

# NOW THAT THE STUDY IS OVER, WHAT DID YOU REALLY TELL US? IDENTIFYING AND CORRECTING MEASUREMENT ERROR IN THE JOB OPENINGS AND LABOR TURNOVER SURVEY PILOT TEST

Karen L. Goldenberg and Mary Anne Phillips, U.S. Bureau of Labor Statistics\*  
Karen L. Goldenberg, BLS, 2 Massachusetts Avenue, N.E., Washington, DC 20212, Goldenberg\_K@bls.gov

## ABSTRACT

This paper describes qualitative research by the Bureau of Labor Statistics (BLS) to identify sources of measurement error in the Job Openings and Labor Turnover Survey Pilot Test. Over the year of pilot data collection, BLS refined the concepts being studied and modified the data collection form. BLS subsequently conducted in-depth debriefing interviews with a small group of Pilot Test participants and compared their data with intended survey concepts. The results of these interviews were used to identify and correct additional sources of measurement error, and to further refine the data collection forms.

**Key Words:** Qualitative research, debriefing interview, job openings, hires, separations

## 1. BACKGROUND

The U.S. Bureau of Labor Statistics (BLS) has a long tradition of measuring the *supply* of labor in the United States. BLS generates monthly employment and unemployment data through the Current Population Survey (CPS) and monthly changes in payroll employment through the Current Employment Statistics (CES) survey. However, there has been no comparable, consistent, and long-term information about the *demand* for labor. BLS developed the Job Openings and Labor Turnover Survey (JOLTS) to remedy this situation. Data collection began in the spring of 2000. A number of forthcoming papers describe the survey's economic underpinnings, general approach, and development (Clark, Cohen, and Hyson; Mueller and Phillips; Crankshaw and Stamas; Mueller and Wohlford). This paper focuses on measurement error aspects of the survey.

### 1.1. Study Design

JOLTS is patterned after the CES survey, but on a much smaller scale. The sample consists of 16,000 business establishments, selected at random within region, major industry division, and establishment size class (number of employees). BLS will collect data from establishments in the JOLTS sample once a month for a period of 18 months.<sup>1</sup> Operationally, BLS staff in the Atlanta Data Collection Center (DCC) contact sample members by telephone, enroll them in the survey, and mail them an information package. Both the interviewers and a letter in the package encourage respondents to record their data on the data collection form provided. For the next 6 months, BLS interviewers contact respondents and collect data using a Computer Assisted Telephone Interview (CATI) system. At the end of the 6-month period, respondents are asked to report their data by Touchtone Data Entry (TDE), which they will continue for a year. New panels of respondents will be phased in on a monthly basis, with most noncertainty respondents phased out after 18 months.

### 1.2. The JOLTS Pilot Test

During the summer of 1998, BLS contracted with Westat for a feasibility study of JOLTS issues.<sup>2</sup> Researchers from both organizations conducted site visits to learn whether turnover data were available at all, the likely quality of those data, and where in the firm to find them. Based on these interviews, the team concluded that the proposed study should be feasible (Levin et al., 1998). Westat then worked with BLS to plan and implement a Pilot Test. The Pilot Test, which began in September 1998 and lasted for a year, was a telephone study designed to test JOLTS data collection and survey procedures. The overall plan was to identify appropriate respondents, enroll them, send them a

---

\* The views expressed in this paper are those of the authors and do not represent official policy of the Bureau of Labor Statistics.

<sup>1</sup> Selected large employers will be assigned to the virtual certainty stratum, and will remain in sample indefinitely.

<sup>2</sup> Leda Kydonieffs of BLS worked with Westat to conduct the feasibility study and implement the Pilot Test. She continued these activities until her death in June 1999.

data collection form, and obtain data on what were then eight data elements. BLS designated a sample of 201 establishments, distributed across industries and size classes, as Panel 1 (Levin et al., 2000).<sup>3</sup>

Selected sample units received a package of advance materials by mail. Within two weeks, trained Westat interviewers called the respondents, attempted to enroll them in the Pilot Test, and collected the first month's data. Interviewers successfully enrolled 137 of the 201 Panel 1 establishments (68 percent), collecting data for the previous month. A year later, at the end of the data collection period, 115 of those 137 firms were still participating (84 percent retention rate).

