AN ANALYSIS OF THE SAMPLING FRAME FOR THE CHEMICAL USE AND FARM FINANCE SURVEY

Susan Cowles and Susan Hicks, USDA/NASS
Susan Cowles, 200 N. High, Room 608, Columbus, OH 43215

I. Abstract

The National Agricultural Statistics Service (NASS) has begun a pilot study of the Chemical Use and Farm Finance Survey (CUFFS). The CUFFS combines parts of Form H of the Objective Yield Survey and the Cost of Production Survey (COPS) versions of the Farm Cost and Returns Survey (FCRS). Most NASS surveys utilize a multiple frame design -- a combination of list and area frames. The purpose of this analysis is to evaluate whether the multiple frame design is necessary for CUFFS. If not, then the sample will be selected from the list frame only.

II. Overview of the Chemical Use and Farm Finance Survey (CUFFS)

The CUFFS design consists of three phases. In the first phase, operations are contacted to determine if they have the commodity of interest. In the second phase, pesticide and fertilizer use information is collected in the Fall from operations that reported having the commodity. Finally, those same operations are recontacted the following Spring to obtain economic data.

The CUFFS was developed by NASS in an effort to:

- reduce respondent burden,
- improve data quality, and
- improve response rates.

The data that CUFFS would collect is currently collected through two other surveys: the Objective Yield Cropping Practices Survey (Form H) and the Farm Costs and Returns Survey's Cost of Production Survey (FCRS-COPS). When, or if CUFFS becomes operational, it would totally replace Form H and the COPS questionnaire would be shortened for crops targetted by CUFFS. The shorter interviews for Objective Yield and FCRS-COPS will reduce respondent burden for these two surveys.

Data quality is expected to improve for the COPS version of CUFFS as a result of collecting information closely following harvest. Currently, the COPS versions of the FCRS collects data six months after harvest. A better response rate is expected for the economic data due to its association with chemical use data. Farmers are more willing to provide data on chemical use due to widespread public concern for farming's effect on water quality and the environment.

III. Why Select a List Only Sample for CUFFS?

The list frame consists of known farm operators while the area frame consists of all land segments. The area frame is a complete frame and thus is used to measure undercoverage in the list frame. Farm operators found in the area frame that are not represented on the list comprise the NonOverlap Sample or NOL. There are three major advantages to a list only sample:

1) reduction in respondent burden for the NOL,

2) cost savings, and

3) reduction in variances.

Respondent burden is a major advantage of a list only sample. The NOL domain is relatively small due to small area frame sample sizes and more complete list frames. However, the relatively small population of NOL operators must be spread across all surveys, with the result that some NOL operators must be interviewed for multiple surveys.

The cost savings due to a list only sample are small compared to total survey costs. However, for less common commodities the NOL produces few if any positive operations. Thus, the cost per positive record is high. If the NOL domain is included, this cost could be reduced by screening operations by telephone for the commodity of interest prior to interview. See table 1.

Table 1

	CV%			June Planted Acres (in Mil)		
Commodity	NOL	LO	MF	NOL	List	MF
Corn	3.1	.6	.7	12.1	68.9	81.0
Soybeans	3.4	.9	.9	9.6	52.0	61.6
Spring Wheat	2.4	1.7	1.9	2.4	16.6	19.0

LO=List Only MF=Multiple Frame

At the U.S. level, the NOL contributes about 15% to total planted acres for major commodities, but contributes about 40% to the total variance.

For most commodities, the CV for a list only sample would be smaller than the multiple frame CV. However, the decrease in variance comes at a cost and that cost is bias. A list only sample introduces an inherent bias into the estimate by excluding some members of the population from the sample universe. However, if farm operators in the NOL domain are similar to farm operators on the list frame, then the bias may be minimal.

IV. Analysis Study to Compare List vs NOL Estimates

The goal of the research was to compare chemical use data from the list and NOL domains to see if there were significant differences. Ideally we would like to compare list to NOL estimates from the CUFFS questionnaire. However, the CUFFS pilot survey, which was conducted in Minnesota, used the proposed list only sample design, thus the NOL component was not available. To obtain a proxy for CUFFS chemical use data, we obtained Minnesota Form H data for corn, soybeans and spring wheat. Form H data is area frame only. The data was divided into overlap (OL) and nonoverlap (NOL) domains to allow comparisons between the two domains.

The OL and NOL domains were determined by classifying operations as OL or NOL to FCRS for 1991. The OL to FCRS group was further divided into groups determined by whether they were in a strata being sampled for CUFFS. If an operation was OL to FCRS and in a CUFFS strata, it was OL to CUFFS. All others were considered NOL to CUFFS.

The Form H summary system was used to obtain the mean rate of application per treatment and mean percent of acres treated for each active ingredient by domain. We then compared the estimates obtained between OL and NOL domains for the twelve most common commodity/chemical combinations.

V. Methodology

Percent acres treated is estimated as:

$$\hat{p}_d = \frac{n_d}{u_d}$$

where:

d = *OL or NOL domain*

n_d = *number of positive responses in domain d*

u_d = *number of usable responses in domain d*

The variances were calculated using the usual formulas for the variance of a proportion when the data are obtained by a simple random sample. In fact, the sample design was more complicated than a SRS, but the effect of this approximation should be a slight overestimate of the variance, which we were willing to accept.

Mean rate of application is estimated as:

$$\hat{R}_d = \frac{\bar{z}_d}{\bar{y}_d}$$

where:

$\bar{z}_d$ = *average rate of applicationfor each commodity/chemical combination in domain d*

$\bar{y}_d$ = *average number of treatmentsfor each commodity/chemical combination in domain d*

For mean rate of application per treatment, we calculated bootstrap-t confidence intervals instead of the usual t-test because of concerns about normality of the statistic being tested. Using the bootstrap methodology we constructed histograms of the distribution of the statistic mean rate of application per treatment. The histograms suggested severe departures from normality for some statistics. See Rao and Wu (1988).

We selected 10,000 bootstrap samples from the combined sample -- OL and NOL domains combined. For each bootstrap sample we calculated the usual t-statistic for a difference. Then based on the distribution of the t-statistics we estimated the 5[th] and 95[th] percentiles of the t-distribution for each commodity/chemical combination. The confidence interval is defined as:

$$\{\ \hat{D} - t_{.95}\hat{\sigma}(\hat{D}), \hat{D} - t_{.05}\hat{\sigma}(\hat{D})\ \}$$

where:

$\hat{D} = \hat{R}_{OL} - \hat{R}_{NOL}$ -- *from full sample*

$\hat{\sigma}(\hat{D})$ = *standard error of difference*

$t_{.05}, t_{.95}$ = *percentiles of the bootstrap t distribution*

VI. Results for Mean Rates of Application per Treatment

Table 2 shows the mean rates of application per treatment, the normal t-test for the difference and the bootstrap-t confidence interval.

The normal t-tests are included for comparison purposes. The bootstrap confidence interval reflects the skewness in the distributions of the differences while the

Table 2

Commodity	Active Ingredient	OL Rate per Treatment	CV(R) (%)	NOL Rate per Treatment	CV(R) (%)	Normal test t(diff)	Bootstrap CI LL	Bootstrap CI UL
Corn	Nitrogen	67.25	2.50	63.07	5.09	1.151	-2.05	9.90
	Dicamba	0.32	2.79	0.25	5.21	4.598*	0.05	0.10*
	Atrazine	0.80	4.35	0.82	12.23	-0.215	-0.23	0.14
	Alachlor	2.24	3.70	2.63	12.18	-1.174	-0.91	0.20
	Metolachlor	2.14	3.18	2.31	5.18	-1.243	-0.38	0.07
Soybeans	Trifluralin	0.77	3.15	0.81	5.09	-0.955	-0.13	0.03
	Imazethapyr	0.05	2.06	0.06	2.30	-1.820*	-0.01	0.00
	Alachlor	2.60	3.26	2.54	6.89	0.310	-0.24	0.42
	Bentazon	0.69	5.44	0.76	8.10	-1.071	-0.19	0.06
Spring Wheat	MCPA	0.29	4.91	0.30	7.76	-0.403	-0.07	0.03
	2,4-D	0.26	16.78	0.31	14.45	-0.761	-0.15	0.06
	Bromoxynil	0.24	5.09	0.19	18.97	1.148	-0.01	0.21

normal t-test relies on the assumed bell-shaped distribution. Therefore, the bootstrap-t confidence interval more accurately reflects the true differences between the OL and NOL groups.

Using the normal t-test, two significant differences would have been found: Dicamba used on corn and Imazethapyr applied to soybeans. Their respective t-values are 4.598 and -1.820 which, in absolute value, are greater than the 1.645 critical value for a 90% confidence test. The bootstrap intervals show only one clear difference between the OL and NOL mean rates of application per treatment of Dicamba on corn. However, we could expect to find one or two significant differences, by chance, even if no true difference exists based on a 90% confidence interval. We conclude that the data do not suggest a difference in mean rate of application between the two domains.

VII. Results for Percent Acres Treated

Table 3 shows the results for a difference between the OL and NOL domains for percent acres treated.

Of the twelve commodity/chemical combinations tested, four showed a significant difference. Also, the difference for spring wheat treated with MCPA had a t-statistic of 1.639 which is quite near the critical value of 1.645. The absolute difference for MCPA was 16.1 percentage points. If this difference is considered significant, all three of the spring wheat/chemical combinations show significant differences. Alachlor applied to corn and Trifluralin applied to soybeans also showed significant differences.

As with rate of application per treatment, based on a 90% confidence interval one or two significant differences could be expected, by chance, when no true difference exists. However, because at least four significant differences were found, we conclude there appear to be differences in percent of acres treated between the OL and NOL domains.

Table 3

Commodity	Active Ingredient	Percent Acres Treated OL	Percent Acres Treated NOL	t
Corn	Nitrogen	97.0	98.3	-0.905
	Dicamba	30.8	26.7	0.898
	Atrazine	32.3	29.3	0.644
	Alachlor	25.8	19.0	1.654*
	Metolach..	25.0	26.7	-0.382
Soybeans	Triflur..	46.6	37.0	1.959*
	Imazeth..	55.5	51.3	0.848
	Alachlor	10.0	12.3	-0.735
	Bentazon	12.4	10.4	0.647
Spring Wheat	MCPA	67.6	51.5	1.639
	2,4-D	27.6	51.5	-2.455*
	Bromoxynil	37.1	21.2	1.866*

IV. Conclusions

We looked at rate per treatment and percent acres treated for twelve commodity/chemical combinations. For rate of application per treatment, the data do not show a consistent statistical difference. However, several differences were found between OL and NOL percent acres treated. While the data suggest some differences exist, we are reluctant to draw conclusions for the nation as a whole based on results for one state, for the following reasons:

- Cropping practices vary by state.
- Commodities vary by state.
- Applications of chemicals vary by commodity.

The next phase of the research will examine 1992 Form H data from Minnesota and Louisiana to determine whether these results are consistent over time and across states. We recommend delaying the decision about whether or not to proceed with a list only sample for CUFFS until that research is complete.

References

Cowles, Susan and Hicks, Susan (forthcoming), "An Analysis of the Sampling Frame for the Chemical Use and Farm Finance Survey," NASS Staff Report.

Rao, J.N.K. and Wu, C.F.J., (1988) "Resampling Inference with Complex Survey Data," JASA, 83, 231-241.

U.S. Department of Agriculture (1983): Scope and Methods of Statistical Reporting Service, Publication No.1308, Washington, D.C.

MEASUREMENT OF PRIVATE TRUCKING IN CANADA

Andrea Mathieson, Statistics Canada
Transporation Division, Jean Talon Bldg., Ottawa, Ont., Canada

KEY WORDS: Transportation, trucking

Purpose

This paper will outline the treatment by Statistics Canada of private trucking in recent years, as well as a brief discussion of possible alternatives for future measurement of this vital component of transportation.

Introduction

Commercial trucking activity in Canada is associated with nearly every type of business in order to get goods to market or to provide some sort of service to customers. Some firms will hire a transportation company to carry out this function. Other firms will perform the delivery of the goods or service in-house. For example, a manufacturing company maintains a fleet of trucks and distributes its own product to market. While the company could have chosen to hire a (for-hire) trucking firm to deliver these goods, the decision was made to do it in-house. These private truck fleets are believed to contribute significantly to total transportation activity.

Statistics Canada conducts an ongoing annual survey of economic production which measures the activity of for-hire trucking companies, that is those which are considered to earn the principal share of their total revenue from trucking. In addition, there is a survey of *private trucking establishments* with annual operating expenses of at least $1 million.

Private trucking is, however, spread over many industries, and identifying all the businesses that perform an in-house transportation function in order to develop a survey frame has proven expensive and complicated. A good frame is the necessary starting point for any survey, if one expects the survey results to accurately describe the target population.

