DETECTING A ROTATIONAL BIAS ON MULTIPLE FRAME SURVEYS

by

Kathy Burstrom Hall

Statistical Research Division
Economics, Statistics, and Cooperatives Service
U.S. Department of Agriculture
Washington, D.C. 20250

December 1979

Detecting A Rotational Bias on Multiple Frame Surveys. By Kathy Burstrom Hall; Statistical Research Division; Economics, Statistics, and Cooperatives Service; U.S. Department of Agriculture; Washington, D.C. 20250; December, 1979.

ABSTRACT

An analysis of hog data from six states shows that a rotational bias had an important effect on list estimates in only one state. One other state had a significant difference in nonresponse rates due to rotation. Thus, although rotational bias does not appear to be a frequent problem for list estimates, even as an occasional problem it may indicate the need for rotating the sample more often on list surveys.

Key words: rotational bias, panel bias, conditioning effect.

ACKNOWLEDGEMENTS

The author thanks the office personnel and enumerator staff in the states of Illinois, Indiana, Iowa, Minnesota, Missouri, and North Carolina for their assistance on this research project.

CONTENTS

DETECTING A ROTATIONAL BIAS ON MULTIPLE FRAME SURVEYS

Background

A rotational bias is the result of using the same sample units for several surveys -- e.g., interviewing the same operators from quarter to quarter to make hog estimates. The conditioning effect of a previous survey may cause a sample to yield different estimates than the estimates from a newly selected sample. Six states cooperated with the Statistical Research Division during the September, 1978 Multiple Frame Hog Survey to determine if a rotational bias affects list estimates of hogs. These six states were Illinois, Indiana, Iowa, Minnesota, Missouri, and North Carolina.

Several previous studes [1,4,7] motivated this research about rotational bias. One study [7] by Hill and Rockwell reported on a research project in Ohio and Wisconsin during the June, 1976 Multiple Frame Hog Survey. The purpose of that study was to evaluate an alternative questionnaire. In Ohio the estimate of total hogs from the alternative questionnaire was significantly larger than the estimate from the operational questionnaire. However, in Wisconsin there was no significant difference. Ohio used a newly selected sample to apply the alternative questionnaire while the sample that received the operational questionnaire had been contacted in two previous quarters. Both samples in Wisconsin were newly selected. Thus, the significant difference in estimates for Ohio may have been the result of a rotational bias.

In another study [4] Ford discussed the effects of rotational bias on the area frame. This study showed that the level of tract data reported by respondents was not affected by the length of time they were included in the sample. However, the nonresponse rate and nonoverlap classification were subject to a rotational bias.

Other survey organizations have also found rotational bias to be a problem. For example, the Bureau of the Census has conducted several investigations of this problem. Finkner and Nisselson [3] discussed these investigations in a report entitled "Some Statistical Problems Associated With Continuing Cross-Sectional Surveys". In that report the work of Bailar [1] on the Current Population Survey and Woltman and Bushery [9] on the National Crime Survey document the effects of a rotational bias on two major surveys.

Collection of Data for Analysis

To assess the level of rotational bias in six states, ESCS collected data on two groups during the September, 1978 Multiple Frame Hog Survey. One group -- called the old group -- was a 40 percent subsample of the operational sample. (The other 60 percent was used for another research experiment [6].) The old group had been contacted on hog surveys for the previous two quarters. The second group -- called the new group -- was newly selected from the list frame, i.e. the group had not been contacted on any previous hog surveys. The new group was one half the size of the old group.

TABLE 2

Mean Number of Hogs and Response Rate by Group

Six States

Group	All Data					Positive Data			
	Total Hogs	Previous Farrowings	Expected Farrowings First Quarter	Expected Farrowings Second Quarter	Nonresponse Rate	Total Hogs	Previous Farrowings	Expected Farrowings First Quarter	Expected Farrowings Second Quarter
	(Mean)	(Mean)	(Mean)	(Mean)		(Mean)	(Mean)	(Mean)	(Mean)
Previously Contacted	105.0	5.7	6.9	5.5	0.17	258.1	14.2	16.6	13.9
First Contact	105.7	6.2	6.7	5.3	0.17	251.6	14.7	15.9	13.1
All Groups	105.3	5.9	6.8	5.4	0.17	254.9	14.4	16.2	13.5

TABLE 3

Significance Levels of the Data When Testing the Hypothesis that There Is No Rotational Bias [Previously Contacted Vs. First Contact]*

State	All Data					Positive Data (Outliers Removed)			
	Total Hogs	Previous Farrowings	Expected Farrowings First Quarter	Expected Farrowings Second Quarter	Nonresponse	Total Hogs	Previous Farrowings	Expected Farrowings First Quarter	Expected Farrowings Second Quarter
	α	α	α	α	α	α	α	α	α
Illinois	.39	.29	.91	.02*	.02*	.74	.58	.23	.23
Indiana	.55	.71	.59	.45	.89	.14	.44	.74	.29
Iowa	.96	.73	.81	.43	.82	.64	.69	.81	.37
Minnesota	.03*	.56	.01*	.41	.93	.02*	.07*	.01*	.25
Missouri	.67	.47	.40	.52	.16	.25	.58	.93	.85
North Carolina	.34	.37	.67	.99	.52	.69	.53	.14	.83
6 States Combined	.83	.51	.30	.86	.44	.30	.40	.19	.31

* Significance is indicated by a value ≤ .10.

