

Quality Assurance of the Consumer Price Index Housing Survey

Carol Lynn Kosary, William D. McCarthy, John Paul Sommers, James M. Branscome
Statistical Methods Division
Office of Prices and Living Conditions
Bureau of Labor Statistics

1.0 Introduction

Given the increasing demands put on data by decision makers, it is essential to improve data quality. With the 1987 Consumer Price Index Revision (CPIR), the Bureau of Labor Statistics (BLS) was mandated to institute a Quality Assurance program for all CPI surveys (Comptroller General of the United States, 1983). This mandate presented a unique opportunity to explore and determine what is needed in a Quality Assurance program for a survey operation. To date, most quality assurance efforts have been in the area of manufacturing and most techniques and models have been designed for that environment. Most standard texts (such as Juran, 1974) on quality assurance speak of comparing the item to be inspected against certain predetermined measures. These could be mean time to failure for a lightbulb, the mean and standard deviation of the weight of boxes of cereal or the number of times a preset tolerance is exceeded in margin width or pitch in printing a book. What all of these examples have in common is that they can be measured in an absolute sense. Measures are made on end product results, with these measures giving information on the conformance of the end product to preset standards. It can also be readily determined what steps to take to bring a non-conforming process into conformance. Many of these techniques can be implemented in a survey operation environment. What this approach fails to reveal are the impact of such areas as data outliers, response effects of respondent comprehension of survey questions, missing data and imputation, and the other myriad of effects which contribute to non-sampling error (Wright, 1983).

A survey operation is similar, in some ways, to a multi-step manufacturing process, with its end product being the data collected. Standards can be easily obtained for such steps in the process as material preparation, keypunching, and data collector conformance to data entry procedures. Evaluation of these steps should be a part of a survey quality assurance investigation. Such evaluations, though, fail to indicate whether the data being collected is actually the data intended. Treating a survey operation as purely a manufacturing process and utilizing

standard quality assurance techniques brings up some interesting questions. What are the standards against which the end product can be measured? What are the sources of non-conformance? What can be done to correct non-conformance?

In an attempt to answer these questions, quality assurance of the CPI Housing survey was designed to be more than just an end process where systematic sources of inconsistencies are identified through a continual data recollection effort. Such an approach, focused only on the area of data collection, is incomplete. The quality of the data collected will not make a difference if the intended purpose of the survey has not been a consideration of the activities which came before and after the data is collected. Quality assurance of the CPI Housing Survey was used as an opportunity to not just measure the "quality" of the system using standard techniques, but also to "build" quality into the overall process. In short, the purpose of the activities developed to evaluate the quality of the CPI Housing Survey have the aim of attempting to ensure that the CPI Housing Survey is implemented in a manner which is in satisfactory agreement with the design of the survey.

Quality assurance of the CPI Housing Survey therefore consisted of two stages. The first stage involved all of the effort and testing which went into the design and development activities before implementation of the CPIR in July 1985. The second stage consisted of ongoing efforts to monitor the all facets of the survey process. The work done in each stage was designed to exert a preventive control, identifying and attempting to minimize potential major sources of error in advance, or process control, identifying and eliminating errors as they occur. The remainder of this paper will be primarily concerned with describing these activities.

2.0 CPI Housing Survey - A Brief Description

The Housing sample consists of 10,930 area clusters (segments) which when processed will result in approximately 40,000 renter and 20,000 owner occupied units which will support both the Rent and Rental Equivalency (REQ) Indexes of the CPI. These housing units are further divided into six subsamples, called panels. Renter occupied units

are priced every six months and owner occupied units are priced once a year.

Interviews are conducted either in person or over the telephone, and rental units are priced either by contacting the occupant or the unit owner/manager. Therefore, in designing the procedures and data collection questionnaire the dual purpose of the data and multi-mode data collection must be taken into consideration.

3.0 First Stage

First stage work involved the sample selection and all efforts in design and development of the data collection forms and procedures. This was basically preventive control work intended to identify and minimize potential errors. In this stage the sample selection process was automated and extensive redesign was done on the procedures and questionnaires used to collect CPI Housing data.