Industry and Activity: Two important concepts

Before examining the private trucking issue specifically, the concepts of "industry" and "activity" must be introduced. The scope of most Statistics Canada business surveys is restricted to a target industry, in order to eliminate duplication in business statistics. Thus, we have the manufacture of furniture separate from the logging industry, sawmills and the retail sale of furniture.

In the Standard Industrial Classification (1980)[1], an industry is defined as "a group of establishments whose production represents a homogeneous set of goods or services". (A statistical establishment is defined in Appendix A) If one wishes to measure economic production **within that industry**, the infrastructure exists at Statistics Canada to do so.

The Business Register contains information on all known businesses (with annual revenues greater than $25,000), and each business therein is coded to an industry (or, in the case of large and complex organizations, each production unit within the organization is coded to an industry). While a business is classified to, say, manufacturing, there may be a unit within the business organization which performs a transportation function, its principal client being the manufacturing unit.

A survey geared to providing industry statistics would not include the transportation unit outlined in the above example, since it is not in the trucking SIC (456), **unless** it was clearly defined as a **production unit** within the organization. Surveys of this nature are conducted across the Business and Trade Statistics Field at Statistics Canada, and provide critical input to the System of National Accounts for most industries found in the Standard Industrial Classification.

An activity survey would attempt to capture all of a given type of activity, regardless of the type of production unit involved. Truck activity statistics are of interest to those who wish to know how much trucking is being done, regardless of the industry controlling the trucks.

Thus, the industry and activity approaches yield data varying in scope and application, yet both are currently in demand in Canada, as will be explained below.

Why measure private trucking?

It has been estimated[2] that for-hire trucking accounts for at least half of all commercial truck movements of commodities, therefore private trucking accounts for the remainder. This constitutes an important component of trucking requiring attention, given the estimated 264 million tonnes moved in 1990 by Canadian for-hire truckers[3]. If one subscribes to the above assumption that private fleets carry, on an annual basis, "at least" as much tonnage as for-hire trucks, then this economic activity should be captured.

Just as there are several possible definitions of private trucking, there are various uses for this type of data. Policy makers and other industry researchers require a picture of "total trucking activity". Those interested in marketing a new technology to truckers are not just interested in for-hire trucking; they want to know about the whole market. In this regard, all commercial trucking is important to measure.

Private trucking of goods (including hazardous materials) constitutes another area of interest, since this traffic can be compared (in volume and composition) to for-hire truck traffic. In this case, the target population would be those private fleets which carry commodities. This goods-moving group is a subset of all commercial trucking, since service vehicles and other trucks not used in moving goods would be out of scope for a commodity-based survey. Having information on private commodity fleets would allow for comparisons of private and for-hire fleet performance, cost structure and efficiency. In addition, since deregulation of the Canadian trucking industry in 1987, private and for-hire fleets are in competition for some routes.

Finally, at Statistics Canada, both for-hire and private (or own-account) transportation data are used in the analysis of transportation, national accounts and business classification. These needs are currently served by the existing annual survey carried out by the Transportation Division. This survey targets private trucking units, identifiable within the organization of a business. This definition is in keeping with the establishment concept applied to other business surveys, which aim to measure unduplicated economic activity. The private trucking establishments thus surveyed currently are only those which are presumably not covered by other business surveys. The focus of the current private trucking survey is on company fleets which carry commodities; utilities and other such services are not covered at present.

What is the population of interest?

The three target populations (all commercial trucking, private carriage of commodities, transportation units within a larger business entity) described above form a hierarchy, in that they become increasingly narrow in scope. Thus, depending on the user, there are two possible frames for the economic activity called "private trucking" (since the first two groups can be identified from a single, truck-oriented universe). They are, in effect, interrelated as shown below.

Activity:	All commercial trucking = **A+E**, as shown in Figure 1 below
	All trucking of commodities = **A**, as shown in Figure 1
Industry:	Private trucking establishments = **C**, in Figure 1

This last group consists of trucking establishments within other types of businesses, provided that they are production units which can provide information on expenses. The private trucking unit may be a cost centre or a profit centre (i.e., revenue may be earned, but earnings will not always be present). **For additional information on the possible configurations of a private trucking production unit, see Appendix A.**

For an **activity survey**, the population is defined in terms of trucks because the private trucking "business" cannot always be identified. For an **industry survey**, the trucking production unit would be identified within the structure of the business. Ideally, *all businesses* would be asked during routine profiling whether they transport their products in-house. All production units (establishments) so identified would form the basis for the survey. The measure obtained would therefore be suitable for input to the System of National Accounts, provided there is no duplication between the private trucking industry and its "parent" non-trucking industry figures.

Determining the Private Trucking Universe

The **activity approach** could target trucks which will be registered to non-transportation companies. Given the current survey structure, the only successful method of creating a frame for private trucking activity has been to use provincial/territorial vehicle registration files. This would allow for sampling of either individual trucks or for identification of truck fleets (i.e., a number of trucks registered to a common owner).

Of the 3.9 million trucks and road tractors ("J" in Figure 1) registered in Canada[4] in 1990, some 167,000 are operated by for-hire trucking companies[5], couriers[6] and owner-operators.

Figure 1: Activity approach - private trucking universe

	Total	For-Hire	Not for-hire meets establishment criteria*:	
			Yes	No
Transport of Goods	A (B+C+D=A) 202 k	B 167 k	C 35 k	D ?
Service/ Farm	E (F+G+H=E) 595 k	F ?	G ?	H ?
Personal/ Other**	I 3,100 k			
Total	J 3,900 k			

* as described in Appendix A: cost centre or similar organizational unit

** may include ambulances, buses, fire trucks, taxicabs and road construction/repair equipment

Some of these trucks are used on farms and for purposes other than trucking commodities. In order to zero in on goods-moving vehicles, farm trucks and service vehicles (such as utility repair vehicles, or "toolboxes on wheels") would be removed from the count. An estimate of these out-of-scope trucks is approximately 595,000[7]. Thus the number of trucks potentially in scope for a commodity-oriented survey would be as many as 3 million, ranging from pick-up trucks to road tractors. Pick-up trucks for personal use are classified as trucks in the registration files. Pick-ups have become very popular personal vehicles in recent years, and are probably creating a lot of "noise" in the truck numbers.

It should be noted that smaller commercial vehicles (pick-up trucks, mini-vans) are used to transport lightweight, high-value items such as pharmaceutical products, so they would be in-scope. These same vehicles would also tend to distort overall statistics such as fuel consumption (miles per gallon), and care should be taken if vehicles smaller than 10 tonnes Gross Vehicle Weight are included in the survey. In developing a private trucking frame, it will be necessary to address this "contamination" of this small vehicle group by personal vehicles.

To sample trucks used in commercial trucking (at both levels, all trucking and commodity movement), it would be necessary to obtain the provincial vehicle registration files. Information on the types of vehicles in scope for each sample will be necessary to identify the in-scope trucks. Once all the commercial trucks are identified, they can be traced to their owners, thus reconstructing the "fleets" in service. A frame can therefore be created to accommodate both types of activity surveys (i.e., including or excluding the vehicles not used for transporting commodities). Small vehicles, as noted above, can be detected, and sampled separately, or accounted for otherwise.

The **industry approach** would target only private trucking "establishments" which meet the criteria in Appendix A. These production units are entities for which production account data[8] can be obtained for the estimation of value added to be used in calculation of transportation margins in the System of National Accounts. About 500 of these establishments have been identified in Canada, accounting for approximately 35,000 vehicles in 1990.[9] This is understood to be an underestimate since not all "non-transportation" companies have been queried as to whether they have private trucking production units within their organizations.

Private Trucking Surveys: The Experience at Statistics Canada

For-hire trucking has been recognized as an important economic activity, which is measured annually. However, it has proved difficult to construct and maintain a survey frame for private trucking (see Figure 2). In practice, the coverage has varied from year to year, which has made analysis of this group of carriers very difficult.

In 1973, the Federal Provincial Committee on Transportation Statistics struck a working group to assess the feasibility of measuring private trucking activity. Taking the broadest definition (i.e., non-for-hire trucking), the study group recommended that the frame for such a survey be drawn from vehicle registration files (the source of the 3.7 million count in Figure 1). While it was a labour intensive exercise and costly to run, these early surveys appear to have achieved the highest degree of coverage and the best approximation of trucking activity. In fact, these first surveys collected primarily vehicle information.

budgetary constraints) is the assumption that the numerous small companies that were excluded accounted for a small proportion of the tonnage carried. The commodity orientation was the main feature of the private trucking survey over this period. In addition to collecting financial and operating data, intercity carriers were asked to summarize their principal routes and provide tonnage and distance travelled. This facilitated the calculation of tonne-kilometres and permitted the comparison of private and for-hire carriers.

The development of the industry approach was a result of a major overhaul of the annual trucking (financial) surveys, in which operating expenses became the criterion for inclusion in the private survey target population. With the introduction of the new Business Register, private trucking establishments were created to reflect the transportation establishments within manufacturing and other non-transportation companies.

Figure 2 **Private Trucking in Canada: Survey coverage in recent years**

Survey Coverage	Number Reporting	Total Expenses $'000,000	Number of Vehicles*	Number of Employees
1990 >$1 million expenses	**505**	**2,331**	**32,802**	**29,384**
1989 >15 veh intercity	**521**	**1,671**	**37,146**	**57,187**
1988 >15 vehicles	**2,487**	**4,623**	**133,482**	**84,091**
1987 >15 vehicles	**2,320**	**4,349**	**127,949**	**84,166**
1984 >15 vehicles	**2,954**	**4,081**	**141,996**	**84,100**
1983 5 to 14 vehicles	**69,653**	**1,333****	**367,630**	**N/A**
1982 >4 vehicles	**52,623**	**9,100**	**376,600**	**287,000**

source: Trucking in Canada, 53-222

* Includes trailers - trailers were counted as vehicles for determining fleet size in the selection process

** Fuel expense was the only financial variable collected in 1983.

In fact, the 1982 and 1983 surveys were attempts at obtaining activity information, while subsequent surveys targeted fleets which carry goods, with a supplemental form to obtain information on intercity movements.

In the mid-eighties, the financial and operating component of the questionnaire was added, and the frame, which was still being drawn from provincial registration files, was limited to fleets of 15 vehicles or more. The reasoning behind this (in addition to

The administrative (vehicle registrations) frame source proved difficult to maintain, as were the internal resources required to manipulate twelve different input files from the various jurisdictions across the country.

Beginning in 1990, operating expenses replaced size of fleet as the threshold for the annual private trucking survey. The threshold was set at $1 million in all provinces and territories. Using the new definition (Appendix A), government enterprises and many construction companies, to name a few, were dropped from the survey as they had annual operating expenses under the $1 million threshold, or they did not meet the requirements necessary for delineation of a private trucking establishment.

The Next Step?

The current measurement of private trucking industry activity appears to be serving a limited number of users. The activity approach would be more appropriate for a wider variety of applications, but, the source for a reliable frame has not yet been identified.

One difficulty in creating an activity-based frame lies in the fact that these entities are not easily identified. Indeed, they are most often "buried" within the structure of a larger organization. For example, a manufacturing company maintains a fleet of trucks and distributes its own product to market. While the company could have chosen to hire a (for-hire) trucking firm to deliver these goods, the decision was made to do it in-house. Prior to the deregulation of the Canadian trucking industry, this was the manufacturer's response to the limited availability of trucking services available to meet its needs. These fleets were believed to contribute significantly to total transportation activity.

A possible course of action would be to ask all businesses whether they maintain a fleet of trucks, possibly extending the question to request a differentiation between service vehicles and trucks used for moving commodities. If this were the case, a description of both types of trucks would have to be provided. Such a question could be incorporated into other surveys of economic production, or collected as a separate undertaking.

It goes without saying that this "top down" identification of truck fleets would require the cooperation of a large number of subject matter areas at Statistics Canada. Within this body of information, the "bona fide" private trucking establishments could be flagged for inclusion on the Business Register, and the entire file would serve as a master file linked to the Business Register. Inclusion of this question at regular intervals would be necessary to update the frame.

The vehicle registration data remains an option for creating a truck frame, although one operational constraint could result in undercoverage. Specifically, unless provided for, the fleets that are leased, and therefore registered to leasing companies, will not be included in the survey unless the recipient of the questionnaire is directed to forward the form to the operator of the trucks. This is the current practice in the United States on the Truck Inventory and Use Survey, in which an inventory of all commercial vehicles is followed up with a detailed questionnaire collected from a sample of trucks, regardless of the industry where they "belong".

Whatever the course of action chosen, it appears that the private trucking statistics require some stabilizing as soon as possible. Continuing discussions with interested users will be useful in determining data requirements for this area of interest.

