Food Marketing C

1973



# OPTIMUM SAMPLING OF DELIVERIES FOR MEASURING CHANGES IN FOOD PRODUCT MOVEMENT*

Thomas L. Sporleder, Robert E. Branson, and Charles E. Gates†

Measurement of food product movement is of interest to those engaged in market research, advertising, and food manufacturing or distribution. A particularly vexing situation exists when measurement of product movement is desired in a metropolitan market for a short time period (say, a month or less). For example, this situation arises when a short-term metropolitan market promotion for a particular product is conducted.1/ Auditing a sample of retail grocery stores in the market typically has been the method used for monitoring product movement. However, audits are not usually feasible when only short-term changes are desired because of the expense in collecting such data.

An alternative to the retail grocery store audit approach is to collect product delivery data. Such records are relatively inexpensive to collect since they may be obtained at chain warehouse level without necessitating the expense of individual store visits as is the case with audits. The geographic confines of a metropolitan market, however, require individual store delivery records rather than warehouse withdrawal data. For nearly any chain warehouse, the latter data would include deliveries to stores in a larger area than a metropolitan market.2/ Since individual store delivery data must be obtained, a sample size problem arises. The problem revolves around determination of the optimum composition and number of stores to include in the sample. Obviously, in any metropolitan market, a number of chains operate, each with a differing number of stores per chain. Should all chains be included in the sample or only

†Thomas L. Sporleder and Robert E. Branson are Associate Professor and Professor, Texas Agricultural Market Research and Development Center, Department of Agricultural Economics. Charles E. Gates is Professor, Institute of Statistics. All are of Texas A&M University.

Contributed paper, AAEA, Edmonton, Aug. 1973

some portion? If any particular chain is included, how many stores of that chain should be included?

The objective of this manuscript is to provide a technique for optimum sampling of individual store delivery data. This essentially involves estimating sample size requirements in terms of stores per chain and chains for some specified level of accuracy in measurement, given an estimate of variability in deliveries. The technique outlined is applied to estimate sample size requirements for cheese and butter deliveries in specific metropolitan markets.

## Methodology

The two basic approaches to any sample size problem are to minimize variance subject to a specified budget allocation for measurement or minimize sampling cost subject to some specified accuracy. The latter approach was chosen in the present instance since measurement accuracy is prerequisite to delivery data as a feasible alternative to store audits.

### The data

To estimate the normal variability of cheese and butter deliveries to supermarkets in metropolitan markets during a short time period, deliveries of both products were obtained by weeks from a selected group of cities representing varied geographic location and size. The cities chosen were Dallas, Texas; Omaha, Nebraska; Terre Haute, Indiana; and Toledo, Ohio. All data were collected for at least an eight week period for selected stores of all major chains in these cities during the fall of 1970.

First differences

The original data in terms of pounds of product delivered per store per week were transformed to first differences, reducing or removing any first-order autocorrelation which might exist. 3/ Eight weekly records for each store were included in the analysis. Seven one-week first differences were calculated for each store by:

$$(1) \qquad Y_{ij_t} = P_{ij_h} - P_{ij_{(h+1)}}$$

where $Y_{ij_t}$ = $t^{\underline{th}}$ first difference for store j of chain i

$P_{ij_h}$ = pounds delivered in period h to store j of chain i

and $t = 1, 2, \ldots, 7.$

$h = 1, 2, \ldots, 7.$

In this case, h corresponds to the number of one week records.

Analysis of variance

To estimate variance components, an unbalanced one-way analysis of variance (ANOVA) was utilized. The appropriate ANOVA model for a given first difference (with the t subscript omitted) is:

$$(2) \qquad Y_{ij} = \mu + c_i + \varepsilon_{ij}$$

where $Y_{ij}$ = first difference of deliveries (cheese, butter) for store j of chain i.

$\mu$ = grand mean

$c_i$ = effect of the $i^{\underline{th}}$ chain

$\varepsilon_{ij}$ = residual

and $i = 1, 2, \ldots, c$

$j = 1, 2, \ldots, s$

All effects in this model are regarded as random effects [8, pp. 2 - 9]. Also, disproportionate subclass frequencies exist since the number of stores per chain differ over chains. The expected mean squares for the ANOVA model incorporate a finite population correction (fpc) factor into the expected mean square calculations, Table 1 [3,5,9]. The fpc is necessary in order to reflect the finiteness of chains and stores within chains for any given metropolitan area.