3.1 Sample Selection

Extensive use of automation was made in the sample selection process. Using 1980 Census information, the Housing Survey Sample was completely computerized. This has meant an improvement in the sample because of the capability to produce more uniformly sized segments than is possible in samples selected by hand. The segments were also more uniform regards to the number of renter and owner housing units, resulting in a potential decrease in the variance of the Rent Index and improvement in REQ by supporting the matching of owner and renter housing units. Monitoring of this sample can also be easily accomplished due to computerization. This meant that the need to augment could be quickly recognized. With more information on file and the process computerized, it will be possible to use these computer routines in future revisions.

3.2 Segment Listing

After the sample was selected, mapping information was generated and sent to the Primary Sampling Units (PSU's) which make up the CPI sample for Segment Listing. In Segment Listing the data collector lists on a specifically developed Segment Listing Form (BLS 222C) all of the units in a specified area. This was done on a one unit per line basis with ineligible units such as commercial establishments and public housing eliminated from sampling consideration. A predetermined sampling rate was then applied to the remaining

units. The sample selection criteria was also computer generated and was specific for each segment, being based on such criteria as expected number of owner and renter units. This use of automation allowed for the timely comparison of expected units to actual units.

Prior to the start of Segment Listing, a number of small scale investigations were made to determine whether the data collection form and procedures for Segment Listing were adequate for the purpose intended. These investigations involved independent listings of selected segments by several listers followed by an analysis of the listing differences and an investigation by another lister into the causes for the differences. These investigations were a great aid in the development of the final data collection form and procedures.

3.3 Screening/Initiation/Repricing

As a result of Segment Listing the selected units were then screened utilizing a Screening Form (BLS 222S). From this screening it was determined whether the unit was eligible for inclusion in the sample (i.e., it was not public housing, a college dormitory, hospital or nursing home, etc.), whether the unit was built prior to 1980, that the unit was intended for year round occupancy and the primary residence of at least one occupant. The unit is also screened on owner/renter status. Depending on the sample selection criteria a unit will be included only if it was a owner or a renter, while some units are included if either, unless the other screening criteria are not met. If the unit does not successfully meet the screening criteria and desired tenure, it is said to have "failed" the screening and the interview process is terminated.

If the screening criteria are met, the unit is said to have "passed" and it will then be initiated into the sample using an Initiation Form (BLS 222I). These units will remain in the sample, being repriced on a regular basis (using a Repricing Form, BLS 222R) until they are replaced.

Data collection is possibly the largest source of error in a survey operation. The survey designer's goal of conveying their requirements through the data collection procedures and survey questions can at times seem similiar to the children's game of Post Office (Federal Committee on Statistical Methodology, 1984). In the design of the CPI Housing Survey data collection forms and procedures, a great deal of work was

involved in the attempt to translate the desired data requirements into questions. These questions were then tested under actual field conditions. The interviews were observed by a member of the Survey Design group utilizing an observational checklist designed to take note of specific behaviors regarding each question (Cannell, et al, 1975). The respondents were then questioned as to their comprehension of specific questions (Committee on National Statistics, 1984). Following this investigation, the information obtained was used to further refine the wording of the questions.

4.0 Second Stage

Second stage work involves all ongoing efforts to insure the continual quality of the data collected. Most of these efforts are designed as process control activities which are being designed to identify problems areas so that corrective steps can be taken. This area includes the monitoring of the sample selection and screening processes, a systemitized method of checking the data collection activities, checking of the data collection staff.

4.1 Housing Sample Initiation Edit and Control System

4.1.1 Purpose

Since extensive work in computerizing the Housing Sample was done for this revision, it was deemed advantageous to create an edit and control system to keep track of the sample as it was initiated. Goals of this system included correcting such errors as erroneously coded addresses for units that were correctly initiated, keeping incorrectly screened and initiated units out of the data base, and identifying units which should have been screened but were not so that they could be reinterviewed. Since its initial development, this system has been enhanced to the point that many of the edits initially entered by hand are now entirely automated. In addition, this system can be used to augment the sample with different housing units from the same sample quickly and efficiently and provide total consistency between files giving the ability to automatically rotate the sample, if desired.

