OPTIMUM SAMPLING FOR MEASURING SHORT-TERM CHANGES IN FOOD PRODUCT DELIVERIES

Thomas L. Sporleder

Robert E. Branson

Charles E. Gates

December 1974

Texas Agricultural Market Research and Development Center
Department of Agricultural Economics

in cooperation with
American Dairy Association, United Dairy Industries Association
Chicago, Illinois

THE TEXAS AGRICULTURAL MARKET RESEARCH AND DEVELOPMENT CENTER

An Education and Research Service
of the
Texas Agricultural Experiment Station
and the
Texas Agricultural Extension Service

The purpose of the Center is to be of service to agricultural producers, groups and organizations, as well as processing and marketing firms in the solution of present and emerging market problems. Emphasis is given to research and educational activities designed to improve and expand the markets for food and fiber products related to or of interest to Texas agriculture.

The Center is staffed by a basic group of professional agricultural and marketing economists from both the Experiment Station and Extension Service. In addition, support is provided by food technologists, statisticians and specialized consultants as determined by the requirements of individual projects.

Robert E. Branson
Coordinator

ACKNOWLEDGEMENTS

The authors wish to acknowledge the many retail business firms and non-profit institutions which supplied delivery data for this study.

Special appreciation is expressed to Dr. Charles Zeis, Institute of Statistics, Texas A&M University, for his efforts in programming. In addition, gratitude is owed Dr. H. O. Hartley, Director, Institute of Statistics, Texas A&M University, for his contributions to the methodology of this research project.

Thanks are also due Dr. Robert Degner, former Research Associate of the Texas Agricultural Market Research and Development Center, for his untiring effort in field operations, data tabulation, and computer programming.

This research was partially supported by a grant from the American Dairy Association, United Dairy Industries Association, Chicago, Illinois.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDIX TABLES

Table		Page

OPTIMUM SAMPLING FOR MEASURING SHORT-TERM CHANGES IN FOOD PRODUCT DELIVERIES

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

INTRODUCTION

The agricultural sector, during the past two decades, has witnessed an increasing number of commodity organizations responsible for generic advertising of their product. These programs are typically funded by producer assessments or check-offs. From time to time the producer-directors of such organizations ask for an evaluation of the effectiveness of specific advertising expenditures. Such evaluation, from a research point of view, involves measurement of changes in generic, instead of brand, food product movement among time periods (for example, among pre-promotion, promotion, and post-promotion time periods).

For example, the American Dairy Association of United Dairy Industries Association is responsible for decisions on generic advertising for butter and cheese, among other dairy products. One part of their promotion effort is directed toward short-term (a month or less) metropolitan market area advertising campaigns. The measurement problem associated with the evaluation of a short-term metro market advertising campaign is particularly difficult. Controlled experimentation, using some variation of a latin-square design, has been successfully employed to measure generic promotion results [2, 7, 8].

---

*Thomas L. Sporleder and Robert E. Branson are Associate Professor and Professor, respectively, Department of Agricultural Economics. Charles E. Gates is Professor of Statistics, Institute of Statistics. All are of Texas A&M University.

However, controlled experimentation is costly in relation to the typical budget allocated for metro market advertising. The ratio of research cost to advertising budget is unfavorable primarily because individual retail store audit data is required. Consequently, a more pragmatic method is needed when changes in product movement among time periods in a single metropolitan market are to be measured.

This report contains research results which explores the potential of measuring changes over time in product delivery data in metro markets. Product delivery data possess a substantial cost advantage over store audit or sales data.[1/] The main cost savings occur from the ability to collect store deliveries from a central location, such as a chain grocery warehouse, without the necessity of individual store visits, as in the case of audits.

A key question must be resolved in judging the feasibility of the delivery data approach. It is whether it is possible to measure with statistical significance differences of only 3 to 5 percent in product deliveries from one time period to another. The research reported herein was designed to investigate the potential of achieving that goal. Data consisted of actual product deliveries in four markets under typical non-promotion conditions. One approach to the question is to see if reasonably accurate measurement of changes in product delivery under normal marketing conditions can be achieved. If so, then one can hope to measure promotion

[1/] Audit data involves obtaining not only records of deliveries to individual stores, but also the expensive process of period inventory measurement.

program results.[2/] Conceptually the factors requiring consideration are 1) the normal variability of product deliveries over time, 2) the relationship of that variability to the amount of measurement error one can tolerate, 3) the influence of market size on the sample size necessary, 4) the effect of the time length of the sampling period on normal variability in deliveries and thereby on the behavior of sample size, and 5) the size of likely changes in store inventory levels in relation to the size of total store deliveries by period time length. Each of these factors, except for the last one, is considered in this report using butter, cheese, and margarine deliveries.

[2/] "Normal" is used here in the context of variability in product deliveries during periods of no promotion.

CONSIDERATIONS IN DELIVERY DATA COLLECTION

One might intuitively surmise, in the context of the problem presented above, that chain warehouse withdrawal data would suffice on a total divisional basis. A major factor mitigates against use of warehouse withdrawals. Food chain distribution warehouses evolve from many considerations besides metropolitan market geographics. Typically several metro areas are served by a distribution center. For example, the Safeway stores center in Dallas metro market, according to Chain Store Guide for 1972, delivers to 182 stores in 57 counties. Kroger's Dallas distribution center located in the suburb of Irving, serves 68 stores in 19 counties. Warehouse withdrawal data from these two centers obviously would not pertain to the same service area. When mass media advertising is involved, the area of dominant influence (ADI) is the relevant market area. To eliminate stores outside the ADI involves collecting individual store data. In short, no readily available aggregate delivery data exist which adequately reflect any given metropolitan market, with the exception of SAMI data. The latter is 1) expensive to purchase and 2) does not cover all products. Since the relevant market area requires that individual store deliveries must be collected anyway, the possibility of sampling among the stores rather than using the total universe arises. However, the sampling error involved when utilizing individual store delivery data must be determined. Two sources of variance are present in the data. One is that among stores during any one time segment. The second is that over the time segments.

Less obvious, yet further complicating the matter, are delivery variations among various types of sales outlets in a metropolitan market. The most important outlets are chain supermarkets, independent or affiliated food stores, public eating establishments (including hotels and motels), and non-profit institutions such as schools, colleges, universities, or hospitals. Aside from these major outlets, complete accounting would consider deliveries through the public welfare food distribution programs operative in a given market. Such deliveries may also affect product movement through normal trade channels.

Product Movement Channels

Expense of obtaining product movement data is closely related to the collection point for that data. Possible collection points depend on the distribution system used for a commodity. Several basic systems were observed in the four research cities. Predominant among them were:

1. Processor to independent or chain retailer
2. Processor to processor's branch house to independent or chain retail distribution center
3. Processor to independent wholesaler representative to independent retailer
4. Processor to affiliated chain warehouse to independent retailer
5. Processor to food chain warehouse to chain retailer

"Independent retailer," as used here, means retail food stores, restaurants and/or cafeterias or specialty eating establishments. All the above distribution systems are encountered in most markets of 100,000 population or over. The first three channels are classified as direct store delivery systems since the product goes directly to retail food outlets. Individual items go directly to a central warehouse for the retail chain rather than direct to the store in channels 4 and 5.

The preferable channel from a market measurement standpoint is one which utilizes a central distribution warehouse for all products stocked by the retail stores. There, data for all products, and brands thereof, for multiple retail outlets can be accumulated simultaneously. Otherwise, all sources of individual brands must be located and necessary monitoring systems set in motion to cover them.

Ways of Intercepting Data by Channel

Food Chains

Three key factors influence the data interception task. These are:

1. How the records are kept--by computer system or by hand.
2. Location at which the records are maintained.
3. How long they are maintained in the file.

Some major food chains have not yet fully converted to computer record systems. In such instances, product shipment data must be developed by going through all the invoices for the stores concerned. This creates a substantial interruption of normal office procedures. Regular office staff cannot be assigned the task because it takes too much time away from normal operations. Assignment of an outsider to the task involves working with confidential internal records, which is seldom permitted.

If records of the food chain are on computer, no major difficulty arises in obtaining product shipment records to individual stores. Permission, of course, must be obtained for record access. Usually the sales measurement group provides the personnel to copy the needed information from the computer print-outs since the food chain personnel cannot afford time for these extra duties.

Location of the warehouse records can influence the point of data collection. The warehouse and central office may be in another city. If several chains serve a given city, this may mean dealing with central offices and warehouses scattered among several outside cities. One means of handling this problem is to obtain records at the individual retail store level. Copies of the shipment invoice are provided to a store for each delivery. Caution and planning are necessary, however, because the length of time stores keep these records varies substantially.

It must be recognized at the outset that food chains are not in the detailed record keeping business. Even the divisional or central office normally keeps full records for only 13 weeks. At the end of that period, a quarter year, individual store records by products are typically discarded. Thus, it is usually impossible to retrieve back records unless a special arrangement with the divisional office of the food chain is developed.

Independent Food Stores

Independent stores that are members of an affiliated group may have comparable record systems to centrally owned chains. In that case the procedures above apply. Unaffiliated stores present a near impossible situation for collection of back data and a difficult one for current or forward information.

Visits with independent stores found essentially two classes. The larger ones were associated generally with an affiliated chain with a central warehouse system of records. Smaller stores were served by multiple independent wholesalers making record retrieval a major obstacle.

Independent wholesalers may not be on a computer system, as found in Terre Haute, Indiana, for example. Special clerical help must be arranged to collect information on sales to individual retail stores.

These conditions can force data collection to the store level among small retailers. Here, too, record keeping usually becomes progressively worse because the so-called mom and pop corner grocery stores have notoriously poor accounting systems. Invoices normally are kept on a simple chronological basis. To find any product one must go painstakingly through the entire set of invoices--a time consuming, expensive procedure. For this reason, the store owner usually refuses to take the time to recover back records. It is possible to get current purchase invoices, but this only allows for recording present movement.

Jobbers

Sales offices may only take orders whereas shipments are made from a central warehouse located in another city. Sales records often are not kept at the local sales office but are forwarded to a regional center.

There the computer system, if one is used, may not be programmed to record other than the dollar value of product sold. When a multi-product line is involved, the usual case, identity of specific product volume is lost in the dollar totals. Arrangements for delivery of current sales flow information requires clearance, in most cases, with the national office of the company.

Independent Wholesalers

The same obstacles exist at this distribution level as are encountered at food chain warehouses. In fact, they are often more severe because of a tendency toward less record computerization. Back records are nearly impossible to secure. Forward records can be intercepted with a minimum amount of labor.

Restaurants and Specialty Eating Establishments

Records found in general menu restaurants, which are usually independently owned, resembled those in independent food stores. An exception arises if a public accountant keeps the business books. Nonetheless, going through this third person creates extra communication needs, time delay and expense be it for either back or forward delivery data.

Delivery information is more available from franchise or chain specialty eating establishments. A comprehensive accounting system is characteristically a part of the franchise package. Therefore, either previous or forward data usually can be obtained.

An exception exists where an entrepreneur invests in several franchises, of the same or mixed type. Central records may be located at the entrepreneur's main office rather than at each unit. This can complicate data collection.

Specialty restaurants, independently operated, were found to be supplied not only by the HRI (hotel, restaurant, and institution) distributor but also direct from processors for cheeses. The latter was especially noted for those with Italian cuisine where special flavors are important.

Schools and Institutions

Attention in the research effort was given to public and parochial schools, as well as colleges and universities. Public and parochial school records are reasonably adequate. The difficulty is that foods are received in large, infrequent deliveries for an item like cheese. Only a use record would reflect weekly, or perhaps even monthly utilization. The infrequent delivery system reflects in part school lunch program distribution methods.

Because of the fixed menu system in schools, there is reason to doubt that public media promotion programs would influence product utilization in these outlets. Excluding them from the analysis is advised. The nature of the promotion program would have to be evaluated, however, since this is not intended as a blanket suggestion.

It is evident from the foregoing descriptions of the various facets of delivery or purchase records that no single delivery measurement approach can be used in a metropolitan market. In the smaller SMSA's data interception usually is best keyed to the retail store unit. Use of the distributor level can involve almost as many data collection points as there are retail outlets in the test market. Furthermore, the distribution points are much more widely dispersed geographically than the stores.

Commercial Market Data Services

A number of commercial market research services collect product movement data on a continuing basis. Two bases are used--store audits and consumer panels. Examples are Neilson store audits and the Market Research Corporation of America consumer panel. A more recent addition is the SAMI service (Sales Area Market Information) providing warehouse withdrawals.