It is also apparent that the issue of own account trucking (just one of many own account business services) should be a topic of discussion at the next revision of the Standard Industrial Classification, due to be implemented for reference year 1997.

Endnotes:

1. Statistics Canada. 1980 Standard Industrial Classification Manual, Catalogue 12-501

2. Statistics Canada, Federal-Provincial Committee on Transportation Statistics. Private Trucking Survey, January, 1984.

3. Statistics Canada. Trucking in Canada, 1990, Catalogue 53-222.

4. Statistics Canada. Road Motor Vehicles - Registrations, Catalogue 53-219.

5. Statistics Canada. Trucking in Canada, 1990, Catalogue 53-222.

6. Statistics Canada Courier study, 1992.

7. The 1991 Census of Agriculture placed the number of farm trucks at 526,808. A trade publication entitled, Today's Trucking" (November, 1991) listed the top 100 private, utility and municipal fleets, which were reported to operate 67,548 vehicles. Thus, farm and non-goods moving would total 594,356. Pick-up trucks for personal use are usually considered "trucks" in the registration data, therefore they would need to be excluded from the count. Inclusions in the various vehicle categories vary among jurisdictions.

8. A production account is a set of data relating to a production entity, which will enable the calculation of value added. Specifically:

   Operating Revenue

   Operating Costs - Salaries and Wages
   - value of inventory change
   - depreciation
   - interest

9. Statistics Canada. Surface and Marine Transport Service Bulletin, Catalogue 50-002, Vol 9, No. 3, May, 1993.

The opinions expressed in this paper are solely those of the author, and in no way reflect the views of Statistics Canada.

Appendix A

TREATMENT OF OWN ACCOUNT TRUCKING

PROPOSAL BY

QUESTIONNAIRE REVIEW GROUP*

I. The case put forward by Transportation Division is that the trucking industries are not properly represented at the establishment level under the current rules which result in treating some own-account trucking as ancillary units. The proposal presented below does not discuss the merits of the argument but the "what" and "how" of deriving the proper set of units for production statistics for the trucking industries as argued by the subject matter area.

II. Background definitions
... For delineation purposes, for implementation of this particular standard statistical unit in the new business register, the **establishment** definition reads:

> One production unit or the smallest grouping of production units which produces as homogeneous a set of goods and services as possible, which does not cross provincial boundaries, and for which records provide data on the value of output together with the cost of principal intermediate inputs used and cost and quantity of labour resources employed to produce the output.

In the way of empirical considerations, the 1980 SIC itself proposes:

> Where the only statistic missing for delimiting an ancillary activity as a separate establishment is a measure of gross output, "imputation" of this by the respondent or the statistical office may be acceptable. Thus, an establishment may be created for a captive transportation unit.

III. Organizationally, in the business world, "own-account" or "private" trucking takes one or another of the following forms:

1. Own-account trucking units which are also profit centres. Both revenues and expenses are in full view, accounted for. Prices fully reflect market transactions. These units may provide both own-account and for-hire transportation services.

2. Own-account trucking units which are cost centres that earn some revenues. A total value of trucking services provided may be available, but prices tend not to reflect market transactions. The actual revenue generation comes from the provision of some "for-hire" trucking services for compensation.

3. Own-account trucking units which are strictly cost centres. These cost centres can provide many of the data elements required to delineate an establishment (i.e. material inputs, purchased services, employment, salaries and wages, inventories), but will not have a value of services provided, not even in the form of an internally imputed value. On the other hand, depreciation is generally available.

4. Own account trucking activity not accounted for by the business as a cost centre but as a series of itemized expenses of a broader organizational unit. Although such an activity may be carried out in a separate physical location, and employment and salaries and wages may be available for the operation, the business accounting practice does not cluster the direct costs of the operation.

IV. An own-account trucking establishment, in order to be a viable, meaningful unit, would have to be an operating entity for which a production account could be constructed from available data.

Under the provision that a "measure of gross output" can be imputed from "otherwise available data", and given that a profit centre readily offers all the characteristics required to delineate an establishment, the first three organizational forms of own-account trucking of the previous paragraph qualify as establishments. The case of own-account trucking activity where the business does not maintain sufficient information to allow an imputation of gross output cannot be delineated as an establishment since it is neither an operating entity nor an accounting entity.

... prepared by: S. Mozes
Transportation Division

G. Côté
Standards Division

date: 26 March, 1990

* The Questionnaire Review Group was formed to review the concepts used in Business Surveys to encourage standardization.

A BUILT-IN EVALUATION SYSTEM FOR A SAMPLE DESIGN WITH AN IMPERFECT FRAME

Masato Asanuma, Amy M. Newman-Smith, U.S. Bureau of the Census[1]
Amy M. Newman-Smith, U.S. Bureau of the Census, Iverson Mall, Room 300-09, Washington, D.C. 20233-6600

Keywords: Sample Allocation; Undercoverage; Matching

1. INTRODUCTION

Each month Construction Statistics Division (CSD) of the U.S. Bureau of the Census conducts the privately owned Nonresidential Value Put-in-Place (VIP) Survey to measure new construction installed or erected in the United States. The sampling frame for this survey consists of lists (files) of construction projects valued at $50,000 or more. CSD purchases these files containing projects whose work has started or will start within 60 days from a private firm. The firm provides us with these files monthly. From the sampled projects, CSD collects monthly VIP from the start to the completion of each project. This paper discusses a continuous evaluation that was set up to measure undercoverage of this frame. As with many surveys, the frame suffers from omissions (under-coverage), duplications, and errors in measures of size and other reported data.

Kish [1965] systematically presents general problems due to incomplete frame and suggests specific solutions. However, we cannot apply his suggested solutions such as redefining the population to fit the frame, correcting the deficiencies in the frame, supplementing the frame or adjusting the weights. Our remedy is to continuously measure the coverage bias and to adjust the survey estimates. Few survey practitioners in the past have chosen this approach because measuring the coverage bias requires formidable labor and designing a separate evaluation survey within the main survey. Among many options available to us, we concluded that a continuous coverage check procedure was the best, practical alternative under our survey conditions. We measure the bias of our sampling frame by obtaining information about new construction projects from a second source and then matching a sample of these projects to our sampling frame. Our second source is building permits for private nonresidential construction.

The building permit lists are not cost-effective as an alternative frame to the main survey or as a dual frame, but it provides sufficient information for an evaluation. Field representatives visit a sample of permit offices and list nonresidential building permits according to our sample plan. This list of permits serves as a sampling frame for the coverage evaluation. We do not use the building permits as a supplement to our sample for several reasons. One reason is that building permits are usually taken out when a project starts. By the time we processed the information the project would be well underway and our estimates would not be very efficient. The private firm on the other hand provides us with the information 60 days prior to start. Another problem with building permits is that one project may have multiple permits and we would have to keep track of these as part of our survey operations. Using building permits to supplement our frame has one other drawback, building permits are not always taken out by the owner and therefore the owner's name may not be on the building permit. To obtain the necessary information for our survey, we need the name of the owner or owner's representative. The private firm provides us with the owner's name or the name of the contact person.

This paper presents the detailed activities, such as sampling and matching of construction projects from the second source, that were set up as an integral part of the monthly survey operations. This paper also presents the results of this continuous evaluation since May 1987 and applications of the results to the national estimates of the privately owned nonresidential VIP values.

Purpose of Survey

The Nonresidential VIP Survey is one of the Construction Progress Reporting Surveys (CPRS) that cover privately owned as well as state, local and federal owned residential and nonresidential construction projects in the United States. Statistics from CPRS represent about half of the monthly measurement of new construction. The Census Bureau publishes these statistics in the monthly statistical series, Current Construction Report (C-30).[2] Government agencies at all levels and the private sector use them to monitor the amount of construction done each month. In addition, many private businesses and trade organizations use these statistics for marketing studies, determining investment choices and a wide range of other purposes.

Survey History

Before 1961, estimates of privately owned nonresidential building projects such as office buildings, warehouses, retail stores, shopping centers, hotels, motels, industrial buildings, etcetera were produced indirectly by applying fixed patterns of monthly construction progress to the monthly series of value of contracts awarded for such projects in the United States. In 1961, the Bureau of the Census began conducting a monthly survey to collect data on construction progress in the 37 eastern states and the District of Columbia. This was done so that directly measured VIP could be obtained from individual project owners. The sampling frame for this survey was created from construction news reports that we purchased from the F.W. Dodge Division of McGraw-Hill.[3] F.W. Dodge provided Census with about 20,000 slips of paper each month. Each slip contained the information for one construction project. At that time, the F.W. Dodge reporting system did not cover the 13 western states.[4] The estimates for these western states continued to be based on indirect measurement.

In 1966, the survey was expanded to cover the remaining 13 western states using building permits as a sampling frame. This design required us to set up two separate survey operations, one for the East and one for the West. We continued to use this design until October, 1986.

Sometime during the 1970's, Dodge started issuing construction news reports in the West except Hawaii. The impact of this coverage expansion was enormous. We were eager to evaluate the quality of the new construction news reporting system in the West because of the high costs of maintaining survey operations for the two frame design and the one month time lag to create the sampling frame from building permits. In 1979 and 1984, we conducted special studies to measure the coverage of the two frames. Both studies evaluated the coverage of news reports in the East and the West and the coverage of building permits in the United States. The overall news report coverage was not as good as the coverage by the building permits system but detailed analysis indicated that news report coverage rates increased dramatically as construction project values increased in both the

East and the West. Still, our major concern was a relatively low coverage rate, about 68 percent, of construction news reports for projects valued at \$500,000 or less and its effect on the estimates of monthly VIP values at the national level.

At that time, we began to investigate new survey designs that would improve frame coverage under the constraints of budget, timeliness, survey errors, and practicality of survey operations. While we were investigating the different alternatives, F.W. Dodge indicated the availability of its \$50,000+ construction news reports on computer tape in 1985. After careful consideration, we decided to use Dodge reports as a frame for the entire United States except Hawaii. We felt this was the most efficient alternative. Specifically, it was no longer necessary to deal with 20,000 slips of paper each month to create a sampling frame as we had done since 1961. Economically, it eliminated the need for maintaining two separate data collections and processing systems for the East and the West. However, this plan would require evaluations of both the Dodge frame as well as the building permit frame. The evaluations could be conducted continuously or periodically at regular intervals.

2. NEW DESIGN

The new design developed and implemented in 1986 uses three sources of information for identifying nonresidential construction projects:

1. Dodge reports on projects valued at \$50,000 or more in the United States except Hawaii.
2. Building permit notifications from the permit-issuing places of Honolulu and Maui to represent Hawaii.
3. Projects in a sample of areas not covered by building permit systems. Dodge virtually does not report on any projects in nonpermit issuing areas.

Projects from source 1 and Honolulu are stratified by type of construction and construction value. Fifty-six strata are created and each stratum is assigned a specific sampling rate. Before sample selection, all strata with the same sampling rate are grouped together, and within each stratum the projects are sorted by value in ascending order. An independent systematic sample of projects is then selected each month from each group. Projects from source 3 and Maui in Hawaii are selected with certainty. Once a project is selected, monthly construction progress reports are requested from the owner and the project remains in the sample until completion. The average number of projects in the survey at any one time is about 4,700. These include newly selected projects as well as projects carried over from previous months.

Monthly estimates of VIP are obtained by multiplying the final weight[3] of each project by the monthly reported construction value and summing up the weighted values. The tabulated results are also increased to account for construction projects not covered by either construction news report or the building permits system.

Built-in Evaluation System for an Imperfect Frame

The new design appears to be similar to the design used in the East with the coverage extended to the West. However, the similarity ends there. In the past, we conducted a frame coverage evaluation study every 10 years with special funding and spent about two years to complete it. For the new design, we allocated funds for the coverage evaluation from the annual survey budget. The new design assures the continuous updating of correction factors for reducing the noncoverage bias in spite of the severely limited budget and maintains experienced personnel with extensive knowledge of the survey and evaluation operations.

The new design tried to reduce the total survey error by allocating the limited resources judiciously between the unadjusted survey estimates and the measurement of the coverage bias. The monthly survey and the coverage study were developed as one design by applying the principle of Neyman allocation in stratified sampling. We created 56 strata defined by type of construction (TC) and construction value for the main survey and 5 strata defined by construction value for the coverage study.

The population of N construction projects is divided into 56 strata for the survey and the population of M permits is divided into 5 strata for the study. The 56 strata collapsed into the 5 strata based on value.