4.1.2 Process

The Housing Sample Initiation Edit and Control System is designed in three phases. Phase one runs an edit of the Housing Survey Segment Listing Forms. These forms provide a line by line listing of the housing units in a selected segment. When the sampling

rates are applied to these listings, the units to be screened are obtained. Phase one consists of a series of edits to verify the proper completion of the Segment Listing Form.

Phase one is rerun until all errors are eliminated. Counts of errors for each run are transmitted to a dataset so that reports indicating the types of errors encountered can be produced at the end of each month. The results of these reports can be used to alert the data collectors to problem areas.

Once the listings are "clean", the sampling rates for owner and renters are applied to the list of uniss in order to identify those units to be screened. These sampling rates were sent to the field earlier to be applied to the listings. The sampling rates are checked at this step to verify that the task was done correctly.

Phase two involves an edit of the passed and failed CPI Housing Schedules. These verify that the answers to the screening questions are consistent with the "passed" and "failed" status coded.

In phase three the "clean" listings and screening files are merged and the final edit program is run. This final edit program also tabulates the following counts.

COUNT OF GOOD INITIATIONS
COUNT OF GOOD RENTER INITIATIONS
COUNT OF GOOD OWNER INITIATIONS
COUNT OF TOTAL SCREENINGS
COUNT OF GOOD SCREENINGS
COUNT OF LISTINGS

These counts and the error counts are output to another data set to be tabulated at the end of the month.

After all edits are run and all errors are corrected, a series of reports are generated. The first report shows the types and numbers of errors encountered along with counts of the number of units listed, number of units screened, the number of renter units which "passed" initiation, the number of owner units which "passed" initiation, and the total number of units initiated along with an indication of keypunch error. A second report presents counts by interviewer type code. The final report shows the number of listings, screenings and initiations that have been received and the total number which were expected. This report will be generated by PSU as well as PSU/strata. This reports indicates areas of the sample that might require augmentation. It is

an aid in determining whether work is being done equally across all PSU's and strata .

4.2 Process Audit

The CPIR quality assurance activity which most closely follows the traditional quality assurance model is the process audit. As part of the CPI revision, a separate group responsible for auditing of the CPI surveys was established. The function of this group is to recollect the data originally collected by the data collectors.

Where an effort of this type differs from typical quality assurance comparisons is in the fact that no "true standard" exists against which the data collected by the Process Auditors and the original data collector can be compared. The recollected data can be used only to establish patterns of discrepancies between the process auditor and the data collector. Tabulating these differences indicates only that the "true standard" is being misinterpreted by either the data collector, the process auditor, the respondent or the process design itself, (i.e., the "reflected standard", as manifested in the survey procedures, questionnaire, and training programs does not adequately convey what was intended by the designers or the designers failed to clearly convey the requirements). Once these areas of potential inconsistencies are discovered further investigation of all possible sources of the problem and possibly further detailed study of the area through formal testing are needed before any correction can be attempted. Process Audit is a tool for identifying but not fixing potential design error in the survey.

To date, two areas of the CPIR Housing Survey have been process audited. These areas are Segment Listing and Initiation.

4.2.1 Segment Listing

Approximately 450 segments from 24 PSU's comprising the CPIR Housing Sample were independently relisted by Process Audit in July, August and October 1985. In order to evaluate the effectiveness and fitness for use of the Segment Listing process, it is important to understand fully the purpose and intended use of Segment Listing. According to the CPI Housing Survey Data Collection Manual: "the purpose of listing is to provide a universe of housing unit addresses from which a sample can be selected for the Bureau of Labor Statistics Consumer Price Index - Housing Survey. As a result of the listing process and

subsequent data collection efforts housing information is available for inclusion in the housing component of the CPI."