Neilson is generally recognized as one of the pioneer organizations in providing retail store audit data. It covers major metropolitan U.S. markets with a bi-monthly service. Stores in the sample are audited regularly on a two-month rotating basis. Data are available thereby six times per calendar year. The audits, however, are for a one month (four week) average within the two month time lapse. Audits provide net sales movement through the retail outlet. Beginning inventories are taken plus store deliveries received during the four weeks minus ending inventories. Inventories include shelf stock plus back room stock. A problem arises for a market test if 1) the reporting periods need to be on a shorter time interval, 2) the regular service does not include the metro market involved or 3) the store sample will not provide data with the required sampling tolerances.

If the above limitations exist, Neilson will provide a special audit program. The cost understandably is more than for the regular service. Also, Neilson audits do not include Safeway stores. Where Safeway is an important market factor, the data may not be adequate.

The Market Research Corporation of America, rather than store audit based, is predominantly a consumer panel service. Panel members report purchases on a weekly basis and market data are available either on a weekly or monthly (four week) schedule. National market and regional data are the main purposes of MRCA. Provided separately, if desired, are reports on the New York, Chicago and Los Angeles metro markets. The MRCA panel comprises a sample of about 7,500 households located coast to coast to give national representation.

There are several advantages to the MRCA household panel. Ability to provide basic demographic information on purchasers is one. Purchases by age of household head, education, income and family size can be tabulated. Size and frequency of purchases are also indicated.

Obviously the MRCA consumer panel is not useful for monitoring short one-time market promotions, except in the three markets noted above. These markets are expensive for promotion program testing because of their size. Promotion effectiveness measurement is typically sought in medium size markets where media and other costs are less expensive. To meet specific market measurement requests, MRCA does offer tailor-made services of the store audit type.

Unfortunately, it is not feasible to initiate a household panel in a specific market for short time periods for two reasons. Recruiting costs become prohibitive in obtaining households, if the expense must be borne for a single market test. Secondly, experience has shown household panels to be biased in their purchase habits for as long as three months after joining a panel. These circumstances preclude household panels being used in most market tests.

The SAMI reporting system on warehouse withdrawals for shipment to retail food stores covers a 28 day (4 week) period. Twenty-seven major markets are covered. Whereas Neilson lacks coverage of Safeway stores, SAMI does not include A&P stores. Also, SAMI does not cover independent food stores. If major food chain warehouse withdrawal, minus A&P, is an adequate market indicator, use of SAMI data is a viable alternative. Also, not all food products are covered by SAMI. For example, cheese, which is central to this research endeavor, is not included.

DELIVERY DATA ANALYSIS--CHAIN GROCERY STORES

The food items concerned in this research are retail store delivery data for three dairy case commodities--butter, margarine and cheese. These products offer a sizable range of difference in sales volume that typify variation found among other foods. Variability of butter, margarine, and cheese deliveries on a weekly, two week, and four week basis to individual chain stores was examined in four markets varying in size and geographic locations. Chosen were Dallas, Texas; Omaha, Nebraska; Terre Haute, Indiana; and Toledo, Ohio. The primary objective was to identify the size of the sampling error obtained in individual store delivery data. The remainder of this report concerns the methodology and results of estimating sample size needs in terms of number of stores and chains required in the test cities in order to achieve selected data accuracy levels.

There are two basic approaches to any sample size problem. One approach is to minimize variance subject to specified cost while the second is to minimize cost subject to specified accuracy. Both approaches were utilized on the delivery data collected for this research. Minimizing variance subject to specified cost is utilized on the absolute level of deliveries per store. These results are reported only in Appendix E.

The second approach, and the one covered in the following report text, minimizes cost of sampling subject to some specified accuracy. Sample size requirements for each product (cheese, butter, and margarine) are calculated for each city. The details of the analytical approach are discussed in the following sections.

Methodology

Minimizing sampling cost subject to some specified measurement accuracy obviously allows pre-selection of the desired accuracy. Sample cost is of secondary concern since any money spent for measurement that is not accurate enough on an a priori basis would obviously be unwise.

The methodology has two basic phases. First is estimation of variance components from an analysis of variance model. Second is constrained optimization utilizing variance components estimated from the ANOVA model to minimize cost subject to some specified accuracy.

Based on First Differences

The overall objective of any measurement system appropriate for monitoring product movement is to detect a change or difference in movement from one period to another. Thus, subsumed in the analysis is an experimental design which would generate data for more than one time period.

By transforming the original data to first differences, autocorrelation is reduced or removed.[1/] Use of first differences in deliveries reduces data variability and, thereby, the sample size necessary for a specified accuracy. Consequently, a difference transformation on the original data from each city was analyzed.

From the 8 weekly records for each store, three sets of first differences were calculated. One was for weekly data, the second for biweekly, and the third for four-week periods. The first set was constituted of 7 one-week first differences calculated for each store by:

(1) $$Y_{ij_t} = P_{ij_h} - P_{ij(h+1)}$$

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 or chain i

[1/] Autocorrelation coefficients were estimated and are reported in Appendix B.

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

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

In this first case, h corresponds to a one-week period.

The second set was constituted of 3 biweekly period first differences where h is, of course, a two-week period, and t = 1, 2, 3. The third difference was calculated by defining h as a four-week period, and t = 1.

Analysis of Variance

To estimate the variance components of MINCOST, an unbalanced one-way analysis of variance (ANOVA) was utilized which incorporates a finite population correction (fpc) factor into the expected mean square calculations. The fpc reflects the number of chains and stores within chains for the metropolitan market area.

The appropriate ANOVA model for a given first difference (with the t subscript omitted) is:

$$Y_{ij} = \mu + c_i + \varepsilon_{ij} \tag{2}$$

where

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

μ = grand mean

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

ε_{ij} = residual

and

$i = 1, 2, ..., c$

$j = 1, 2, ..., s$

As with the ANOVA of MINVAR model, all effects are regarded as random [11, 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 (fpc) factor into the expected mean square calculations, Table 1 [4, 6, 12].

Variance components for chains and stores/chain were estimated for each of the 7 first differences calculated from one week data periods. The 7 variance components for chains were then averaged to produce one variance component for chains based on one week data periods. The same technique was utilized to produce one variance component for stores/chain for one week data periods.

Similarly, the variance components for chains and stores/chain were estimated (3 each) for two-week data periods. The estimates were then averaged into one. The four-week data period allowed estimation of only one variance component for chains and one for stores/chain.

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) minimize $c(K_c + sK_s)$

subject to [1]:

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 [3, 10]:

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

and where

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

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

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

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

c = sample number of chains

s = sample average number of stores/chain

N = universe number of chains

M = universe average number of stores/chain[1/]

K_c = cost of adding a chain to the sample[2/]

K_s = cost of adding a store within a chain to the sample[2/]

$V(\bar{\bar{y}})$ = 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. The equation reflects a cost for obtaining chain cooperation (K_c) as well as the cost of actually

---

1/ 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

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

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

obtaining delivery records for individual stores within that chain (K_s). Thus, 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 the grand mean of deliveries per store per period adjusted for a finite population. The fact that a finite number of chains exist in any one metropolitan market area is reflected by the correction factor $(1 - \frac{c}{N})$, while the factor $(1 - \frac{s}{M})$ is the finite population correction factor for stores within chains. Note that as the sample number of chains approach the universe number of chains ($c \rightarrow N$), the variability attributable to that component approaches zero. Similarly, as the sample average number of stores per chain approaches the universe average number of stores per chain ($s \rightarrow M$), 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 specification of V. The derivation of the specified accuracy is from a 95 percent confidence interval on mean deliveries per store unit time. Let $\overline{X}$ represent some mean level of delivery per store per unit time, then a 95 percent confidence interval (C.I.), two-tailed, is:

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

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

(6) $$(m\%)^2\bar{X}^2 = (1.96)^2\hat{\sigma}^2_{\bar{X}}$$

or

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

The Hartley-Hocking algorithm. Solution to the constrained optimization 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 [5, 9]. Essentially, the problem is a nonlinear programming problem (in this case, the objective function and one constraint is nonlinear in c and s). Restated as such, let $X_1 = c$ and $X_2 = s$ then the problem is to:

(8) $$\text{minimize } F = K_cX_1 + K_sX_1X_2 \qquad \text{subject to:}$$

(9) $$\begin{cases} \dfrac{1}{X_1}\left(\hat{\sigma}^2_c + \dfrac{\hat{\sigma}^2_{s(c)}}{X_2} - \dfrac{\hat{\sigma}^2_{s(c)}}{m}\right) \leq V + \dfrac{\hat{\sigma}^2_c}{N} \\ X_1 \leq N \\ X_2 \leq m \end{cases}$$

and the usual nonnegativity requirements. Note that both the objective function and the first constraint are nonlinear in X_1 and X_2. Detail of the algorithm and the boundary constraints as formulated for this particular problem are found in Appendix A.

Utilizing convex programming for a particular specification of accuracy (V) yields solutions in terms of the number of chains (X_1) and the number of stores per chain (X_2) necessary to obtain the specified accuracy and minimize cost. Obviously, the sample size required (i.e., the magnitude of X_1 and X_2) 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 in the previous section. This creates the opportunity for simulated solutions to sample size requirement via perturbation of V. This is investigated under the "Sensitivity Analysis" section below.

Results

Basic analyses were performed with one specified level of accuracy. In these analyses, specified accuracy remained constant over cities and products. That is, for each city a total of eight weekly records for each sample store for each product was included in the analysis with specified accuracy identical. Accuracy specified for these is a 95 percent C.I. within 3 percent of mean deliveries per store per unit time.

Results of prime interest from these analyses are the sample size, in terms of number of chains (c) and average number of stores per chain (s), necessary in order to obtain the stipulated accuracy. To aid in interpretation of the 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 would need to be sampled in order to obtain a 95 percent C.I. within 3 percent of mean deliveries per store per unit time. Universe values are given by city in Table 2.

As previously noted, the variance components estimated for chains and stores within chains are based upon first differences of the eight weekly records available for each store included in the sample [equation (1)]. Besides first differences computed from the original weekly records, there exist two aggregations of weekly records which utilize all eight observations per store. These are aggregation of adjacent weeks to: 1) two-week periods and 2) four-week periods. These aggregations are logical alternative analyses of the data since delivery records could be collected for either time unit. Of course, the four-week aggregation is of greater interest than the two week for practical purposes of obtaining records since single monthly delivery figures by store are more likely to be available than single bimonthly figures.

The statistical rationale for aggregation revolves around the extent to which autocorrelated observations exist. If deliveries by store over time are autocorrelated, then aggregation to longer than one-week periods will reduce the sample size required to obtain the same accuracy. In essence, if substantial autocorrelation exists in the data then results from the three different time period analyses should be markedly dissimilar. However, since the variance components are estimated from first differences the effects of aggregation should be somewhat mitigated.

The procedure for this section of the report is to present, in turn, results from each of the three alternative time units of analysis for each product and city with specified accuracy invariant. Then, the sensitivity

Table 2. Universe number of chains and stores per chain, by city.

City	Chains	Stores/Chain[1/]	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

[1/] Average number of retail stores in chain in the individual city.

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

of sample size requirement from these basic analyses is presented. The sensitivity analysis is conducted by perturbation of V through changes in the m parameter of equation (7), by utilizing various $\bar{X}$'s, and by subsequently combining the resultant V perturbations with the various time units. The sensitivity analysis allows for a comparison of results from various combinations of assumptions concerning parameters. This, then, permits conclusions to be drawn about the relative significance of key parameters in affecting results.

One Week First Differences, Actual Means

The variance components for chains and stores/chain for cheese, butter, and margarine differ markedly among cities, Table 3. There is no general pattern exhibited by the various components for any 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. Actual weekly mean delivery by product and city also varies widely, Table 4.

For cheese, the sample size required for a 95 percent confidence interval within 3 percent of mean deliveries per store per week varies from 91 percent of the universe in Toledo up to 100 percent of the universe in Terre Haute and Omaha, Table 5. A total of 93 stores are in the Toledo metropolitan area, Table 2. These stores belong to 5 chains, for an average of 18.6 stores per chain. Required for the specified accuracy on mean cheese deliveries would be a sample of, on the average, 17.0 stores/chain from each of the 5 chains, or a total of 85 stores. Since 85 of the

Table 3. Estimated variance components with fpc for one week first differences, by city and product.

