TABULATION OF MULTIPLE INDUSTRY COMPANIES IN THE ANNUAL CAPITAL EXPENDITURES SURVEY

Miriam D. Rosenthal, Mark S. Sands, U.S. Bureau of the Census
Miriam D. Rosenthal, AGFS, U.S. Bureau of the Census, Washington, D.C. 20233[1]

Key Words: Multiple industry companies, multiple industry classification, Lavallée-Hidiroglou method

Abstract

In most sample designs for company surveys, a company is classified into a single industry category according to certain industry classification and size constraints. For the Annual Capital Expenditures Survey (ACES), companies are classified into a single industry category for stratification and sample selection, but may be tabulated in multiple industry categories based on response. First, we examine the effect on the ACES industry estimates of tabulating companies in multiple industry categories. Subsequently, we evaluate the stratification methodology, a modification of the Lavallée-Hidiroglou method, for meeting reliability criteria.

I. Survey Description

The Annual Capital Expenditures Survey (ACES) is a new survey for the Census Bureau. The goal of the survey is to provide annual estimates of business spending for structures and equipment by nearly all nonfarm domestic companies with five or more employees. Exclusions from the survey include government-owned operations, foreign-owned operations of domestic companies, establishments located in United States territories, establishments engaged in agricultural production, and private households. A pilot survey was conducted to collect 1991 data to test the survey instrument and determine the collectability of the requested data (Champion, 1993). A preliminary survey was conducted to collect 1992 data (Champion, Funk, and Berry, 1994). The first full-scale ACES was conducted to collect 1993 data.

The ACES is on a five year cycle (see Table 1). The basic survey is conducted every year, and collects data on spending for new and used structures and equipment for each of 94 ACES defined business activities within a company. In addition, in the first, third, and fifth years of the cycle, capital expenditures data are collected for companies with fewer than five employees; in year two of the cycle, questions are added to the basic questionnaire to collect detailed data on expenditures for structures; and in year four of the cycle, questions are added to the basic questionnaire to collect detailed data on expenditures for equipment.

Table 1: ACES Five Year Cycle

	Basic Survey 5+ Employees, Structures, Equipment	Supplemental Survey < 5 Employees, Structures, Equipment	Additional Questions on Structures	Additional Questions on Equipment
Year 1	X	X	--	--
Year 2	X	--	X	--
Year 3	X	X	--	--
Year 4	X	--	--	X
Year 5	X	X	--	--

X = Data collected
-- = Data not collected

The research presented in this paper deals with the basic ACES (Table 1, column 1). Any descriptions of the current methodology or proposed methodologies will pertain only to that part of the survey. All sample sizes and estimates presented pertain to the 1993 ACES unless otherwise noted.

The Census Bureau's establishment-based database, the Standard Statistical Establishment List (SSEL), is used as the base for developing the ACES sampling frame. A company can be comprised of either one single establishment (a single unit), or two or more establishments (a multiunit). Since the ACES is a company-based survey, the establishment level records on the SSEL must be collapsed into company level records to form the frame. The collapsing procedure maintains payroll and employment data at the four-digit Standard Industrial Classification (SIC) code level. Then, when the collapsing is complete, the company is assigned first to an Industry Division (i.e., manufacturing, construction, etc.), then to a Major Group (i.e., two-digit SIC), then to a three-digit SIC, and finally to an ACES industry code based on that three-digit SIC. The 1993 ACES universe contained approximately 2 million companies.

The ACES frame is divided into four strata for sampling purposes. Stratum 1 contains all companies with 500 or more employees. All companies in this stratum are selected in the sample with certainty. In the 1993 ACES, there were 14,295 companies in stratum 1.

[1]This paper reports the general results of research undertaken by Census Bureau staff. The views expressed are attributed to the authors and do not necessarily reflect those of the Census Bureau.

The remaining companies, those with from 5 to 499 employees, are grouped by ACES industry code. Then within each ACES industry, the companies are placed into one of three noncertainty strata based on payroll. The stratification methodology used resulted in minimizing the sample size subject to a desired level of reliability for each industry (a coefficient of variation, or *cv*, ranging from 1 percent to 3 percent). Since capital expenditures data were not available in the sampling frame, the target reliability levels were based on payroll. Approximately 15,200 companies out of nearly 2 million were selected in the noncertainty strata sample.

