EVALUATION OF A FIVE-STATE PILOT FOR AN SIC REFILING AND QUALITY MEASUREMENT SYSTEM

Susan C. Hostetter* and Karen Vlasek O'Conor, Bureau of Labor Statistics

INTRODUCTION

The Standard Industrial Classification (SIC) code on a universe frame is one of the most important characteristics, and also one of the most significant potential sources of error for sample surveys. Inaccurate SIC's may affect: the determination of whether or not a unit is considered in-scope of a survey, the stratification for the sample design, the statistical properties of the estimator, and the publication cells for the survey data.

A recently published statistical policy paper from the Office of Management and Budget points out that most Federal agencies that maintain industry coding systems have limited information about specific quality assurance measures for their systems.[1] This two-year study of major coding systems found no information on quality measurement results of the accuracy of SIC codes or coding procedures. Since most economic indicators are based on data from one or more of these coding systems, there is currently a need for specific research testing in this area. This paper describes recent pilot studies and improvements to the SIC coding system for the Bureau of Labor Statistics (BLS) Universe File.

The BLS Universe File is based on Unemployment Insurance (UI) administrative records supplied by States to BLS under a Federal/State Cooperative Program. Under the SIC refiling program, States mail a questionnaire to each employer, covered by unemployment insurance, on a three-year cycle. The questionnaire requires information on: economic activity for SIC coding, geographic location for county coding, type ownership, auxiliary status, and multi-unit status. This will be referred to as the classification method of refiling.

As a step toward maintaining a regular three-year SIC refiling cycle to establish current and accurate SIC's, BLS has conducted a detailed evaluation of its system for refiling and the actual procedures used by the States. Concurrent with this evaluation, BLS developed and tested an alternative verification method for refiling SIC's. The goals of the new method are to:

- provide high quality SIC codes,
- provide the capability for objective measurement and control of the quality of SIC codes,
- reduce State costs for maintaining an accurate three-year refiling cycle, and
- reduce respondent burden.

Last year BLS described its verification method for obtaining current SIC codes and its plans for testing the viability of the new method.[2] This year BLS has the results of four of the five pilot tests which were described last year. All results currently indicate that the verification method is a reliable means of updating SIC codes with the added bonus of reducing respondent burden and State costs.

Initially, the verification method was indended as an alternative methodology limited to States unable to maintain a full SIC refiling cycle using the classification method. However, based on the favorable results from the pilot studies and the significant cost and burden reductions, well over half of the States already use or are planning to use the new method.

This paper will present results of the pilot studies which were conducted in Maine, Oklahoma, South Carolina, and Texas. The following sections of the paper will:

- Describe the verification pilot tests' objectives and design.
- Present the pilot tests results.
- Describe current improvements developed to control nonsampling error in SIC coding.
- Describe future improvements including a proposed automated system for State SIC refiling and the quality implications of the system for SIC codes.

VERIFICATION PILOT TESTS

Objectives

The objectives of the verification pilot tests were:

- to provide a framework for comparing the verification method of refiling with the classification method with respect to response rates, respondent burden, and State resources required.
- to develop measures of response and nonresponse error for the verification method through a Quality Measurement (QM) reinterview study.

Design of the Pilot Tests

Scope - BLS contracted with five States; Maine, Michigan, Oklahoma, South Carolina, and Texas to carry out pilot tests of the verification method. The pilot tests contained two activities: refiling by mail using the verification method and conducting a telephone QM reinterview survey on the verification method refiling results.

The target population for each state consisted of the industries the State was currently scheduled to refile. Maine and Oklahoma refiled wholesale trade (SIC's

50-51) and retail trade (SIC's 52-59) establishments. South Carolina refiled finance, insurance, and real estate (SIC's 60-67); and services (SIC's 70-89) 70-89). Texas refiled agriculture (SIC's 01-09); transportation, communications, electric, gas and sanitary services (TCPU) (SIC's 40-49); and services (SIC's 70-89). Michigan refiled mining (SIC's 10-14); construction (SIC's 15-17); manufacturing (SIC's 20-39); finance, insurance, and real estate (SIC's 60-67); and all units with 500 or more employees (all SIC's). Michigan conducted its pilot test this year and the results are not currently available.

