THE VERIFICATION METHOD AS A SOLUTION TO AN INDUSTRY CODING PROBLEM

Susan C. Hostetter, Bureau of Labor Statistics

INTRODUCTION

The Bureau of Labor Statistics (BLS) and some State Employment Security Agencies (SESA's) have worked together for two years on a project which has great potential for reducing agency cost and employer reporting burden. The project is the development and implementation of the verification method of reclassifying business units with industry codes. This paper describes the importance of industry coding, one of the ongoing problems in industry coding, the verification method as a solution to this problem, testing the verification method, and projected gains if the method is a success.

Standard Industrial Classifications, the SIC codes that United States Government statistical agencies use to classify the primary activities of establishments or enterprises, are important. Why are they important? Why is it important that they be accurate? They, and their accuracy, are important because the great majority of economic policy decisions made by the Federal government, by State or local governments, and by the private sector are based on data derived from the assignment of industry codes to establishments or enterprises.

As an example, a psychiatric hospital incorrectly coded as other than a hospital would cause average hospital wages for that area to be miscalculated. Because psychiatric hospitals generally have more specialists and a higher average wage than other hospitals, omission of such data would probably reduce the payments that all hospitals in the area would receive from the Health Care Finance Administration for their Medicare patients.

BLS industrial coding is conducted cooperatively with SESA's through the Covered Employment and Wages (ES-202) Program. The ES-202 Program creates a quarterly report by four-digit SIC for county, State, and National levels. The data are used for administration of many programs and as input to economic data at local, State and National levels. One of many major uses is the wage and salary component of both the National personal income and the Gross National Product.

INDUSTRY CODING PROBLEMS

Correct identification and coding of establishment location and activity are essential to effective use of ES-202 and other establishment data. Statistical agencies have been trying to maintain, improve, and perfect coding systems for years, but each has found it to be an endless process.

The cost of monitoring industry codes for all establishments and reporting units is great not only in dollars and in burden for respondents, but also for BLS, in hours required for SESA staff to review and code information. Some of the major problems faced by agencies which conduct industry coding are the following.

1. Defining the economic unit to be assigned the industry code -- The goal is to assign the industry code to the establishment as it is defined in the Office of Management and Budget Standard Industrial Classification Manual (SIC Manual). BLS uses a "reporting unit" that may include more than one establishment in the same industry and the same county.

2. Defining the criteria for industry coding -- BLS follows the SIC Manual standards in that it assigns codes for each activity, designating the SIC based on the combination of products or services that produce the largest percent of total receipts or value of production under one SIC.

3. Using a source document that requests all information necessary for determining an industry code -- BLS uses a separate questionnaire for each industry division so that all questions necessary to determine a code in a given industry are asked.

4. Defining the procedure for coding -- BLS adheres to standard manual review coding and supports and monitors such a system for all States. The State-level coding, which is unique to BLS, provides excellent accuracy due to local or regional understanding of industries and of local terminology used to describe activities. SESA's systematically use telephone follow-up to obtain missing data or clarifications.

5. Maintaining coding staff who are adequately trained -- BLS has been able to retain experienced State staff in industry coding activities by maintaining an ongoing refiling cycle in each State. BLS recently initiated an SIC training program to assure use of standard methods and concepts throughout the States. This training package, the related workbook, and other materials assisted BLS staff to improve, to standardize, and to monitor State industry coding procedures. Because the refiling cycle includes an even workload each year, staff is maintained at a constant level with little turnover. It is not uncommon for State coders to remain for 20 years or more.

6. Maintaining an update process to assure current industry codes in the system — A major problem in industry coding is to maintain a regular updating process for an industry coding system. Without this updating, codes become obsolete and inaccurate. BLS has worked consistently with SESA's to maintain the regular three-year refiling cycle for industrial codes. This is the subject of this paper.

During the past two years reduced Federal

statistical budgets have rendered maintaining regular SIC refiling programs particularly difficult.