II. Multiple Classification Categories

For the ACES, companies are classified into one ACES industry for stratification and sampling, but may be tabulated in multiple industries based on response. This methodology presents two difficulties:

1. The stratification methodology used to assign a company to one industry for sampling purposes may not provide the best placement of the company for estimation purposes. This is partially a "measure of size" question and is partially a classification question. For example, a company may have activity in two different industries. Suppose the distribution of payroll is as follows:

Industry	Company Payroll	Total Industry Payroll	Company Contribution to Industry Total (%)
220	30,000	300,000	10%
310	10,000	50,000	20%

This company would be classified in the industry with the most payroll, ACES industry 220, for stratification and sampling. However, this company accounts for more of the industry 310 payroll (20 percent), while accounting for only 10 percent of the payroll in its sample industry of 220.

2. Since the desired reliability levels are based on counting the entire company in the sampled (assigned) industry, and since the company may respond and be tabulated in multiple industries, the resulting *cvs* will, most likely, have little relation to the design *cvs*.[2]

To partially deal with these two problems, for each company in stratum 1, the payroll was maintained in all industries for which that company had activity. Then, the payroll was accumulated within industry across all stratum 1 companies. When the allocation and sample sizes were determined for the noncertainty strata, this accumulated payroll across industries was used as the stratum 1 contribution to the allocation algorithm. So, in effect, the stratum 1 companies were classified in multiple industries for stratification purposes.

Before modifying the survey methodology further, we identified the companies in the noncertainty strata with multiple industry activity and examined these companies' contributions to the industry totals.

A. Frame Characteristics

The 1993 ACES frame was split into several files based on each company's assigned industry and size. Summary counts are presented in Table 2.

[2]Adding to these two problems is the fact that the reliability levels used for sampling are based on payroll, whereas the data of interest are expenditures. Unfortunately, no expenditures data (from either a prior survey or some other source), and no other data closely related to expenditures, such as assets, were available at the time the 1993 ACES sample was selected.

Table 2: 1993 ACES Frame Counts by Stratum and Amount of Industry Activity

	5 to 499 Employees (Noncertainty, Strata 2, 3, & 4)	500 or more Employees (Certainty Stratum 1)	Total
Single Units[a]	1,774,487 (89.4%)	2,996 (0.2%)	1,777,483 (89.6%)
Multiunits	195,503 (9.9%)	11,299 (0.6%)	206,802 (10.4%)
one industry[b]	147,800	2,893	150,693
multi- industry[b]	47,703	8,406	56,109
Total	1,969,990 (99.3%)	14,295 (0.7%)	1,984,285 (100%)

[a] By definition, single units are one industry companies.
[b] Figures in these rows are estimates based on results from 1994 ACES frame creation

All single unit and all multiunit one industry companies do not present a problem since they are one industry companies; so, where these companies are placed for sampling is straightforward. Furthermore, since the payroll for companies with 500 or more employees (stratum 1) is maintained for all industries in which a company has activity, these companies do not present a problem. However, the multiunit multi-industry companies with from 5 to 499 employees are a concern. From the table above, 47,703 or 24.4 percent of the noncertainty multiunits (2.4 percent of all noncertainty companies) have activity in 2 or more industries. Where these companies are placed for stratification and sampling purposes can have an effect on the overall allocation and sampling results.

B. Multiple-Industry Companies' Contribution to the Estimates

From the above discussion, the multi-industry noncertainty companies account for only 2.4 percent of all noncertainty companies. Expending a large amount of resources to allocate this small proportion of the universe may not be cost efficient. But, if these companies contribute a significant amount to the total capital expenditures estimate, it may be a worthwhile endeavor. A summary of responding company counts is presented in Table 3, and a summary of the contribution to the 1993 ACES total capital expenditures estimate is presented in Table 4.

Table 3: Distribution of Responding Companies to the 1993 ACES[a]

		Stratum 1 (Certainty)	Strata 2,3, & 4 (Noncertainty)	Total
Responded in Multiple Industries	Total	4,110 (19.1%)	666 (3.1%)	4,776 (22.2%)
	Multiunits	4,075 (19.0%)	574 (2.7%)	4,649 (21.7%)
	Single Units[b]	35 (0.2%)	92 (0.4%)	127 (0.6%)
Responded in a Single Industry	Total	7,499 (34.9%)	9,192 (42.8%)	16,691 (77.8%)
	Multiunits	5,558 (25.9%)	2,824 (13.2%)	8,382 (39.0%)
	Single Units	1,941 (9.0%)	6,368 (29.7%)	8,309 (38.7%)
All Responding Companies	Total	11,609 (54.1%)	9,858 (45.9%)	21,467 (100%)
	Multiunits	9,633 (44.9%)	3,398 (15.8%)	13,031 (60.7%)
	Single Units	1, 976 (9.2%)	6,460 (30.1%)	8,436 (39.3%)

[a] Responding companies are companies that met minimum reported data requirements.
[b] Single units may respond in multiple industries. Reasons include: company growth, misclassification on the frame, and misreporting by the company.