Refiling Survey - Each of the five States conducted a full SIC refiling using the verification method for the industries specified above. The verification method of refiling SIC's provides a computer-generated, four-digit, SIC Manual-based industry description printed on a specially designed form. The description mailed to each establishment is based on the SIC code currently on file for that establishment. The form requests employers to verify the industry description as an accurate indicator of their primary economic activity. If the description is correct the employer checks the appropriate box, answers some additional questions on ownership and multi-establishment status, and returns the questionnaire. Forms on which the respondent identified the description as correct do not require recoding the industry classification. However, if the industry description is incorrect the respondent is asked to provide detailed product/activity information so that the correct industry can be coded. The mail results from the refiling survey provided the population/universe frame for the QM reinterview survey.

QM Survey - In each State the QM reinterview survey consisted of a probability sample of 500 establishments from the in-scope units for the refiling. The target population was grouped into the following three strata based on respondent answers to the survey: refiled units which responded yes, the industry description is correct (C-SIC); refiled units which responded no, the industry description is not correct (N-SIC); and the survey non-respondents (NR-SIC). A sample of 400 establishments was selected from the C-SIC, and a sample of 100 establishments from the NR-SIC. The samples were selected systematically. Samples were not drawn from the N-SIC units since these units were reviewed and recoded by the same method as in the classification refiling.

Sample units were contacted and interviewed by telephone. For the C-SIC units, the returned forms contained the name and telephone number of the company official who completed the form. For the NR-SIC units interviewers were provided with a list of resources to help them locate the telephone number. The interviewer collected the information required to determine and assign an SIC code.

The QM survey's estimates are percentages and ratios. The principal estimates are the percentage and number of establishments by response categories based on the respondent's verification of the SIC (industry description) and the telephone interviewer's determination of the SIC, and by nonrespondents with SIC incorrect, and nonrespondents with SIC correct. The estimates are intended to measure the response error, nonresponse error, and the percentage of units with incorrect SIC's remaining on the file. Sampling errors were calculated for all estimates. The sample allocation was designed to give standard errors of 2.5 percent at 2 sigma for all estimates of percent of C-SIC units and of 5 percent for estimates of percent of NR-SIC units. The larger sample size for the C-SIC units is consistent with the major emphasis of the pilot test - to study the effect of the verification form on the respondents and the processing of their forms.

VERIFICATION PILOT TEST RESULTS

Response Rates

The four QM States showed high response rates and a high percentage rate of responses in which the respondent verified the SIC as correct. BLS staff also compared the pilot states' response rates with those of other states' that had used the verification method. Most States with large increases (10 percent or more) had a previous response rate of 68 or less. The Maine and New Hampshire surveys and Missouri's manufacturing and FIRE surveys show a smaller increase. The systematically higher response rates obtained in the verification method enhance the quality of the SIC coding since under the classification method all non-respondents retained their previous SIC designation. See Table 4 for the relative SIC error from the verification and classification methods.

Respondent Burden

The respondent required, on average, considerably less time to complete the verification questionnaire than to complete the classification questionnaire. BLS estimates two minutes to complete the verification questionnaire if the respondent verifies the industry description as correct (C-SIC) and 10 minutes to complete the product/activity information on the questionnaire if the respondent feels the industry description is incorrect (N-SIC). Under the classification method each respondent is treated as an N-SIC case and thus also requires approximately

10 minutes to complete the product/activity information.[2]/

Table 1 shows the projected percent decrease in respondent burden for the verification method compared to the classification method based on the response rates for each state. The average projected decrease in response burden using the verification method is 66.7 percent.

State Resource Requirements

The reduction in State staff hours required to review and code SIC questionnaires mirrors the respondent burden reduction. Estimates from SESA staff of the reduction in required resources confirmed the projections made last year.[2]/

Using staff time estimates of two minutes to review a verification questionnaire on which the respondent has verified the SIC description as correct (C-SIC), and fifteen minutes to review those questionnaires for which the respondent did not verify the description and instead supplied the product/activity information (N-SIC), there is a considerable reduction in State resources required under the verification method. Table 1 shows the percent decrease for verification compared to the classification method. The average decrease in cost for the questionnaire review and SIC assignment to States using the verification method is 72.3 percent.