Allocation for Fixed Sample Size

We decided we needed to be able to make reliable estimates of the undercoverage adjustment factor every two years. Based on the optimum design as discussed next, we are able to achieve a coefficient of variance of about 3.5 percent for the monthly National estimate from two years worth of data. The form of the estimator for the design is

$$\hat{Y}=\sum_{i=1}^{8}\frac{y_i}{r_1}+\sum_{i=9}^{24}\frac{y_i}{r_2}+\sum_{i=25}^{32}\frac{y_i}{r_3}+\sum_{i=33}^{48}\frac{y_i}{r_4}+\sum_{i=49}^{56}\frac{y_i}{r_5}$$

where:

$\hat{Y}$ = monthly national estimate of VIP;

y_i = monthly total VIP estimate for stratum i from the survey, $i = 1, 2, \ldots, 56$;

r_j = estimated frame coverage rate for stratum j from the study, $j = 1, 2, \ldots, 5$.

To determine the optimal design, we needed to develop a cost function and the variance for the above estimator. The cost function we used has the following form:

$$C-c_o=12\sum_{i=1}^{56}c_i n_i+\sum_{j=1}^{5}k_j\frac{m_j}{2}$$

where the first sum exclude certainty strata and

C = annual budget for taking the measurement on $\left[12\sum_{i=1}^{56}n_i+\sum_{j=1}^{5}\frac{m_j}{2}\right]$ units = \$424,969;

c_o = annual fixed cost for taking the measurements for the survey; i.e., measuring projects in a priori certainty strata $i = 1, \ldots, 8, 13, 14, 22$, and 24 = \$145,738;

c_i = the average cost per unit over the life of a project in stratum i for the survey;

$c_1, \ldots, c_{16}$ = \$74.28

$c_{17}, \ldots, c_{35}$ = \$55.71

$c_{36}, \ldots, c_{56}$ = \$37.14

k_j = the average cost per unit to evaluate coverage in stratum j for the study;

$k_1, \ldots, k_5$ = \$ 9.33

n_i = the average sample size per month in stratum i;

m_j = the average sample size over 2 years in stratum j.

Because we are interested in minimizing the variance of $\hat{Y}$, we use the following approximation:

$$Var(\hat{Y})=\frac{(\sum_{i=1}^{8}y_i)^2}{r_1^2}\left[\frac{\sum_{i=1}^{8}Var(y_i)}{(\sum_{i=1}^{8}y_i)^2}+\frac{Var(r_1)}{r_1^2}\right]+\frac{(\sum_{i=9}^{24}y_i)^2}{r_2^2}\left[\frac{\sum_{i=9}^{24}Var(y_i)}{(\sum_{i=9}^{24}y_i)^2}+\frac{Var(y_2)}{r_2^2}\right]$$

$$+ \frac{\left(\sum_{i=25}^{32} y_i\right)^2}{r_3^2} \left[\frac{\sum_{i=25}^{32} Var(y_i)}{\left(\sum_{i=25}^{32} y_i\right)^2} + \frac{Var(r_3)}{r_3^2} \right] + \frac{\left(\sum_{i=33}^{48} y_i\right)^2}{r_4^2} \left[\frac{\sum_{i=33}^{48} Var(y_i)}{\left(\sum_{i=33}^{48} y_i\right)^2} + \frac{Var(r_4)}{r_4^2} \right]$$

$$+ \frac{\left(\sum_{i=49}^{56} y_i\right)^2}{r_5^2} \left[\frac{\sum_{i=49}^{56} Var(y_i)}{\left(\sum_{i=49}^{56} y_i\right)^2} + \frac{Var(r_5)}{r_5^2} \right]$$

where:

$$Var(y_i) = \begin{cases} N_i^2 \dfrac{S_i^2}{n_i} \left(1 - \dfrac{n_i}{N_i}\right) & i=9,\ldots,12,\ 15,\ 17,\ldots,20,\ 23,\ 25,\ldots,56; \\ & N_i = \textit{average monthly number of projects in stratum } i; \\ & S_i^2 = \textit{population variance of monthly VIP for projects in stratum } i\textit{; including the effect of systematic sampling.} \\ 0 & \textit{Otherwise (These are certainty strata)} \end{cases}$$

For each j,

$$Var(r_j) = Var\left(\frac{x_j}{z_j}\right) \doteq \left(\frac{1}{z_j}\right)^2 \left[Var(x_j) - r_j^2 Var(z_j)\right]$$

$$\doteq \left(\frac{1}{z_j}\right)^2 \frac{M_j^2 P_j (1-P_j)}{m_j} \left(S_j^2 + \bar{z}_j^2\right)$$

$S_j =$ standard deviation of permit values over two years in stratum j;

$x_j =$ total permit value matched to projects reported by construction news reports over 2 years in stratum j;

$z_j =$ total permit value over 2 years in stratum j;

$\bar{z}_j =$ mean permit value in stratum j;

$M_j =$ total number of permits issued in stratum j;

$m_j =$ the average sample size over 2 years in stratum j;

$P_j =$ probability of matching a permit to a construction news report in stratum j, assuming all permits in stratum j have an equal chance of being matched.

Using the method of Lagrange multipliers, we select n_i, m_j and the multiplier λ to minimize

$$Var(\hat{Y}) + \lambda \left(12 \sum_{i=1}^{56} c_i n_i + k \sum_{j=1}^{5} \frac{m_j}{2} - C + c_o\right)$$

Here, we are assuming $k = k_1 = \ldots = k_5$.

We obtained

$$n_i \sqrt{\lambda} = \frac{N_i S_i}{r_j \sqrt{12 c_i}} \quad \textit{for all noncertainty strata } i;$$

$$m_j \sqrt{\lambda} = \frac{\left(\sum_i y_i\right)}{r_j^2 z_j} \sqrt{\frac{2M_j^2}{k} P_j(1-P_j)\left(S_j^2 + \bar{z}_j^2\right)} \quad \textit{for } j=1, 2, \ldots, 5;$$

λ can be found by substituting these in the cost equation.

The summation in the second equality is over all survey strata that correspond to the selection value range defined for the study stratum j. For example, for value range $j=2$, we sum the monthly total VIP values from strata $i=9, \ldots, 24$.

This allocation method gives us about 400 sample cases per month for the survey and about 150 sample cases per month for the study. The number of sample projects for the study over 2 years from the five study strata are: $m_1=894$, $m_2=441$, $m_3=722$, $m_4=698$, *and* $m_5=750$. To accomplish this desired allocation, we had to consider two constraints. First, the study was planned to piggyback the Survey of Construction (SOC); second, we could afford to list nonresidential building permit authorizations only twice a year from each SOC permit office (place). We compared the required annual sample sizes for each stratum with the estimated numbers of SOC listing units from two months. This comparison showed that all projects with a permit value greater than \$3 million listed by SOC field representatives should be selected for the study and projects with a permit value less than \$3 million should be subsampled. Also, this comparison showed that projects with a permit value greater than \$10 million needed to be supplemented.

Before explaining how we accomplish the subsampling of small projects and the supplementation of large projects, a brief description of the SOC survey design is in order. The SOC is a monthly survey that measures the estimates of new privately owned housing units started, under construction and completed. The sample for the SOC is a stratified three-stage cluster design. Each state is divided into geographic areas called primary sampling units (PSUs). These PSUs throughout the United States are grouped into 169 strata. One PSU is selected from each of the 169 strata. Within each of these 169 sample PSUs, places (building permit offices) are stratified into six size classes based on the past building activity. From each of these size classes, a systematic sample of places is selected independently. We expect to list 4,928 nonresidential building permits per month from these SOC sample places.

3. COVERAGE EVALUATION

In this section, we discuss in detail how we implemented the study design that selects a sample of new construction projects from building permits and matches these sample cases against the survey sampling frame.

Selecting Building Permits

We use a subset of field representatives who list new privately owned residential building permits for the SOC each month to prepare a list of authorized privately owned nonresidential building permits. The listing is done on a rotating basis among the 12 regional offices, with field representatives from two regional offices listing each month. Since we are using $1/6^{th}$ of the SOC sample each month, it takes us 6 months to collect a representative sample for the United States. While listing, the field representatives look for permits which are part of the same project and list them together as one project.

Once we have received the list of permits from field representatives we systematically select a subsample of projects for the study.

As was previously discussed, the sample of \$10+ million permits from two months of SOC listings per year is insufficient. In addition to the list of projects provided by the SOC field representatives, we obtain a list of projects valued at \$10 million or more from the monthly Building Permits Survey (BPS). We supplement our sample with the $10+ million BPS projects. We accomplish this by selecting with certainty $10+ million projects located in SOC permit offices areas which are in the BPS sample. We unduplicate $10+ million projects listed by SOC and also reported in the BPS survey. The BPS is a mail survey which covers approximately 9,000 sampled permit offices monthly and covers the entire universe once a year through an annual survey of the remaining 8,000 nonsampled permit offices. The monthly BPS collects data on permit authorizations for residential and nonresidential construction and detailed information about permits valued at $500,000 or more. Although BPS provides more detailed information for the permits valued at $500,000 or more, it is not as informative as the SOC listing.

The sampling scheme supplemented by the monthly BPS provides us with about 150 cases per month (or 3,500 cases over two years) for the coverage study.

Procedures for Matching Building Permits (BP) with F.W. Dodge Reports (DR)

The procedures described below are collaborative efforts between F.W. Dodge and the Census Bureau to estimate the coverage of our sampling frame. The payoff resulting from this joint quality enhancement program will be an improved product for Dodge and a better estimate with reduced bias for the Census Bureau.

Using Human Judgement for Match Decisions

Matching construction projects seems straightforward at first but it can be complex. In the first place, no unique common identifier between Building Permits and Dodge Reports exists, so computer matching algorithms are difficult to develop. We make a series of match decisions and combine the results of the decisions to determine a match status. We contrast the following identifying variables to determine whether or not two records describe the same construction project:

a. Project location (street address, city, county, state)
b. Project description
c. Ownership of the project
d. Permit issuance date and Dodge start date
e. Permit value and Dodge value

In general, Building Permit and Dodge Report records are classified as matches if variables **a**, **b**, and **c** are the same or if all three variables carry similar values, and there is a reasonable period of time between permit date and Dodge start date. For example, the intersection of "Silver Hill Road and Branch Avenue" can be reported as "3737 Branch Avenue" or "Iverson Mall." "Hotel" may be reported as "Casino Hotel and Parking Facility." On occasion, a project that matches on these three variables to a building permit may be classified as a questionable match or nonmatch if the common project description says "tenant improvement" or "alteration" and the period between the permit date and the Dodge start date is more than 6 months. BP and DR records are classified as questionable matches if only one or two of these three variables are the same. These three variables are supplemented by the other two variables, **d** and **e**, and other information for decision making. Other information provided on the Dodge report may include detailed data on structure, size of project, general contractor, architect, etc.

Implementation of Matching Process

Step I - Dodge Matching

We transcribe the key building permit data discussed in the previous section for each of the selected cases onto a CSD-201 form and then send the forms to Dodge for matching. The Dodge statistical department enters data from the forms into their internal computerized database and identifies the area newsmanager for each permit form. There are 140 Dodge news areas throughout the United States. The statistical department sends the forms to the proper newsmanager for matching. The newsmanagers use all available resources to provide CSD with a Dodge report that matches the permit. The resources include:

- internal database of the Dodge projects;
- reporters familiar with project location;
- contact with owners or general contractors;
- internal project files;
- contact with permit issuing offices.

Once matching is completed by the newsmanagers, the senior newsmanager reviews the forms and attaches Dodge slips to the match cases. The senior newsmanager sends the forms to the statistical department. The statistical department reviews the forms and sends the questionable matches back to the senior newsmanager for reassessment. The statistical department researches the nonmatch cases further. The statistical department keeps track of the match projects that haven't advanced to the start stage. Dodge sends the CSD-201 forms along with the proper Dodge reports to CSD and gives CSD a summary of the match results specifying the number of cases matched, the number of cases not matched and the explanation for cases excluded from their matching. Included in the summary is a list of the previous months match cases that have reached the start stage.

Step II - Check-in and Match Verification

We check in the cases making sure that Dodge has sent back all of the CSD-201 forms. If a form is missing, we send another CSD-201 form to Dodge. After the check-in, we begin the match verification. We verify that the Dodge match cases are truly matches using the guidelines previously explained. We then group the forms into the following categories:

1. Match - verified Dodge match;
2. Nonmatch - rejected Dodge match or a Dodge nonmatch;
3. Questionable - Dodge match that needs further investigation;
4. Out-of-scope - case that Dodge said is not a private nonresidential construction project and CSD agreed.