In order to evaluate the effectiveness and fitness for use of the listings, it was determined that the following minimum requirements must be met:

A. The data collector must be able to use the listings in order to locate the selected units.

B. The listings must be legible for use by both the data collector and the keypunchers. The information must also conform to instructions designed to aid in keypunching.

C. The information collected must conform to instructions designed to aid the Housing Sample Initiation Edit and Control system. These instructions include the use of proper abbreviations.

D. The listings must provide the information required by the sample designers. This includes;

- 1) conformance to procedures designed to minimize listing errors.
- 2) presenting accurate and complete coverage of the selected sample frame.
- 3) conformance, on average, to the Census information from which the sample was selected. That is, how closely do the housing units listed in a segment correspond to the number of housing units expected from the Census.
- 4) listing the sample accurately enough that selected units can be located using the listings, local maps and initiation procedures.

Results indicated that the data collector conformed to instructions. The number of housing units listed, on average, was slightly lower than the count made by Census. Some of this undercount may be due to non-apparent units (hidden units not apparent to the observer) and units considered out-of-scope by BLS but not Census such as certain religious and health institutions. Other areas where differences between the two listers occurred were in determining number of units in multi-unit structures, especially in the case of houses which have been converted to several apartments, and determining segment boundaries in segments where a boundary does not consist of a fixed feature. No modifications have been made to the

procedures to better handle those cases where the number of units in a structure is unknown. In the case of non-physical boundaries, there does not appear to be any bias in one lister always listing fewer or more units than the other. So long as the errors in boundary determinations are random they should have no detrimental effect on the CPI.

4.2.2 Initiation

At initiation it is determined whether the selected units have "passed" or "failed". Those which "pass" are initiated into the sample. Once screened and initiated, a housing unit remains in the sample until it is replaced. It is the information collected regularly on these units which is used in the calculation of the rent and rental equivalency indexes of the CPI.

Approximately 1200 housing units in 37 PSU'S were recollected by Process Audit between October, 1985 and March, 1986. The data collected by Process Audit was compared to the data originally collected by the data collector. Different tabulations of discrepancies were then performed.

Although the analysis of this data is not complete, there are some general patterns which, although not large, do point to survey design and operational issues which need to be explored in detail. Answers to some questions differ systematically by the type of respondent (tenant, manager, etc.). Answers to other questions appear to vary between the first and second asking. Discrepancies also seem to be related to structured features in the survey instrument, such as ordered multiple-choice responses and questions dependent on earlier responses.

4.3 Monitoring Reports

Extensive use has been made of a series of standard reports, generated at fixed intervals, designed to monitor the progress of the sample and the conformance of the data collectors. In the case of monitoring data collector performance these reports are the first efforts made to increase uniformity in the evaluation of the data collectors. Previously, data collectors were monitored solely by the Region Office under which they worked. In the past each of the eight regions had differing standards against which they compared their data collectors. By monitoring data collector performance on a national level, levels of performance and standards can be established to assure that all data collectors are performing

at some minimum overall level. This will insure that all regions are evaluating their data collectors using the same standards and that these standards are relevant to the requirements of the survey.

Four types of reports are now being generated. Two of these monitor the progress of the sample to insure that counts are close to expected values. The other two monitor the performance of the data collector.

4.3.1 Reports to monitor sample

These reports are designed to monitor the progress of the listing and initiation process and to forecast the final counts from the work already completed.

The Listing Report indicates how many segments there were to start with. It also reports how many units were listed and screened compared to how many were expected from Census. This report can be examined down to the PSU/panel level. From these breakdowns, a good indication of the progress of Segment Listing and early indications of where undersampling may be occurring can be obtained.

The Initiation Report looks at number of owner and renter units initiated compared to number expected. This report also forecasts what the final counts will be based on the work completed. Breakdowns to the PSU/panel and PSU/panel/strata level indicate where undersampling may be occurring and where augmentation may be required.