QM Reinterview Survey Results

The final activity was an assessment of the quality of the data obtained under the verification method; this was obtained through the QM reinterview survey. Table 2 shows results of the QM reinterview survey by response category. The results of the Quality Measurement Surveys are given in Table 3. Table 3 shows for each pilot State the estimates for the percentage of the original SIC's incorrect and the percentage of incorrect SIC's after verification by nonresponse and response categories. Michigan pilot tests results are excluded from this report. Preliminary results reveal some

Table 1. Projected Percent Respondent Burden and State Resource Reduction with Verification Method

State	% Reduction Respondent Burden[3]/	% Reduction State Resources[4]/
Maine	59.4	64.3
Oklahoma	55.9	60.6
South Carolina	76.1	82.4
Texas	68.4	74.1
Total	66.7	72.3

questionable data due to procedural error. Analysis of the data shown in Table 3 identified the following three interesting issues.

First: Is the error rate for nonrespondents higher than the percentage of SIC's incorrect in the original population? The error rate is not significantly higher for the nonrespondents than the original population. Nonrespondents do not have a higher percentage of incorrect SIC's than the entire population before refiling. What is the significance to the SIC refiling process? Nonrespondents to the verification refiling are not systematically avoiding refiling or reporting to the State or Federal Government because of incorrect SIC's. This result also means models using estimates for nonrespondents would not have significant bias.

Second: Is there an important difference in the SIC incorrect rate before and after verification refiling? The verification mail survey corrected the following percents of incorrect SIC's: Maine-69%; Oklahoma-32%; South Carolina-4%; and Texas-55%. All States except South Carolina showed an important difference. The greatest proportion of incorrect SIC's remaining after verification refiling is in the response error category except in Oklahoma. This is the area over which the Bureau can be most effective in reducing error. A review of nonsampling error and plans for improve-

Table 2. Results of QM SIC Reinterview Survey

Telephone Response	Maine		Oklahoma		South Corolina		Texas	
	C-SIC	NR	C-SIC	NR	C-SIC	NR	C-SIC	NR
No. in Sample...400	100		400	100	434	108	407	108
Correct SIC.....369	71		379	68	362	86	378	71
Incorrect SIC....17	1		20	12	17	3	10	6
Out of Business...4	17		1	20	10	4	6	15
Nonresponse.....10	11		0	0	45	15	13	16

Table 3. Number and Percent Incorrect SIC, Before and After Refiling

State	Original SIC Incorrect		Number Incorrect After Verif.	Percent Incorrect After Verif.		
	Number	Percent		Total	Non-Response	Response Error
Maine.....916		9.5	283	2.9	.2	2.7
Oklahoma...2158		10.3	1470	7.0	4.4	2.6
South Carolina....912		4.3	877	4.0	.6	3.4
Texas.....4139		6.5	1862	2.9	1.0	1.9
Average of the 4 States		7.7		4.2	1.6	2.6

ments appear in later sections of this paper under Current Improvements and Future Improvements.

Third: Is there a bias for respondents to incorrectly identify an industry description as correct? This issue was identified by BLS last year as a key point to determine the success of the verification method. The answer is no. Approximately three fourths of the respondents whose SIC was incorrect, correctly identified their status based on the SIC description provided. For the 64 units which were incorrectly classified as correct (see Table 2) 90 percent were incorrect only at the fourth digit. Improvements in design and classification are expected to reduce error for this detailed level of reporting.

Using an average of response and nonresponse errors calculated for each State, Table 3 shows an average decrease from 7.7 percent incorrect SIC's to 2.6 percent incorrect for the response population. Because of nonresponse the final average SIC error rate was 4.2 percent.

During the reinterview process respondents whose SIC changed were asked questions to determine the cause of their incorrect identification. In all 66 cases, respondents indicated that they did not understand or had misinterpreted the SIC description. There was no indication that respondents had "taken the easy way out" to avoid completing the form.