Step III - Computer-assisted matching

After completing the check-in and reviewing Dodge matching, CSD staff matches the nonmatch and questionable cases by computer. The computer-assisted matching consists of matching these cases to the Dodge Major Project (DMP) file. This is one of the two computer files that we use as a sampling frame for the Nonresidential VIP Survey. The DMP includes all Dodge projects valued at $750,000 or more that Dodge reports as expected to start within the next 60 days. The computer-assisted match enables us to match some of the cases that Dodge failed to match. The matching program is set up so we can search for all projects reported to us by location, owner, project description, or permit value. The screen displays both the permit information and the possible match information so they can be easily compared. When a possible match or a definite match is found we can print the information from both and do further verification if necessary. Up to three possible matches may be saved for each Dodge Report.

The other computer tape we use in sampling is called Dodge Construction Potentials (DCP) and this is used primarily for selecting projects with values less than $750,000. We do not use the DCP in the computer assisted match because it contains limited, coded information on construction projects. No key variables necessary for matching are included.

Since the DMP only provides information on projects valued at $750,000 or more, we do one final match on the remaining questionable and nonmatch cases. We search through all of the printed Dodge slips by hand.

Step IV - Hand Matching

During this stage of our matching, we hope to match the cases that Dodge couldn't find and any of the cases that weren't on the DMP files. We search through all Dodge reports which were issued within one year before the permit issue date and up to six months after the issue date. CSD staff searches for the remaining nonmatch or questionable cases at least once. If a match is found it is verified and then registered in the computer as a match. A second search is done for the remaining nonmatch and questionable cases. We usually do this manual matching in teams of at least two people so all cases are verified at least once.

Step V - Followup of Questionable Cases

We send all of the questionable cases to our field representatives for match verification. They must determine the match status by either visiting the site, speaking to the permit official, or talking to the owner. Sometimes, we resolve questionable cases by looking at maps and telephoning the owner from our Washington office.

Step VI - Checking for Duplicates

Since we obtain the $10 million plus projects from two different sources as described in "Selecting Building Permits" section, we must make sure that a permit does not fall into our sample more than once. During our monthly sample selection, we check for duplicates and reconcile those cases then. We reconcile duplicate cases by keeping the BPS-based project if a permit is listed on the BPS form and the SOC-554 form. However, some BPS-based cases come in late and may not be unduplicated until all the matching is complete. We have set up a system that checks for these cases. We first check to see if two or more cases match the same Dodge report. If two projects match the same Dodge number, we determine if they are the same permit or not. If they are the same permit, we unduplicate them by keeping the BPS-based case and deleting the SOC-554 case. If both are listed on the BPS form, we delete the last one listed. If they are not the same permit we keep both of them in the study. After we check the match cases, we check for duplicates among match and nonmatch cases. It is necessary to check every case because Dodge may match one case and not match its duplicate. We look at all $10 million plus projects in the area covered by the same permit office. If we find a duplicate, we delete one of the cases in the same manner as previously described.

Step VII - Tape Verification

For our processing to be complete we must check all the match cases against the DMP and the DCP files. A case cannot be considered as covered by the survey if it did not have a chance of selection. If a match case is not found in either of these files, and it can't be reconciled by CSD or Dodge, then it is treated as a nonmatch for our coverage estimates. One reason for the tape verification is that Dodge may issue a report after the project has started and therefore, not include it on the tapes. Another reason for the verification is that Dodge may misclassify the project either as government owned or as residential and may not include it in the nonresidential files. Dodge reconciles some of these cases by providing us with a new Dodge report number or updated project information. We verify the new Dodge report number against the DMP and DCP files. Further investigation by Dodge sometimes reveals that a project has not started or has been abandoned. If Dodge says a project has not started, we verify it with our field representative. Status verification is the next step of our processing.

Stage VIII - Status Verification

All of the $10+ million nonmatch cases and all cases that Dodge says haven't started are sent to field for one last verification of the existence of the projects. We ask the field representative to determine if a project has started or not. At this time our field representatives may find out that the project has been abandoned. All abandoned projects are removed from our files and are not included in our tabulation of the coverage rates. We do find that almost all the cases have started.

__Estimation__

We estimate the match rate using the following formula:

$$Matchrate = \frac{\sum_{Dodge} W_i \, Pval_i}{\sum_{Sample} W_i \, Pval_i}$$

where

$\sum_{Dodge} W_i \, Pval_i$ = *sum of the weighted permit values matched by Dodge and CSD*

$\sum_{Sample} W_i \, Pval_i$ = *sum of the weighted permit values over the entire sample (match and nonmatched, excluding the out-of-scope cases)*

4. RESULTS OF COVERAGE STUDY

Dodge reports were found for 74 percent of the 6231 in-scope projects listed by building permits; these matches accounted for about 81 percent of the weighted valuation of the building permits in the United States excluding Hawaii.

Essentially, the match rate hasn't changed significantly over the 4 1/2 year period covered by the study except a low estimate of 75% (standard error of 2%) from the May 1988 to October 1988 sample and a high estimate of 86% (standard error of 2%) from the November 1990 to April 1991 sample.

Although the overall match rate has not changed significantly in 4 1/2 years, we have accumulated enough data to see a significant difference in the match rate by type of construction. The "Other" construction category with a match rate of 88% (standard error of 2%) is significantly higher than "Industrial" and "Commercial" with match rates of 78% (standard error of 3%) and 79% (standard error of 1%), respectively.

Also, the match rate differs significantly by permit value. The lower valued permits have a significantly lower match rate than the larger valued permits. However, this important finding can not be translated into an immediate application to the survey estimates. Dodge reported values and construction values reported by respondents are available for projects sampled for the CPRS. Permit values are not available. A lower valued permit may be matched to a project with a large Dodge value, since the permit may have been issued for only part of a project. The match rate for lower valued permits are mixture of the coverage rates for higher and lower valued Dodge projects. Thus the rates for lower valued permits cannot be applied to lower valued Dodge projects. Therefore, we cannot estimate the coverage rates for the five strata as had been planned.

Another interesting finding is that frame coverage rates for the western states and the eastern states, discussed in the "Survey History" section, are 79% and 81% respectively. This result strongly supports the correctness of the sample design decision made in 1986 to go with a single frame approach rather than a dual frame approach.

To derive the adjustment factor, we assume that building permits and our sampling frames collectively cover all new construction projects in permit issuing areas. This is not a realistic assumption, but our data cannot estimate the amount of construction missed by both frames. We cannot assume independence of the sampling frame because Dodge also uses permit offices as sources. We also assume that the 1979 study of the coverage of building permit systems is still accurate. The 1979 building permit study was similar to our coverage study except reversed. Instead of selecting a sample of permits, this study selected a sample of Dodge reports and matched them to building permits.

Let $V(B)$ = valuation covered by building permits
$V(D)$ = valuation covered by the Dodge frame
$V(B \cap D)$ = valuation covered by building permits and the Dodge frame

From this coverage study, we estimate $V(B \cap D) = .81V(B)$. From the 1979 study we estimated $V(B \cap D) = .95V(D)$. Therefore, $.81V(B) = .95V(D)$

We want to estimate $V(B \cup D)$ using the undercoverage adjustment factor f'

$$f' = \frac{V(B \cup D)}{V(D)} = \frac{V(B) + V(D) - V(B \cap D)}{V(D)} = 1.22$$

Finally, to account for the 1% of projects with a value less than \$50,000 that Dodge does not cover, the adjustment factor is

$$f = \frac{f'}{.99} = 1.24$$

5. UNSOLVED PROBLEMS

One of our assumptions for the coverage study is that the respondents to the VIP survey report only on the projects that are reported to us by Dodge. This is not always the case. The effect of misidentification of a project by a respondent is not controlled or measured now. For example, a respondent who is constructing more than one project may report on more than the requested project, including projects not covered by Dodge. Therefore, a project may not be covered by our sampling frame but it could be covered in the VIP survey. This is overreporting. In theory, our coverage factor could be too high. On the other hand, our factor could be too low if a respondent doesn't report on all that the Dodge report encompasses. One example of this is a respondent reports on a section of a new plaza and not on the whole thing, although the Dodge report includes the whole plaza. In this case, we have underreporting and the undercoverage factor would be too small. Also, a respondent may report on a different project than the one we selected. Thus the selected project is not covered by the survey. The match rates cannot reflect these cases.

To solve these problems of reporting error, we are setting up a process that will investigate cases where the selection value and the reported value differ by a significant amount. The process involves finding out what a respondent is reporting on and comparing it to the sampled Dodge report. This measures a portion of the response bias. We will decide how to adjust our coverage rates derived from the study based on this investigation. This new process will be part of the VIP survey operations and should allow for better VIP estimates. As far as the coverage of the VIP survey is concerned, we may introduce a study that looks at survey coverage as opposed to this study that looks at frame coverage.

A second problem that we have not solved is using the information on differential coverage rates by permit value group. As described before, we can't apply adjustment factors by value group because we do not believe that the relationship between permit value and Dodge value for matched permits would hold for nonmatched permits. In fact, no Dodge value exists for nonmatched permits. Although the permit value and the Dodge value are highly correlated, we can't make the assumption that a high Dodge value would always correspond to a high permit value. In fact, lower valued permits are sometimes matched to higher valued Dodge projects. Also, the opposite happens. We do not believe that a nonmatched lower valued permit would be as likely to correspond to a high "Dodge value" as the matched permits do. Thus we cannot assume that the joint distribution of Dodge value and permit value is the same for matched and nonmatched permits. Without this assumption or more information about the construction value for nonmatched permits, the match rates by value group cannot be applied. Ultimately, we would like to have adjustment factors by value group and type of construction. A study to examine the Dodge value, building permit value and reported final construction value is under consideration but no funding or personnel are assigned.

NOTES

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 authors would like to thank all members of the Research and Methods Staff of the Construction Statistics Division for implementing the frame coverage study since 1988 and the strong continuous commitment of Senior Management of Construction Statistics Division and F.W. Dodge for their continuous support.
2. U.S. Bureau of the Census, Current Construction Report Value of New Construction Put in Place: Month/Year U.S. Government Printing Office, Washington, D.C.
3. This firm started as an independent company at the end of the 19th century. It was acquired by McGraw-Hill, Inc. in 1961. With data on construction activity recorded in the Dodge data bank, the F.W. Dodge Division is by far the largest, most experienced, most comprehensive collector of construction information in the United States.
4. Thirteen western states are: Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington and Wyoming.
5. The final weight for each project reflects the product of the reciprocal of the probability of selection, the ratio adjustment factor which is to reduce the contribution to the variance arising from the sampling of noncertainty projects within each type of construction, the adjustment for prorating of architectural, engineering and miscellaneous costs, and a duplication control factor to adjust for projects that are included in the sampling frame more than once.

REFERENCE

Kish, L. (1965), Survey Sampling, John Wiley and Sons, New York

TIME RELATED COVERAGE ERRORS AND THE DATA ADJUSTMENT FACTOR (DAF)

Jeffrey T. Bailey, USDA/NASS
Research Division/3251 Old Lee Hwy., Fairfax, VA 22030

ABSTRACT

The National Agricultural Statistics Service (NASS) conducts quarterly surveys to estimate crop acreage, grain stocks and hog inventories. Sample replicates from the stratified sample design are surveyed on a rotating basis to allow for quarter to quarter overlap while bringing other operations into the survey. With this design, farming operations may be enumerated from one to four quarters in a particular year's survey cycle.

Operations are sometimes reported as "out-of-business" in one of the quarterly surveys when they were in business during a previous quarter. While this is not a problem if the questionnaires are correctly coded, a review of survey data reveals a significant number of coding errors in one quarter or the other. This between-quarter discrepancy in an operation's business status can change the coverage of the population (particularly if the change is due to incorrect coding) and have a major impact on the resulting indications.

This study looked at the effect of the coverage change on the indications and the reasons for questionnaires being coded as "out-of-business". From this research we hope to determine: 1) the extent to which those "out-of-business" changes represent data collection errors rather than real operation changes, 2) how to reduce the number of operations incorrectly being coded "out-of-business" and, 3) whether the data are increasing for operations remaining in business to offset operations legitimately going "out-of-business".

SUMMARY

The National Agricultural Statistics Service (NASS) conducts quarterly surveys to estimate crop acreage, grain stocks, and hog inventories. The replicated, stratified sample design results in sampled operations being surveyed in a rotating fashion, allowing for some quarter to quarter overlap while reducing respondent burden. A new sample begins in June with quarterly surveys in the following months of September, December, and March.

A dilemma arises as the year's survey cycle progresses beyond the June base survey, because the percentage of "out-of-business" operations increases. This creates a situation where the indications from the survey decrease and the population coverage may become incomplete. Observations show that approximately 4 to 6 percent of operations change from "in business" one quarter to "out of business" the next. Reviewing the questionnaires indicates that a substantial number of these were inaccurately coded or lacked complete information.