After enrollment, the interviewers entered the establishment's first month of data on a data collection form and faxed or mailed the form to the respondent. For the next 6 months, interviewers continued to update the forms and fax them to respondents each month, following each fax with a telephone call to collect the information. After 6 months, interviewers stopped the faxes and collected all data by telephone.

The Pilot Test was also a vehicle for exploring conceptual and data quality issues. Occasional extra questions covered the use of outside employees and independent contractors, frequency of employer payrolls, data sources, and terminology. Westat's monthly meetings with interviewers also contributed to our knowledge of respondent misunderstandings, difficult topics, and potential sources of measurement error.

### **1.3. Survey Evolution**

By the end of the Pilot Test, the BLS team believed that labor turnover data could be collected by telephone. However, the Pilot Test no longer represented the plans for JOLTS. The team had made a number of significant decisions that affected the survey. First, we changed the sample unit from the Unemployment Insurance (UI) account to an establishment. This change did not affect single units, but shifted the basis of multi-unit collection from an aggregation of units to a single location or work site. We also changed the number of data elements and some key survey concepts. Because of these changes, we reworked the entire data collection form. Since the form serves as both a place to record information and a reference for compiling it, we addressed the limitations we identified in the Pilot Test version. We added more detail to column headings, used bullets to make information more visible, clarified wording, and increased the amount of explanatory information in the instructions on the back. Finally, we developed separate data collection forms for two industries: units in Education (SICs 821 and 822) and units in the Help Supply Services industry (SIC 7363). In the former case, we tailored instructions to the academic school year, while we needed the latter version because the basic JOLTS form instructs employers to exclude employees of temporary agencies, employee leasing companies, and independent contractors, instructions that are not appropriate for employers of these otherwise-omitted employees.

## **2. EVALUATING THE PILOT TEST**

While the Pilot Test was winding down, we conducted face-to-face debriefing interviews with a small number of participants. Because so many aspects of the survey had changed, the focus of these interviews reflected both the Pilot Test data and current plans for the survey. First, we explored whether the data provided on the Pilot Test conformed to our specifications, both in concept and for the appropriate reference periods. Second, we obtained feedback from respondents about the content and layout of the new data collection form. Working from a semi-structured interview guide, we asked about use of the data collection form and instructions, reference periods for each data element, and sources of the data. Finally, we described hypothetical situations in vignettes of a sentence or two and obtained respondent reactions to them.

We present qualitative observations based on 11 interviews with respondents in a variety of industries and sizes of business. Seven of the respondents were in Human Resources or Personnel offices, and one was the firm's comptroller. Others were administrative or managerial personnel. All interviews were conducted by the first author with another team member present, and were audiotaped and transcribed. Interviews ranged in length from 50 to 89 minutes, averaging 67 minutes. Responding units ranged in size from 26 to approximately 1200 employees.

The number of interviews is small, and the interview results can only be viewed as suggestive. Nevertheless, for many issues the responses were strongly consistent, which is a good indication that the survey has—or has not—met its objectives. Where the in-depth interviews pointed to specific wording or layout problems with the form, we note how we addressed them.

---

<sup>3</sup> We created an Education Panel with a small sample of schools (SIC 8211) because it quickly became clear that some of the concepts and definitions did not fit their situation. Two other panels addressed other research questions.

## 2.1. Survey Mechanics

While the Pilot Test was conducted by telephone, respondents were sent copies of the data collection form for compiling their monthly data. In spite of the regular faxing and mailing of forms, among the debriefing respondents use of the form was very limited. Five of the 11 interviewees said they completed the form monthly, one did "some months," and the rest did not use it at all. Asked why they did not use it, one respondent said "they call me." Two initially claimed they had never received a form, although in subsequent discussion they recalled having seen it. We have since modified survey procedures to emphasize the use of the form.

