

EVALUATION OF COVERAGE OF THE ADMINISTRATIVE RECORDS FRAME
FOR THE 1977 ECONOMIC CENSUSES - EMPLOYER SEGMENT

Paul S. Hanczaryk and John M. Sullivan
U.S. Bureau of the Census

I. Introduction

In assembling the mailing and control lists for the economic censuses, the Bureau of the Census relies on a variety of administrative record information from the Internal Revenue Service and Social Security Administration. Through their use, along with regularly maintained Bureau lists of multiestablishment firms, the population of active businesses to be included in the economic censuses is defined.

The primary purpose of this evaluation was to assure that the group of establishments classified as out-of-scope, and therefore not given a chance for coverage in the censuses, were indeed out-of-scope. Results from the evaluation show the reliability of administrative records in determining census coverage. This paper presents the results of that evaluation.

II. Overview

The out-of-scope universe is composed of eleven distinct exclusion categories, each of which was tabulated separately to measure the composition of the universe. The following list describes these exclusion categories:

Exclusion Category A - Government organizations that elect social security coverage.

Exclusion Category B - Establishments located in American Samoa.

Exclusion Category C - Establishments involved in international operations, for example, American citizens employed in foreign countries.

Exclusion Category D - Establishments which are foreign subsidiaries of domestic corporations.

Exclusion Category E - Entities of the Federal Government not covered by social security.

Exclusion Category F - Entities of a state or local government not covered by social security.

Exclusion Category G - Establishments which are located in foreign areas.

Exclusion Category H - Activities that are out-of-scope of the economic censuses, but classified within inscope industry codes. These are identified by type-of-operation (T/O) codes, specifically T/O 55 and T/O 90. The major use of T/O 55 is to exclude from census coverage those establishments in the transportation industry that are regulated by the Interstate Commerce Commission (ICC). T/O 90 represents government-owned and operated establishments.

Exclusion Category I - Establishments that have no name and no address on the record.

Exclusion Category J - Establishments in Puerto Rico, Virgin Islands, or Guam with industrial classifications that are out-of-scope of the economic censuses for those areas.

Exclusion Category K - Establishments with industrial classifications that are out-of-scope of the economic censuses. Certain industries are not included in census coverage. These include:

Agricultural production and services
Forestry, fishing, and trapping
Certain industries in transportation
Public utilities
Finance, insurance, and real estate
Religious organizations
Private households
Public administration

As shown in the following table, the total out-of-scope employer universe consisted of 557,884 establishments. Over 75% of these establishments originated from category K, cases with an out-of-scope SIC. As such, the evaluation naturally concentrated on this area.

Exclusion Type	Total Employer Universe	
	Establishments	% of O/S Universe
A - Government	55,188	9.89
B - American Samoa	154	0.02
C - Int. Operations	1,008	0.18
D - Foreign Subsidiary	281	0.05
E - Federal Government	1,235	0.22
F - State, Local Gvt.	64,999	11.65
G - Foreign Areas	56	0.01
H - T/O 55 and 90	8,865	1.58
I - No name/address	---	---
J - P.R., V.I., Guam	5,932	1.06
K - Out-of-scope SIC	420,166	75.31
TOTAL	557,884	

III. Initial Processing

The objectives of the evaluation were to document any existing misclassification in the out-of-scope universe, to show the corresponding census undercoverage rate and to identify areas of concentrated misclassification. In order to meet these objectives, each sample case was examined to assure that it was properly excluded from census coverage. For the most part, this was accomplished by a mailout of the NC-X4 form, which is the 1977 Economic Censuses General Schedule. The NC-X4 form is used principally to determine the industrial classification of an establishment; in some instances it was not adequate to verify all of the exclusion categories. For example, there was no question on the physical location of the establishment, so businesses excluded because of location could not be verified by the NC-X4 form. Because of this, and also because certain groups did not require mailing, the sample was reviewed prior to mailout to exclude two types of cases.