The Data Adjustment Factor (DAF) adjusts the data for duplication and eliminates data that should not be summarized. When an operation is "out-of-business" the DAF is zero. Calculations of the average DAF show that it continually decreases the further you get from June. The DAF reduced the December expansions relative to June by about 2 percent in 1991 and 1 percent in 1992. This drop from June is substantial, but how much of it reflects a legitimate change in the target population? What led to the reduction of the DAF impact in 1992 and how can we further reduce its effects?

The DAF should continue to be monitored and efforts be made to reduce its artificial impact upon the survey indications. Some suggestions to reduce the DAF decline are more training, changes in coding old replications, and the use of historic data to confirm "out-of-business" operations. These suggestions will likely not completely eliminate the DAF problem and more ideas should be developed and studied to lessen and monitor the DAF impact.

INTRODUCTION

The National Agricultural Statistics Service (NASS) conducts many surveys to estimate inventory and production of various agricultural commodities. As a part of its Agricultural Survey Program, NASS conducts quarterly surveys to estimate crop acreage, grain stocks and hog inventories. Analysis of December 1991 Agricultural Survey data showed that the December crop indications for planted acres were always lower than the June indications. Within a growing season the reported planted acreage of a crop should not change, unless intentions were reported in June and the crop was never actually planted. It was discovered that many operations which reported crops in June were now "out-of-business" in December.

Reviewing the data of these "out-of-business" operations focused attention on the Data Adjustment Factor (DAF). The DAF is used to adjust for duplication and to eliminate any data reported on an "out-of-business" operation. The value of the DAF is always inclusively between zero and one. The average DAF was calculated for successive quarterly surveys and found to decline as time passed. Several reasons can account for this and many ideas have been expressed.

This paper will begin with a description of the multiple frame surveys at NASS and how coverage errors can occur as time passes. Then the analysis of the DAF will be presented.

NASS MULTIPLE FRAME SURVEYS

NASS conducts many surveys and for each it is necessary to define the sampling population or frame of units to sample. For most NASS surveys the target population is all operations with the agricultural commodities of interest. NASS maintains a list frame of names thought to be farm operators in each state for its sampling. Considerable time and resources are spent in the state offices updating and maintaining these lists. In addition to the samples drawn from these lists, samples are drawn from an area frame of all land in the U. S. from which estimates are generated to measure list incompleteness. Together the two frames form a multiple frame survey design which NASS uses in many of its surveys.

This study focuses on NASS's quarterly multiple frame Agricultural Surveys. The list sample is selected in the spring with the surveys conducted during June, September, December and March. During the base survey in June a complete area sample is enumerated. For this survey, every operation in the U. S. has a chance to be sampled either from the list and area frame or the area frame alone. Names found in the area frame during June that are not on the list frame (NOL) will be used in subsequent quarters to represent those operations which had no chance of list frame selection.

The list sample consists of several replications which are selected each spring for use during the course of the survey year. These replications are rotated in and out from survey to survey to provide quarter to quarter comparability and to relieve respondent burden. With the rotation scheme used, farming operations may be enumerated from one to four quarters in a particular year's survey cycle.

TIME RELATED COVERAGE ERRORS

The samples for the Agricultural Survey are selected in the spring of each year. Before some samples are surveyed they will go "out-of-business". If an "out-of-business" operation is taken over by a new operation, this new operation must have a chance of selection. Any new operations taking over an "out-of-business" operation before June 1, will have a chance of inclusion in the area frame sample during the June Agricultural Survey. New operations starting up after June 1 can only be accounted for by substitution procedures, since there is no complete area frame survey done after June.

These substitution procedures provide a means to give everyone a chance of being selected to assure population coverage. Substitutions should be made when sampled units are "out-of-business" and the new operator was not farming on June 1, but there is concern that the procedures are not always executed properly and all needed substitution is not being done (Jones 1988). Furthermore, substitution only occurs when an operation is completely "out-of-business". If an operation sells off only part of its land to a new operator, that operation is not eligible for substitution and does not have a chance of selection (Dillard 1993). NASS is currently researching how effectively substitution procedures are being followed and the impact of the substitution process on survey indications.

For the follow-on quarterly surveys of September, December, and March, about 40% of the sample is from new replicates, with the remaining from old replicates that were surveyed in a previous quarter. For old replicate samples only those operations that were in business in the previous quarter will be surveyed in a following quarter.

Figure 1 shows the percentage of active samples from old replications that were coded "out-of-business". While over the course of time it is natural for some operations to go "out-of-business", the percentage coded as "out-of-business" is questionably high. It is doubtful that all operations so coded actually went "out-of-business" since the earlier quarter contact; some may be miscoded and others may have been refusals in a previous quarter.

This study looked at the errors of reporting and coding "business" status and their effect on coverage. While some operations legitimately go "out-of-business" between quarters, and these can be substituted for, a substantial number of changes from quarter to quarter are errors in coding. For example, an operation is

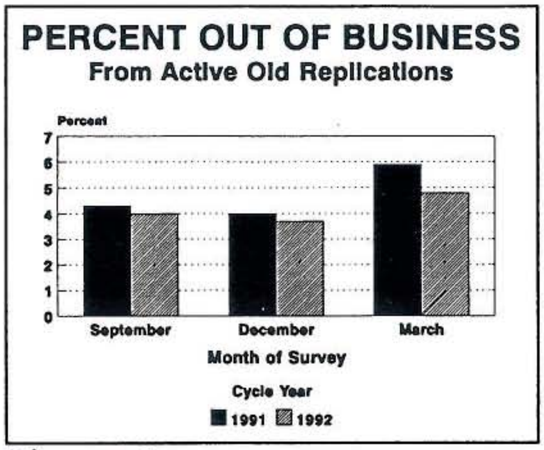


Figure 1

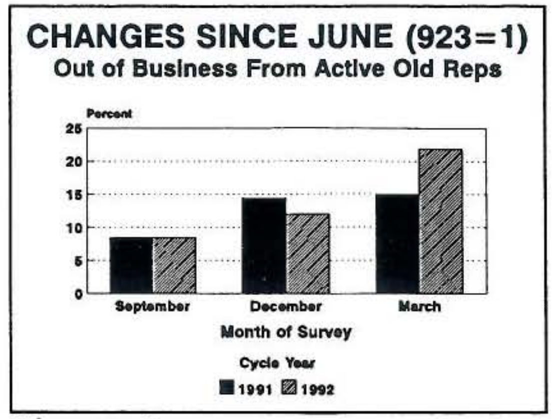


Figure 2

coded as "out-of-business" in a current quarter but "in business" for a previous quarter, when in fact it should have been recorded as "out-of-business" during the first quarter because the sample unit was a landlord. The converse can also happen when an operation is coded as "out-of-business" when it is really in business, since it continues to have potential for agricultural production.

In addition to being coded as "out-of-business", questionnaires are coded as to whether the sampled operation has changed since June 1. When an operation has gone "out-of-business" since June 1 item code box 923 on the face page of the questionnaire is coded a 1. Figure 2 shows the surprisingly low percentage of "out-of-business" operations from active old replicates that were coded as a change since June 1. Since all old replicates were reported in business during a previous quarter, we would expect nearly all current survey "out-of-business" reports to be changes since June 1. Therefore, if the current survey coding is correct, most operations were reported erroneously during the previous quarter. However, it is believed that code box 923 is frequently left uncoded. The coding of this box may be overlooked for old replications in part because it does not need to be coded for new replications.

Any operation that is reported as "out-of-business" is not surveyed again during that year's survey cycle. By NASS's definition, an "out-of-business" operation does not have any agricultural commodities and has no potential for agriculture during the rest of the year. Therefore, if correctly reported, it will have nothing to report in the following quarters and need not be surveyed. Each quarter more of these known zeros are accumulated, which creates problems when an operation is misreported as "out-of-business." State Statistical Offices (SSO) are instructed to review the known zero operations, but since not all are enumerated again some previous survey errors may go undetected. Any undetected misreporting of business status will cause a downward bias in the indications.

ANALYSIS OF THE DATA ADJUSTMENT FACTOR

In NASS's Agricultural Surveys, the Data Adjustment Factor (DAF) adjusts reported data for duplication and eliminates any positive data for operations that should not be summarized. Under normal situations the DAF is one, but it can have other values between zero and one. Common situations where the DAF is not one are: 1) an operation is duplicated in the same stratum (DAF=.5), 2) an operation is duplicated in a higher stratum (DAF=0), and 3) an operation is "out-of-business" (DAF=0). Table 1 shows the weighted (by the expansion factor for each design stratum) average of the DAF during the last two cycles of the Agricultural Surveys. The pattern of a decline is clear. One would expect to see some decline as operations go "out-of-business", but the amount of decline is of concern since it can have a large impact on survey results.

To determine the effect of the DAF on the expanded data, analysis was done comparing June to December expansions (Tables 2 & 3). The effects of the DAF, reported data, and the tract/farm weight factors were separated to assess the magnitude of each. This was done by calculating the normal June expansion, then using the information from those reporting in December to recalculate the June expansion. For example, the expanded data for an operation that was in business in June but not in December, would be positive in June

Table 1. Average Data Adjustment Factor

Cycle Year	Month of Survey			
	June	September	December	March
1991	.926	.899	.871	.849
1992	.946	.933	.907	.87

Table 2: Data Adjustment Factor (DAF) Effect on the Corn Planted Acreage Expansion for Survey Years 1991 and 1992.

Factor	June to December Comparable Reports for Factor					
	1991			1992		
	Ratio June to Dec.	Difference June - Dec. (000)	Difference as % of US	Ratio June to Dec.	Difference June - Dec.(000)	Difference as % of US
DAF	.95	-1,554	-2.0	.96	-1,108	-1.4
List Data	.99	-192	-0.2	1.00	-63	-0.1
Area Data and Weight	.99	-95	-0.1	1.07	464	0.6

Table 3: Data Adjustment Factor (DAF) Effect on the Total Hog Inventory Expansion for Survey Years 1991 and 1992.

Factor	June to December Comparable Reports for Factor					
	1991			1992		
	Ratio June to Dec.	Difference June - Dec.(000)	Difference as % of US	Ratio June to Dec.	Difference June - Dec.(000)	Difference as % of US
DAF	.95	-1,271	-2.3	.98	-615	-1.0
Data	.97	-685	-1.2	1.03	807	1.4
Area Weight	.99	-122	-0.2	.99	-99	-0.2

and zero for the recalculated June expansion with the December information. Comparable reports for a particular factor had to have usable factor information from both the June and December surveys.

Additionally, comparable reports for data and weight had to be in business both quarters. For the corn planted acreage expansion the area data and tract/farm weight factors can not be separated, because in June only tract data are reported while in December only farm data are reported. For total hogs, farm data are reported in both June and December, so comparisons between June and December of both data and weights can be made.

From Tables 2 & 3, we can see that in 1991 the DAF factor had a greater impact upon the difference in expansions between June and December than did the data or the weight. For example, the DAF factor resulted in a decrease in the U. S. expansion of 2 percent for corn planted acreage while the list/area data and weight factors decreased the expansion by only 0.2 and 0.1 percent, respectively. The situation for total hog inventory was similar, with the DAF decreasing the hog expansions by 2.3 percent. The size of the decrease due to the DAF factor is larger than the coefficient of variation for both estimates, illustrating the substantial effect the DAF has.

When we look at the 1992 analysis in Tables 2 & 3, we see that the effect of the DAF is about one half the size it was in 1991. This is encouraging, but the reason for the change in results is hard to determine. It is possible that training to make people aware of the DAF concerns has had a positive impact. One possible reason for the drop is the new list sampling unit/reporting unit association procedures, which half of the states used in 1992. These new procedures for associating reported data with sampled list names are called "operator dominant," as compared to the previous procedures which are referred to as "operation dominant." To see if this procedural change reduced the DAF impact, the effect of the DAF was compared between the two groups of states. Analysis showed there is only slight evidence that the DAF effect was smaller in the group with the new list dominant procedures.

To learn why operations were being coded as "out-of-business" we began to collect reasons. Observations made in Missouri during June 1992 were used to compile a preliminary list of these reasons. This list was used in Kansas during the December 1992 Agricultural Survey to code all questionnaires for which the reporting unit was coded "out-of-business" (i.e. item code 921=9). All old replications so coded were in business a previous quarter, while new replicates had not been surveyed. The reasons to be used in the coding were designed to differentiate between the situations expected between old and new replicate samples. The resulting list of reasons, while a starting point, turned out to be inadequate since too many reasons were grouped as "other."

To improve upon the reason coding, listings were sent to selected states after the December 1992 Agricultural Survey. State office personnel were to write out the reasons that operations changed their business status to "out-of-business". Table 4 is a compiled list of the reasons from four states. The most common reason was that incomplete information was obtained during the prior survey, because the respondent either refused or did not provide information about a partner involved in the operation.