City	Variance Component for: Cheese		Butter		Margarine	
	Chains	Stores/Chain	Chains	Stores/Chain	Chains	Stores/Chain
Dallas	9,572	249,627	4,284	20,372	43,681	233,016
Omaha	45,880	1,019,570	10,827	192,890	a/	a/
Toledo	33,283	109,984	24,328	66,311	a/	a/
Terre Haute	1,373,139	332,266	57,545	12,510	797,099	1,301,633

a/ No usable data were available in this city.

Source: Computed.

Table 4. Actual weekly mean deliveries by product and city, all chain stores. a/

City	Mean Cheese	Mean Butter	Mean Margarine
		--pounds per store--	
Dallas	469.2	78.2	750.8
Omaha	807.3	394.6	a/
Toledo	697.7	242.1	a/
Terre Haute	1,095.4	189.7	1,586.8

a/ No usable data were available in this city.

Source: Computed.

universe total of 93 stores are required, this amounts to a sample size requirement of 91.4 percent of the universe. A similar interpretation may be given to the data for other cities and for other products over cities, Table 5.

The percent of the universe required for the same accuracy on butter is generally higher than that for cheese, Table 5. In this case, the lowest percent of universe is nearly 98 percent (Toledo) with the highest again being 100 percent (Omaha and Terre Haute). Actual mean deliveries for butter ranged from a low of 78.2 pounds per store per week in Dallas to a high of 394.6 pounds per store per week in Omaha, Table 2.

Delivery data for margarine were collected in only two of the four cities, Dallas and Terre Haute. The percent of universe required for margarine was 89 in Dallas but 100 in Terre Haute, Table 5. Actual mean deliveries per store per week were 750.8 pounds and 1,586.8 pounds in the two cities, respectively, Table 2.

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.

Two Week First Differences, Actual Means

In order to investigate the effect of data aggregation on the sample size requirement, weekly delivery records were first combined into two-week periods. This was accomplished utilizing equation (14) where $t = 1, 2, 3$; $h = 1, 2, 3$; thus, the resultant Y_{ij_t} represents the t^{th} first difference of two-week periods for store j of chain i. There are three first differences for each store since the original eight weekly records by store aggregate to four two-week records by store.

The rationale for aggregation is that variability should be reduced if a significant amount of autocorrelation exists in the original weekly delivery records. The aggregation allows investigation of the autocorrelation phenomenon.

As with one week records, no perceptible pattern exists for variance components over products or cities, Table 6. The same specified accuracy, a 95 percent C.I. within 3 percent of the mean, is utilized for each product and city as with the one week differences. Results from the Convex program for two week differences are presented in a similar format, Table 7.

Results for cheese indicate that the percent of universe that must be sampled to attain the stipulated accuracy ranges from 81.7 percent in Toledo up to 100 percent in Omaha and Terre Haute, Table 7. The actual means used in computation for each city are twice the actual mean deliveries for a one-week period in all cases. The aggregation of one week records to two did reduce the sample size requirement in Dallas and Toledo for cheese. The sample size requirement in Omaha and Terre Haute was unchanged by the aggregation--100 percent sample required for both one and two week differences.

The analysis on butter data for the two week aggregation reveals a range in sample size requirement of from 90.3 percent of the universe up to 100 percent, Table 7. Some small improvement in sample size requirement was achieved by the aggregation in Dallas, Omaha, and Toledo, Table 7 compared with Table 5. Sample size requirement still remains higher than

Table 6. Estimated variance components with fpc for two week first differences, by city and product.

City	Variance Component for: Cheese		Butter		Margarine	
	Chains	Stores/Chain	Chains	Stores/Chain	Chains	Stores/Chain
Dallas	18,517	267,44	6,504	33,095	111,560	261,301
Omaha	29,584	638,640	31,423	296,413	a/	a/
Toledo	53,608	182,398	11,146	45,093	a/	a/
Terre Haute	2,163,802	446,671	123,367	33,243	1,708,545	1,482,142

a/ No usable data were available in this city.

Source: Computed.

Table 7. Two week differences, percent of universe required for a 95% C.I. within 3% of actual mean deliveries, all cities and products.

Product and City	Sample Chains	Sample Stores/Chain	Sample Total	Percent of Universe Required
	-number-	-number-	-number-	-percent-
Cheese:				
Dallas	6.0	31.4	189	85.5
Omaha	5.0	11.2	56	100.0
Toledo	5.0	15.1	76	81.7
Terre Haute	7.0	1.9	14	100.0
Butter:				
Dallas	6.0	35.4	213	96.4
Omaha	4.9	11.2	55	98.2
Toledo	5.0	16.7	84	90.3
Terre Haute	6.9	2.0	14	100.0
Margarine:[a/]				
Dallas	6.0	25.4	153	69.2
Terre Haute	7.0	2.0	14	100.0

[a/] No usable margarine data for Omaha or Toledo were obtained.

Source: Computed.

is desirable for practical data collection procedures, particularly in large metro areas.

Margarine sample size requirement in the two cities, Dallas and Terre Haute, for the two week differences are 69.2 and 100 percent of the universe in the two cities, respectively, Table 7. Compared with the one week difference results, the sample size requirement in Dallas was reduced by about 20 percent, from 89 to 69 percent, but Terre Haute was 100 percent in either case.

This suggests thatthe Dallas data exhibited relatively greater autocorrelation than did the Terre Haute data, since aggregation would reduce sample size requirements substantially only in autocorrelated records. Referring to the estimated autocorrelation coefficients of Table B-3, Appendix B, the coefficient for Dallas is significantly different from zero at the 5 percent level while the Terre Haute coefficient is not. This is consistent with the result of sample size requirement for margarine shown in Table 5 and 7 for the two cities.

For Dallas, as previously noted, sample size for margarine based on two week differences was decreased by 20 percent compared with the sample size based on one week differences. Note, however, that the entire reduction came from a reduction in the sample size requirement of the stores per chain component, a reduction from an average of 32.7 stores per chain to an average of 25.4 stores per chain. This suggests that aggregation of weekly records may reduce sample size in terms of the number of stores per chain required relatively more than reducing the number of chains required.

Four Week First Differences, Actual Means

The basic data for each store included in the sample were eight weekly delivery records. This allows convenient aggregation to two and four week periods. The first four week difference is computed in a manner similar to the two week first difference. Utilizing equation (1) where t = 1, h = 1, 2 yields Y_{ij_t} as the singular first difference of two four week periods for each store. Most variance components for the four week aggregation are slightly proportionally lower compared to the one week variance components, Table 8.

As with the two week aggregation, specified accuracy remained unchanged for the four week aggregation compared to the basic one week analysis. For cheese, some improvement in sample size requirement is evident in Dallas and Terre Haute, Table 9. For Toledo, sample size requirement drops from about 82 percent of the universe for the two week aggregation to around 20 percent for the four week. However, the variance component for Toledo chains for the four week aggregation is only 750 which must be regarded as an outlier. Variance components are subject to sampling variation which likely accounts for the dramatic decrease in size of this variance component. Repeated sampling would likely yield a higher variance component for chains and, therefore, a larger sample size requirement.

Results of the four week aggregation on butter data reveal no improvement over the two week aggregation, Table 9 compared to Table 7. In fact, sample size requirement actually increased in Omaha and Toledo. In either

Table 8. Estimated variance components with fpc for four week first differences, by city and product.

City	Variance Component for: Cheese		Butter		Margarine	
	Chains	Stores/Chain	Chains	Stores/Chain	Chains	Stores/Chain
Dallas	30,429	490,923	38,080	60,960	72,344	856,154
Omaha	128,967	2,104,997	62,455	753,683	a/	a/
Toledo	750	258,308	15,162	160,289	a/	a/
Terre Haute	5,322,731	395,446	40,543	67,377	876,089	2,325,196

a/ No usable data were available in this city.

Source: Computed.

Table 9. Four week differences, percent of universe required for a 95% C.I. within 3% of actual mean deliveries, all cities and products.

Product and City	Sample Chains	Sample Stores/Chain	Sample Total	Percent of Universe Required
	-number-	-number-	-number-	-percent-
Cheese:				
Dallas	6.0	26.8	161	72.9
Omaha	5.0	11.2	56	100.0
Toledo	1.0	18.6	19	20.4
Terre Haute	7.0	1.8	13	92.9
Butter:				
Dallas	6.0	34.0	204	92.3
Omaha	5.0	11.2	56	100.0
Toledo	5.0	18.6	93	100.0
Terre Haute	7.0	2.0	14	100.0
Margarine:[a/]				
Dallas	6.0	23.8	143	64.7
Terre Haute	7.0	2.0	14	100.0

[a/] No usable margarine data for Omaha or Toledo were obtained.

Source: Computed.

case, this is due to a proportionally higher variance component for stores/ chain, Table 6 compared with Table 8. Some slight improvement is apparent in Dallas but certainly not enough for any overall improvement.

Results for the margarine data are also essentially unchanged from utilization of four week first differences compared with two week, Tables 7 and 9. Terre Haute sample size requirement remains at 100 percent of the universe while Dallas drops slightly from around 69 percent to 65 percent.

In general, regardless of product or city, slight if any improvement results from aggregation to four week periods. This suggests that significant decreases in sample size requirement are likely only by lowering specified accuracy or by increasing the absolute size of mean deliveries per store per period by selective sampling.

Sensitivity Analysis

The purpose of this section is to investigate the sensitivity of results to changes in various parameters specified and/or estimated for the basic analyses from the preceding section. 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 (7) and the mean level of deliveries per store per unit time ($\overline{X}$)

The cogency of this type analysis is enforced by results from the basic analyses. For most practical purposes, a large proportion of the universe was required to obtain the accuracy stipulated under the basic analyses. A logical procedure, therefore, would be to successively relieve the stringent accuracy previously stipulated. This is accomplished by requiring a 95 percent C.I. to be within only 5 or 7 percent of mean deliveries per store per unit time. This amounts to allowing the m parameter of equation (7) to successively have the values .05 and .07.

Another important question which may be answered by simulated procedures is the sensitivity of results to changes in mean deliveries per store per unit time ($\overline{X}$ of equation (7)). 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 the mean level may be investigated by utilizing a normative mean.[1/]

The procedure of this section is to present sample size requirement in a manner identical to the previous presentation of results. Differences in the analysis lie in the relaxation of stipulated accuracy. All of the results presented here are for variance components estimated from first differences on one week data, Table 3. Results from the same analysis except with variance components estimated from first differences on two and four week time periods are presented in Appendix C.

[1/]The term "normative mean" is used here simply to distinquish a hypothetical mean delivery from the actual mean delivery calculated from the particular sample drawn for this study.

<u>C.I. within 3 Percent</u>. Sensitivity of sample size requirements to mean delivery level is investigated first. This involves maintaining the same stipulated accuracy as before except utilizing a normative mean in equation (7). That is, 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 cheese an arbitrary normative mean level of 700 was used across cities rather than the actual mean, Table 4. As would be expected, sample size requirements are decreased for only Dallas and Toledo (Table 10 compared to Table 5. For Dallas, the sample size requirement reduces from about 97 percent of the universe to about 92 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, Table 10. 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. For Dallas, sample size requirement decreased from 98.2 percent of the universe to 97.7 percent with an increase in mean level from 78.2 to 100 (Table 10 compared with Table 5). Thus, for Dallas butter, an increase of nearly 28 percent in the mean level was associated with a decrease of only one-half of one percent in the sample size required.

Table 10. Normative mean deliveries, percent of universe required for a 95% C.I. within 3%, all cities and products, one week first differences.[a]

Product and City	Sample Chains	Sample Stores/Chain	Sample Total	Percent of Universe Required
	-number-	-number-	-number-	-percent-
Cheese:				
Dallas	5.7	36.3	207	93.7
Omaha	5.0	11.2	56	100.0
Toledo	5.0	17.0	85	91.4
Terre Haute	7.0	2.0	14	100.0
Butter:				
Dallas	6.0	35.9	216	97.7
Omaha	5.0	11.2	56	100.0
Toledo	5.0	18.5	93	100.0
Terre Haute	7.0	2.0	14	100.0
Margarine:[b]				
Dallas	6.0	30.1	181	81.9
Terre Haute	7.0	2.0	14	100.0

[a] Normative means delivery per store per week was held invariant across cities. By product they were: cheese 700, butter 100, margarine 1,000.

[b] No usable margarine data for Omaha or Toledo were obtained.

Source: Computed.