References

1. 1975 Bailar, Barbara, "The Effects of Rotation Group Bias on Estimates from Panel Surveys", Journal of the American Statistical Association, Volume 70, pages 23-30.

2. 1976 Barr, Anthony J., James H. Goodnight, John P. Small and Jane T. Helwig, A User's Guide to SAS 76, SAS Institute, Raleigh, North Carolina.

3. 1977 Finkner, A.L. and Harold Nisselson, "Some Statistical Problems Associated with Continuing Cross-Sectional Surveys", presented at the Second Symposium on Survey Sampling at the University of North Carolina.

4. 1975 Ford, Barry, Rotation Group Effects in SRS Surveys, Economics, Statistics, and Cooperatives Service, Washington, D.C.

5. 1976 Gleason, C.P., A Rotation Sampling Plan For ESCS Quarterly Hog Surveys, Economics, Statistics and Cooperatives Service, Washington, D.C.

6. 1979 Hall, Kathy Burstrom and Barry Ford, The Effects of Data Collection Methods, Economics, Statistics, and Cooperatives Service, Washington, D.C.

7. 1977 Hill, George W. and Dwight A. Rockwell, Associating a Reporting Unit With a List Frame Sampling Unit In Multiple Frame Surveys -- Ohio and Wisconsin, Economics, Statistics, and Cooperatives Service, Washington, D.C.

8. 1975 Timm, Neil H. Multivariate Analysis with Applications in Educational Psychology, Brooks/Cole Publishing Co., Monterey, California.

9. 1975 Woltman, H. and J. Bushery, "A Panel Bias Study in the National Crime Survey", Proceedings of the Social Statistics Section, American Statistical Association, pages 159-167.

Appendix A

Variable means by state:

The following tables (A1-A6) provide the sample means for total hogs, previous farrowings, expected first quarter farrowings, expected second quarter farrowings, and the nonresponse rates in each state. Means are given for both "All Data" and for "Positive Data".

TABLE A1

Mean Number of Hogs and Response Rate by Group

Illinois

Group	All Data					Positive Data			
	Total Hogs (Mean)	Previous Farrowings (Mean)	Expected Farrowings First Quarter (Mean)	Expected Farrowings Second Quarter (Mean)	Nonresponse Rate	Total Hogs (Mean)	Previous Farrowings (Mean)	Expected Farrowings First Quarter (Mean)	Expected Farrowings Second Quarter (Mean)
Previously Contacted	128.4	7.2	9.7	5.6	0.19	334.7	18.5	23.3	15.6
First Contact	140.8	8.7	9.5	8.1	0.13	337.4	20.3	21.1	18.2
All Groups	134.6	8.0	9.6	6.9	0.16	336.0	19.4	22.1	16.9

TABLE A2

Mean Number of Hogs and Response Rate by Group

Indiana

Group	All Data					Positive Data			
	Total Hogs (Mean)	Previous Farrowings (Mean)	Expected Farrowings First Quarter (Mean)	Expected Farrowings Second Quarter (Mean)	Nonresponse Rate	Total Hogs (Mean)	Previous Farrowings (Mean)	Expected Farrowings First Quarter (Mean)	Expected Farrowings Second Quarter (Mean)
Previously Contacted	90.1	5.3	5.1	5.8	0.12	244.8	13.6	12.9	14.5
First Contact	79.2	4.9	4.8	4.6	0.15	191.3	11.9	12.4	11.9
All Groups	84.1	5.1	5.0	5.2	0.13	220.7	12.6	12.6	13.2

TABLE A3

Mean Number of Hogs and Response Rate by Group

Iowa

Group	All Data					Positive Data			
	Total Hogs (Mean)	Previous Farrowings (Mean)	Expected Farrowings First Quarter (Mean)	Expected Farrowings Second Quarter (Mean)	Nonresponse Rate	Total Hogs (Mean)	Previous Farrowings (Mean)	Expected Farrowings First Quarter (Mean)	Expected Farrowings Second Quarter (Mean)
Previously Contacted	185.1	9.1	11.0	8.9	0.25	382.3	19.1	22.6	19.0
First Contact	189.9	10.2	11.5	7.7	0.26	380.6	20.9	23.6	16.7
All Groups	187.5	9.7	11.2	8.3	0.25	381.4	20.0	23.1	17.9

TABLE A4

Mean Number of Hogs and Response Rate by Group

Minnesota

Group	All Data					Positive Data			
	Total Hogs	Previous Farrowings	Expected Farrowings First Quarter	Expected Farrowings Second Quarter	Nonresponse Rate	Total Hogs	Previous Farrowings	Expected Farrowings First Quarter	Expected Farrowings Second Quarter
	(Mean)	(Mean)	(Mean)	(Mean)		(Mean)	(Mean)	(Mean)	(Mean)
Previously Contacted	75.5	4.7	5.5	4.7	0.18	180.8	10.8	13.0	11.3
First Contact	62.6	4.5	3.5	4.4	0.23	145.5	9.3	7.6	9.6
All Groups	69.1	4.6	4.5	4.5	0.20	163.1	10.0	10.3	10.4