In early 1982 BLS staff felt growing concern that several States would disrupt this regular cycle. The staff assessed the magnitude of the anticipated reduction and estimated that seven percent of units change SIC's over a three year period. This reinforced concern that reductions in refiling would be harmful. BLS staff listed possible alternatives and, after rejecting several suggestions, accepted a plan to refile using a verification method which BLS believes is a breakthrough for the refiling problem.

THE VERIFICATION METHOD

What is the verification method? Simply stated, the industry verification method of refiling uses a computer-generated, four-digit industry description printed on a specially designed form, the Industry Verification Form (see Exhibit 1). The description generated is based on the SIC code 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 simply checks the appropriate box, answers some other standard refiling questions on ownership and multi-establishment status, and returns the questionnaire. Basically, the SESA staff need not recode industry classifications for forms with descriptions marked correct. If the industry description does not correctly describe the economic activity, the employer is asked to provide detailed product/activity information so that the correct industry can be coded.

IOWA TEST

Under agreement with BLS, the Iowa Department of Job Service (the Iowa SESA) tested the verification method of industry refiling in finance and retail trade, using its own forms and descriptions. Table 1 shows the final useable response rate as 87 percent and the percent of units that changed SIC to be 5 percent.

The Iowa SESA reported that respondents were favorably disposed toward the new method and the reduced reporting burden. The SESA attributed some of the increased response to respondents' desires to have the SESA or Department of Labor classify them in the correct industry. The SESA also reported that the verification method considerably reduced the State resources needed for review and coding of returned forms.

RESPONDENT BURDEN

Iowa's report agreed with BLS research which showed that respondent time required to complete the Industry Classification Statement, a full industry classification form, averaged ten minutes, while time required to complete the Industry Verification Form averaged two minutes. BLS estimated that about 12 percent of verification respondents would complete the

FIGURE A. RESPONSE RATES OF OLD AND NEW METHODS COMPARED.

OLD METHOD
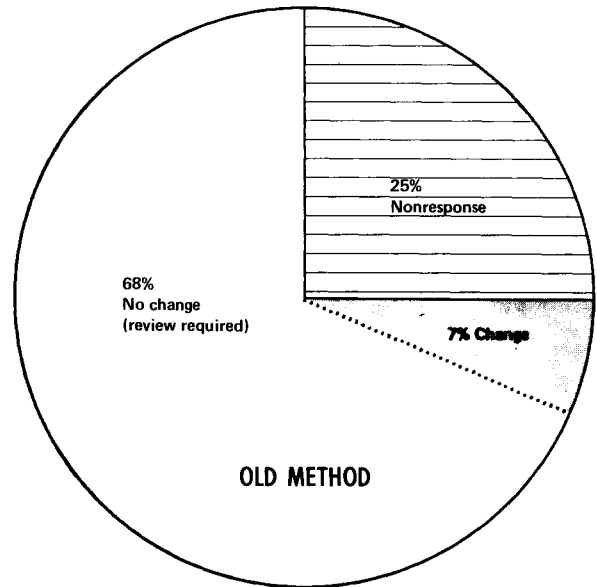
50,000 units mailed; 37,500 (75%) useable responses
all useable response required product/activity review and SIC assignment

