meet resistance requiring skill, tact, and determination on the part of the interviewer.

Careful testing before initiating a new survey is both desirable and difficult. Our low overall response rates were caused by many problems in trying to mount the survey itself and an evaluation survey at the same time. With nonresponse follow-up, reasonable cooperation rates were achieved on both the main survey and the reinterview, but getting both done within our time and resource constraints proved difficult.

For re-collection of data, the reinterviews must be timely or involve data not liable to early deterioration. The number of variables collected on reinterview should be limited, recognizing time constraints in the interview and sample size constraints usually placed on the design. With independent collection, followed by reconciliation, we were able to achieve our goal of improved reinterview values in two of four states.

With respect to data quality for job openings, limited conclusions can be drawn from the QM data alone. We have numbers, but given our low overall response they are best treated as indicators rather than as estimates. We will still be addressing collectibility and quality issues by examining Unit Profile and QM statistics together. In 1980, recruiting/hiring patterns have been collected in a diary survey and telephone collection has been tested in small firms. Reports covering all these areas are planned by the end of the year.

Table 2. QM Statistics

		Unit Counts - Agreements and Disagreements				Total Units	Selected Relative Net Differences(%)
		$X_O = X_R = 0$	$X_O = X_R > 0$	$X_O > X_R$	$X_O < X_R$		
Fla							
	0-49	25	4	1	0	30	
	50-249	5	8	3	0	16	10%
	250+	7	8	4	5	24	-3%
	Total	37	20	8	5	70	
Utah							
	0-49	32	3	2	1	38	
	50-249	8	6	2	7	23	-32%
	250+	3	4	2	5	14	-26%
	Total	43	13	6	13	75	

APPENDIX. Survey variables with brief definitions.

Period Covered--Job openings information should be reported as of the close of the last business day of the month of March, 1979.

Number of Current Job Openings (Vacancies)--A current job openings is an existing vacant job that is immediately available for filling and for which your firm is actively trying to find or recruit someone from outside your firm (i.e., a "new" worker--not a company employee).

"Actively trying to find or recruit" means current efforts to fill the job opening through such means as orders listed with public or private employment agencies and school placement offices; "help wanted" advertising (newspaper, posted notice, etc.); recruitment programs, interview and selection of applicants.

Current Job Openings Remaining Unfilled for a Month or Longer.

Current Job Openings that are for Part-Time Positions--Current job openings that are for positions expected to provide less than 35 hours of work a week.

Number of Openings with Future Starting Dates--

Openings for which your firm is actively trying to find or recruit someone from outside your firm (i.e., a "new" worker), but which relate to jobs that are currently occupied or unavailable for immediate occupancy for such reasons as: job unavailable until the present incumbent leaves; work will not start until a future date; new branch to be opened or anticipated increase in business.

REFERENCES

Hansen, M.H., Hurwitz, W.N., and Pritzker, L. (1964), "The estimation and interpretation of gross differences and the simple response variance", Contributions to Statistics, Pergamon Press, pp. 111-136.

Konstant, Raymond A. and Wingard, Irwin F.O. (1968), "Analysis and use of Job Vacancy Statistics: Part I", Monthly Labor Review 91, 22-31.

Ostry, Sylvia and Sunter, Alan (1970), "Definitional and design aspects of the Canadian Job Vacancy Survey," JASA 65, 1059-1070.

Schreiner, Irv (1980), "Reinterview results from the CPS independent reconciliation experiment", Bureau of the Census, 5/7/80 internal memo.