The first type of case involved obvious government entities. Certain establishments were exempted from mailing because they were clearly government affiliated. Included were such cases as the U.S. Department of Commerce or the State of Vermont which, in addition to the government name, had specific government indicators within their record. These establishments were considered verified and accepted as out-of-scope. Because of the strict criteria used to exclude a government case from mailing, it was felt that the evaluation results would not be compromised.

The second type of case excluded involved establishments with foreign addresses. Certain establishments are considered out-of-scope based on their physical location. Since the NC-X4 does not have a physical location question, it would not have been productive to mail these cases. Special handling was required for these cases, either telephone contact or a special mailing which verified the physical location.

Because of these two exclusions from mailing, it was possible to limit the mail panel to 5,505 while keeping a larger sample for overall evaluation purposes. This allowed a more optimal allocation of the mailed cases, i.e., a greater number of cases from category K were included in the NC-X4 mail panel. Category K represents establishments with out-of-scope SIC's, and because of the subjective nature of SIC coding it was certain that misclassification would occur within this category.

The sample for this study was selected from the out-of-scope employer cases on the Master Sample for the 1977 Economic Censuses. The Master Sample is a 1/50 sample of all legal entities, except for some large establishments which were selected with certainty. A subsample of the Master Sample was drawn 1/, and from these sample cases the misclassification rate and other characteristics were calculated for the population.

IV. Tabulations

The tabular presentation for the study focused on misclassification within the out-of-scope employer universe, the corresponding census undercoverage rate, and the identification of concentrated areas of misclassification. In that virtually all misclassification was derived from cases with an out-of-scope industrial classification, the tables emphasized this area. The results were separated into the following three sections.

- o Summary of Out-of-Scope Misclassification - The misclassification for each exclusion category as well as for the total O/S universe was shown.
- o Characteristics of Inscope Establishments - The derivation of misclassification for a given inscope division was outlined. In other words, for each inscope division, the initial exclusion category and SIC of the misclassified establishments were listed.

Also, this section included tabulations highlighting concentrated areas of misclassification.

- o Comparison of Resulting Inscope Activity to Published Totals - The resulting inscope activity was compared to 1977 census totals to show the effect of undercoverage.

Summary of Out-of-Scope Misclassification

The summary of out-of-scope misclassification showed the resulting inscope activity for each exclusion category and the total, and the corresponding misclassification rate (resulting inscope expressed as a percent of original O/S). As can be seen in Table 1, misclassification occurred in only three of the eleven exclusion categories; international operations, out-of-scope T/O codes, and establishments with out-of-scope SIC's. The overall misclassification in the out-of-scope universe was 17,203 establishments, 124,169 employees and over one billion dollars in annual payroll. As previously mentioned, the majority of this misclassification came from category K, with 15,661 establishments.

The total misclassification rate for the out-of-scope universe was 3.1% for establishments, and only 0.4% for employment and 0.3% for annual payroll. The reason that the percentage was relatively low for employment and annual payroll as compared to establishments was that no government establishments were misclassified, and although they represent less than a quarter of the establishments, these establishments make up the majority of the out-of-scope employment and annual payroll.

Misclassification in category C, international operations, was due to foreign entities which operated in the United States. Since this comprised only 79 establishments, an analysis of this category is not included here.

Misclassification within category H was quite high, with 16.5% of the establishments misclassified. Category H represents establishments with out-of-scope T/O exclusions, and includes T/O 55 and T/O 90. The resulting inscope cases that began in category H came exclusively from T/O 55 and not T/O 90. T/O 55 is used to indicate ICC regulation of establishments in the transportation industry, but many establishments not regulated by the ICC were wrongly assigned T/O 55, and should have been included in the economic censuses. On closer examination of these records it was found that misclassification was limited to those cases that were assigned T/O 55 during the 1971 and 1972 processing.