4.3.2 Reports to monitor data collector

The questions on the CPI Housing Schedule follow a fairly complex pattern, with the flow of the interview being determined by the answers to previous questions. These questionnaires are even more complex because they were designed to serve both renter and owner interviews, with certain sections of the questionnaires used exclusively for either a renter or owner, while other sections are dual purpose. The first of the two reports of data collector performance was designed to indicate whether a data collector was correctly following the questions on the questionnaire. These "skip pattern" reports can help to identify data collectors who are experiencing problems. Averages, by question and overall, at the national and regional level can be automatically generated. This makes it possible to look at data collectors who are several standard deviations beyond these averages. This report can also look for

problems caused by the questionnaire. If particular questions have a higher than average error rate this could indicate a problem with the question such as its location on the questionnaire, its wording, or the procedures regarding its completion.

The second report looks at response codes. Response codes indicate the completion status of an individual questionnaire (i.e., the questionnaire is totally complete, the questionnaire is partially complete, the respondent refused, there was no contact, the unit was vacant, etc.). This report can be analyzed in the same manner as the skip pattern report. Data collectors can be compared against national and regional averages and those who indicate a higher number of no contacts or refusals, for example, can be more closely examined.

4.4 Certification

Certification was designed to test a minimum level of proficiency of the data collector, after on the job training but before the actual start of independent data collection. Certification was instituted to insure that the data collector could adequately collect the data required. Separate requirements were developed for Segment Listing and Initiation/Repricing.

4.4.1 Segment Listing

The purpose of Segment Listing Certification is to demonstrate that every data collector involved in listing can complete an assignment in a satisfactory manner, following established procedures. In Segment Listing Certification, the data collector is assigned several preselected segments to list independently. These segments are evaluated for conformance to procedures (i.e., legibility, proper abbreviations, etc.), and for the accuracy of the work (i.e., listed complete segment, listed every unit) using a standardized checklist. On the basis of the results obtained for this evaluation, the data collector will either be certified or required to undergo additional on the job training and reassessment.

4.4.2 Initiation/Repricing

The purpose of Initiation/ Repricing Certification is to demonstrate that every data collector involved in Initiation/Repricing can conduct a data collection interview, following established procedures. In Initiation/ Repricing Certification the data collector is observed while conducting actual interviews. These interviews are evaluated, using a standardized

checklist, to insure that the data collector followed the procedures and probed, when necessary, in a non-leading manner, etc. The interviewer's overall behavior and the manner in which they completed the questionnaire are also evaluated. On the basis of the results of this evaluation, the data collector can be certified or required to undergo additional on the job training and reassessment.

5.0 Concluding Remarks

As with any large project, work is continual. The CPI Revision will be concluded in January 1987. At the present time the CPI Housing Survey staff are planning various projects which will further impact the quality of the survey. Some of these plans call for the production of detailed reporting back to the data collection staff on problem areas encountered. The impact of data editing during final review will also be examined along with the feasibility of automating many of the data reviews and edits currently conducted by hand. Finally, the use of Computer Assisted Telephone Interviewing

is being studied to determine its potential impact on the data.

BIBLIOGRAPHY

Cannell, Charles F., Lawson, S.A., and Hausser, D.L., (1975), A Technique for Evaluating Interviewer Performance, Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor.

Committee on National Statistics, (1984), Cognitive Aspects of Survey Methodology: Building Bridges Between Disciplines, National Academy Press, Washington, D.C.

Comptroller General of the United States, (1983), Funds Needed to Develop CPI Quality Control System, General Accounting Office, GGD-83-32, April 1, 1983.

Federal Committee on Statistical Methodology, (1984), Statistical Working Paper 10 - Approaches to Developing Questionnaires, U.S. Government Printing Office, Washington, D.C., PB-84-105055.

Juran, J.M.(ed.), (1974), Quality Control Handbook, McGraw-Hill, New York.

Wright, Tommy (ed.), (1983), Statistical Methods and the Improvement of Data Quality, Academic Press Inc. Orlando.