Several of the reasons for operations being coded as "out-of-business" are related to the (small) size of operations and to whether they have agricultural potential. NASS defines as "out-of-business" an operation which has no potential for agricultural inventory or production during the remainder of the survey year. With this definition, no operation with potential for agricultural commodities should be coded as "out-of-business". While these operations may have nothing to report for any particular quarter they may have agricultural inventory or production during a subsequent quarter.

From the Table 4 list we can not tell directly whether the change in business status occurred after June 1 or was simply not picked up during a previous quarter. We can presume that some reasons, like 'landlord only', reflect situations which were not picked up in a previous quarter. Others, like 'sold farm', may or may not represent actual changes since June 1. If the change occurred after June 1 then the selected unit would be a candidate to be substituted for. If there is not an actual operation change, then there is a mistake in one quarter or the other. This may result from the respondent failing to answer correctly, some recording error, erroneous office coding, or one of many other possibilities.

Table 4: Detail of Reasons for Old Replications Coded as "Out-of-Business"

Number Times Occurred	Reason
37	Previously refusal and status not determined
15	Partner reported in higher strata
12	Partner reported in same strata
11	June with potential only.
8	Landlord only: incorrectly reported in previous quarter
7	Turned over to someone else
7	Sold farm
6	Name on label does not farm
5	Reported crops or livestock earlier, and reported none now

4	Minor crops or a few livestock only in previous survey
4	Turned over to son
4	Deceased
4	Retired
3	Land is now idle
3	Valid "out-of-business" (reason unknown)
3	Box 921 coded in error in current survey
3	Land is now rented, operated it previous quarter
2	CRP operator which should not be coded "out-of-business"
2	Miscoded multiple operations
2	Operator lied on previous report
2	Farm operated by someone else
2	Previously reported as 2 operations, actually only 1
2	Name correction on area frame, now OL
2	Partner strata boxes coded incorrectly
1	Chicken contractor only
1	Works on another farm only
1	Wrong name collected on June tract
1	Grain Co. only

DISCUSSION AND RECOMMENDATIONS

There are many causes for the DAF decline. Some of the decrease is valid and expected since operations will always be going "out-of-business", but some is due to survey error. The many causes increase the complexity of determining what needs to be done. The evidence suggests that the DAF decrease is large, meriting further analysis. Education and awareness can reduce errors. Procedural changes in coding to distinguish the difference between reporting errors and valid changes may provide better indications. Collecting more reasons for operations coded as "out-of-business" may give further insight, while measuring and adjusting for the DAF and the use of ratio estimates based on operations whose DAF did not change may need to continue.

There already have been efforts to educate people about the DAF. During the 1991 Midyear Survey Training, a session was conducted which provided DAF averages and comparisons between June and December expansions. This awareness may have made a difference since the decrease in the DAF in 1992 was about one half what it was in 1991.

Based on these results, I recommend continued, enhanced training with each state to examine their unique problems and further reduce the DAF dilemma. This education could be done during the advanced mid-year workshops. Statisticians in each state office could compile a list of reasons why some of their operations were coded "out-of-business". This list could then be the subject of small group discussions, probing for solutions.

I recommend the coding scheme for the "change since June box" be modified to improve the accuracy of its coding. Procedures that would require it be coded for all "out of business" operations would prevent it from being ignored. Once accurate information is obtained, ratios to a previous quarter could exclude illegitimate changes.

Another way to reduce the number of old replicate samples inappropriately being coded as "out-of-business" is by using historic data. When a respondent responds that they do not have the items of interest, we could then verify that they no longer have the items reported previously. This would be especially beneficial on CATI/CAPI.

I recommend we look more closely at the "out-of-business" operations and assess whether data compensation is being realized through the use of the current substitution procedures. This is the thrust of a separate research activity currently being addressed in NASS.

BIBLIOGRAPHY

Dillard, Dave. (1993) "Design Specifications to Standardize Survey Procedures." National Agricultural Statistics Service, United States Department of Agricultural, February 1993.

Jones, Ned. (1988) "QAS Substitution Analysis." Estimates Division, National Agricultural Statistics Service, United States Department of Agricultural, NASS Staff Report Number SMB-88-04.

ESTIMATING AGRICULTURAL LABOR TOTALS FROM AN INCOMPLETE LIST FRAME

D. Scot Rumburg, Charles R. Perry, Raj S. Chhikara*, and William C. Iwig, USDA/NASS
Charles R. Perry, Research Division, 3251 Old Lee Hwy., Room 305, Fairfax, VA 22030

KEY WORDS: Post-stratification, Multiple Frame, ratio estimate

ABSTRACT

The National Agricultural Statistics Service previously has conducted monthly labor surveys to estimate the number of total agricultural laborers. It employs a multiple-frame approach, using both a list and area frame. The list frame is highly efficient in sampling the target population of agricultural operations but does not have complete coverage of that population. The area frame covers all agricultural operations but is relatively inefficient in sampling those operations. An approach utilizing population count estimates from an initial area sample and post-stratified estimates from the monthly list sample has been investigated as a method for improving the precision of the survey estimate while reducing area frame respondent burden. Preliminary results indicate that survey to survey ratios of post-stratified list-only estimates can produce estimates which are comparable to current multiple frame estimates in both level and variance.

1. INTRODUCTION

A multiple frame approach, employing both a list and an area frame, has long been a cornerstone for many of the agricultural surveys which are conducted by the National Agricultural Statistics Service (NASS). Area frame responses often account for a majority of the total variance for multiple frame estimates but only a small part of the total indication. For this reason and others, it was recommended that a study be initiated to investigate alternatives to the current multiple-frame approach for administering surveys. A post-stratification approach whereby list respondents could be used to represent the entire target population, was recommended for consideration (Vogel, 1990a, 1990b and 1991). Kott (1990a and 1990b) elaborated on the proposal and outlined the two model-based estimators, their variance and potential bias. Perry, et al. (1993) provide an estimation method for the variance of a generalized post-stratification estimator based on its linear approximation using a Taylor Series expansion. Survey data from the California Agricultural Labor Survey series from July 1991 through June 1992 were used to investigate the alternative estimators.

2. METHODOLOGY

2.1 NASS Survey Methodology

NASS conducts numerous surveys with regard to agricultural commodities and related subjects. The majority of these surveys employ a multiple-frame (MF) methodology using both a list frame and an area frame. The list frame is stratified based on known data about agricultural operations with regard to the survey item(s) of interest. The list frame is not a complete listing of all agricultural operations. For the 1992 survey year beginning in June, the entire NASS list frame is estimated to contain 56% of all agricultural operations (often referred to simply as farms) and 81% of all land in farms. The area frame is stratified based on the agricultural intensity of a region. Unlike the list frame it has complete coverage of all agricultural operations in the U.S.

All reporting units (agricultural operations) in the June area survey (JAS:A) are classified as either overlap (OL) or as non-overlap (NOL) with the list frame. All operations found to be NOL are divided into several sampling pools to be used in follow-on surveys for the year. The list frame takes precedence over all OL operations when a MF estimate is calculated. A MF estimate is obtained by summing the list frame sample component estimate with the area frame's NOL sample component estimate. In most cases, the list frame provides about 75% of the total MF estimate while the NOL component adds only the remaining 25%. However, the NOL estimate is often a major contributor to the overall variance of the MF estimate, due to both the high variability of sampled units for many commodities and the sizable sample weights associated with small sampling fractions. The post-stratification approach investigated in this paper is an attempt to improve the reliability of the NOL component of MF estimates.

2.2 Post-Stratification Methodology

The proposed list-only estimator based on modeling of the NOL population represents a departure from the present NASS survey design and estimation methodology. Three factors motivate use of list only estimators: (1) the NOL sample units are highly burdened, (2) the current NOL estimates are often

* Raj S. Chhikara is Professor, Division of Computing and Mathematics, University of Houston - Clear Lake, Houston, Texas 77058

unreliable, and (3) the presence of NOL sample units increases the complexity of a survey.

Post-stratification for the Agricultural Labor Survey (ALS) was based on three classification variables: (1) The peak number of agricultural workers an operation expected to have over the course of a year (Peak), (2) the annual farm value of sales for agricultural goods (FVS), and (3) the type of farm operation (FType). These classification variables were selected based on their ability to describe distinct post-stratum populations and to correlate with the number of hired agricultural workers, which is the variable of interest. Basic strategy to obtain homogeneous post-strata populations involved selecting class boundary values for the two numerical classification variables (Peak and FVS), and creating combinations of the third categorical variable (FType). No more than twelve total post-strata could be created in order to maintain adequate sample counts for all post-strata across all surveys. Depending on cutoff values and FType groups selected, fewer post-strata could be constructed. An attempt was made to maintain a minimum of 20 respondents per post-stratum for all post-strata, though this was not always possible.

2.2.1 Post-Stratified Estimators

Post-stratification is often used as a variance reduction tool in a design unbiased survey. It can also compensate for the undercoverage of a target population by a particular selected sample. Both uses are employed for the approach explored in this paper. First it is hoped more homogeneous populations are produced with post-strata, resulting in variance reduction. Second, the list frame is used exclusively as the selected sample for follow-on surveys, resulting in undercoverage (actually non-coverage) of the NOL.

Once selected, the list sample is post-stratified to obtain post-stratum estimates. In the case of unweighted list responses, the estimator of the characteristic of interest Y is of the form:

$$\hat{Y}^{PS}_{(unex)} = \sum_{\substack{all\ k \\ post\text{-}strata}} (\hat{N}_{k(J)}) \cdot \left(\frac{1}{n_k} \sum_{i \in U_k} y_i \right) \qquad (Eq.1)$$

where

$\hat{N}_{k(J)}$ = k^{th} post-stratum population size estimate from the June survey (JAS:A),

n_k = k^{th} post-stratum sample size, and

U_k = the set of all useable sample reporting units in the k^{th} post-stratum.

Similarly, a weighted estimator of Y is of the form:

$$\hat{Y}^{PS}_{(expd)} = \sum_{\substack{all\ k \\ post\text{-}strata}} (\hat{N}_{k(J)}) \cdot \left(\frac{\sum_{i \in U_k} w_i y_i}{\sum_{i \in U_k} w_i} \right) \qquad (Eq.2)$$

where

w_i = i^{th} sample reporting unit weight, and other variables are defined as in Equation 1

For each choice of $\hat{Y}^{PS}$ one can compute a ratio and ratio expansion based on a combined ratio.

3. RESULTS

3.1 Preliminary Research - Simulation Studies

Simulation studies provided a theoretical perspective on several aspects of the post-stratification methodology. Approximate variance estimates were derived and evaluated. These numerical evaluations showed that the performance of a post-stratified estimator is largely a function of the sample size used to estimate the post-stratum sizes, the sample size used to estimate the post-stratum means of the variable of interest, and the ratio of these two sample sizes. The relative efficiency of the post-stratified estimators all increased as the ratio of the two sample sizes increased. Given the sample size for the follow-on survey, the sample size for the base survey should be at least twice as large for gains in efficiency. Moreover, for post-stratification to be effective, the entire sample size in the follow-on survey should be at least 50 (preferably much larger) with the sample size in all post-strata at least 10 (preferably 20 or more).

3.2 Comparison of List and NOL Respondents

Table 1 shows that the NOL has a lower average estimate within nearly all post-strata for California, whether one compares weighted or unweighted responses. Particularly troubling are the large FVS post-strata with open-ended peak workers (5 or more) and specifically the fruit, nut and vegetable post-stratum. The few NOL respondents which fell into this category had many fewer hired workers than did their list counterparts. The high FVS post-stratum, with Peak **5+** and FType **Crop & Misc** produced a larger NOL average hired workers than did the list and was due to one large NOL respondent reporting 391 hired workers.

Table 1 also characterizes the difference between weighted and unweighted averages. Unweighted averages are consistently higher than weighted averages for both list and NOL respondents for nearly all post-strata. Since operations with larger numbers of hired workers are sampled at a higher rate, and because operations with larger numbers of workers tend to represent fewer number of farms, the sampling weights are negatively correlated with the number of hired workers, the variable of interest. This situation occurs even within post-strata. The negative correlation of weights and number of hired workers within post-strata suggests that the unweighted average will tend to overestimate the number of hired workers per farm for both frames.