A similar analysis with margarine data again yields similar results. By using an arbitrary normative mean level of 1,000 for margarine, sample size required would be expected to decrease again only for Dallas, Table 4. Increasing the Dallas mean margarine level from 750.8 to 1,000 resulted in a decrease in sample size required of from 89.1 percent of the universe to 81.9 percent (Table 10 compared with Table 5). This represents an increase in mean level of about 33 percent for a reduction of about 8.1 percent in sample size required.

Briefly considering the elasticities of response for each of the three products is interesting. In the first instance, for Dallas cheese a 10 percent increase in mean level yields 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. For the Dallas margarine instance, a 10 percent increase in mean level yields a 2.5 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, the elasticities are in each instance inelastic--a one percent increase in mean level yields a less than one percent decrease in sample size requirement. This suggests that for any of the three products, 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. Thus, sample size is relatively insensitive to mean level.

C.I. within 5 Percent. The above analysis is repeated here except that stipulated accuracy is relaxed through requiring a 95 percent C.I. to be within 5 percent rather than 3 percent of the mean level deliveries per store unit time. The 5 percent stipulation on accuracy is performed utilizing exactly the same normative mean level deliveries as reported above. This facilitates comparison of results.

Sample size resulting from increasing the C.I. on mean level for cheese reveals relatively greater sensitivity than that obtained by perturbation of mean level, Table 11. Dallas sample size requirement is reduced from 93.7 percent of the universe to 84.2 percent, from 100 to 98.2 percent for Omaha, from 91.4 to 79.6 percent for Toledo, but remains at 100 percent for Terre Haute (Table 11 compared with Table 10). This represents a range of from zero to 12.9 percent reduction in sample size by changing the stipulated accuracy from a 95 percent C.I. within 3 percent to one that is within 5 percent.

The same analysis on butter yields a sample size requirement which is smaller in Dallas only, Table 11. For Dallas, the decrease in sample size required fell from 97.7 percent of the universe to just under 94 percent. All of the decrease was attributable to a reduction in the average number of stores per chain required (Table 11 compared with Table 10). The reduction in Dallas sample requirement is just over 4 percent. Thus, butter sample size requirement is not as sensitive to relaxation of stipulated accuracy as is cheese sample size requirement.

For margarine data, a significant decrease in sample size requirement is obtained in Dallas by relaxing stipulated accuracy, Table 11. The sample size requirement dropped from 81.9 percent of the universe to 62.0

Table 11. Normative mean deliveries, percent of universe required for a 95% C.I. within 5%, all cities and products, one week first differences.[a]

Product and City	Sample Chains	Sample Stores/Chain	Sample Total	Percent of Universe Required
	-number-	-number-	-number-	-percent-
Cheese:				
Dallas	5.1	36.3	186	84.2
Omaha	4.9	11.2	55	98.2
Toledo	5.0	14.7	74	79.6
Terre Haute	7.0	2.0	14	100.0
Butter:				
Dallas	6.0	34.4	207	93.7
Omaha	5.0	11.2	56	100.0
Toledo	5.0	18.5	93	100.0
Terre Haute	7.0	2.0	14	100.0
Margarine:[b]				
Dallas	6.0	22.8	137	62.0
Terre Haute	7.0	2.0	14	100.0

[a] Normative mean delivery per store per week was held invariant across cities. By product, they were: cheese 700, butter 100, margarine 1,000.

[b] No usable margarine data for Omaha or Toledo were obtained.

Source: Computed.

percent in Dallas by changing from a 95 percent C.I. within 3 percent to one within 5 percent (Table 11 compared with Table 10). All of this smaller sample size requirement (a 24.3 percent decrease) is a result of a decrease in the average number of stores per chain required. For Terre Haute, no decrease in sample size requirement is realized from the change in accuracy required. The difference between the results obtained from Dallas and Terre Haute may be explained by comparing their respective components of variance.

C.I. within 7 Percent. The final sensitivity analysis is performed by relaxing stipulated accuracy even more. Results reported in this section 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 5 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, would be expected to further decrease sample size required over the previous stipulated accuracy.

Utilizing the 95 percent C.I. within 7 percent of normative mean level yields a smaller sample size requirement for cheese in every city except Terre Haute, Table 12. Comparing the results obtained from the relaxation from 5 to 7 percent reveals that the percent decrease in sample size requirement ranges from zero in Terre Haute to 17.6 percent in Toledo. The percent decrease for Dallas is 14.0 while it is only 3.7 for Omaha. Comparing Tables 10, 11, and 12 discloses that Toledo results are most sensitive followed by Dallas and Omaha with no sensitivity obtained for Terre Haute. This is roughly the relationship of the cities with respect to the absolute magnitude of the variance components.

Table 12. Normative mean deliveries, percent of universe required for a 95% C.I. within 7%, all cities and products, one week first differences.[a]

Product and City	Sample Chains	Sample Stores/Chain	Sample Total	Percent of Universe Required
	-number-	-number-	-number-	-percent-
Cheese:				
Dallas	4.4	36.3	160	72.4
Omaha	4.7	11.2	53	94.6
Toledo	5.0	12.2	61	65.6
Terre Haute	7.0	2.0	14	100.0
Butter:				
Dallas	6.0	32.3	194	87.8
Omaha	5.0	11.2	56	100.0
Toledo	5.0	18.3	92	98.9
Terre Haute	7.0	2.0	14	100.0
Margarine:[b]				
Dallas	6.0	16.7	101	45.7
Terre Haute	7.0	2.0	14	100.0

[a] Normative mean delivery per store per week was held invariant across cities. By product, they were: chese 700, butter 100, margarine 1,000.

[b] No usable margarine data for Omaha or Toledo were obtained.

Source: Computed.

For cheese data, successively relaxing stipulated accuracy successively reduced the number of chains required in Dallas and Omaha while the number of average stores per chain required remained stable. In Toledo, however, the reverse is true--the average number of stores per chain required is successively 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 3 times greater than the chain variance component.

The stipulated accuracy of a 95 percent C.I. within 7 percent of mean level using butter deliveries yields decreased sample size requirement only in Dallas and Toledo, Table 12. The Dallas decrease is from 93.7 percent of the universe to 87.8 percent, a reduction of 6.3 percent (Table 12 compared with Table 11). The slight Toledo decrease to 98.9 percent of the universe from the previous 100 represents only a 1.1 percent change. No decrease is obtained for Omaha or Terre Haute by relaxing stipulated accuracy from 5 to 7 percent. In fact, the Dallas reduction in sample size obtained from relaxing accuracy from a 95 percent C.I. within 3 percent to one within 7 percent is only 10.1 percent. In general, sensitivity of butter sample size requirements is relatively less than sensitivity of cheese sample size requirement to changes in stipulated accuracy.

The same analysis on the margarine data of Dallas and Terre Haute produces a further reduction in sample size requirement in Dallas, but not Terre Haute, Table 12. Dallas sample size required decreased from

62.0 percent of the universe to 45.7 percent, a reduction of 26.3 percent (Table 12 compared with Table 11). As before, all of this reduction is attributable to a smaller average number of stores per chain required. The difference in results obtained in the two cities may again be explained by comparison of the magnitude of their respective variance components.

Sensitivity Analysis Summary. The sensitivity analysis suggests a number of important relations. Results are not as sensitive to changes in mean level compared with changes in the m parameter of equation (7). 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 three products, butter sample size requirement is the least sensitive to changes in mean level or C.I. while cheese sample size requirement is the most sensitive to such changes. Also, even by relaxing accuracy from the standard 95 percent C.I. within 3 percent to one within 7 percent, the smallest sample size required is still nearly 66 percent of the universe for any product or city. This suggests that the magnitude of variability is so large that needed accuracy simply cannot be obtained from anything less than nearly the entire universe. And, in some instances, like Terre Haute, accuracy 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 sampling the entire universe.

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 the store in the market, a large proportion of the universe would need to be sampled in order to obtain accuracy necessary for measurement of promotion effects in that market.

Results reported in the text with respect to sensitivity are based entirely on one week first differences. If the same type analysis is performed on aggregated week (two and four week periods) first differences, results in terms of sample size requirement are reduced (Appendix C). These results suggest that monthly records would be superior to either two or one week periods in terms of sample size requirement. However, the final decision on longer than one week periods ultimately depends on the individual situation regarding the amount of time a promotion campaign may cover.

INDEPENDENT GROCERY AND HRI DELIVERY RECORDS--
A COMPARISON

As previously noted, delivery records were obtained from independent and affiliated food stores, public eating establishments, and non-profit institutions such as schools and hospitals. Even though product movement of butter and cheese through these establishments is not likely to be significantly influenced by short term promotion programs, delivery data from these establishments were necessary for complete market accountability.

A rough approximation of the importance of the HRI distribution channels nationally can be calculated by the ratio of disappearance to HRI movement. The latter data by product is available for 1969 from The Foodservice Industry [15]. The 1969 domestic civilian disappearance of cheese, butter, and margarine for food was obtained from Food Consumption, Prices, and Expenditures [14]. Using these data, approximately 21 percent of total cheese except cottage moved through HRI. Only about 11 percent of butter, exclusive of individual portions, moved through HRI. For margarine, about 8 percent moved through the HRI channels, exclusive of baker's margarine.

Of course, since these figures are for only one year and are national approximations, variation would exist over time or among metropolitan markets. However, the figures do indicate the relative importance of HRI distribution channels among the products of concern as well as indicate the general absolute importance of these channels.

HRI is a more important distribution channel for cheese than either butter or margarine yet nearly 80 percent of the cheese moves through non-HRI outlets (sold for home consumption).

Added to this consideration is the relative inelasticity of sales response to consumer promotion of the HRI sector compared to the retail grocery sector. The absolute level of short term sales in a market is not likely to be significantly influenced by a consumer promotion program in that market since most establishments have a "fixed" menu. Nearly all institutions and many restaurants (except cafeterias) utilize cheese, butter, and margarine in relation to items on their menus and thus would tend not to fluctuate usage due to market consumer promotion.

In light of these considerations, delivery records from the independent grocery and HRI sectors is given a cursory treatment relative to the chain grocery delivery data. Primary concern is with a comparison of variability in deliveries of chain grocery and other outlets through the coefficient of variation. In order to remain consistent with the preceding analysis of chain store deliveries, the coefficient of variation is computed for one week first differences and first differences on aggregated periods of two and four weeks.

Cheese

Data on the mean of first difference deliveries, the standard error of the mean and the ratio of the standard error to the mean were computed for each city for independent grocery stores and two major categories of HRI--restaurants and institutions, Tables 13 through 15. For comparative purposes, the same statistics for chain grocery stores are also included. No pattern or relationships exists for the coefficients of variation among cities or type of outlet for a given time period. However, comparing the same type outlet for a city over the one, two, and four week time periods reveals that variability substantially dampens down from aggregation of weekly periods. The standard error to mean ratio generally steadily declines as the measurement time period lengthens from one week to four weeks. This, of course, is consistent with the findings of the previous analysis which pertained only to chain grocery delivery records.

Butter and Margarine

The mean of first differences, standard error of the mean and standard error to mean ratios were also computed for butter and margarine deliveries for one week periods, as well as the two and four week aggregated periods, Tables 16 through 21. As with these statistics for cheese deliveries, no pattern is apparent for a given time period either among establishment types or among cities. Standard errors in relation to the means do generally dampen however, with the same type outlet for a city, as the measurement time period is increased from one to two, and then four week time periods. This again indicates that longer than one week time periods would reduce sample size required somewhat for a given level of accuracy.