Table 4: Summary of Company Contribution to the 1993 ACES Estimates of Total Capital Expenditures

		Stratum 1 (Certainty)	Strata 2,3, & 4 (Noncertainty)	Total
Responded in Multiple Industries	Total	49.5%	2.6%	52.0%
	Multiunits	49.3%	2.3%	51.6%
	Single Units	0.2%	0.3%	0.5%
Resonded in a Single Industry	Total	25.9%	22.0%	48.0%
	Multiunits	23.1%	7.1%	30.2%
	Single Units	2.8%	14.9%	17.8%
All Responding Companies	Total	75.4%	24.6%	100.0%
	Multiunits	72.4%	9.4%	81.8%
	Single Units	3.0%	15.2%	18.2%

From Table 3, only 574 or 2.7 percent of the responding companies are noncertainty multiunit multi-industry companies. Weighting the responses to estimate universe counts gives 22,205 or 1.1 percent of the noncertainty universe being multiunit multi-industry companies. And, from Table 4, these companies contribute only 2.3 percent to the total estimate.

Since the noncertainty multiunit multi-industry companies constitute such a small portion of the universe (2.7 percent, or 1.1 percent weighted) and contribute such a small amount to the total estimate for capital expenditures (2.3 percent), it appears that any further investigation of different methods for allocating and sampling these cases will not significantly impact the efficiency of the overall sampling methodology. Instead, our research will concentrate on evaluating the current stratification methodology.

III. Evaluation of Current Stratification Methodology

Under the current stratification methodology, the frame is divided into four strata. Stratum 1 is a certainty stratum and contains all companies with 500 or more employees. The remaining companies are divided into three noncertainty strata based on ACES industry and payroll size. The boundaries between these three strata are determined using a modification of the procedure developed by Pierre Lavallée and Michael Hidiroglou. This modification allows for all strata determined by the procedure to be noncertainty strata. How well did this stratification methodology do at meeting the reliability constraints?

We compared the *cvs* achieved on the survey estimates of total capital expenditures to the target *cvs* used in the stratification and allocation algorithm. Since the target *cvs* are based on payroll, and the achieved *cvs* are based on total expenditures, one cannot infer how good or bad the algorithm performed. So, we estimated the total payroll using administrative data[3] for the responding companies and calculated the *cvs* for payroll to see if the resulting achieved *cvs* for payroll were close to the target *cvs*.

From the table in the Appendix, we see that the achieved *cvs* for payroll are very close to the target *cvs* for the survey. This indicates that the stratification algorithm performed satisfactorily. However, the achieved *cvs* for total capital expenditures tend to be much larger than the target *cvs*. This is an indication that payroll (the design variable) did not perform very well as an allocation variable for estimating total capital expenditures.

IV. Conclusions and Further Research

The initial direction of our research focused on determining a better allocation methodology in order to meet reliability constraints for the 94 ACES industries. We thought this approach would yield the most gain in improving the methodology because we thought that one of the things that affected the estimates the most was the fact that for ACES, a company is assigned one industry code for stratification and sampling, but can respond and is tabulated in multiple industries. Upon examination, we found that the number of noncertainty companies that actually respond in multiple industries is relatively small, and that the contribution of these companies to the total estimate of capital expenditures is also small. So, our research refocused on evaluation of the current stratification methodology.

This evaluation indicates that the basic stratification and allocation method used can produce an adequate stratification of the sample. However, indications are that payroll is not a good variable to use for allocation. Future research will examine several possible alternatives including the following:

1. continue to use payroll to determine the stratification boundaries, and, since prior survey data will now be available, use the total expenditures results from a prior ACES to determine the allocation,

2. if the Census Bureau can access other data items that are more closely related to expenditures, such as total assets, then use total assets to determine the stratification boundaries, and use the total expenditures results from a prior ACES to determine the allocation, and

3. continue to use payroll to determine the boundaries for stratification, but use the total assets (if available) to determine the allocation.