## 2.2. JOLTS Concepts and Data Elements

The JOLTS survey now obtains six data items for a specific establishment or work location. As part of survey enrollment, respondents receive a package of materials including a data collection form. The form is preprinted with the name and address of the survey contact, and requests information about job openings and employee turnover for the sample unit at the physical location also preprinted on the form. Instructions on the back clarify that we want information "about employees *on your payroll* at the location shown in Section 1 on the front of this page."<sup>4</sup>

The six data elements and their reference periods are: (1) **Total Employment** for the pay period that includes the 12th of the month; (2) **Job Openings** on the last business day of the month; (3) **Hires** for the entire month; (4) **Quits** for the entire month; (5) **Layoffs and Discharges** for the entire month; and (6) **Other Separations** for the entire month. Each data element has two separate components: the concept and a reference period. There is potential for measurement error if the concept is misunderstood or misreported, or if the correct concept is reported for the wrong time period. Because of space limitations, we present only the highlights of the research findings here. See Goldenberg and Phillips (2000) for more detailed results.

## 2.3. Employment

*Concepts.* We define Total Employment as all employees who worked or received pay for any part of the pay period that includes the 12th of the month. Who should be counted as an employee? During the Pilot Test, one concern was whether employers who use contractors and other outside employees would report those workers in their employee counts. During the debriefing interviews, we asked respondents if they had any outside employees working at their reporting units. Not only did all of the relevant employers exclude them from their counts, but most of the employers volunteered that these workers were not on their payrolls and thus not counted as employees.

Another concern was employment in Education during school vacation periods. While we wanted respondents to include teachers and professors during the summer, we were unclear about whether or not they would. During Pilot Test data collection for July, we asked Education Panel respondents if their employment counts included the employees who would be returning in the fall. Some respondents only included their returning employees if they actually worked during the reference period, omitting those who did not receive pay. We learned from the debriefing interviews that many school workers were under contract during the summer, even if they were not paid. Based on our discussions, we determined that the "worked or received pay" provision of the definition was too limiting. On the education form, we redefined employment as the total of "faculty under contract" and "all other full- or part-time employees who worked or received pay for the pay period that includes the 12th of the month."

*Reference period.* JOLTS follows CES procedures to obtain employment data, and uses the pay period that includes the 12th of the month as the reference period. Misreporting of the reference period is problematic because it introduces measurement error into estimates and the error is likely to add a bias. However, if the data are reported consistently, the overall effect should be relatively small.

One issue associated with the reference period is whether respondents understand the concept of the pay period that includes the 12th of the month and report accordingly. In the debriefing interviews, we asked respondents if the employment figure might refer to a more current pay period, or if they always went back to the specified pay period. Only 4 of the 11 respondents explicitly reported for the correct pay period. Others told us that the pay period was later some or all of the time, or that the data were for the entire month. One small, low-turnover respondent gave employment data as of the time of the interviewer's call. The effect of this misreporting is mitigated by the fact that most of these respondents have very low turnover, so the numbers for the incorrect pay period were actually the same as those for the correct pay period. However, we developed procedures to emphasize the correct pay period. Interviewers review the form with respondents prior to collecting any data, offer to help respondents identify that pay period, and refer to the pay period when they collect data each month.

---

<sup>4</sup> The Help Supply Services industry (SIC 7363) is an exception. Instructions on the JOLTS form for this industry cover their special circumstances, and differentiate between headquarters staff and employees at client locations.