Table 13. Means and their standard errors for first differences, one week time periods, cheese deliveries

	Dallas			Omaha			Toledo			Terre Haute		
Retail Outlet	Mean	SE	$SE/\bar{X}$	Mean	SE	$SE/\bar{X}$	Mean	SE	$SE/\bar{X}$	Mean	SE	$SE/\bar{X}$
Chains	11.28	40.95	3.6	-38.99	123.81	3.2	10.73	61.23	5.7	33.10	348.40	10.5
Independents	11.17	32.57	2.9	-26.98	75.78	2.8	3.61	193.7	53.7	-6.21	20.29	3.3
Restaurants	-0.82	1478.74	a/	-1.28	15.35	12.0	-16.36	64.36	3.9	3.12	26.03	8.3
Institutions	8.54	35.66	4.2	0.39	4.72	12.1	67.33	219.1	3.3	11.21	129.04	11.5

a/ Exceeds 100

Source: Computed

Table 14. Means and their standard errors for first differences, two week time periods, cheese deliveries

	Dallas			Omaha			Toledo			Terre Haute		
Retail Outlet	Mean	SE	$SE/\bar{X}$	Mean	SE	$SE/\bar{X}$	Mean	SE	$SE/\bar{X}$	Mean	SE	$SE/\bar{X}$
Chains	12.08	53.71	4.4	-218.15	146.24	0.7	34.89	78.5	2.2	207.45	429.08	2.1
Independents	26.11	49.36	1.9	-55.95	71.84	1.3	20.23	549.14	27.1	-11.67	33.61	2.9
Restaurants	-9.08	16.47	1.8	-2.29	15.26	6.7	-33.88	98.97	2.9	12.57	25.77	2.1
Institutions	31.75	62.39	2.0	1.30	5.67	4.4	155.55	332.17	2.1	-8.00	276.02	34.5

Source: Computed

Table 15. Means and their standard errors for first differences, four week time periods, cheese deliveries

	Dallas			Omaha			Toledo			Terre Haute		
Retail Outlet	Mean	SE	SE/$\bar{X}$	Mean	SE	SE/$\bar{X}$	Mean	SE	SE/$\bar{X}$	Mean	SE	SE/$\bar{X}$
Chains	-40.16	72.3	1.8	-830.11	268.22	0.3	-140.20	113.99	0.8	1056.73	662.72	0.6
Independents	86.50	50.34	0.6	-346.57	135.83	0.4	-326.60	388.33	1.2	5.83	30.3	5.2
Restaurants	-34.88	15.2	0.4	13.20	18.48	1.4	-22.64	19.48	0.9	1.59	38.13	24.0
Institutions	68.25	77.46	1.1	-2.41	5.87	2.4	453.00	582.98	1.3	-9.00	500.5	55.6

Source: Computed

Table 16. Means and their standard errors for first differences, one week time periods, butter deliveries

	Dallas			Omaha			Toledo			Terre Haute		
Retail Outlet	Mean	SE	SE/$\bar{X}$	Mean	SE	SE/$\bar{X}$	Mean	SE	SE/$\bar{X}$	Mean	SE	SE/$\bar{X}$
Chains	-1.81	16.0	8.8	-40.04	93.31	2.3	0.01	18.18	a/	39.0	74.32	8.4
Independents	0.26	7.7	29.6	-2.48	14.46	5.8	0.38	17.63	46.4	0.60	6.55	10.9
Restaurants	-0.53	783.76	a/	-1.67	10.41	6.2	-0.38	5.52	14.5	0.55	6.5	11.8
Institutions	6.43	25.08	3.9	1.4	0.61	0.4	b/	b/	b/	37.79	150.07	4.0

a/ Exceeds 100

b/ Sample size insufficient for computation

Source: Computed

Table 17. Means and their standard errors for first differences, two week time periods, butter deliveries

	Dallas			Omaha			Toledo			Terre Haute		
Retail Outlet	Mean	SE	SE/$\bar{X}$	Mean	SE	SE/$\bar{X}$	Mean	SE	SE/$\bar{X}$	Mean	SE	SE/$\bar{X}$
Chains	-34.66	20.66	0.6	-312.80	124.30	0.4	-109.89	44.58	0.4	50.47	115.04	2.3
Independents	4.24	9.37	2.2	-8.78	23.77	2.7	4.07	32.9	8.1	0.50	11.0	22.0
Restaurants	2.73	11.93	4.4	-9.63	16.47	1.7	-27.75	32.71	1.2	5.95	9.50	1.6
Institutions	15.17	75.17	5.0	3.33	0.92	0.3	a/	a/	a/	186.58	230.3	1.2

a/ Sample size insufficient for computation

Source: Computed

Table 18. Means and their standard errors for first differences, four week time periods, butter deliveries

	Dallas			Omaha			Toledo			Terre Haute		
Retail Outlet	Mean	SE	SE/$\bar{X}$	Mean	SE	SE/$\bar{X}$	Mean	SE	SE/$\bar{X}$	Mean	SE	SE/$\bar{X}$
Chains	-103.01	32.91	0.3	-990.46	161.85	0.2	-342.13	68.24	0.2	0.20	84.12	420.6
Independents	21.09	25.75	1.2	-9.33	10.67	1.1	26.00	31.33	1.2	4.50	16.05	3.6
Restaurants	11.88	21.63	1.8	-6.10	16.50	2.7	-92.75	35.2	0.4	24.47	14.98	0.6
Institutions	-44.50	31.37	0.7	10.00	1.60	0.2	a/	a/	a/	940.50	815.97	0.9

a/ Sample size insufficient for computation

Source: Computed

Table 19. Means and their standard errors for first differences, one week time periods, margarine deliveries

	Dallas			Terre Haute		
Retail Outlet	Mean	SE	SE/$\bar{X}$	Mean	SE	SE/$\bar{X}$
Chains	6.01	52.14	8.7	60.75	304.98	5.0
Independents	20.14	43.96	2.2	-2.40	53.68	22.4
Restaurants	3.15	12.4	3.9	-0.96	6.96	7.3
Institutions	27.71	105.44	3.8	-11.57	100.38	8.7

Source: Computed

Table 20. Means and their standard errors for first differences, two weeks time periods, margarine deliveries

	Dallas			Terre Haute		
Retail Outlet	Mean	SE	SE/$\bar{X}$	Mean	SE	SE/$\bar{X}$
Chains	-20.44	60.17	2.9	92.06	473.54	5.1
Independents	65.40	66.65	1.0	4.93	93.42	18.9
Restaurants	15.58	16.25	1.0	-4.28	20.12	4.7
Institutions	76.00	123.88	1.6	67.44	108.88	1.6

Source: Computed

Table 21. Means and their standard errors for first differences, four weeks time periods, margarine deliveries

Retail Outlet	Dallas Mean	Dallas SE	Dallas $SE/\bar{X}$	Terre Haute Mean	Terre Haute SE	Terre Haute $SE/\bar{X}$
Chains	-47.83	96.22	2.0	-357.64	422.70	1.2
Independents	239.60	71.52	0.3	-173.80	65.47	0.4
Restaurants	63.32	25.85	0.4	22.67	8.70	0.4
Institutions	222.00	280.83	1.3	440.00	367.97	0.8

Source: Computed

SUMMARY AND CONCLUSIONS

The Research Objective

The purpose of the study was to determine the feasibility of measuring product movement through food chain stores, in a given market, by using delivery data to stores instead of employing the traditional store auditing methods. The latter system, store audits, requires taking periodic individual store level inventories. Costs of obtaining individual store inventories represent the major component of the field research expenditures.

Optimum sampling of delivery data to individual food chain stores was conducted in four markets. These were Dallas, Texas; Toledo, Ohio; Omaha, Nebraska; and Terre Haute, Indiana. These cities were selected because they provided a reasonable variation with regard to geographic location as well as market size within the national market.

Delivery records for individual stores on a weekly basis were obtained for a period of eight weeks from cooperating food chains in the four cities. Aside from butter and cheese, data on margarine also were obtained in Dallas and Terre Haute for comparative purposes.

The Analysis

The methodology reported here indicates the optimum sampling allocation among component sampling strata to achieve a designated accuracy at minimum cost. In the present application this means the optimum division of sampling among 1) food chains and 2) stores within chains for a given market and level of measurement error.

Application of the methodology may be made to a measurement situation associated with market-wide promotion programs. Such promotion programs are sponsored by American Dairy Association, as well as other major food commodity producer groups and processors. However, any given short-run mass media advertising program for a market, without a price special, couponing or related incentives, may not change product movement (the total of all brands) by more than a small percent. Under these circumstances, measurement accuracy must be stringent (such as a 95 percent confidence level within 3 or 5 percent of mean deliveries).

A finite population correction factor was applied in the methodology. This permits sample size determination consonant with the size of the market involved. Since two sources of variation exist, chains and stores within chains, correction factors were used for both.

In nearly all cases, measurement of changes within 3 percent accuracy for weekly or biweekly periods required the full universe of chains and stores (see text, Tables 5 and 7). The product movement data became somewhat more stable on a 4-week period basis. As a result, the required sample for cheese reduces to 73 percent of the universe in the Dallas market and to about 21 percent of the universe in Toledo, Table 9.

The Toledo case, however, is considered to be an atypical situation due to possible data error and not one from which to draw conclusions.

Further indications of the effect of lengthening the sampling period from 1-week to 4-week units are found in Tables 22 to 24. Sample size requirements shown are associated with an accuracy of ± 5 percent in mean change in deliveries per store. Perceptible lowering of sample size requirements occurs in the Dallas market. Little, if any, lowering of sample size as a proportion of the universe occurs in Omaha, Toledo or Terre Haute.

A countervailing influence is present in sample size requirements noted above. Whereas it takes 100 percent of the universe in Terre Haute, that represents only 14 stores. In Omaha, the universe is 56 stores and in Toledo 93 stores. Since the cost of delivery data is greater for the inclusion of a chain ($500) than for a store within a chain ($34), the data collection cost among the four cities would not be substantially different. Estimates for cheese measurement using 4-week data are Dallas, $6,002; Terre Haute, $3,905; Omaha, $4,404; and Toledo, $3,132.[1/] The significance of the cost estimates is that one may select the test market on other considerations than market measurement costs. Advertising would cost more in absolute dollars in a city the size of Dallas, so total research costs including advertising expense should be less expensive in Terre Haute.

[1/] Again, actual figures are not used from Table 24 for Toledo because results represent some unknown data error.

Table 22. Effect of time period length on sample size requirements to obtain no more than five percent error in measuring changes in deliveries of butter, cheese and margarine to chain food stores.1/ Dallas Metropolitan Market.

Product and Time Period	Food Chains Sample	Food Chains Proportion of Universe	Stores per Chain Sample	Stores per Chain Proportion of Universe	Total Sample as Proportion of Universe
	No.	%	No.	%	%
Butter					
1-week	6	100	35.3	96	94
2-week	6	100	31.4	85	86
4-week	6	100	26.7	73	73
Cheese					
1-week	5	83	36	98	84
2-week	6	100	18.7	51	51
4-week	5	83	20.6	56	47
Margarine					
1-week	6	100	23	62	62
2-week	6	100	11.5	31	31
4-week	5	83	16	43	36

1/Confidence level of 95 percent

Source: Computed

Table 23. Effect of time period on sample size requirement to obtain no more than 5 percent error in measuring changes in deliveries of butter, cheese and margarine to chain food stores.1/ Terre Haute Market.

Product and Time Period	Food Chains		Stores per Chain		Total Sample as Proportion of Universe
	Sample	Proportion as Universe	Sample	Proportion of Universe	
	No.	%	No.	%	%
Butter					
1-week	7	100	2	100	100
2-week	7	100	2	100	100
4-week	7	100	2	100	100
Cheese					
1-week	7	100	2	100	100
2-week	7	100	1.9	95	100
4-week	7	100	1.7	85	86
Margarine					
1-week	7	100	2	100	100
2-week	7	100	2	100	100
4-week	7	100	2	100	100

1/Confidence level of 95 percent.

Source: Computed

Table 24. Effect of time period on sample size requirements to obtain no more than 5 percent error in measuring change in deliveries of butter and cheese to chain food stores.1/ Omaha and Toledo Markets.

Product and Time Period	Food Chains		Stores per Chain		Total Sample as Proportion of Universe
	Sample	Proportion of Universe	Sample	Proportion of Universe	
	No.	%	No.	%	%
			OMAHA		
Butter					
1-week	5	100	11.2	100	100
2-week	5	100	11.2	100	100
4-week	5	100	11.2	100	100
Cheese					
1-week	5	100	11.2	100	100
2-week	5	100	11.2	100	100
4-week	5	100	11.2	100	100
			TOLEDO		
Butter					
1-week	5	100	18.5	99	100
2-week	5	100	17.6	95	95
4-week	5	100	18.6	100	100
Cheese					
1-week	5	100	14.7	79	80
2-week	5	100	11.3	61	61
4-week	1	20	18.6	100	20

1/ Confidence level of 95 percent.

Source: Computed

Finally, it is useful to examine the effect of relaxing the accuracy of measurement requirements. Levels of 3, 5, and 7 percent were examined as to their impact upon sampling requirements. Results for 4-week periods are given in Tables 25 through 28. Definite sample reduction is possible with more allowed measurement error in a market like Dallas, Table 25. Sample size as a proportion of the universe declines as error allowed increases from ± 3 percent to ± 7 percent. For the other three cities, no consistent pattern emerged.

General Conclusions

Several general conclusions are warranted by the analysis. These are:

1. Varying the required accuracy level from 3 to 7 percent definitely reduces sample size in a large market like Dallas. For smaller, i.e., less populated markets, a change in the stipulated accuracy level has little effect on sample size in relation to the total universe of stores in the market.

