United States
Department of
Agriculture

Statistical
Reporting
Service

**SRS Staff
Report
Number AGESS 810615**

An Analysis of 1979 Nebraska Rotation Plan for the Quarterly Hog Survey

Chapman P. Gleason

AN ANALYSIS OF 1979 NEBRASKA ROTATION PLAN FOR THE QUARTERLY HOG SURVEY; by Chapman P. Gleason, Statistical Research Division, Statistical Reporting Service, U.S. Department of Agriculture, Washington, D.C. SRS Staff Report NO. AGESS810615. June 1981.

ABSTRACT

This paper summarizes procedures and estimators used in the analysis of a rotation sampling plan in Nebraska for the 1979 quarterly hog surveys.

Key words: Rotation sampling, successive sampling, panel studies, regression or ratio estimation, change estimation, double sampling.

This paper was prepared for limited distribution to the research community outside the U.S. Dept. of Agriculture. The views expressed herein are not necessarily those of SRS or USDA.

CONTENTS

An Analysis of 1979 Nebraska Rotation Plan for the Quarterly Hog Surveys

Chapman P. Gleason

INTRODUCTION

This paper is a summary of the procedures and estimates used in the analysis of a rotation sampling plan in Nebraska for the 1979 quarterly hog surveys (QHS). The plan grew out of earlier research documented in [2]. The 1979 Nebraska Quarterly Hog Surveys were the first time that this rotation plan was used on an operational basis for the list frame in the Hog Multiple Frame Survey.

The objectives (documented in [2]) of the proposed rotation plan are:

1. Provide consistent and stable estimates of both level and change,
2. Reduce respondent burden, and
3. Reduce sample sizes by using more efficient estimators.

This report will compare the different estimators to the direct expansion for the list frame only in the Quarterly Hog Multiple Frame Survey for the 1979 series of quarterly surveys in Nebraska.

The rotation plan for the 1979 series of quarterly surveys in Nebraska is described below. Four independent replications (S_1, S_2, S_3, S_4) were selected within each stratum. For any survey period, two replications were used to make up the following rotation pattern for the entire 1979 survey period.

December '78	March '79	June '79	September '79
S_1			S_1
S_2	S_2		
	S_3	S_3	
		S_4	S_4

For any survey period this replicated sampling plan assures a 50 percent overlap in the sample from quarter to quarter.

Table 1 gives the population size, sample size, and the replicate size for each stratum in Nebraska.

Table 1: Population Size, Sample Size and Replicate Size, 1979 Nebraska

Stratum Number	Stratum Description (no. hogs)	Population Size	Sample Size	Replicate Size
81	no livestock	11,171	164	82
82	no hogs	23,913	292	146
83	1-99	9,427	300	150
84	100-199	5,542	366	183
85	200-299	1,889	152	76
93	300-599	2,308	308	154
94	600-1249	623	124	62
95	1250-2999	113	56	28
96	3999-4999	47	24	12*
98	5000 +	15	15	15

* For June and September, the replication size is 11.

ANALYSIS

Estimators

Several different estimators were computed and compared for each survey period. The estimators were:

E1. The direct expansion using the entire list sample,

E2. The direct expansion using only the units (list names) matched between quarters,

E3. The direct expansion using only the units which were newly rotated into the sample,

E4. A separate composite ratio estimate in each stratum,

E5. A separate composite regression estimate in each stratum, and

E6. A combined composite ratio estimate over all strata.

A composite estimator is an estimator which combines the matched sample estimate of the mean/unit and the unmatched sample estimate of the mean per/unit together into a composite estimate of the mean/unit. Estimators E4 through E6 are all composite estimators.

The computational formulas for these six estimators and their corresponding variances are given in [2]. However, to facilitate further discussion, a

brief description of each of these estimators is given below. Consider the expression:

(1) $\Sigma N_h \bar{y}^*_h$, where N_h = population size in h^{th} stratum, and

$\bar{y}^*_h$ is defined in one of the following ways or each of the

estimators E1 through E5:

E1: $\bar{y}^*_h$ is the h^{th} stratum mean computed over the entire sample.