Since the 1994 ACES universe was created and stratified, and the sample selected before any data from the 1993 ACES were available, the same stratification and allocation methods were used for 1994 as in 1993. Hopefully, time and resources will allow the alternatives mentioned above to be researched and the results incorporated into the universe preparation, stratification, and allocation for the 1995 ACES.

---

[3]Administrative payroll data were used for this evaluation because payroll data are not collected on the ACES questionnaire.

REFERENCES

Champion, E. J. (1994). *Update on Investment Surveys: Launching the Annual Capital Expenditures Survey and A New Measure on Investment Plans*, presented to the Census Advisory Committee of Professional Associations, April 14, 1994.

Champion, E. F., Funk, C. A., and Berry, J. H. (1993). *Preliminary Evaluation of the Annual Capital Expenditures Survey (Pilot)*, presented to the Census Advisory Committee of the American Economic Association at the Joint Advisory Committee Meeting, April 15-16, 1993.

Krenzke, T. (1994, December). *Multiple Classification Categories in Company Surveys*. Internal memorandum, U.S. Bureau of the Census, Agriculture and Financial Statistics Division.

Lavallée, P. and Hidiroglou, M. A. (1988). "On the Stratification of Skewed Populations," *Survey Methodology*, Vol.14, No. 1, pp. 33-43.

Slanta, J. and Krenzke, T. (1994). "Applying the Lavallée and Hidiroglou Method to Obtain Stratification Boundaries for the Census Bureau's Annual Capital Expenditures Survey," Paper presented at the 1994 Joint Statistical Meetings of the American Statistical Association, August 14-18, 1994.

U.S. Bureau of the Census (1994, May). *Annual Capital Expenditures: 1992*. Washington, DC: U.S. Government Printing Office (Annual Capital Expenditures: 1992, Item ACE(92)-1).

U.S. Bureau of the Census (in press). *Annual Capital Expenditures: 1993*. Washington, DC: U.S. Government Printing Office (Annual Capital Expenditures: 1993, Item ACE(93)-1).

APPENDIX

ACES Industry	Achieved CV (%) (expend.)	Target CV (%) (payroll)	Achieved CV (%) (payroll)
090	12	3	5
100	4	3	3
120	8	3	3
131	3	2	2
138	12	2	9
140	10	3	4
150	16	3	4
160	19	3	3
170	15	3	3
208	26	3	2
209	6	2	3
210	11	3	3
220	7	3	4
230	10	3	4
240	11	3	3
250	16	3	3
260	5	2	2
271	7	2	3
275	7	2	2
283	3	2	1
289	4	2	1
290	4	2	1
300	8	3	3
310	10	3	4
320	6	3	3
331	4	3	1
335	9	3	2
339	14	3	3
340	5	2	2
357	8	3	8
359	9	2	2
360	6	2	2

ACES Industry	Achieved CV (%) (expend.)	Target CV (%) (payroll)	Achieved CV (%) (payroll)
371	4	2	1
372	7	3	1
376	16	3	2
379	13	3	2
380	6	3	2
390	7	3	3
400	4	3	3
410	9	3	6
420	8	2	4
440	11	3	4
450	14	2	8
460	5	3	1
470	14	3	6
481	5	1	1
483	8	3	4
491	2	1	1
492	3	2	1
493	3	1	1
499	9	3	3
501	6	2	3
509	8	2	2
514	15	3	3
517	9	3	3
519	9	3	4
530	6	2	1
540	4	2	4
560	5	2	4
599	5	2	2
601	0	3	1
602	8	2	3
603	8	2	2
606	10	2	3
609	15	2	4

ACES Industry	Achieved CV (%) (expend.)	Target CV (%) (payroll)	Achieved CV (%) (payroll)
610	7	3	3
620	19	3	2
631	5	3	1
639	5	3	2
640	15	3	5
650	11	3	5
670	7	3	4
700	22	3	4
720	11	3	4
735	11	1	4
737	6	2	3
739	7	2	2
751	16	1	5
759	14	3	4
760	21	3	4
780	11	2	3
790	9	2	3
801	7	2	2
805	11	1	2
806	2	2	1
809	11	2	2
810	14	3	4
820	6	2	3
830	17	3	3
840	8	2	2
860	29	3	4
870	7	2	2
890	6	2	3
999*	n/a	3	n/a

* This code represents a "not classified" category that was used for sampling. Companies classified in this category were tabulated in appropriate ACE code industries based on their response.