Variance components for chains and stores/chain were estimated for each of the seven first differences calculated from one week data periods following Searle [8, pp. 34 - 67]. The seven variance components for chains were then averaged to produce one variance component for chains based on one week data periods.

Constrained optimization

Once variance component estimation is accomplished, an objective function defining costs of obtaining records can be minimized subject to a specified accuracy in terms of variability in the grand mean. Conceptualization of the problem is as follows:

(3) $$\text{minimize } c(K_c + sK_s)$$

subject to [1]:

(4) $$V(\bar{\bar{y}}) = \left(\frac{\hat{\sigma}_c^2}{c}\right)\left(1 - \frac{c}{N}\right) + \left(\frac{\hat{\sigma}_{s(c)}^2}{sc}\right)\left(1 - \frac{s}{M}\right) \leq V$$

where $0<c<N$, $0<s<M$, and where:

Table 1. ANOVA, random effects, unequal subclass numbers with finite population correction*

| Source of Variation | df | MS | E(MS) |
|---|---|---|---|
| Chains | $c - 1$ | $A_1$ | $\sigma^2_{s(c)}\left(1 - \frac{k}{M_1}\right) + k\sigma^2_c$ |
| Stores/Chain | $\Sigma(s_i - 1)$ | $A_2$ | $\sigma^2_{s(c)}$ |

*Notation in table:

$A_1$ = observed mean square for chains

$A_2$ = observed mean square for stores/chain

$k$ = coefficient for variance component for chains where [2,7]:

$$k = \frac{\left[\Sigma s_i - \frac{\Sigma s_i^2}{\Sigma s_i}\right]}{(c - 1)}$$

and where

$s_i$ = number of stores in the $i^{th}$ chain

$M_i$ = universe number of stores in the $i^{th}$ chain.

$c$ = sample number of chains

$s$ = sample average number of stores/chain

$N$ = universe number of chains

$M$ = universe average number of stores/chain [4/]

$K_c$ = cost of adding a chain to the sample [5/]

$K_s$ = cost of adding a store within a chain to the sample [5/]

$V(\bar{\bar{y}})$ = variance of the grand mean of period differences in deliveries per store per unit time

$\hat{\sigma}_c^2$ = variance component for chains

$\hat{\sigma}_{s(c)}^2$ = variance component for stores/chain

$V$ = specified accuracy in terms of variance of mean in deliveries per store per unit time.

Equation (3) reflects the total cost of generating records from a sample of chains and stores within those chains. Given c chains included in a sample from some metropolitan market area, the cost of obtaining chain cooperation in the market would be $cK_c$. Actually obtaining individual store records from an average of s stores per chain would add another $csK_s$ dollars to sampling cost.

The constraint, equation (4), is the expression for the variability of mean deliveries per store per period adjusted for a finite population. The fact that a finite number of chains exist in any one metropolitan market is reflected by the correction factor $(1 - \frac{c}{N})$, while the factor $(1 - \frac{s}{M})$ is for stores within chains. As c approaches N the variability attributable to that component approaches zero. Similarly, as s approaches M the variability attributable to that component also approaches zero.

Specifying accuracy

The constraint $V(\bar{\bar{y}}) \leq V$ of the above objective function requires specfication of V. The derivation of the specified accuracy is from a 95 percent confidence interval on mean deliveries per store per unit time. Let $\bar{X}$ represent some mean level of delivery per store per unit time, then a 95 percent confidence interval (C. I.), two-tailed, is:

(5) $$\bar{X} \pm (1.96)\ \hat{\sigma}_{\bar{X}}$$

Since specified accuracy is in terms of variance, and a change in $\bar{X}$ of no more than m percent is specified as the magnitude of change desired to be detected, V may be derived from equation (5) as:

(6) $$V = \hat{\sigma}^2_{\bar{X}} = \frac{(m\%)^2\ \bar{X}^2}{3.8416}$$

The Hartley-Hocking algorithm

Solution to the constrained optimization problem of equations (3) and (4) in the parameters c and s may be accomplished by utilizing convex programming which employs the Hartley-Hocking algorithm of tangential approximation [4,6]. Essentially, the problem is a nonlinear programming problem. Detail of the algorithm and the boundary constraints as formulated for this particular problem is found in [10].