## 2.4. Job Openings

*Concepts.* JOLTS obtains a snapshot of job openings at the specified location on the last business day of the month. The definition covers three explicit conditions. First, a position must exist, with work available for that position. Second, the job *could* start within the next 30 days, even if it will not actually start that quickly because the employer is unable to fill it. Third, the employer must be actively recruiting to fill the position from outside of the sample location

The job openings item has a number of issues associated with it that could lead to measurement error. For one thing, to say that "a position exists and there is work available" is not necessarily the same thing as saying "a position is vacant." Although most of the theoretical measures of job openings have this expectation (e.g., Abraham, 1983), the same cannot be said of the respondents. Through vignettes, we described a situation where, toward the end of one month, employees gave notice or announced an intent to depart early in the next month. We asked if these upcoming departures would be included as job openings for the month in which the employees gave notice. Depending on the question, 4 or 5 of the 11 respondents said that a position was not open if the incumbent was still in it. However, more than half disagreed, generally because they knew they would have a vacancy to fill. Unless the interviewers make a concerted effort to exclude any still-occupied positions, misreporting and measurement error are likely.

The 30-day time period for starting work can also be a source of confusion, because we are looking for current rather than future job openings. We had to explain that "30 days" referred to the job and not the presence of a candidate, and subsequently added this explanation to the data collection form. This can be a problem in education, where anticipated vacancies are announced well ahead of the 30-day time frame, but should only appear as JOLTS job openings if the positions are still open 30 days before school starts. Using vignettes, we explored whether respondents would count positions as job openings if they were recruiting for jobs that would start in more than 30 days. In these vignettes, respondents generally went along with the definition, if sometimes grudgingly.

We were reassured to see that respondents were reporting other aspects of job openings as we defined them. All respondents were actively recruiting for the job openings they listed, and all of those positions could start within 30 days. Also, in response to a vignette, most respondents said they would not report a position as a job opening once a candidate had accepted a position.

*Reference period.* The reference period for job openings is intended to be a snapshot of job openings on the last business day of the month. Do respondents actually report for this reference period? Five debriefing respondents specifically said they reported their job openings for the last business day, although there were a number of qualifications and questions associated with this observation. Several respondents told us their job openings figures were for the whole month, but the jobs reported were still open on the last business day. As long as all of those jobs are still open, they are reporting correctly. We learned, however, that not everyone *can* report for the last business day. Three respondents obtained their data from external sources, and had no control over its timing or its content.

## 2.5. Hires

*Concepts.* Hires are additions to the payroll at the specified location. The concept includes all newly-hired and returning employees, whether they are hired for full-time or part time-work, and regardless of whether the positions are permanent, seasonal, or for other temporary or short-term activities. Hires also include recalls from layoff. Hires specifically exclude employees returning from leaves of absence or strikes, and transfers or promotions within the sample location.

The hires concept tends to be relatively straightforward, especially since none of the debriefing respondents had experience with layoffs and recalls. One situation that we explored, using a vignette, was of an employee who was hired and fired during the same week. Nine of the 11 respondents told us they would include such an employee in the new hires counts. One who said the employee would not be counted referred to the fact that the person would not appear on the payroll. Another talked about how the position would be open again. The directions on the form instruct respondents to include as a hire someone who was hired and separated during the month.

A few of the multi-unit respondents had transfers during the course of the year. Under the Pilot Test sample unit, transfers from one unit to another were not hires, and they were not treated as hires by respondents. With establishment-based reporting, however, a transfer from one location to another does become a hire. This could be a cause for concern if what constitutes a hire is newly generated employment paperwork, but we did not pursue the topic in the debriefing interviews.

*Reference period.* We collect hires data for the entire reference month, and all of the debriefing respondents reported for the correct reference period. We probed to see if "month" meant the first to the last day of the month, or if it actually ended earlier. Only one person said the month ended earlier, because if an employee started at the end of the month she might not know about it until the next month. We asked about the hire date, because we thought

that hire information might be based on the date an offer was made or accepted. Everyone told us that hires referred to people who had actually started employment during the month.