TABLE 1. Counts and Mean Number of Hired Workers Within Post-Strata For the California July 1991 Agriculture Labor Survey

Survey Post-strata Definitions			Counts		Weighted Mean		Unweighted Mean	
FVS	FType	Peak	List	NOL	List	NOL	List	NOL
$1-50K	Crops&Misc	0-4	49	70	0.28	0.12	0.24	0.24
$1-50K	Crops&Misc	5+	1	1	0.00	0.00	0.00	0.00
$1-50K	Veg,Frt&Nut	0-4	70	79	0.21	0.11	0.17	0.13
$1-50K	Veg,Frt&Nut	5+	28	15	2.03	0.27	6.89	0.40
$1-50K	Dairy,Poultry, GrnHse&Nursry	0-4	4	0	1.36	.	0.75	.
$1-50K	Dairy,Poultry, GrnHse&Nursry	5+	0	0	.	.	.	.
$50K+	Crops&Misc	0-4	56	35	1.11	0.39	0.93	0.69
$50K+	Crops&Misc	5+	59	17	7.90	15.10	13.80	35.10
$50K+	Veg,Frt&Nut	0-4	57	15	0.70	0.48	0.77	0.67
$50K+	Veg,Frt&Nut	5+	249	38	15.30	5.29	38.80	16.30
$50K+	Dairy,Poultry, GrnHse&Nursry	0-4	30	0	1.15	.	1.30	.
$50K+	Dairy,Poultry, GrnHse&Nursry	5+	63	5	22.90	16.30	33.70	19.40

***Cell counts and means for the weighted and unweighted response values by frame.** Note that the NOL cell averages tend to be smaller than the list averages and that the weighted cell averages tend to be smaller than the unweighted averages.*

3.3 Overall Performance of the Estimators

3.3.1 Post-Stratified Estimators

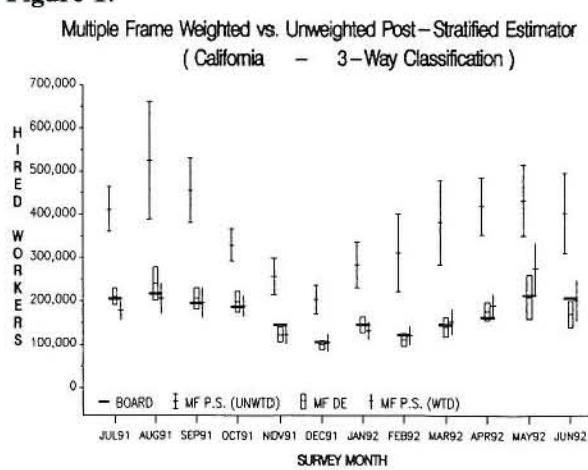
The combinations provided by selecting unweighted or weighted averages and an ability to select for list-only, NOL-only or both respondent types, produced six possible post-stratification estimators to study and evaluate. The NOL-only estimators were used only in conjunction with list-only estimators to provide comparative differences between the two frames on a state level basis. The MF post-stratified estimators were used to evaluate changes in variance due to list-only post-stratification.

Not surprisingly, it was found that the unweighted estimator consistently overestimated the actual labor force by a large margin (recall Table 1). The estimators using unweighted survey values produced the largest biases of all the estimators. Use of weighted survey values produced adequate, though somewhat more variable, estimates when compared to MF survey design direct expansion (MF DE) estimates. Since much of the post-stratification information is included in the list survey design (FVS and FType) and because the bulk of the ALS estimate comes from the list, it is not surprising the weighted MF post-stratified and the MF DE estimates are comparable.

Figure 1 depicts the level of bias produced by using a strictly unweighted post-stratified estimator and compares survey estimates across the 1991 ALS series year. For this and all succeeding graphs of this type, the vertical length of each estimate represents one standard error from the survey estimate in either direction. In some extreme cases the length in one or both directions has been truncated.

Also not surprising, given the post-stratum mean differences as shown in Table 1, it was found that the list-only estimator consistently overestimated the actual number of laborers while the NOL-only underestimated the actual labor force number. Figure 2 illustrates graphically the problems inherent in the weighted list and NOL-only post-stratified estimators, again comparing survey estimates to the Agricultural

Figure 1.



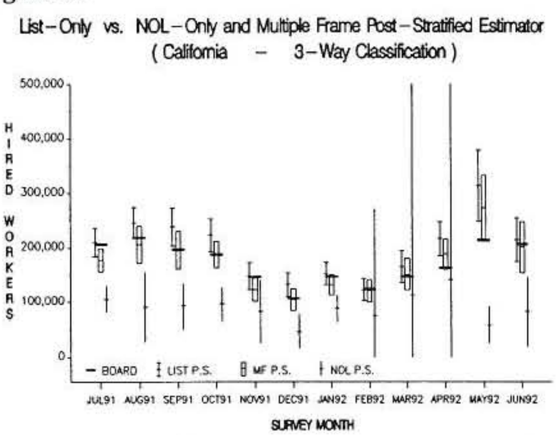
(Vertical Symbol Length Represents Two Standard Errors)

Comparison of Weighted versus Unweighted Post-Stratified Estimators.

Statistics Board as well as to the combined MF estimate across the 1991 ALS year.

Overall, list-only post-stratification CVs for California were mostly comparable with original MF DE CVs. This occurs for the most part because list-only post-stratified estimates generally are larger than the survey indication and have more variance introduced through the use of estimated June population counts. This leaves the overall percentage error of the total (CV) roughly equal to the MF DE CV. One must remember however, that the computed variance underestimates actual variance by as much as 10% resulting in a CV increase of approximately 5% since the $\sqrt{1.1} = 1.049$. For purposes of this report however, all CVs displayed will be the actual value computed with no compensation for bias. For California, the average CV for the weighted list-only post-stratification estimate for the survey year 1991 averaged 15.6%. This compares to an average MF DE CV for California of 14.2%.

Figure 2.



Comparison of the List-Only, NOL-Only and Multiple Frame Post-Stratified Estimators.

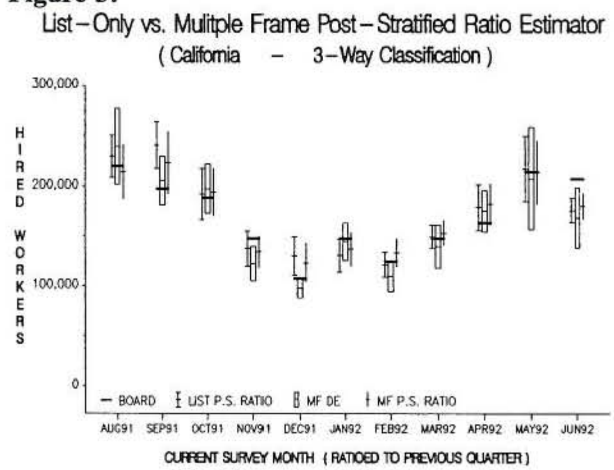
3.3.2 Ratio Estimators

Post-stratified combined ratio expansion estimates were calculated using MF and list-only data. In addition, a combined survey design ratio expansion estimate was computed using list-only data. Eleven monthly estimates were produced over the survey year for each estimator since a ratio estimate for July 1991 was not feasible. The ratio estimators were produced using only matched useable reports from both surveys.

3.3.2.1 Post-Stratified Ratio

Ratio expansion estimates were obtained using a combined post-stratified ratio estimator and the three-way post-strata classification scheme. Post-stratified survey total estimates using either list-only or MF respondents were constructed, and the results are shown in Figure 3 alongside the actual MF DE and the Board number for that month. The weighted list-only post-stratified survey total ratio tracks well with the Board estimate and, in fact, seven of the eleven ratio expansion estimates obtained for California were closer to the Board estimate than the MF DE indication. The average CV for California was 11.3% for the list-only ratio expansion estimate, which was less than the MF DE average CV of 14.6% over the same eleven surveys.

Figure 3.

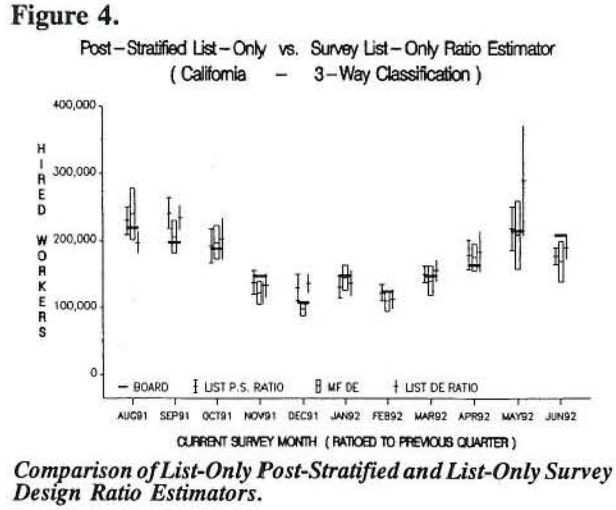


Comparison of the List-Only and Multiple Frame Post-Stratified Ratio Estimators.

3.3.2.2 Survey Design List-Only Ratio

Figure 4 compares the three-way post-stratified list-only combined ratio expansion with the survey design list-only combined ratio expansion. The post-stratified ratio estimator uses a weighted ratio, accounting for differences in farm numbers across post-strata. It is this difference which makes the post-stratified combined ratio estimator a more accurate estimator than the survey design combined

Figure 4.



Comparison of List-Only Post-Stratified and List-Only Survey Design Ratio Estimators.

ratio estimator. The two estimators have about the same precision.

3.4 Summary of Results

The combined three-way post-stratified list-only ratio estimator seemed to provide a viable estimate for the total number of hired workers indication. Though the post-stratification model is somewhat complex and would have to be optimized for each state or region, it does fulfill the objective of using a sample which ignores a subgroup, specifically the NOL.

4. CONCLUSIONS

For the Agricultural Labor Survey, there appear to be differences in mean values of list and NOL respondents within post-strata. Also, the sample design produces negative correlations between the sample weight and the response within post-strata. These two factors make the unweighted post-stratified estimator biased. Though bias is reduced in the case of the weighted post-stratified estimator, differences between weighted list and NOL respondents still exist within post-strata. Ratio expansion estimators, however, appear to avoid these problems and may have potential within the NASS framework.

The list-only combined ratio expansion estimator using three-way post-stratification appears to model the NOL adequately, while reducing variances on average. However, development of post-strata for individual states and regions would be a time consuming job and would involve reworking of the current survey summary system. Additionally, an estimator that uses only list respondents will probably be biased and must be cautiously approached and monitored if any list-only estimator were to become operational.

One problem with the post-stratified estimators investigated here is the estimated farm counts from the JAS:A. These counts are estimated using the area weighted estimator and tend to be quite variable. The inaccuracies can be corrected to some degree by using the MF population estimate. Any variability in the counts translates to higher overall variances of the post-stratified estimates. The post-stratified ratio estimators reduce the magnitude of this problem, but more accurate population estimates would surely help these estimators also.

5. REFERENCES

Flores-Cervantes, Ismael; 1990a; **Simulation Results: A Preliminary Study of the "Strawman" Estimator and Derived Formulas**; November 21, 1991; NASS Internal Correspondence.

Flores-Cervantes, Ismael; 1990b; **Effect of an Incomplete List on the Design Based "Strawman" Estimator**; November 26, 1991; NASS Internal Correspondence.

Flores-Cervantes, Ismael; 1990c; **Estimating the Number of Farms in a Post-Stratum using the June Agricultural Survey**; December 2, 1991; NASS Internal Correspondence.

Kott, Phil; 1990a; **Some Mathematical Comments on Modified "Strawman" Estimators**; NASS Internal Discussion Paper.

Kott, Phil; 1990b; **Comparing "Strawman" and "Weighted Tract" Estimators in June**; NASS Internal Discussion Paper.

Nealon, John Patrick; **Review of the Multiple and Area Frame Estimators**; SF&SRB Staff Report No. 80; Wash., D.C. March 1984.

Perry, Charles R., Chhikara, Raj S., Deng, Lih-Yuan, Iwig, William C., and Rumburg, D. Scot; 1993; **Generalized Post-Stratification Estimators in the Agricultural Labor Survey**; SRB Research Report Number SRB-93-04, Washington, D.C., July 1993.

Rumburg, D. Scot, Perry, Charles R., Chhikara, Raj S., and Iwig, William C.; 1993; **Analysis of a Generalized Post-Stratification Approach for the Agricultural Labor Survey**; SRB Research Report Number SRB-93-05; Washington, D.C., July 1993.

NASS Program Planning Committee, 1992; **Minutes of the Program Planning Committee of the National Agricultural Statistics Service**; July 14-16, 1992.

Turner, Cheryl L.; **Evaluation of Estimation Options for the Monthly Farm Labor Survey**; SRB Research Report Number SRB-91-12; Washington., D.C., December 1991.

Vogel, Frederic A; 1990a; **"Strawman" Proposal for Multiple Frame Sampling**; October 3, 1990; NASS Internal Memo.

Vogel, Frederic A; 1990b; **"Strawman" Proposal**; November 20, 1990; NASS Internal Memo.

Vogel, Frederic A; 1991; **A New Approach to Multiple Frame Sampling and Estimation**; January 16, 1991; NASS Internal Discussion Paper.