Utilizing convex programming for a particular specification of accuracy (V) yields solutions in terms of number of chains (c) and number of stores per chain (s) necessary to obtain the specified accuracy and minimize cost. Obviously, the sample size required (i.e., the magnitude of c and s) depend upon the variance components and the specified accuracy, V. Also, the magnitude of V depends upon mean deliveries per store per period as noted above. This creates the opportunity for simulated solutions to sample size requirement via perturbation of V. Due to space limitations, data aggregation to larger sampling units will not be examined here but is reported in [10].

## Results

The procedure is to present results for cheese and butter deliveries by city with specified accuracy invariant. Then, sensitivity analysis is conducted by perturbation of V through changes in the m parameter of equation (6), and by utilizing various $\overline{X}$'s. The sensitivity analysis allows for comparison of results from various combinations of assumptions concerning parameters. This permits conclusions to be drawn about the relative significance of key parameters in affecting results.

### One week first differences

The variance components for chains and stores/chain for cheese and butter for one week first differences vary markedly among cities, Table 2.

Table 2. Estimated variance components with fpc for one week first differences, cheese and butter, by city

| City | Variance Component For: | | | |
|---|---|---|---|---|
| | Cheese | | Butter | |
| | Chains | Stores/Chain | Chains | Stores/Chain |
| | - number - | | | |
| Dallas | 9,572 | 249,627 | 4,284 | 20,372 |
| Omaha | 45,880 | 1,019,570 | 10,827 | 192,890 |
| Toledo | 33,283 | 109,984 | 24,328 | 66,311 |
| Terre Haute | 1,373,139 | 332,266 | 57,545 | 12,510 |

Source: Computed.

There is no general pattern exhibited by the variance components for either product over cities or for any city over products. These variance components suggest that generalization is impossible concerning variability of deliveries over products, cities, chains, or stores within chains.

Analyses utilizing the above variance components and constrained optimization methodology for one specified level of accuracy, constant over cities and products, are reported first. The accuracy specified for these is a 95 percent C.I. within 3 percent of mean deliveries per store per unit time (m of equation (6) is 3 percent).

Results of prime interest from these analyses are sample size in terms of number of chains (c) and average number of stores per chain (s) necessary to obtain the stipulated accuracy. To aid interpretation of results, a total sample size in terms of number of stores is computed by simply multiplying c times s. This total sample size requirement is then compared with the universe number of stores (N times M) by computing the percent of the universe total that the sample total represents. Thus, if $\frac{cs}{NM}$ equals 50 percent, this suggests that one-half of all the chain stores in that market must be sampled to obtain a 95 percent C.I. within 3 percent of mean deliveries per store per unit time. Universe values are presented in Table 3.

For cheese, the sample size required varies from 91 percent of the universe in Toledo to 100 percent of the universe in Terre Haute and Omaha, Table 4. Toledo would require, on the average, 17.0 stores from each of the

Table 3 . Universe number of chains and stores per chain, by city

| City | Chains | Stores/Chain | Total |
|---|---|---|---|
| | | - number - | |
| Dallas | 6.0 | 36.8 | 221 |
| Omaha | 5.0 | 11.2 | 56 |
| Toledo | 5.0 | 18.6 | 93 |
| Terre Haute | 7.0 | 2.0 | 14 |

Source: Supermarket News, 1971 Distribution of Food Store Sales in 288 Cities, (Fairchild Publications, Inc., New York, 1971), pp. 43, 93, 117, 125.