Within category K, the misclassification rate was 3.7% for establishments, 3.2% for employees, and 2.8% for annual payroll. Analysis of category K will be covered in a later section.

Characteristics of Inscope Establishments

Table 2 outlined the resulting inscope activity from category K, and was intended to show the

source of misclassification for a given inscope division. As can be seen in the table, the inscope divisions are listed in the columns on the top of the table, and the original 2-digit SIC's are included in the rows on the left. Using this table format, misclassification for a given inscope division can be traced to the original O/S 2-digit SIC. The complete evaluation report details all O/S 2-digit SIC's, but for the following table SIC 65 is listed as an illustration.

Overall misclassification in category K was 15,661 establishments, 108,618 employees, and over \$866 million in annual payroll. The majority of this misclassification originated from SIC 65, which represents Real Estate establishments. Notice in the table that of the total number of 3,608 establishments that resulted in inscope Construction, 2,854 originated from SIC 65. This was due in part to the interrelationship of the construction and real estate industries. Often, such establishments were comprised of distinct but related activities, i.e., construction and real estate. The SIC system provides for treating each activity as a separate establishment if: no one SIC covers the combined activities; the employment in each activity is significant; and if separate reports can be prepared. These conditions were not in evidence and, as is the practice, such cases were coded to the industry with the highest percentage of total receipts. For example, a case with 55% of receipts attributable to construction and 45% attributable to real estate sales was assigned an inscope construction classification. At the time the SIC code was originally assigned, the percentage distribution of receipts could easily have been different and probably was.

Special tables outlining specific areas of misclassification were also included in this section. Each establishment has associated with it various control file information, including a Principle Industrial Activity (PIA) code, which is the industrial classification code from IRS. It was determined that for those O/S cases which had an inscope PIA code, the misclassification rate was quite high. Within category K 3.7% of the establishments were misclassified. If those establishments with an inscope PIA code were isolated, the misclassification jumped to 20.2%. This data confirms a correlation between misclassified establishments and establishments with an inscope PIA code.

Comparison of Resulting Inscope Activity to Published Totals

This section of the tabulations compared the resulting inscope activity to 1977 census totals to show the effect of undercoverage. As Table 3 indicates, the overall undercoverage rate for census data was 0.5% for establishments, 0.2% for employees and 0.2% for annual payroll.

The three inscope divisions which contained the highest undercoverage rates were Construction, Transportation, and Services. The undercoverage rate in Construction was 0.8% for establishments,

0.7% for employees, and 0.5% for annual payroll, and was mainly the result of SIC 15, building construction, and SIC 6552, subdividers and developers. The undercoverage rate in Transportation was 1.2% for establishments, 2.8% for employees, and 3.7% for annual payroll. As stated previously, this was largely due to the assignment of T/O 55 to many nonregulated companies that should have been included in the census. The undercoverage rate in Service was 0.7% for establishments, 0.3% for employees, and 0.3% for annual payroll, with no particular concentration in any SIC.

V. Summary

In summary, the primary objectives of the evaluation report were to assess the misclassification within the out-of-scope employer universe, to show the corresponding census undercoverage, and to identify areas of concentrated misclassification. As noted earlier, the overall misclassification was 17,203 establishments. The majority of this misclassification originated from category K, establishments with industrial classifications which are out-of-scope of the economic censuses. Of the total misclassification, 15,661 were from category K.

Throughout the evaluation ^{2/}, an attempt was made to identify specific areas of misclassification, in order that they might be corrected in future economic censuses. Table 4 summarizes some significant areas of misclassification and their effect on total misclassification.

Two areas that produced considerable misclassification were establishments with a T/O code of 55 and a year code = 1 or 2; and cases with an inscope PIA code. If both these areas were refiled, which would include 30,326 establishments, the overall misclassification rate theoretically would drop from 3.08% to 1.93% for establishments. In addition, the undercoverage rate for the inscope divisions of Transportation Retail and Services would drop considerably, and the overall census undercoverage would drop from .5% to .3% for establishments.