A related issue raised last year by BLS is - do the industry descriptions communicate to respondents accurately for most industries? Most (from 69 to 96 percent) of respondents that said the description did not correctly describe their activity were in fact correctly classified. All questionnaires where the respondent checked no and completed the information were reviewed in detail. States provided both excellent comments and recommendations for changes to descriptions and forms. SIC's covering a wide range of activities are more difficult to describe and more likely to have incorrect responses. Some SIC descriptions in the Services division yielded a particularly high error rate. However, the answer is yes, most descriptions do correctly describe industries; but, clearly, the industry descriptions can and will be evaluated and improved over time.

Comparison to Classification Method

Table 3 shows response error for the verification method. This response error represents a type of nonsampling error. A similar figure has not been measured for the classification method, but that method is more likely to be subject to at least two types of nonsampling error: (1) coder error because it requires more review, and (2) survey procedure error because it is more dependent upon survey procedures over a longer period of time than is the verification method.

Using nonresponse results from each State's previous refiling of the same industries, Table 4 shows a comparison of the effectiveness of the two methods (verification/classfication) for reducing incorrect SIC's.

The comparison is as follows:

Table 4.Relative SIC Error after Refiling Verification vs. Classification

State	Verif. NR Error	Class. NR Error
Maine	.2	.3
Oklahoma	4.4	5.3
South Carolina	.6	1.5
Texas	1.0	2.5

CURRENT IMPROVEMENTS

With the anticipated shift of more States to the verification method, BLS has begun a program to improve a wide range of SIC refiling activities aimed at reducing major sources of nonsampling error. This section will review some of the current improvements in this area.

Coder Error

BLS has continued to provide formal training in SIC coding to State agencies through a standardized training program that was fielded with considerable success in 1981. The objectives of this training are to improve new coders rapidly and to refresh the skills of experienced coders. BLS plans to modify, improve, and continue to deliver this course.

BLS developed specifications for, and plans to implement, a quality assurance program designed to monitor new coders during the learning period and to check periodically on experienced coders.

An additional benefit of the verification method, that was reported by the States was the absence of coder "burnout". With the classification method every questionnaire required review by a State coding technician. Since over 90 percent of units are correctly coded before the survey starts, review work was voluminous, repetitive and tedious. With the verification method, coders review only the questionnaires having an expectation of change (ie., N-SIC units). This type of review appears to result in a higher level of coder attentiveness and more incisive decision making.

Questionnaire Design

BLS uses a different SIC questionnaire for each industry division. BLS staff continually study the forms for possible revision. A network has been established among regional offices and States to solicit recommendations for change. Such recommendations are reviewed in Washington and implemented if approved. An evaluation of all comments and suggestions will be completed this year before requesting renewed approval from OMB.

SIC Description

Nonsampling error also stems from weakness in the SIC descriptions used in the verification process. BLS plans a regular annual update of the descriptions to include any improvements identified by States during the previous refiling process. In addition to improving individual descriptions, BLS also plans to use a new survey processing and control system to identify selected SIC's for different treatment. Certain industries are too broad in scope for numerous respondents to be able to recognize their activity among the many included in the industry; these industries will continue to be handled under the classification method.

Survey Procedures

Along with the implementation of the verification method will be an increasing use of monitoring techniques. The purpose of monitoring is to implement the most effective procedures and to have standard procedures to describe the program in key areas. With documentation collected from the pilot test States and supplemented where possible, BLS is developing descriptions of model procedures to recommend to the States. As previously stated, this points toward the need for further pilot studies to test new procedures.

Recent improvements to standardize procedures include:

- The SIC Refiling Status Report (a regular State level tracking system to monitor the industries refiled, processing cycle, response rates and future plans for refiling). This supersedes previous reporting requirements that provide only partial information.
- A requirement to use the OMB approved, BLS SIC refiling forms.
- A recommendation for identifying and mailing questionnaires to respondents representing the most detailed establishment level available.
- Development of methodology and procedures for reporting micro-level SIC changes to the national office on the annual code change report.
- An improved edit and review process for the annual code change report.
- Improved definitions and clarifications to the ES-202 Manual (the technical standard between BLS and States for SIC activities).