Table 4. Sample size required for specified accuracy within 3 percent of actual mean deliveries, cheese and butter, by city

| Statistic | Dallas | | Omaha | | Toledo | | Terre Haute | |
|---|---|---|---|---|---|---|---|---|
| | Cheese | Butter | Cheese | Butter | Cheese | Butter | Cheese | Butter |
| Chains | 5.9 | 6.0 | 5.0 | 5.0 | 5.0 | 5.0 | 7.0 | 7.0 |
| Stores/Chain | 36.3 | 36.2 | 11.2 | 11.2 | 17.0 | 18.2 | 2.0 | 2.0 |
| Sample Total | 215 | 217 | 56 | 56 | 85 | 91 | 14 | 14 |
| Percent of Universe Required | 97.3 | 98.2 | 100.0 | 100.0 | 91.4 | 97.8 | 100.0 | 100.0 |

Source: Computed.

five chains, or a total of 85 stores for the sample. Since 85 of 93 stores are required, this represents a sample size requirement of 91.4 percent of the universe. A similar interpretation may be given to the results for other cities and for butter over cities. The percent of the universe required for the same accuracy on butter is generally higher than that for cheese.

These results indicate that the variability in delivery data is great enough so as to require almost the entire universe to be contained in the sample, given the specified accuracy. This is true regardless of product or city.

Sensitivity analysis

The purpose of sensitivity analysis is to investigate the sensitivity of results to changes in parameters specified. As previously noted, simulated results may be obtained by changing the parameters which are functionally related to V. Of prime interest are the m parameter of equation (6) and the mean level of deliveries per store per unit time ($\overline{X}$).

One procedure is to relieve the stringent accuracy previously stipulated by allowing the m parameter of equation (6) to have the arbitrary value 7.

Another area of concern is the sensitivity of results to changes in mean deliveries per store per unit time ($\overline{X}$ of equation (6) ). Of course, as the level of mean deliveries changes, the magnitude of V will change in the opposite direction (all else constant) for the same stipulated accuracy. There is no logical prerequisite that the $\overline{X}$ utilized in determining V be an actual mean; thus, the sensitivity of mean level may be investigated by utilizing a normative mean. 6/

Sample size requirement is presented in a manner identical to the previous presentation. All results are based on the previous variance components, Table 2.

C.I. within 3 percent. Sensitivity of sample size requirements to mean delivery level is investigated first. Stipulated accuracy remains at a 95 percent C.I. within 3 percent of mean deliveries per store per period while the mean level utilized is normative rather than actual. For comparative purposes, actual mean level deliveries are given, Table 5.

For cheese an arbitrary normative mean level of 700 was used across cities rather than the actual mean, Table 5. As would be expected, sample size requirements are decreased for only Dallas and Toledo. For Dallas, the requirement reduces from about 97 percent of the universe to about 94 percent when the mean is increased from the actual of 469 pounds to 700 pounds. Thus, for an increase of 49.3 percent in the mean, a reduction in sample size requirement of only 3.7 percent was realized.

Much the same result is obtained for butter. Using a normative mean level of 100 across cities no decrease in sample size requirement would be expected except in Dallas, since 100 is below the actual mean for the other three cities. In this instance, an increase of 27.9 percent in the mean level was associated with a decrease of only one-half of one percent in the sample size required.

Table 5 . Actual mean deliveries in pounds per store per week, by product and city

| City | Mean Deliveries For | |
|---|---|---|
| | Cheese | Butter |
| | - pounds - | |
| Dallas | 469.2 | 78.2 |
| Omaha | 807.3 | 394.6 |
| Toledo | 697.7 | 242.1 |
| Terre Haute | 1,095.4 | 189.7 |

Source: Primary data.

Table 6. Sample size required for specified accuracy within three percent of normative mean deliveries, cheese and butter, by city

| | Dallas | | Omaha | | Toledo | | Terre Haute | |
|---|---|---|---|---|---|---|---|---|
| Statistic | Cheese | Butter | Cheese | Butter | Cheese | Butter | Cheese | Butter |
| Chains | 5.7 | 6.0 | 5.0 | 5.0 | 5.0 | 5.0 | 7.0 | 7.0 |
| Stores/Chain | 36.3 | 35.9 | 11.2 | 11.2 | 17.0 | 18.5 | 2.0 | 2.0 |
| Sample Total | 207 | 216 | 56 | 56 | 85 | 93 | 14 | 14 |
| Percent of Universe Required | 93.7 | 97.7 | 100.0 | 100.0 | 91.4 | 100.0 | 100.0 | 100.0 |

Source: Computed.