E2: $\bar{y}^*_h$ is the h^{th} stratum mean computed only on the units matched between survey periods.

E3: $\bar{y}^*_h$ is the h^{th} stratum mean computed for units that are newly "rotated-in".

E4: $\bar{y}^*_h$ is a separate ratio estimate of the h^{th} stratum mean computed only on matched units between survey periods.

E5: $\bar{y}^*_h$ is a separate regression estimate in each stratum computed from matched units.

The combined composite ratio estimate is computed as follows: Let

$\hat{R} = \frac{\Sigma N_h \bar{y}_{2mh}}{\Sigma N_h \bar{y}_{1mh}}$ where, for example using December and March data,

$\bar{y}_{2mh}$ = March mean of matched operations for h^{th} stratum

$\bar{y}_{1mh}$ = December mean of matched operations for h^{th} stratum, and

N_h = Population size for the h^{th} stratum

$\hat{R}$ is defined to be the estimate of the (change) parameter in the combined ratio estimate which is:

E6: $\hat{Y}_{ratio,\ comb} = \hat{W}\ \hat{R}\ \hat{Y}_{1m} + (1-W)\ \hat{Y}_{2u}$

where

$\hat{Y}_{1m}$ is the estimate of the total from the matched December sample,

$\hat{Y}_{2u}$ is an independent estimate of the March total from the new rotated in sample and

W is a weight (determined to be near ½) which is used to combine the two independent estimates of the total.

Comparison of Estimates

Each of the estimators was computed for the different survey periods and is presented in Table 2. Note that estimates are prepared only for the March, June and September quarters, since December is an entirely new sample. The composite estimators (E4 - E6) agree quite closely with the entire sample direct expansion for the list (E1). Sampling errors for the composite estimates ranged from 6 to 9 percent below the direct expansion sampling errors (list frame only).

Taking the E1 estimate as a fixed parameter, the difference between the levels of the half sample estimates (E2, E3) is about 4 percent for the March quarter, 3 percent for the June period, and 7 percent for the September period. This is well within sampling errors (4.3 percent) for the March and June period, but outside the entire-sample's sampling error (5.2 percent) for the September quarter. However, the 66.7 percent confidence intervals ($\pm$ 1 standard error) overlap for estimators E1 and E3 for the September survey period, which indicates the estimators are not significantly different. Thus, the "current" September half sample replicate (S_1) is more variable than the other replicates.

Cochran [1, pp. 170-172] shows that the precision of both the combined ratio and the separate ratio (and hence the separate regression) are better than the direct expansion. The separate regression is generally more precise than the combined ratio unless the stratum-by-stratum ratios are nearly equal. It can be seen in Table 3 that the stratum ratios (entered as percentages) are not nearly equal for any quarter, indicating that the separate regression would be the more precise estimator. The sampling errors in Table 2 support this; however, the gain in precision is marginal.

A measure of change which was computed was the ratio of the previous quarter's data to the current quarter's data for matched units on a stratum-by-stratum basis. This was computed for each survey period. This ratio estimates the percent change from the previous quarter on a stratum-by-stratum basis for those units responding both quarters. Table 3 gives the percent change in stratum means based on matched units.

Table 2: Estimates of Total Hogs and Pigs -- March, June and September 1979, Nebraska List Frame

Type of Estimator	MARCH			JUNE			SEPTEMBER		
	Est. (000)	S.E. (000)	C.V. (%)	Est. (000)	S.E. (000)	C.V. (%)	Est. (000)	S.E. (000)	C.V. (%)
E1. Direct Expansion (whole sample)	3,237	138.7	4.3	3,481	148.7	4.3	3,265	171.0	5.2
E2. Direct Expansion (Prev. half sample)	3,172 S_2	171.8	5.4	3,532 S_3	204.4	5.8	3,150 S_4	190.4	6.0
E3. Direct Expansion (Cur. half sample)	3,305 S_3	195.8	5.9	3,431 S_4	190.5	5.6	3,385 S_1	265.3	7.8
E4. Separate Ratio Composite	3,285	129.8	4.0	3,486	139.0	4.0	3,317	162.2	4.9
E5. Separate Regression Composite	3,248	126.5	3.9	3,490	137.1	3.9	3,338	160.2	4.8
E6. Combined Ratio	3,247	130.8	4.0	3,504	141.6	4.0	3,317	164.0	4.9