NEW METHOD
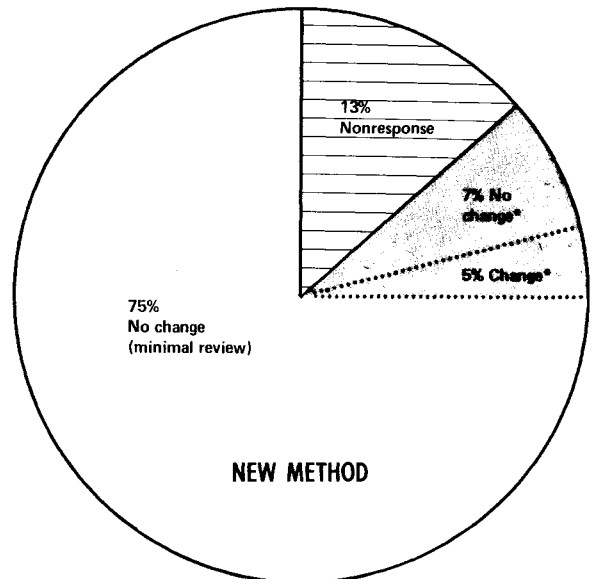
50,000 units mailed; 43,500 (87%) useable responses
6000 (14%) of useable response required product/activity review and SIC assignment

TABLE 1. RESULTS OF 1982 IOWA STATE VERIFICATION REFILING

SIC	Number Mailed	Useable Responses	Response Rate	Percent of Useable Responses Changed
Total...............9235	8033	87%	5%	
55 Automotive Dealers and Gas Stations..........3265	2864	88%	5%	
56 Apparel and Accessory Stores......1418	1024	72%	3%	
59 Miscellaneous Retail...............2946	2600	88%	6%	
60 Banking...............660	656	99%	2%	
61 Credit Agencies Other than Banks.......552	533	97%	5%	
62 Security and Commodity Banks........137	120	88%	21%	
63 Insurance.............257	236	92%	4%	

TABLE 2. ESTIMATED RESPONDENT BURDEN AND SESA STAFF HOURS REQUIRED FOR INDUSTRY CLASSIFICATION

Number of Forms/Hours	Industry Classification Statement (ICS)	Industry Verification Form (IVS)
Number of forms mailed........................1000		1000
Number of completed forms returned.............870*		870
Returned forms with all questions completed by respondent		
Respondent burden		
Number of forms completed...............870		104
Respondent burden hours..................145		17
SESA staff hours required		
Number of forms reviewed................870		104
SESA staff hours........................218		26
Returned forms with industry description checked as correct by respondent		
Respondent burden		
Number of forms completed...................................766		
Respondent burden hours.....................................26		
SESA staff hours required		
Number of forms received....................................766		
SESA staff hours..26		
Total respondent burden hours.................145		43
Total SESA staff hours required................218		52

*A respondent rate of 87 percent would not be projected for the BLS Industry Classification Statement, but was used in this Table to allow comparison of respondent burden and State staff hours required. There is some indication that use of the IVF leads to an increase in response rate of employers, thus reducing costs and improving the quality of industry codes.

entire form because they believed the description incorrect, but that the remaining 88 percent would be in the two-minute category.

Correspondingly, the workload for the SESA staff would be proportionately reduced. If a fully completed Industry Classification Form required an average of fifteen minutes to review and code, the verification form checked as being correct would require about two minutes to review and code.

Table 2 shows the estimated difference in respondent burden and SESA staff resources necessary for a mailing of 1,000 forms with an assumed response rate of 87 percent. Using the old classification form, respondent burden is estimated at 145 hours and SESA staff hours necessary for classification are estimated at 218. However, for the same number mailed and the same number returned, with the verification form the respondent burden is estimated at 43 hours and the SESA staff hours are estimated at 52.

BLS estimates of response rate and industry code changes from the Iowa test, and some preliminary results from a later study by the State of Texas, produced the data shown in Figure A.

For both old and new methods a mailing of 50,000 units is assumed. The upper figure is based on using the old classification form and shows a response rate of 75 percent, which is what BLS normally expects with this traditional method. Since all of the useable responses must be reviewed for product/activity information and SIC assignment, the results show that all forms completed by respondents required complete review by SESA staff. Only seven percent of these forms required a change in SIC.

The lower figure is based on the new verification form where the respondent checks "yes" if the description provided accurately describes his primary activity. For this form a response rate of 87 percent was assumed. Only 14 percent of useable responses had "no" checked and product/activity information provided. This 14 percent required complete review and SIC assignment. The remaining 86 percent of useable responses required only a brief review and check for county and ownership codes.