Briefly considering the elasticities of response for each product is interesting. In the first instance, for Dallas cheese a 10 percent increase in mean level yields about a 0.8 percent decrease in sample size requirement. For Dallas butter a 10 percent increase in mean level yields less than a 0.2 percent decrease in sample size requirement. Of course, these elasticities between mean level and sample size requirement cannot be taken as a general relationship which retains validity over a range of mean level increases or decreases. However, response is inelastic which suggests that for either product, changes in the size of deliveries (changes in average store size) have a relatively insignificant effect on sample size required to obtain a specified accuracy.

C.I. within 7 percent. Results reported here involve the same normative mean levels as used in the above analysis but the accuracy required is relaxed from a 95 percent C.I. within 3 percent of mean level delivery per store per unit time to a 95 percent C.I. within 7 percent of mean level. This, of course, should decrease sample size required.

Utilizing the 7 percent accuracy yields a smaller sample size requirement for cheese in every city except Terre Haute, Table 7. Comparing the results obtained from the relaxation from 3 to 7 percent reveals that the percent decrease in sample size requirement ranges from zero in Terre Haute to 28.2 percent in Toledo. The percent decrease for Dallas is 22.7 while it is only 5.4 for Omaha. Toledo results are most sensitive followed by Dallas and Omaha with no sensitivity obtained for Terre Haute. This is

Table 7. Sample size required for specified accuracy within seven percent of normative mean deliveries, cheese and butter, by city

| Statistic | Dallas | | Omaha | | Toledo | | Terre Haute | |
|---|---|---|---|---|---|---|---|---|
| | Cheese | Butter | Cheese | Butter | Cheese | Butter | Cheese | Butter |
| Chains | 4.4 | 6.0 | 4.7 | 5.0 | 5.0 | 5.0 | 7.0 | 7.0 |
| Stores/Chain | 36.3 | 32.3 | 11.2 | 11.2 | 12.2 | 18.3 | 2.0 | 2.0 |
| Sample Total | 160 | 194 | 53 | 56 | 61 | 92 | 14 | 14 |
| Percent of Universe Required | 72.4 | 87.8 | 94.6 | 100.0 | 65.6 | 98.9 | 100.0 | 100.0 |

Source: Computed.

roughly the relationship of the cities with respect to the absolute magnitude of the variance components.

Relaxing stipulated accuracy reduced the number of chains required in Dallas and Omaha while the average number of stores per chain required remained stable. In Toledo, however, the reverse is true--the average number of stores per chain required is reduced while the number of chains remains stable. This may be explained by the relative magnitude of the two variance components in each city. For Dallas and Omaha, the stores/chain variance component is 26 and 22 times greater than the chain variance component, respectively. The relationship for Toledo, however, is a variance component of stores/chain only about three times greater than the chain variance component.

For butter, decreased sample size requirement is obtained only in Dallas and Toledo. The Dallas decrease is 10.2 percent. The slight Toledo decrease to 92 stores from the previous 93 represents only a 1.1 percent change. No decrease is obtained for Omaha or Terre Haute by relaxing accuracy from 3 to 7 percent. In general, sensitivity of butter sample size requirement is relatively less than sensitivity of cheese sample size requirement to changes in stipulated accuracy.

## Conclusions

The sensitivity analysis suggests that results are not as sensitive to changes in mean level as to changes in the m parameter of equation (6). This means that results obtained are not as sensitive to average store size changes as they are to changes in accuracy via stipulations on the confidence interval.

Of the two products, butter sample size requirement is the least sensitive to changes in mean level or C.I. while cheese sample size requirement is more sensitive to such changes. By relaxing accuracy from a 95 percent C.I. within 3 percent to one within 7 percent, the smallest sample size required was still nearly 66 percent of the universe for either product or any city. This suggests the magnitude of variability in individual store delivery data. In some instances, like Terre Haute, small sampling error can be obtained only by a complete accounting of all stores and chains. For this case, however, only 14 stores constitute the universe which makes a market of this size manageable in terms of including the entire universe in the sample.

The relative inelasticity of sample size requirement to changes in store size suggests that the large requirement on sample size in relation to universe size is rather stable for various market sizes. This implies that whether a market is large or small in terms of either number of stores in the universe or average size of store in the market, a large proportion of the universe would need to be sampled in order to obtain the accuracy specified.