## 2.6. Separations

*Concepts.* The final component of labor turnover is separations. *Voluntary* separations are quits, resignations, or other departures initiated by the employee. *Involuntary* separations include layoffs or terminations with no intent to rehire employees, "temporary" layoffs (i.e., layoffs with intent to recall), discharges or firings for cause, terminations of seasonal employees, and staffing reductions or downsizing resulting from mergers or acquisitions of other businesses. In JOLTS, we further break out a third category, *other separations*. Other separations can be either voluntary or involuntary, such as retirements and transfers to a location outside of the sample unit, or can occur for non-economic reasons such as employee disability and death.

In the Pilot Test, respondents reported separately for quits, layoffs lasting over 30 days, and all other separations. "Layoffs over 30 days" was intended to capture layoffs with intent to recall. However, the column heading on the front of the data collection form said simply "Layoffs." In addition, the "over 30 days" stipulation made the information hard to collect. Respondents had to determine when the layoff took place and not report it until the 30 days had passed. Depending on the timing of the interview, two months might elapse between the time the layoff actually occurred and the time it satisfied the definition for capture. For the actual survey, we shortened the layoff period to "more than 7 days," giving us consistency between layoffs and recalls and making it easier to report layoffs. Another problem in the Pilot Test was that discharges for cause were part of other separations, although they are involuntary. For the actual survey, we renamed the category Layoffs and Discharges. The new group includes all types of involuntary separations, permanent and temporary.

In the debriefing interviews, we asked respondents to tell us their understanding of the words "quit" and "layoff," and asked what terms they used for terminations where no rehire was anticipated. Most respondents described *quits* with the word "voluntary," and respondent comments suggest that most of the reported layoffs were permanent. On the revised data collection form, we refer to quits as "employees who left voluntarily" and to layoffs and discharges as "involuntary separations initiated by the employer."

One of the many challenges in designing this form was to find a way to convey all three types of separations to respondents on the front of the form (matrix), because many people will not look at the detailed instructions on the back. Our goal was to present the three different types of separations in an overall topic heading and explain what to include in each of the three columns. As part of the debriefing interview, we showed respondents a draft of the revised form. We found that some respondents were missing the idea that each type of separation went into a separate column. Our solution was to change the information in the topic heading to more clearly indicate that each topic went into a different column (Goldenberg and Phillips, 2000).

*Reference periods.* We collect separations for the entire reference month. All of the debriefing respondents reported separations correctly for the entire month. The only real issue here is how the employer defines the separation date, which could be the last day on the job or the last day the employee was paid. In most cases, if an employee was to be paid for accumulated vacation time, the employer paid the lump sum on the last day the employee worked. One employer noted that in rare situations, a person might get a severance package, so the separation date would be defined by the end of the severance payout rather than the last day worked.

## 3. IMPLICATIONS FOR JOLTS

### 3.1. Measurement Error in the JOLTS Survey

Measurement error refers to the difference between the true value of a data item, as defined by the researcher, and that which is actually reported by the respondent. Survey methodologists generally attribute measurement error to four interdependent sources: the respondent, the interviewer, the questionnaire or data collection form, and the mode of data collection (Groves, 1989). The information system from which respondents obtain data is an additional source in establishment surveys (Edwards and Cantor, 1991). Examples of measurement error—which we also refer to in this paper as response error or reporting error—might be variations in a respondent's answer to the same question because of the way an interviewer collects the data; incorrect responses because a question is ambiguous and the respondent has to guess what is being asked; or responses on TDE that differ from what they would be by mail or CATI.

Over the course of the Pilot Test, we learned both accidentally and systematically of potential or actual measurement errors. Although we revamped the data collection form because of decisions governing survey design, we made additional changes because of knowledge gained in monthly post-data collection interviewer meetings, from answers to the extra questions, and from our growing experience and familiarity with the subject matter.

Development of separate forms for Education and for the Help Supply Services industry were a direct result of findings from the Pilot Test.