Footnotes

^{1/} The entire sample design is included in the complete evaluation report, which is available on request from the Program Development Branch, Economic Surveys Division, Bureau of the Census.

^{2/} The complete evaluation report, including a comprehensive set of tabulations, is available on request from the Program Development Branch, Economic Surveys Division, Bureau of the Census.

Table 1. Summary of O/S Misclassification

Exclusion Category	Original Out-of-Scope		Resulting Inscope		Misclassification Rate	
	# or Value	Standard	# or Value	Standard	%	Standard
		Error		Error		Error - %
<u>Total O/S Universe</u>						
Establishment	557,884	4,114	17,203	1,751	3.1	.3
Employees	29,639,907	3,809,898	124,169	23,774	0.4	.1
Annual Payroll (000)	317,735,288	38,687,587	1,033,462	177,556	0.3	.1
<u>Cat. A (Government)</u>						
Establishments	55,188	1,484	---	---	---	---
Employees	8,586,100	257,583	---	---	---	---
Annual Payroll (000)	66,977,637	2,009,329	---	---	---	---
<u>Cat. B (American Samoa)</u>						
Establishments	154	148	---	---	---	---
Employees	4,501	2,524	---	---	---	---
Annual Payroll (000)	19,158	3,705	---	---	---	---
<u>Cat. C (Int. Operations)</u>						
Establishments	1,008	271	79	78	7.9	7.4
Employees	68,249	6,560	1,662	1,646	2.4	2.4
Annual Payroll (000)	574,473	49,252	25,714	25,455	4.5	4.2
<u>Cat. D (Foreign Subsidiary)</u>						
Establishments	281	124	---	---	---	---
Employees	7,909	1,681	---	---	---	---
Annual Payroll (000)	237,596	35,237	---	---	---	---
<u>Cat. E (Federal Government)</u>						
Establishments	1,235	178	---	---	---	---
Employees	4,122,663	11,248	---	---	---	---
Annual Payroll (000)	50,684,624	198,696	---	---	---	---
<u>Cat. F (State, Local Gov.)</u>						
Establishments	64,999	842	---	---	---	---
Employees	13,050,498	3,807,537	---	---	---	---
Annual Payroll (000)	163,806,604	38,671,272	---	---	---	---
<u>Cat. G (Foreign Areas)</u>						
Establishments	56	49	---	---	---	---
Employees	2,072	247	---	---	---	---
Annual Payroll (000)	9,751	819	---	---	---	---
<u>Cat. H (T/O 55, 90)</u>						
Establishments	8,865	701	1,463	294	16.5	3.0
Employees	272,656	24,403	13,889	6,371	5.1	2.3
Annual Payroll (000)	3,903,233	397,694	141,055	73,463	3.6	1.9
<u>Cat. J (P.R., V.I., Guam)</u>						
Establishments	5,932	1,110	---	---	---	---
Employees	112,168	38,998	---	---	---	---
Annual Payroll	930,146	384,521	---	---	---	---
<u>Cat. K (O/S SIC)</u>						
Establishments	420,166	3,788	15,661	1,724	3.7	.4
Employees	3,412,099	94,744	108,618	22,845	3.2	.7
Annual Payroll (000)	30,592,139	807,806	866,693	159,628	2.8	.5