QUALITY MEASUREMENT SURVEY

At this point BLS was convinced of the considerable value of the verification method. It wanted to test the method in several States using standard descriptions and forms. It wanted to develop a bank of information based on standard materials and not to rely exclusively on the Iowa experience. Of course, with the information BLS had developed, there was no difficulty in finding interested States. Currently 15 States are using the verification method of refiling.

BLS looked for flaws and weaknesses in the verification system. Two were identified.

One is, to avoid the work of completing the questionnaire, a respondent may be tempted to check "yes," even when the description does not accurately describe his primary activity. The other is that some industry descriptions may not provide adequate information for the respondent to correctly check "yes" or "no."

BLS staff developed a quality measurement survey (QMS) to test several aspects of the verification refiling. BLS currently has contracts with Maine, South Carolina, Texas, Oklahoma, and Michigan to conduct the QMS. The States in the QMS will use BLS-prepared industry descriptions and Industry Verification Forms approved by the Office of Management and Budget. After the refiling is completed the State will sort units as shown in the following list:

1. Total units mailed
2. Out-of-business and out-of-scope units
3. No change units -- industry description checked "yes"
4. Units indicating change -- respondent completed product/activity information
   a. those with correct SIC
   b. those with incorrect SIC
5. Nonresponse

Units in groups 3 and 5 are candidates for this quality measurement survey, and are to be sampled separately. Each State, using a simple systematic selection, will select 400 units from group 3 and 100 units from group 5.

Each of the 500 respondent units will be telephoned and asked for complete product/activity information. Trained coders will evaluate the product/activity information to determine the appropriate SIC code. For group 3 units (most of the units in the QMS), a respondent who provides information that differs from that supplied on the questionnaire (e.g. his SIC is incorrect and he checked "yes," that it was correct) will be asked about the industry description using questions designed to determine what caused the incorrect answer.

Group 4 units are not included in the QMS because each of these respondents will have completed the product/activity statement which will be reviewed and coded by trained SESA staff. For purposes of this study it is assumed that the codes for this group of units will be correct. QMS States will provide BLS with mail survey tabulations by four-digit SIC for each of the five groups. Based on the telephone survey they will submit separate listings for the "no change" units and the nonresponse units. Each of the latter listings will show:

1. Unit number
2. SIC when mailed
3. SIC verified by telephone
4. Comments

Comments may indicate whether the description, or a bias toward checking "yes" to avoid completing the form, led to an incorrect response.

Figure B. CATEGORIES OF RESPONSE TO SIC REFILING

Response Status	Original SIC Correct	Original SIC Incorrect
Response		
Correct Response	A. Respondent Indicated No Change	B. Respondent Indicated Change
Incorrect Response	C. Respondent Indicated Change	D. Respondent Indicated No Change
Nonresponse	E. No Response	F. No Response

From the mail survey tabulations and the telephone survey listings BLS staff will develop the following estimates which are based on categories shown in Figure B , above.

1. What percent of "no change" units have an incorrect SIC? $\dfrac{D}{D+A}$

2. What percent of "incorrect SIC" units indicated no change or were nonrespondents?

   For <u>respondents</u> with incorrect SIC $\dfrac{D}{B+D}$

   For <u>all units</u> with incorrect SIC $\dfrac{D}{B+D+F}$

3. What percent of "incorrect SIC" units did not respond? $\dfrac{F}{B+D+F}$

4. What percent of "incorrect SIC" units did the verification method fail to correct? $\dfrac{D+F}{B+D+F}$

Preliminary results from Texas, the first State to provide results, are shown in Table 3. They cover all units in agriculture, transportation, and services for Texas.

Inital estimates from QMS tabulations indicate that less than 3 percent of units identified as "no change" had an incorrect SIC. Of 407 units selected for the first telephone QMS of "no change" units 97 percent were correctly identified.