FUTURE IMPROVEMENTS: AUTOMATED SIC REFILING WITH A QUALITY MANAGEMENT SYSTEM

To enhance quality, uniformity and standardization, BLS has developed specifications for an SIC survey maintenance and update system with quality management components. Specifications were completed more than a year ago, and the development of the system has now begun under a contract with the Idaho SESA.

The software design will provide modules for the following functions:

a. creating the control file
b. setting parameters for sample selection
c. updating the control file
d. selecting and printing of SIC descriptions
e. processing carry-over units
f. editing files

g. selecting samples for quality control
h. selecting sample for quality measurement
i. generating a series of standard monitoring reports

The quality management system will include components for both Quality Control (QC) and Quality Measurement (QM). The purpose of the QC component is to monitor and control potential sources of nonsampling error through stepwise review and validation of all refiling activities. Reports from the automated system will support this validation process. This monitoring will allow States to identify errors quickly and correct them.

The QM component provides the States with the ability to measure sources of nonsampling error using telephone interviews of a probability sample of refiling units.

The QM measures are intended to provide the States with the ability:

- to assess the quality of their individual refiling operations.
- to identify areas requiring higher priority or need for quality improvements and,
- to compare and measure their refiling results/problems to those of other States.

In addition, the QM measures are intended to provide the BLS with the ability:

- to monitor the overall status and quality of the States refiling process,
- to develop a profile of specific problem areas requiring special procedures or priority and,
- to develop national measures of the quality of the SIC coding for UI file.

After Idaho develops and tests the survey management and update system under contract with BLS, it will be available to introduce to other States. When State and National resources permit a large scale implementation, BLS Washington will be able to implement on a National scale many of the quality goals outlined above.

ACKNOWLEDGMENTS

The authors are grateful to George Werking and Alan Tupek for technical advice and comments, to Tawanna Neal for typing preparation, and to Wendy Alvey for editorial comments.

NOTES AND REFERENCES

*Currently employed by Statistics of Income Division, Internal Revenue Service.

1/ See A Review of Industry Coding Systems: Statistical Policy Working Paper 11: prepared by the Industry Coding Working Group, Administrative Records Subcommittee, Federal Committee on Statistical Methodology for the Statistical Policy Office, Office of Information and Regulatory Affairs, Office of Management and Budget.

2/ Hostetter, Susan C. "The Verification Method as a Solution to an Industry Coding Problem" 1983 American Statistical Association Proceedings, Section on Survey Research Methods; pp.499-504.

3/ The respondent burden hours for the verification method were projected using 2 minutes of respondent burden for C-SIC units and 10 minutes of respondent burden for N-SIC units. The respondent burden hours for the classification method were simulated using 10 minutes of respondent burden for all responders and assuming the same response distribution actually achieved in the verification test within each state.

$$\text{Verification Hours} = \frac{1}{30} \text{ hr} \times \#\text{C-SIC} + \frac{1}{6} \text{ hr} \times \#\text{N-SIC}$$

$$\text{Classification Hours} = \frac{1}{6} \text{ hr} \times (\#\text{C-SIC} + \text{N-SIC})$$

4/ The resource requirements were projected using estimates from State agencies for time required to review and code verification questionnaire responses for C-SIC responses (2 minutes) and N-SIC responses (15 minutes). The projected State resource requirements under the classification method were based on estimates from States and assume 15 minutes to review each responder. These projections assume the same response distribution actually achieved in the verification test within each state.

$$\text{Verification Hours} = \frac{1}{30} \text{ hr} \times \#\text{C-SIC} + \frac{1}{4} \text{ hr} \times \#\text{N-SIC}$$

$$\text{Classification Hours} = \frac{1}{4} \text{ hr} \times (\#\text{C-SIC} + \#\text{N-SIC})$$

5/ The sampling error was calculated using $K \times \left(\frac{PQ}{n-1}\right)^{\frac{1}{2}}$. For the nonresponse estimates P is the proportion of "Original SIC Correct" units of the Y-SIC sample; Q is 1-P; n is the number of units in the Y-SIC sample; K is the proportion of Y-SIC units to the total number of respondents. For the nonresponse estimates, P is the proportion of "Original SIC Correct" units of the respondents sampled; Q is 1-P; n is the number of nonrespondents sampled.