In general, for either product or city, a sample of stores from each chain in a metropolitan market must be included in the sample, regardless of product or metropolitan market. Beyond this, in most instances, over three-fourths of the universe average number of stores per chain are also required to be sampled in order to attain a 95 percent C.I. within 3 percent on mean deliveries. Apparant from the analysis is that delivery data collected for a metropolitan market are subject to large sampling error.

The general significance of results from this study for actually measuring changes in food product deliveries for a market is that:
1) individual store deliveries are highly variable from week to week
2) autocorrelation will likely exist in delivery data, but may be dealt with by appropriate statistical techniques and 3) accuracy specified will impact on sample size required relatively more than average size of city or store. Beyond this, the methodological procedures indicated above may be used to determine optimum sample size and composition for any product where individual store delivery data may be desired to monitor changes in product movement.

## Footnotes

*Technical Article No. 10568, Texas Agricultural Experiment Station. This research was partially supported by a grant from the American Dairy Association. The authors acknowledge contributions to this research by Charles Zeis, Institute of Statistics, and H. O. Hartley, Director, Institute of Statistics, Texas A&M University.

1. One time short-term promotions for a particular food product in a metropolitan market are quite common. For example, the American Dairy Association of United Dairy Industries, Inc. funds one time promotions for a month in metropolitan markets aimed at increasing the consumption of cheese. Food processors frequently conduct short-term promotion programs in a metropolitan market which involve bonuses to the retail trade.

2. Typically several metro areas are served by a chain distribution center or warehouse. For example, the Safeway Distribution Center in Dallas, according to Chain Store Guide, serves 182 stores in 57 counties. Kroger's Dallas Distribution Center, located in the suburb of Irving, serves 68 stores in 19 counties. Withdrawal data from these two warehouses would obviously not pertain to the same service area.

3. Autocorrelation coefficients were estimated prior to transformation and are reported in Sporleder, et. al. [10]. In general, these estimated coefficients were significantly different from zero ($\alpha = 0.05$) but nearly all were below 0.40.

4. Since the actual universe number of stores/chain varies from chain to chain, the complexity of the problem was reduced by regarding M as the average number of stores/chain for the universe computed as

$$M = \sum_{i=1}^{c} M_i/N.$$

5. $K_c$ was estimated at $500 and $K_s$ at $34. Of course, these could be changed to reflect different costs in different markets.

6. The term "normative mean" is used here simply to distinguish a hypothetical mean delivery from the actual mean delivery calculated from the particular sample drawn for this study.

## References

[1] Bennett, C. A. and N. L. Franklin, Statistical Analysis in Chemistry and the Chemical Industry (New York: John Wiley & Sons, Inc., 1954).

[2] Gates, Charles E. and C. Shinue, "The Analysis of Variance of the S-State Hierarchal Classification," Biometrics, 18:529-536, December 1962.

[3] Gaylor, D. W. and T. D. Hartwell, "Expected Mean Squares for Nested Classifications," Biometrics, 25:427-430, June 1969.

[4] Hartley, H. O. and R. R. Hocking, "Convex Programming by Tangential Approximation," Management Science, 9:600-612, July 1963.

[5] Henderson, C. R., "Estimation of Variance and Covariance Components," Biometrics, 9:226-252, June, 1953.

[6] LaMotte, L. R. and H. H. Oxspring, CONVEX: A Computer Program for the Hartley-Hocking Convex Programming Algorithm, Institute of Statistics, Texas A&M University, 1971.

[7] McHugh, R. B. and P. W. Mielke, Jr., "Negative Variance Estimates and Statistical Dependence in Nested Sampling," Journal of the American Statistical Association, 63:1000-1003, September 1968.

[8] Searle, S. R., "Topics in Variance Component Estimation," Biometrics, 27:1-76, March 1971.

[9] Searle, S. R. and R. F. Fawcett, "Expected Mean Squares in Variance Components Models Having Finite Populations," Biometrics, 26:243-254, June 1970.

[10] Sporleder, T. L., R. E. Branson and C. E. Gates, Optimum Sampling for Measuring Short-Term Changes in Food Product Deliveries, Texas Agricultural Market Research & Development Center, Texas A&M University, Technical Report to be published.