2. Lengthening the measurement time period from one-week up to four-week intervals reduced sample size required in a market like Dallas. Reflected is higher autocorrelation of weekly deliveries compared to a four-week period. Extending the time length of the measurement period was not uniformly effective in lowering sample requirements in the other three cities. Cheese in Toledo was the exception.

3. Average store size differences among the cities had little effect on sample requirements.

Table 25. Effect of accuracy stipulation on sample size requirements to measure changes in 4-week period deliveries of butter, cheese and margarine. Dallas Metro Market Area.

Product and Measurement Error Level 1/	Food Chains		Stores per Chain		Total Sample as Proportion of Universe
	Sample	Proportion of Universe	Sample	Proportion of Universe	
	No.	%	No.	%	%
Butter					
3%	6	100	32	87	88
5%	6	100	27	73	73
7%	6	100	21	57	58
Cheese					
3%	6	100	21	57	56
5%	5	83	21	57	47
7%	3	50	21	57	28
Margarine					
3%	6	100	19	51	51
5%	5	83	16	43	36
7%	4	67	16	43	29
Universe:					
(chains or stores)	6		36.8		221

1/ Confidence level of 95 percent.

Source: Computed

Table 26. Effect of accuracy stipulation on sample size requirements to measure changes in 4-week period deliveries of butter and cheese. Omaha Market Area.

Product and Measurement Error Level 1/	Food Chains		Stores per Chain		Total Sample as Proportion of Universe
	Sample	Proportion of Universe	Sample	Proportion of Universe	
	No.	%	No.	%	%
Butter					
3%	5	100	11.2	100	100
5%	5	100	11.2	100	100
7%	5	100	11.2	100	100
Cheese					
3%	5	100	11.2	100	100
5%	5	100	11.2	100	100
7%	4	80	11.2	100	80
Universe:					
(chains or stores)	5		11.2		56

1/ Confidence level of 95 percent.

Source: Computed

Table 27. Effect of accuracy stipulation on sample size requirements to measure changes in 4-week period deliveries of butter and cheese, Toledo Market Area.

Product and Measurement Error Level 1/	Food Chains: Sample	Food Chains: Proportion of Universe	Stores per Chain: Sample	Stores per Chain: Proportion of Universe	Total Sample as Proportion of Universe
	No.	%	No.	%	%
Butter					
3%	5	100	18.6	100	100
5%	5	100	18.6	100	100
7%	5	100	18.6	100	100
Cheese					
3%	1	20	18.6	100	20
5%	1	20	18.6	100	20
7%	1	20	18.6	100	20
Universe:					
(chains or stores)	5		18.6		93

1/ Confidence level of 95 percent.

Source: Computed

Table 28. Effect of accuracy stipulation on sample size requirement to measure changes in 4-week period deliveries of butter, cheese and margarine, Terre Haute Market.

Product and Measurement Error Level 1/	Food Chains		Stores per Chain		Total Sample as Proportion of Universe
	Sample	Proportion of Universe	Sample	Proportion of Universe	
	No.	%	No.	%	%
Butter					
3%	7	100	2	100	100
5%	7	100	2	100	100
7%	7	100	2	100	100
Cheese					
3%	7	100	1.9	95	100
5%	7	100	1.7	85	86
7%	7	100	1.4	70	71
Margarine					
3%	7	100	2	100	100
5%	7	100	2	100	100
7%	7	100	2	100	100
Universe:					
(chains or stores)	7		2		14

1/ Confidence level of 95 percent.

Source: Computed

4. Delivery data, for the three test products, were far more variable among chains than among stores within a chain. Consequently, in most instances all chains remained in the required sample and reductions occurred in the average number of stores required within a chain. Reflected here are differences in merchandising policies for the products over chains.

5. Sample size requirements across cities, as a percent of the universe, were found to be essentially similar for one-week first differences.

6. Whereas autocorrelation in delivery data does exist in some data, use of longer time periods dampens its level significantly. However, in no case were estimated autocorrelation coefficients over 40 percent, though they were generally all significantly different from zero (see Appendix B).

7. Inventory changes are a smaller percent of deliveries in large cities than in smaller ones; also in large stores than in smaller ones (see Appendix F).

8. Ability to measure deliveries with a given accuracy will vary regionally with the rate of consumer usage of the product. Thus, butter is more difficult to measure in Dallas, where butter is not as widely used, than in Toledo which has a higher per capita butter consumption rate.

9. It is inadvisable to endeavor to use any market as a test area unless prior information is obtained and analyzed as to the existing variance in the rate of product movement among and within chains.

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] Clement, W.E., P.L. Henderson, and C.P. Eley, The Effect of Different Levels of Promotional Expenditures on Sales of Fluid Milk, U.S. Department of Agriculture, Economic Research Service Report 259, October, 1965.

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

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

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

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

[7] Henderson, P.L. and S.E. Brown, Effectiveness of a Special Promotional Campaign for Frozen Concentrated Orange Juice, U.S. Department of Agriculture, Marketing Research Report No. 356, March, 1961.

[8] Henderson, P.L., S.E. Brown, and J.F. Hind, Special Promotional Programs for Apples, Their Effects on Sales of Apples and Other Fruits, U.S. Department of Agriculture, Marketing Research Report No. 446, 1962.

[9] 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.

[10] 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, Sept. 1968.

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

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

[13] Twining, C.R. and P.L. Henderson, Promotional Activities of Agricultural Groups, U.S. Department of Agriculture, Economic Research Service, June, 1965.

[14] U.S. Department of Agriculture, Food Consumption, Prices, and Expenditures, Supplement to Report No. 138, Economic Research Service, Supplement for 1970, November, 1971.

[15] U.S. Department of Agriculture, The Foodservice Industry: Type, Quantity, and Value of Foods Used, Statistical Bulletin No. 476, Marketing Economics Division, Economic Research Service, Washington, D. C., November, 1971.

APPENDICES

APPENDIX A

The Hartley - Hocking Algorithm

Minimization of cost subject to specified accuracy, MINCOST presented the following problem:[1/]

(A-1) minimize $c(K_c + sK_s)$

subject to:

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

Hartley and Hocking's convex programming by tangential approximation was designed to solve problems of this type [5]. A Convex Program employing the Hartley - Hocking algorithm has been documented by LeMotte and Oxspring [9].

The general form of problems amenable for the Hartley - Hocking convex programming algorithm is:

(A-3) maximize $g(x)$

subject to:

(A-4) $f_i(x) \leq 0,\ i = 1, \ldots, r$

where $g(x)$ is concave and the $f_i(x)$ are convex, real valued functions of the n-vector x for all real x, and the functions are differentiable. Of course, minimization of the objective function, (A-3), can be accomplished by maximizing $-g(x)$.

In order to render the problem stated by equations (A-1) and (A-2) conformable to the H-H convex program, those equations need to be transformed.

[1/] See text page 20 for definition of variables.

To make the constraint globally convex, the following transformations are performed:

(A-5) $X_1 = c$

(A-6) $X_2 = s\ c$

The transformed problem statement becomes:

(A-7) maximize $-(X_1 K_c + X_2 K_s)$

subject to:

(A-8) $$\left(\frac{\hat{\sigma}_c^2}{c} - \frac{\hat{\sigma}_c^2}{N}\right) + \left(\frac{\hat{\sigma}_{s(c)}^2}{sc} - \frac{\hat{\sigma}_{s(c)}^2}{cM}\right) \leq V$$

from equation (A-2),

or

(A-9) $$\left(\frac{\hat{\sigma}_c^2 - \frac{\hat{\sigma}_{s(c)}^2}{M}}{x_1}\right) + \left(\frac{\hat{\sigma}_{s(c)}^2}{x_2}\right) \leq V + \left(\frac{\hat{\sigma}_c^2}{N}\right)$$

from equations (A-5) and (A-6),

where the bounds are:

(A-10) $o < c < N \Rightarrow 0 < x_1 < N$

(A-11) $0 < s < M$

which together imply:

(A-12) $o < sc < Mc$

(A-13) $\Rightarrow sc - Mc \leq 0$

(A-14) $\Rightarrow x_2 - Mx_1 \leq 0$

Also,

(A-15) $0 < sc < Mc$ and $0 < c < N$

(A-16) $0 < x_2 < MN$

which provides the upper bound for X_2.

The bounds on equation (A-9), for purposes of the H-H convex program, become:

(A-17) $\quad X_2 - MX_1 \leq 0$ from equation (A-14)

(A-18) $\quad 0 < X_1 < N$ from equation (A-10)

and

(A-19) $\quad 0 < X_2 < NM$ from equation (A-16)

Even though equation (A-19) is redundant it is used as a boundary constraint on equation (A-9) in the H-H convex program.

APPENDIX B

Autocorrelation

In any time series data, the possibility exists for autocorrelation (i.e., the correlation of a variable in time period t and time period t + h). The ANOVA model yields no information concerning autocorrelation. In order to investigate its existence in deliveries per store over time, first, second, and third order autocorrelation coefficients were estimated. Let $\hat{\rho}_h$ be the h order autocorrelation coefficient, then:

$$(B\text{-}1) \qquad \hat{\rho}_h = \frac{\Sigma_k \left[\Sigma_i X_i X_{i+h} - \frac{(\Sigma_i X_i)^2}{N} \right]}{\Sigma_k \left[\Sigma_i X_i^2 - \frac{(\Sigma_i X_i)^2}{N} \right]}$$

where

X_i = deliveries in pounds (cheese, butter, or margarine) during the i^{th} week for store K

X_{i+h} = deliveries in pounds during the i^{th} + h week for store k.

h = value of lag (h = 1 is first order correlation, etc.)

and

i = 1, 2, . . ., w

k = 1, 2, . . ., s

h = 1, 2, 3

Estimates of these autocorrelation coefficients are presented with their respective t values by establishment type, product, and city, Tables B-1 through B-12. In general, the estimated autocorrelation coefficients were significantly different from zero but were low (under 0.40). Thus, even though autocorrelation in the original delivery data does exist, it is not a serious problem.

Table B-1. Estimated autocorrelation coefficients, cheese, chain supermarkets, by city

City	Lag in Weeks		
	1	2	3
Dallas	-0.2037 (4.8262)[a/] *	-0.1297 (3.0580) *	-0.0967 (2.2680) *
Omaha	-0.2352 (2.9799) *	-0.1174 (1.4490)	-0.0324 (0.4988)
Terre Haute	-0.0764 (0.5012)	-0.2589 (1.9955) *	-0.2066 (1.5669)
Toldedo	-0.2507 (3.5725) *	-0.0589 (0.7864)	-0.2383 (3.3925) *

Table B-2. Estimated autocorrelation coefficients, butter, chain supermarkets, by city

City	Lag in Weeks		
	1	2	3
Dallas	-0.0847 (1.9929)[a/] *	-0.2066 (4.9208) *	-0.1529 (3.6237) *
Omaha	-0.0162 (0.1338)	-0.1476 (1.8414)	-0.2324 (2.9420) *
Terre Haute	-0.0096 (0.4564)	-0.5104 (4.0542) *	-0.0866 (0.5851)
Toledo	-0.2502 (4.8911) *	-0.1212 (2.3425) *	-0.0554 (1.0426)

[a/] t values for estimated coefficients in parentheses.
*Significantly different from zero at 5 percent.
Source: Computed from delivery data.

Table B-3. Estimated autocorrelation coefficients, margarine, chain supermarkets, by city

City	Lag in Weeks 1	2	3
Dallas	-0.2324 (5.5404)[a/]*	-0.1610 (3.8248)*	-0.0543 (1.2611)
Terre Haute	-0.0207 (0.4526)	-0.2259 (1.7249)	-0.1123 (0.7957)

Table B-4. Estimated autocorrelation coefficients, cheese, independents, by city

City	Lag in Weeks 1	2	3
Dallas	0.1697 (1.5316)[a/]	-0.2104 (1.6255)	-0.1435 (1.0696)
Omaha	-0.2371 (2.5821)*	0.1238 (1.4845)	-0.2126 (2.3066)*
Terre Haute	-0.2201 (1.1701)	-0.1009 (0.4450)	-0.2727 (1.4901)
Toledo	0.4240 (3.4420)*	-0.0940 (0.6109)	-0.4796 (3.6158)*

[a/] t values for estimated coefficients in parentheses.
*Significantly different from zero at 5 percent.
Source: Computed from delivery data.