Table 3: Nebraska Total Hogs and Pigs Percent Change in Stratum Means for Matched Units

Stratum	December to March	March to June	June to September
	- - - - - - - - - -percent - - - - - - - - - -		
81	75	131	52
82	231	152	85
83	90	106	90
84	92	119	96
85	101	95	87
93	97	105	97
94	106	97	90

As can be seen from Table 3, there are considerable differences in the ratios between strata. For example, the 231 percent in stratum 82 indicates a sharp increase in the average number of hogs from December to March; whereas, stratum 81 showed a moderate decrease. This indicates that the combined ratio estimate (E6) would likely be less precise than the separate ratio estimate E4. This loss in precision, though not great, may be seen in Table 2 by comparing the standard errors of the estimates. Note that large changes between quarters on individual reports need to be checked to verify these large overall stratum changes.

The overall pattern in the ratios was upward from March to June, and downward from June to September for most strata. However, stratum 82 no hogs has the most amount of change; going from a 131 percent increase to 15 percent decrease in the December/March, June/September survey periods, respectively. This may be an indication that better survey controls and procedures are needed for this stratum or it may be just the nature of the "no hog" stratum, i.e., more variability of operations that go in and out of the hog business on a small scale. The reduced list concept where the "no hog" stratum is not sampled would eliminate this problem.

Table 4 gives overall percent change computed two ways:

1. Taking the ratio of the current quarter's direct expansion to the previous quarter's direct expansion, and

2. Estimating change as the change parameter, R, in the combined ratio estimate E6.

Table 4: Indications of Overall Percent Change from Previous Survey Period

Method	Dec/March	March/June	June/Sept.
		-percent	
Cur. Dir. Expan. / Pre. Dir. Expan.	-.5	8.0	-6.3
Combined Ratio Estimate ($\hat{R}$)	-2.0	11.0	-9.0

These overall change indications are within 1.5, 3.0 and 2.7 percentage points of one another for each survey period, respectively. Both methods of computation of change moved in the same direction for each survey period. For example, the ratio of the direct expansions indicated a higher inventory from March to June (8 percent higher), and the combined ratio estimate $\hat{R}$ also indicated a higher inventory (11 percent higher).

The combined ratio estimator of change ($\hat{R}$) showed a -2 percent change from December to March, compared to a -.5 percent change in the ratio of the direct expansions; a +11 percent change compared to a +8 percent from March to June; and a -9 percent change compared to a -6.3 percent from June to September. Thus, the $\hat{R}$ estimator of percent change indicates a greater degree of change in the inventory levels from survey period to survey period than the ratio of the direct expansions.

Hansen, Hurwitz, and Madow [4, pp. 175] state that generally if the correlations are high between survey periods (see Table 5) that a ratio estimate is more precise than taking the ratio of the current quarter's direct expansion to the previous quarter's direct expansion. Since the correlations in Table 5 are all high between survey periods, the estimated overall percent change with the combined ratio estimate given in Table 3 is more precise than the ratio of the direct expansions. In the long run, the $\hat{R}$ estimator of change would be more stable if "charted".

Optimum Percentage to Match

One of the primary components to estimate totals and change for periodic surveys using rotation sampling is how much of the sample to match between survey periods. Previous results [1] indicated a 50% matching percentage for the Quarterly Hog Surveys. However, this matching percent considered only one survey period to estimate the optimum percentage to match. We now have a sequence of four quarterly surveys in one state to investigate how stable this matching percentage is from quarter to quarter. The matching percentage is a function of the stratum correlation coefficient.