TABLE A5

Mean Number of Hogs and Response Rate by Group

Missouri

Group	All Data					Positive Data			
	Total Hogs (Mean)	Previous Farrowings (Mean)	Expected Farrowings First Quarter (Mean)	Expected Farrowings Second Quarter (Mean)	Nonresponse Rate	Total Hogs (Mean)	Previous Farrowings (Mean)	Expected Farrowings First Quarter (Mean)	Expected Farrowings Second Quarter (Mean)
Previously Contacted	61.0	3.7	4.6	3.2	0.16	183.8	10.9	12.9	9.9
First Contact	61.4	4.0	4.8	3.4	0.11	200.5	10.9	14.6	10.2
All Groups	61.2	3.8	4.7	3.3	0.13	192.2	10.9	13.7	10.1

TABLE A6

Mean Number of Hogs and Response Rate by Group

North Carolina

Groups	All Data					Positive Data			
	Total Hogs	Previous Farrowings	Expected Farrowings First Quarter	Expected Farrowings Second Quarter	Nonresponse Rate	Total Hogs	Previous Farrowings	Expected Farrowings First Quarter	Expected Farrowings Second Quarter
	(Mean)	(Mean)	(Mean)	(Mean)		(Mean)	(Mean)	(Mean)	(Mean)
Previously Contacted	22.5	1.6	1.9	2.0	0.03	110.4	7.9	8.9	8.2
First Contact	27.0	2.0	1.8	1.9	0.05	118.9	8.6	7.8	8.5
All Groups	24.8	1.8	1.8	2.0	0.04	114.7	8.3	8.4	8.4

Appendix B

Data Analysis Techniques

1. Weighting the data:

Weights were assigned to each variable to reflect the relative expansion factors. Weights indicate relative importance of each stratum in each state across all six states.

2. Replication:

The data were ordered by state, group, stratum, crop reporting district, county and reporter. This ordering was systematically divided into ten replicates for analysis purposes. Mean values for each group were calculated within each replicate and combined in the following manner:

$\bar{x}_{sr}$ = sample estimate of the weighted mean in the r^{th} replicate, $r = 1, 2, \ldots 10$, in the s^{th} state, $s = 1, 2, \ldots 6$

$\bar{x}_{s}$ = sample estimate of the weighted mean in state s

$$\bar{x}_{s} = \frac{\sum_{r=1}^{10} \bar{x}_{sr}}{10}$$

$\bar{x}_{r}$ = sample estimate of the weighted mean of the r^{th} replicate over all states

$$\bar{x}_{r} = \frac{\sum_{s=1}^{6} \bar{x}_{sr}}{6}$$

$\bar{x}$ = sample estimate of the weighted mean over all states and strata

$$\frac{\sum_{s=1}^{6} \bar{x}_{s}}{6} = \frac{\sum_{r=1}^{10} \bar{x}_{r}}{10}$$

The unbiased estimate of the standard error for each state estimate is:

$$SE(\bar{x}_{s}) = \left[\frac{\sum_{r=1}^{10} (\bar{x}_{sr} - \bar{x}_{s})^2}{9(10)} \right]^{1/2}$$

An unbiased estimate of the standard error for an estimate of a 6-state total is:

$$SE(\bar{x}) = \left[\frac{\sum_{r=1}^{10} (\bar{x}_r - \bar{x})^2}{9(10)} \right]^{\frac{1}{2}}$$

Besides simplifying the calculation of standard errors, the use of replicate values in the statistical analysis:

1: assured equal cell sizes in the analysis of variance
2: yielded distributions which are fairly normal.

These benefits of replication make the analytical exploration of data from a complex survey design much more straightforward and accurate.

Univariate and multivariate tests were conducted using the SAS computer package. The processes involved in this analysis are outlined below, although more details are available in the SAS 76 User's Guide [2].

A general linear model was constructed to fit the values from the replications: $\underline{Y} = \underline{X}\,\underline{B} + \underline{E}$, where $\underline{E}$ is the residual error matrix, $\underline{Y}$ is the vector of estimated means for four (quantitative) hog variables, $\underline{X}$ is the data matrix and $\underline{B}$ is the vector of model parameters. The model parameter consist of two effects: the state and the group. The group effect refers to the impact of previous contacts on respondents. This is discussed in the background section of this paper.

The MANOVA option of the GLM procedure in SAS was used to test the hypotheses that there is no difference in the data between groups. Both univariate and multivariate tests were executed on the four hog variables, and an univariate test was executed on the nonresponse rates.

Wilk's Λ criterion, which was used for the multivariate tests, is essentially a multivariate extension of the F test used in univariate analysis of variance. Full details on Wilk's Λ criterion are in a book on multivariate analysis by Timm [8].