Table B-5. Estimated autocorrelation coefficients, butter, independents, by city

City	Lag in Weeks 1	2	3
Dallas	0.0246 (0.3378)[a]	-0.0257 (0.1311)	0.0586 (0.6553)
Omaha	-0.1137 (0.5684)	-0.0309 (0.3766)	-0.1303 (0.6745)
Terre Haute	0.0241 (0.3156)	-0.2029 (1.0657)	-0.2747 (1.5028)
Toledo	0.0935 (0.8309)	-0.0413 (0.1692)	-0.2896* (2.0110)

Table B-6. Estimated autocorrelation coefficients, margarine, independents, by city

City	Lag in Weeks 1	2	3
Dallas	0.1181 (1.1752)[a]	-0.0499 (0.3365)	-0.1484 (1.2233)
Terre Haute	-0.1198 (0.4816)	-0.1034 (0.3906)	-0.2754 (1.3486)

[a] t values for estimated coefficients in parentheses.
*Significantly different from zero at 5 percent.
Source: Computed from delivery data.

Table B-7. Estimated autocorrelation coefficients, cheese, restaurants, by city

City	Lag in Weeks 1	2	3
Dallas	-0.1202 (2.3639)[a/] *	-0.1110 * (2.3588)	-0.2246 * (4.4589)
Omaha	-0.3466 (3.3830)*	-0.2294 (2.2055)*	-0.1333 (1.2397)
Terre Haute	-0.5066 (6.6252)*	-0.3781 (5.0782)*	-0.4217 (5.5028)*
Toledo	-0.3285 (2.6885)*	-0.3279 (2.6831)*	-0.3270 (2.6751)*

Table B-8. Estimated autocorrelation coefficients, butter, restaurants, by city

City	Lag in Weeks 1	2	3
Dallas	0.3389 (4.3266)[a/] *	0.2000 (2.5859)*	0.0121 (0.2316)
Omaha	-0.1900 (1.4806)	-0.2809 (2.2472)*	-0.2953 (2.3683)*
Terre Haute	-0.1667 (1.7265)	-0.2091 (2.1890)*	-0.2812 (2.9757)*
Toledo	0.6436 (4.1852)*	0.3848 (2.5683)*	0.3383 (2.2776)*

[a/] t values for estimated coefficients in parentheses.
*Significantly different from zero at 5 percent.
Source: Computed from delivery data.

Table B-9. Estimated autocorrelation coefficients, margarine, restaurants, by city

City	Lag in Weeks 1	2	3
Dallas	-0.0315 (0.2550)[a]	0.0454 (0.5908)	-0.1464 (1.5193)
Terre Haute	0.1429 (1.3838)	-0.5978 (5.1995)*	-0.1436 (1.1625)

Table B-10. Estimated autocorrelation coefficients, cheese, institutions, by city

City	Lag in Weeks 1	2	3
Dallas	0.0673 (0.5552)[a]	-0.2185 (0.9850)	-0.2237 (1.0128)
Omaha	-0.3428 (7.3126)*	-0.1540 (3.2601)*	0.0164 (0.3987)
Terre Haute	0.5882 (4.3445)*	0.3692 (2.7803)*	0.3499 (2.6419)*
Toledo	-0.1428 (0.3810)	-0.1431 (0.3820)	-0.1425 (0.3797)

[a] t values for estimated coefficients in parentheses.
*Significantly different from zero at 5 percent.
Source: Computed from delivery data.

Table B-11. Estimated autocorrelation coefficients, butter, institutions, by city

City	Lag in Weeks 1	2	3
Dallas	0.2595 (1.2405)[a/]	-0.2928 (0.7581)	-0.4583 (1.3568)
Omaha	-0.1429 (0.2154)	-0.1429 (0.2154)	-0.1429 (0.2154)
Terre Haute	0.8403 (5.4138)*	0.7623 (4.9265)*	0.6978 (4.5238)*
Toledo	-0.3333 (1.2127)	-0.3333 (1.2127)	-0.3333 (1.2127)

Table B-12. Estimated autocorrelation coefficients, margarine, institutions, by city

City	Lag in Weeks 1	2	3
Dallas	0.0006 (0.2321)[a/]	-0.0415 (0.0390)	0.0772 (0.5335)
Terre Haute	0.7213 * (3.6808)	0.5035 * (2.6352)	0.2813 (1.5689)

[a/] t values for estimated coefficients in parentheses.
*Significantly different from zero at 5 percent.
Source: Computed from delivery data.

APPENDIX C

SAMPLE SIZE WITH NORMATIVE MEAN

Table C-1. Sample size and percent of universe required for a 95 percent confidence interval within 3 percent of mean deliveries per store per week, and related statistics, two week difference, cheese, all cities.

Statistic	City			
	Dallas	Omaha	Toledo	Terre Haute
Variance Component for:				
Chains	18,517	29,584	53,608	2,163,802
Stores/Chains	267,444	638,640	182,398	446,671
Universe Chains	6.0	5.0	5.0	7.0
Universe Stores	36.8	11.2	18.6	2.0
Universe Total	221	56	93	14
Sample Chains	6.0	5.0	5.0	7.0
Sample Stores	26.7	11.2	15.1	2.0
Sample Total	161	56	76	14
Mean*	1,400	1,400	1,400	1,400
Percent of Universe Required	73.8	100.0	81.7	100.0

*Normative mean delivery in pounds per store per period.

Source: Computed.

Table C-2. Sample size and percent of universe required for a 95 percent confidence interval within 5 percent of mean deliveries per store per week, and related statistics, two week difference, cheese, all cities.

Statistic	City			
	Dallas	Omaha	Toledo	Terre Haute
Variance Component for:				
Chains	18,517	29,584	53,608	2,163,802
Stores/Chains	267,444	638,640	182,398	446,671
Universe Chains	6.0	5.0	5.0	7.0
Universe Stores	36.8	11.2	18.6	2.0
Universe Total	221	56	93	14.0
Sample Chains	6.0	5.0	5.0	7.0
Sample Stores	18.7	11.2	11.3	1.9
Sample Total	113	56	57	14
Mean*	1,400	1,400	1,400	1,400
Percent of Universe Required	51.1	100.0	61.3	100.0

*Normative mean delivery in pounds per store per period.

Source: Computed.

Table C-3. Sample size and percent of universe required for a 95 percent confidence interval within 7 percent of mean deliveries per store per week, and related statistics, two week difference, cheese, all cities.

Statistic	City			
	Dallas	Omaha	Toledo	Terre Haute
Variance Component for:				
Chains	18,517	29,584	53,608	2,163,802
Stores/Chains	267,444	638,640	182,398	446,671
Universe Chains	6.0	5.0	5.0	7.0
Universe Stores	36.8	11.2	18.6	2.0
Universe Total	221	56	93	14
Sample Chains	5.0	4.0	5.0	7.0
Sample Stores	18.7	11.2	8.2	1.8
Sample Total	94	45	41	13
Mean*	1,400	1,400	1,400	1,400
Percent of Universe Required	42.5	80.4	44.1	92.9

*Normative mean delivery in pounds per store per period.

Source: Computed.

Table C-4. Sample size and percent of universe required for a 95 percent confidence interval within 3 percent of mean deliveries per store per week, and related statistics, two week difference, butter, all cities.

Statistic	City			
	Dallas	Omaha	Toledo	Terre Haute
Variance Component for:				
Chains	6,504	31,423	11,146	123,367
Stores/Chains	33,095	296,413	45,093	33,243
Universe Chains	6.0	5.0	5.0	7.0
Universe Stores	36.8	11.2	18.6	2.0
Universe Total	221	56	93	14
Sample Chains	6.0	5.0	5.0	7.0
Sample Stores	34.6	11.2	18.2	2.0
Sample Total	208	56	91	14
Mean*	200	200	200	200
Percent of Universe Required	94.1	100.0	97.8	100.0

*Normative mean delivery in pounds per store per period.

Source: Computed.

Table C-5. Sample size and percent of universe required for a 95 percent confidence interval within 5 percent of mean deliveries per store per week, and related statistics, two week difference, butter, all cities.

Statistic	City			
	Dallas	Omaha	Toledo	Terre Haute
Variance Component for:				
Chains	6,504	31,423	11,146	123,367
Stores/Chains	33,095	296,413	45,093	33,243
Universe Chains	6.0	5.0	5.0	7.0
Universe Stores	36.8	11.2	18.6	2.0
Universe Total	221	56	93	14
Sample Chains	6.0	5.0	5.0	7.0
Sample Stores	31.4	11.2	17.6	2.0
Sample Total	189	56	88	14
Mean*	200	200	200	200
Percent of Universe Required	85.5	100.0	94.6	100.0

*Normative mean delivery in pounds per store per period.

Source: Computed.

Table C-6. Sample size and percent of universe required for a 95 percent confidence interval within 7 percent of mean deliveries per store per week, and related statistics, two week difference, butter, all cities.

Statistic	City			
	Dallas	Omaha	Toledo	Terre Haute
Variance Component for:				
Chains	6,504	31,423	11,146	123,367
Stores/Chains	33,095	296,413	45,093	33,243
Universe Chains	6.0	5.0	5.0	7.0
Universe Stores	36.8	11.2	18.6	2.0
Universe Total	221	56	93	14
Sample Chains	6.0	5.0	5.0	7.0
Sample Stores	27.5	11.2	16.8	2.0
Sample Total	165	56	84	14
Mean*	200	200	200	200
Percent of Universe Required	74.7	100.0	90.3	100.0

*Normative mean delivery in pounds per store per period.

Source: Computed.

Table C-7. Sample size and percent of universe required for a 95 percent confidence interval within 3 percent of mean deliveries per store per week, and related statistics, two week difference, margarine, two cities.

Statistic	City	
	Dallas	Terre Haute
Variance Component for:		
Chains	111,560	1,708,595
Stores/Chains	261,301	1,482,142
Universe Chains	6.0	7.0
Universe Stores	36.8	2.0
Universe Total	221	14
Sample Chains	6.0	7.0
Sample Stores	20.5	2.0
Sample Total	123	14
Mean*	2,000	2,000
Percent of Universe Required	55.7	100.0

*Normative mean delivery in pounds per store per period.

Source: Computed.

Table C-8. Sample size and percent of universe required for a 95 percent confidence interval within 5 percent of mean deliveries per store per week, and related statistics, two week difference, margarine, two cities.

Statistic	City	
	Dallas	Terre Haute
Variance Component for:		
Chains	111,560	1,708,545
Stores/Chains	261,301	1,482,142
Universe Chains	6.0	7.0
Universe Stores	36.8	2.0
Universe Total	221	14
Sample Chains	6.0	7.0
Sample Stores	11.5	2.0
Sample Total	69	14
Mean*	2,000	2,000
Percent of Universe Required	31.2	100.0

*Normative mean delivery in pounds per store per period.

Source: Computed.

Table C-9. Sample size and percent of universe required for a 95 percent confidence interval within 7 percent of mean deliveries per store per week, and related statistics, two week difference, margarine, two cities.

Statistic	City	
	Dallas	Terre Haute
Variance Component for:		
Chains	111,560	1,708,545
Stores/Chains	261,301	1,482,142
Universe Chains	6.0	7.0
Universe Stores	36.8	2.0
Universe Total	221	14
Sample Chains	6.0	7.0
Sample Stores	6.9	2.0
Sample Total	42	14
Mean*	2,000	2,000
Percent of Universe Required	19.0	100.0

*Normative mean delivery in pounds per store per period.

Source: Computed.

Table C-10. Sample size and percent of universe required for a 95 percent confidence interval within 3 percent of mean deliveries per store per week, and related statistics, four week difference, cheese, all cities.

Statistic	City			
	Dallas	Omaha	Toledo	Terre Haute
Variance Component for:				
Chains	30,429	128,967	750	5,322,731
Stores/Chains	490,923	2,104,997	258,308	395,446
Universe Chains	6.0	5.0	5.0	7.0
Universe Stores	36.8	11.2	18.6	2.0
Universe Total	221	56	93	14
Sample Chains	6.0	5.0	1.0	7.0
Sample Stores	20.6	11.2	18.6	1.9
Sample Total	124	56	19	14
Mean*	2,800	2,800	2,800	2,800
Percent of Universe Required	56.1	100.0	20.4	100.0

*Normative mean delivery in pounds per store per period.

Source: Computed.

Table C-11. Sample size and percent of universe required for a 95 percent confidence interval within 5 percent of mean deliveries per store per week, and related statistics, four week difference, cheese, all cities.