### 3.2. Reducing Measurement Error: The Debriefing Interviews

The debriefing interviews provided us with an in-depth way to see exactly how and what respondents were reporting. These interviews pointed to measurement error deriving from the structure of the respondent's organization or work environment, and communication problems that can be reduced or eliminated through forms design, clear questions, and appropriate survey procedures. The overwhelming majority of errors we identified, both through the debriefing interviews and in the Pilot Test, were communication problems.

*Organizational effects on data.* Firms of all sizes and structures can have policies and procedures that affect the timing of hires and separations. During the debriefing interviews, we learned about some types of measurement or reporting error that originate in the employer's internal procedures or result from the organizational structure in which the respondents work. For example, respondents in multi-establishment organizations may have some role, such as recruiting, but may not make hiring or termination decisions. As a result, the respondents use data collected for purposes other than JOLTS, and these data may or may not fit our definitions and reference periods. The change to establishment-based sampling will not eliminate this problem, because it is fairly common in multi-establishment firms for a single contact to collect and compile information for all units in the company.

*Communication problems.* The largest sources of measurement error that we identified in the Pilot Test resulted from communication problems that interfered with respondent comprehension. We addressed forms-based measurement error by changing the form layout, by clarifying and expanding upon the descriptive information, and by providing respondents with instructions to cover as many likely situations as possible. Where adding specific instructions made the form too crowded, as it did for education, we created an industry-specific form.

Another source of error stemmed from respondents not using the form. We addressed this issue by developing survey procedures that stress respondent use of the data collection form and the information on it. We developed a brochure with Frequently Asked Questions and illustrative examples which goes to respondents as part of the enrollment package. Finally, the DCC has trained interviewers to ensure that they thoroughly understand JOLTS survey concepts and procedures. Overall, we believe the information we gained from the Pilot Test and the debriefing interviews, and the changes we made as a result of this research, will result in easier reporting and more accurate data.

## 4. REFERENCES

- Abraham, K.G. (1983), "Structural/Frictional vs. Deficient Demand Unemployment: Some New Evidence." *American Economic Review*, **73**, pp. 708-724.
- Clark, K., J. Cohen, and R. Hyson (forthcoming), "Measuring the Demand for Labor in the United States: The Job Openings and Labor Turnover Survey." To be presented, Joint Statistical Meetings, Indianapolis, August 2000.
- Crankshaw, M. and G. Stamas (forthcoming), "Sample Design in the Job Openings and Labor Turnover Survey." To be presented, Joint Statistical Meetings, Indianapolis, August 2000.
- Edwards, W.S. and D. Cantor (1991), "Toward a Response Model in Establishment Surveys," in P. Biemer et al., (eds.) *Measurement Errors in Survey*, New York: Wiley, pp. 211-233.
- Goldenberg, K.L. and M.A. Phillips (2000). "Using Qualitative and Quantitative Data to Study Measurement Error in the Job Openings and Labor Turnover Survey Pilot Test." Unpublished report, Bureau of Labor Statistics.
- Groves, R.M. (1989), *Survey Errors and Survey Costs*, New York: Wiley.
- Levin, K., C. Gimbel, T. Hagerty, S. Heltemes, A. Becher, and L. Kydoniefs (1998), "Job Openings and Labor Turnover Statistics (JOLTS) Feasibility Study: Report of Site Visit Findings," Rockville, MD: Westat.
- Levin, K., T. Hagerty, S. Heltemes, A. Becher, and D. Cantor (2000), "Job Openings and Labor Turnover Study (JOLTS) Pilot Study: Final Report." Rockville, MD: Westat.
- Mueller, C. and M.A. Phillips (forthcoming), "The Genesis of an Establishment Survey: Research and Development for the Job Openings and Labor Turnover Survey at the Bureau of Labor Statistics." To be presented, Joint Statistical Meetings, Indianapolis, August 2000.
- Mueller, C. and J. Wohlford (forthcoming), "The Debut of a New Establishment Survey: The Job Openings and Labor Turnover Survey at the Bureau of Labor Statistics." To be presented, Joint Statistical Meetings, Indianapolis, August 2000.