Table 5 presents the estimated correlation coefficients for total hogs and pigs between survey periods. For the majority of strata, the correlations remained fairly stable and strong between time periods by stratum. However, correlations from the March/June period compared to the June/September period were consistently smaller except for stratum 83, and 96. Something very unusual happened in stratum 81 between the December/March period when the correlation was .98 and then fell to .11 in the June/September period. Inspecting Table 3 we note that for the June/September time period the ratio was .52 and for March/June it was 1.31. These indicate there was at least one outlier in the June data for stratum 81. I did not go into the data to detect and delete the outlier(s) since this was the exact same data that the Nebraska SSO was using to set its estimate of inventory.

Table 5: Stratum Correlation Coefficient Between Total Hogs and Pigs

Stratum	Dec./Mar.	Mar./June	June/Sept.
81	.98	.83	.11
82	.71	.88	.80
83	.81	.61	.61
84	.81	.68	.64
85	.74	.71	.40
93	.91	.80	.78
94	.84	.93	.88
95	.95	.99	.63
96	.86	.98	.99
98	.99	1.00	1.00

To estimate the optimum percent to match for the rotation plan, Table 6 presents the optimum percentage to match between survey periods for total hogs and pigs. Computational formulas are given in [2]. As can be seen from the table, on a stratum-by-stratum basis 19 to 55 percent of the units should be retained between quarters with 37 percent the median over all survey periods for the regression estimate; 40 percent is the median over all survey periods for the ratio estimate. Assuming a survey response rate of 80 percent retention of 50 percent of the sample between quarters provides near optimum percent matched for either type of estimator. This verifies previous research done in [2] for one survey period. It can also be seen from Table 6 that the median optimum percentage to match remains fairly stable between survey periods.

Table 6: Optimum Percentage to March

Stratum	REGRESSION ESTIMATE			RATIO ESTIMATE		
	Dec/Mar	Mar/Jun	Jun/Sept	Dec/Mar	Mar/Jun	Jun/Sept
81	19	31	48	21	31	55
82	42	32	36	48	33	36
83	37	44	41	39	46	42
84	35	44	41	36	49	42
85	40	43	42	43	44	47
93	28	36	36	40	42	39
94	35	30	28	39	30	28
Median	35	36	41	39	42	42
Overall Median		37			40	

CONCLUSIONS/RECOMMENDATIONS

Discussions with the Nebraska SSO indicated that they liked the new sampling plan since it was easily managed in the field office and reduced respondent burden from 4 contacts per year to 2. The rotation plan is being further evaluated in the five western corn belt states (Nebraska, Iowa, Minnesota, Kansas and South Dakota) for the 1980 series of surveys.

Indications of quarter-to-quarter change using composite estimators were consistent in the direction of change with the ratio of the direct expansions. The rotation sampling plan provides additional estimators of both level and change in the off quarters (March and September), with no additional respondent burden; however, there is very little gain in precision of the composite estimators over the direct expansion estimator.

Analysis by stratum showed that ratios were not equal by stratum; hence, a combined ratio estimate would not be the best estimator in this situation. Since the separate regression estimate is more precise than the separate ratio estimate for a fixed sample size, the separate regression estimate is preferred.

Optimum percentage to match was also computed for both the separate regression and separate ratio estimates for a number of survey periods. The match percentage (50 percent) recommended in [2] was found to be nearly optimum for the separate regression as well as the separate ratio estimator.

In conclusion, I recommend that ESS adopt the rotation plan in the 14 hog multiple-frame states. Further, I recommend that the separate composite regression estimate be computed and charted for all states in the multiple frame program, especially in the March and September quarters where additional survey indications are lacking.

REFERENCES

[1] COCHRAN, W.G., "Sampling Techniques", 2nd Edition, John Wiley & Sons, New York, 1963.

[2] GLEASON, C.P., "A Rotation Sampling Plan for ESCS Quarterly Hog Surveys", Statistical Research Division, Economics, Statistics and Cooperatives Service, U.S. Department of Agriculture, Washington, D.C. 20250.

[3] GLEASON, C.P., "An Estimator of Change from an Updated Frame", Proceedings of the Survey Research Methods Section, American Statistical Association Meetings, San Antonio, TX. August 1980.

[4] HANSEN, M.H., Hurwitz, W.N., and Madow, W.G., "Sample Survey Methods and Theory", Volume 1, John Wiley & Sons, New York, 1953.