Table 3. TEXAS REFILING TABULATIONS

Respondent Status	Number	Percent of Mailing
Total Mailed.........66,721		100
Useable Response.....56,092		84
Checked Correct....47,865		72
Checked Incorrect...8,227		12
SIC Changed.........2,311		4
Out-of-business.........896		1
Nonresponse..........9,733		15

In addition to data from the five States under contract, BLS has asked the other 10 States currently using the verification method to provide information on industry descriptions that may need improvement to obtain correct respondent understanding.

Included in the data from Texas are tabulations of errors that have potential for correction. One type of respondent error is caused by a description that could be improved to help the respondent correctly identify his activity. Another potential error is using the verification method for certain SIC's that are too broad to allow the respondent to easily recognize his activity. The BLS system would allow units in problem SIC's to be mailed the standard form rather than the verification form.

If the BLS study shows there is little or no bias of respondents to incorrectly identify an industry description as correct; and if industry descriptions can be further developed to communicate to respondents accurately for most industries, BLS will have found a viable solution to a major industry coding problem. This solution could save 131,000 respondent hours per year and could save 135 man years for SESA's. BLS will have a practical, expedient, and cost saving method to maintain the industry refiling process and insure accurate industrial codes.

503

U.S. Department of Labor Exhibit 1

The information collected on this form by the Bureau of Labor Statistics and the State agencies cooperating in its statistical programs will be held in confidence and will be used for statistical purposes only.	*This report is authorized by law 29 U.S.C. 2. Your voluntary cooperation is needed to make the results of this survey comprehensive, accurate, and timely.*	Form Approved O.M.B. No. 1220-0032 Approval Expires 12/31/85

OFFICE USE ONLY

1972 SIC	Ownership	County Code	Aux.

TO AVOID DUPLICATE MAILINGS, please return as soon as possible in the enclosed envelope, to the address below.

The Big D Telephone Company
1234 Alamo Drive
Dallas, Texas   78704

Economic Research and Analysis
Texas Employment Commission
117 Trinity Street
Austin, Texas   78701

Telephone Number   512-397-4555

This report will be used to ensure the proper industrial classification of your regular **QUARTERLY EMPLOYER'S TAX REPORT** and should cover the entire activity of the same establishment for the most recent calendar year.

A.   The Standard Industrial Classification typed below describes, in general terms, the **principal** activity, product(s) or service(s) produced by your establishment. Please read the description below and check the appropriate box to the right.

Furnishing telephone communication service by either wire or radio. Includes domestic, international, marine, mobile, and aeronautical services.

PLEASE CHECK THE APPROPRIATE BOX BELOW

1. ☑ Correct—Skip to question C.

2. ☐ Incorrect—Complete all questions below.

B. Principal Products or Activities during most recent calendar year in this establishment(s)	Percent of total sales (value or receipts) this period	For State use
1. Transportation (specify below)		
2. Communications (specify below)		
3. Public Utilities (specify below)		
4. Other Products or Activities (specify below)		

C. Type of ownership (check one)

Private

1. ☑ Corporate
2. ☐ Noncorporate

Government

3. ☐ Federal
4. ☐ State

5. ☐ Local
6. ☐ International

D. Is the establishment primarily engaged in performing services for other units of the company?   ☐ Yes   ☑ No

If yes, indicate nature of activity of this establishment.

1. ☐ Central administrative office
2. ☐ Research, development, or testing
3. ☐ Storage (warehouse)
4. ☐ Other (specify)

_____

_____

E. Is this establishment part of a multi-unit company?

1. ☑ Yes   2. ☐ No   If no, skip to question G on the back of this form.

If yes, enter name and location of controlling company on the space provided below and complete the questions on the back of this form.

The Big D Telephone Co.

_____

_____

BLS 3023-B6 (Rev. Dec. 1982)