Table 2. Original O/S and Resulting Inscope ^{1/}

Category K, O/S SIC - Major Industry Group	Original O/S	Standard Error	R e s u l t i n g I n s c o p e								
			Total Inscope	Standard Error	Mining	Constr.	Manuf.	Trans.	Whlse.	Retail	Service
.											
.											
.											
<u>65 - REAL ESTATE</u>											
Establishments	162,931	3,134	8,009	1,525	---	2,854	1	---	---	1,863	3,291
% of Original			4.9	0.9							
Employees	677,652	69,358	63,899	21,096	---	12,185	350	---	---	24,215	27,149
% of Original			9.4	3.0							
Annual Payroll (000)	5,047,336	404,545	496,605	143,696	---	173,321	3,752	---	---	119,120	200,411
% of Original			9.8	2.7							
.											
.											
.											
<u>TOTAL (O/S SIC)</u>											
Establishments	420,166	3,788	15,661	1,724	240	3,608	527	120	180	3,167	7,819
% of Original			3.7	0.4							
Employees	3,412,099	94,744	108,618	22,845	540	22,244	4,435	660	900	34,634	45,206
% of Original			3.2	0.7							
Annual Payroll (000)	30,592,137	807,806	866,693	159,628	7,955	229,574	42,626	3,625	9,016	194,934	378,963
% of Original			2.8	0.5							

^{1/} Standard errors for the detailed cells are available from the Program Development Branch, Economic Surveys Division, Bureau of the Census.

Table 3. Census Undercoverage

Inscope Division	1977 Census Totals ^{1/}	Changes Resulting from Evaluation (Inscope Activity)	Standard Error	Undercoverage Rate - %	Standard Error - %
Total					
Establishments	3,813,758	17,203	1,751	0.5	---
Employment	56,066,036	124,169	23,774	0.2	---
Annual Payroll (000)	605,588,663	1,033,462	177,556	0.2	---
Mining					
Establishments	29,606	240	119	0.8	0.4
Employment	779,200	540	308	0.1	---
Annual Payroll (000)	12,820,500	7,955	5,843	0.1	---
Construction					
Establishments	480,014	3,687	971	0.8	0.2
Employment	4,272,659	23,906	8,728	0.6	0.2
Annual Payroll (000)	54,980,063	255,287	91,334	0.5	0.2
Manufacturing					
Establishments	358,502	527	219	0.1	0.1
Employment	19,505,800	4,435	2,009	---	---
Annual Payroll (000)	263,020,600	42,626	19,091	---	---
Transportation					
Establishments	69,608	851	225	1.2	0.3
Employment	424,618	11,746	6,301	2.8	1.6
Annual Payroll (000)	3,394,872	124,221	72,966	3.7	2.3
Wholesale					
Establishments	382,837	241	119	0.1	---
Employment	4,397,089	1,204	647	---	---
Annual Payroll (000)	58,289,573	13,014	7,274	---	---
Retail					
Establishments	1,303,621	3,289	820	0.3	0.1
Employment	13,040,082	34,756	17,953	0.3	0.1
Annual Payroll (000)	85,853,987	195,532	89,007	0.2	0.1
Services					
Establishments	1,189,570	8,368	1,181	0.7	0.1
Employment	13,646,588	47,581	11,175	0.3	0.1
Annual Payroll (000)	127,229,068	394,826	98,101	0.3	0.1

^{1/} Data are 1977 Census preliminary statistics for employers only (data figures for Transportation represent 1972 Census of Transportation data, and data figures for Services represent 1977 County Business Patterns data).

Table 4. Summary

Areas of Concentrated Misclassification	Impact if Refiled						Change in Overall Misclassification Rate	Std. Error	Impact on Undercoverage
	Required Mailing	Std. Error	Expected Corrections	Std. Error	Percent Corrected	Std. Error			
Category H, T/O Exclusion									
T/O 55 assigned to nonregulated companies year code = 1 or 2	5,748	701	1,463	2,941	25.5	3.6	3.08 to 2.82	.05	Transportation 1.2 to 0.2 for estabs.
Category K, O/S SIC									
Inscope PIA	24,578	2,087	4,968	978	20.2	3.6	3.08 to 2.19	.17	Retail: 0.3 to 0.1 for estabs. Service: 0.7 to 0.5 for estabs.
TOTAL	30,326	2,202	6,431	1,021	21.2	3.6	3.08 to 1.93	.18	Total: 0.5 to 0.3 for estabs.