Statistic	City			
	Dallas	Omaha	Toledo	Terre Haute
Variance Component for:				
Chains	30,429	128,967	750	5,322,731
Stores/Chains	490,923	2,104,997	258,308	395,446
Universe Chains	6.0	5.0	5.0	7.0
Universe Stores	36.8	11.2	18.6	2.0
Universe Total	221	56	93	14
Sample Chains	5.0	5.0	1.0	7.0
Sample Stores	20.6	11.2	18.6	1.7
Sample Total	103	56	19	12
Mean*	2,800	2,800	2,800	2,800
Percent of Universe Required	46.6	100.0	20.4	85.7

*Normative mean delivery in pounds per store per period.

Source: Computed.

Table C-12. Sample size and percent of universe required for a 95 percent confidence interval within 7 percent of mean deliveries per store per week, and related statistics, four week difference, cheese, all cities.

Statistic	City			
	Dallas	Omaha	Toledo	Terre Haute
Variance Component for:				
Chains	30,429	128,967	750	5,322,731
Stores/Chains	490,923	2,104,997	258,308	395,446
Universe Chains	6.0	5.0	5.0	7.0
Universe Stores	36.8	11.2	18.6	2.0
Universe Total	221	56	93	14
Sample Chains	3.0	4.0	1.0	7.0
Sample Stores	20.6	11.2	18.6	1.4
Sample Total	62	45	19	10
Mean*	2,800	2,800	2,800	2,800
Percent of Universe Required	28.1	80.4	20.4	71.4

*Normative mean delivery in pounds per store per period.

Source: Computed.

Table C-13. Sample size and percent of universe required for a 95 percent confidence interval within 3 percent of mean deliveries per store per week, and related statistics, four-week difference, butter, all cities.

Statistic	City			
	Dallas	Omaha	Toledo	Terre Haute
Variance Component for:				
Chains	38,080	62,455	15,162	40,543
Stores/Chains	60,960	753,683	160,289	67,377
Universe Chains	6.0	5.0	5.0	7.0
Universe Stores	36.8	11.2	18.6	2.0
Universe Total	221	56	93	14
Sample Chains	6.0	5.0	5.0	7.0
Sample Stores	32.4	11.2	18.6	2.0
Sample Total	195	56	93	14
Mean*	400	400	400	400
Percent of Universe Required	88.2	100.0	100.0	100.0

*Normative mean delivery in pounds per store per period.

Source: Computed.

Table C-14. Sample size and percent of universe required for a 95 percent confidence interval within 5 percent of mean deliveries per store per week, and related statistics, four week difference, butter, all cities.

Statistic	City			
	Dallas	Omaha	Toledo	Terre Haute
Variance Component for:				
Chains	38,080	62,455	15,162	40,543
Stores/Chains	60,960	753,683	160,289	67,377
Universe Chains	6.0	5.0	5.0	7.0
Universe Stores	36.8	11.2	18.6	2.0
Universe Total	221	56	93	14
Sample Chains	6.0	5.0	5.0	7.0
Sample Stores	26.7	11.2	18.6	2.0
Sample Total	161	56	93	14
Mean*	400	400	400	400
Percent of Universe Required	72.9	100.0	100.0	100.0

*Normative mean delivery in pounds per store per period.

Source: Computed.

Table C-15. Sample size and percent of universe required for a 95 percent confidence interval within 7 percent of mean deliveries per store per week, and related statistics, four week difference, butter, all cities.

Statistic	City			
	Dallas	Omaha	Toledo	Terre Haute
Variance Component for:				
Chains	38,080	62,455	15,162	40,543
Stores/Chains	60,960	753,683	160,289	67,377
Universe Chains	6.0	5.0	5.0	7.0
Universe Stores	36.8	11.2	18.6	2.0
Universe Total	221	56	93	14
Sample Chains	6.0	5.0	5.0	7.0
Sample Stores	21.2	11.2	18.6	2.0
Sample Total	128	56	93	14
Mean*	400	400	400	400
Percent of Universe Required	57.9	100.0	100.0	100.0

*Normative mean delivery in pounds per store per period.

Source: Computed.

Table C-16. Sample size and percent of universe required for a 95 percent confidence interval within 3 percent of mean deliveries per store per week, and related statistics, four week difference, margarine, two cities.

Statistic	City	
	Dallas	Terre Haute
Variance Component for:		
Chains	72,344	876,089
Stores/Chains	856,154	2,325,196
Universe Chains	6.0	7.0
Universe Stores	36.8	2.0
Universe Total	221	14
Sample Chains	6.0	7.0
Sample Stores	18.7	2.0
Sample Total	112	14
Mean*	4,000	4,000
Percent of Universe Required	50.7	100.0

*Normative mean delivery in pounds per store per period.

Source: Computed.

Table C-17. Sample size and percent of universe required for a 95 percent confidence interval within 5 percent of mean deliveries per store per week, and related statistics, four week difference, margarine, two cities.

Statistic	City	
	Dallas	Terre Haute
Variance Component for:		
Chains	72,344	876,089
Stores/Chains	856,154	2,325,196
Universe Chains	6.0	7.0
Universe Stores	36.8	2.0
Universe Total	221	14
Sample Chains	5.0	7.0
Sample Stores	16.0	2.0
Sample Total	80	14
Mean*	4,000	4,000
Percent of Universe Required	36.2	100.0

*Normative mean delivery in pounds per store per period.

Source: Computed.

Table C-18. Sample size and percent of universe required for a 95 percent confidence interval within 7 percent of mean deliveries per store per week, and related statistics, four week difference, margarine, two cities.

Statistic	City	
	Dallas	Terre Haute
Variance Component for:		
Chains	72,344	876,089
Stores/Chains	856,154	2,325,196
Universe Chains	6.0	7.0
Universe Stores	36.8	2.0
Universe Total	221	14
Sample Chains	4.0	7.0
Sample Stores	16.0	2.0
Sample Total	64	14
Mean*	4,000	4,000
Percent of Universe Required	29.0	100.0

*Normative mean delivery in pounds per store per period.

Source: Computed.

APPENDIX D

Table D-1 Number of establishments providing usable delivery data by type of establishment and city [a/]

Establishment Type	City [b/] Dallas	Omaha	Toledo	Terre Haute
Chain Grocery				
Chains	5	2	2	7
Average Stores/Chains	19.2	14.5	17.5	1.6
Total Stores	96	29	35	11
Independents	11	24	10	9
Restaurants	71	20	14	28
Institutions	4	78	2	5

[a/] Delivery records were obtained from more establishments than listed in the table but some records were not usable due to incorrect unit reporting (such as dollar value of deliveries rather than pounds, or monthly rather than weekly basis). Also, some establishments reported less than eight weekly delivery records, in which case the establishment was excluded from the sample.

[b/] Variability in sample size over city by establishment type occurred primarily because of variability in record keeping, ease of obtaining "back records", and extent of cooperation in providing records.

Source: Primary Data

APPENDIX E

METHODOLOGY AND RESULTS FOR MINIMIZING VARIANCE SUBJECT TO FIXED COST

Methodology

Another methodological approach is to minimize variance of mean deliveries subject to specified cost of sampling. The methodology has two basic phases. The first is estimation of sample size based on an analysis of variance model. The second is a constrained optimization of store - week data requirements. Each phase is discussed in turn.

Analysis of Variance

The analysis of variance (ANOVA) model utilized for analysis of cheese, butter, and margarine deliveries per store per week is:

$$(1) \qquad Y_{ijk} = \mu + w_i + c_j + s_{jk} + \varepsilon_{ijk}$$

where

Y_{ijk} = deliveries in pounds (cheese, butter or margarine) during the i^{th} week for store k of chain j.

μ = grand mean

w_i = effect of the i^{th} week

c_j = effect of the j^{th} chain

s_{jk} = effect of the k^{th} store within the j^{th} chain

ε_{ijk} = residual

and

$i = 1, 2, \ldots, w$

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

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

All effects in the model are regarded as random effects with the stores within chains effect nested. The ANOVA table for this model is given in Table E-1. Although the above described model is not the only formulation that could be used, it is the simplest model which could be used in the absence of a priori information.

Determining Sample Size

Assuming a 95 percent confidence interval (C.I.) on mean deliveries per store per week, the appropriate formula is:

(2) $\bar{X} \pm (1.66)\,\hat{\sigma}_{\bar{X}}$

The constant 1.66 is the table value for a t-statistic with 5 percent probability in one tail. Since only increases or decreases in deliveries are of interest in the analysis, a one-tail t is justified. This constant changes to 1.29 for a 90 percent C.I. (10 percent probability in one tail).

Given an estimate of σ and μ, $\hat{\sigma}$ and $\bar{X}$, respectively, sample size for a 95 percent C.I. within m percent of $\bar{X}$ can be derived. This restriction is:

(3) $(m\%)\,\bar{X} \geq (1.66)\,\hat{\sigma}_{\bar{X}}$

Thus

(4) $\dfrac{\hat{\sigma}}{\sqrt{n}} \leq \dfrac{(m\%)\,\bar{X}}{1.66}$

or (5) $\sqrt{n} \leq \dfrac{(1.66)\hat{\sigma}}{(m\%)\bar{X}}$ or (6) $n \geq \dfrac{2.76\hat{\sigma}^2}{((m\%)\bar{X})^2}$

Since v(store mean) = v(week mean) = $\hat{\sigma}^2$, sample size is determined utilizing either v(store mean) or v(week mean) in the above formula and utilizing $\bar{X}$ as an estimate of μ. The resultant n is interpreted as being the number of stores for one week or the number of weeks for one store needed for a 95 percent C.I. within m percent of $\bar{X}$.

Table E-1. ANOVA, random effects, equal subclass numbers*

Source of variation	df	E(MS)
Total	csw	
Mean	1	
Chains	c-1	$\hat{\sigma}^2_{\varepsilon} + w\hat{\sigma}^2_{s(c)} + sw\hat{\sigma}^2_{c}$
Stores/Chain	c(s-1)	$\hat{\sigma}^2_{\varepsilon} + w\hat{\sigma}^2_{s(c)}$
Week	w-1	$\hat{\sigma}^2_{\varepsilon} + cs\hat{\sigma}^2_{W}$
Residual	(w-1) (cs-1)	$\hat{\sigma}^2_{\varepsilon}$

*Mean square expectations would be somewhat different in the case of unequal subclass numbers.

As an example, suppose the resultant n equaled 30 from formula (6). The conservative interpretation of this would be that 900 or 30^2 store weeks of data would be needed in order to attain the stipulated accuracy. This is true because v(store mean) = v(week mean) = $\hat{\sigma}^2$ which implies that if w = 30 then s = 30. Since the total sample is ws, total sample size is 30^2. This would be a conservative estimate of the data required for the given accuracy since n^2 overestimates store-week requirements. Although the exact magnitude of overestimation is unknown, a "best guess" would be somewhere between 20 and 30 percent.

Constrained Optimization of Store-Week Data Requirements

There exist many combinations of chains, stores within chains, and weeks that would generate a given level of accuracy. The combination which minimizes the overall variance of mean deliveries per store per unit time (minimizes $V(\hat{\mu})$) subject to the cost incurred to collect the data should be chosen. If, for example, a total of $B can be allocated to data collection, then the degree of accuracy that can be attained by exhaustion of the fixed budget is of key operational importance.

The principle of optimum allocation depends on minimizing the variance of the mean subject to some budget allocation for data collection as follows:

(7) minimize: $L = V(\hat{\mu})$

(8) subject to $cK_c + csK_s + cswK_w \leq B$

where

$V(\hat{\mu})$ = the variance of the overall mean

B = budget allocation for data collection

K_c = cost of adding a chain to the data

K_s = cost of adding a store to a chain

K_w = cost of adding a week to a store

c = number of chains in sample

s = number of stores per chain

w = number of weeks per store

Thus, formulating the problem with a Lagrangean multiplier yields:

(9) $$L = V(\hat{\mu}) + \lambda(B - cK_c - csK_s - cswK_w)$$

The constrained minimization reformulated in equation (9) requires the simultaneous solution to four non-linear equations as follows:

(10) $$\frac{\partial L}{\partial c} = -\left\{\frac{1}{c^2}\left[\frac{\hat{\sigma}_\varepsilon^2}{sw} + \hat{\sigma}_c^2 + \frac{\hat{\sigma}_s^2}{s}\right] + \lambda(K_c + sK_s + swK_w)\right\} = 0$$

(11) $$\frac{\partial L}{\partial s} = -\left\{\frac{1}{s^2}\left[\frac{\hat{\sigma}_\varepsilon^2}{cw} + \frac{\hat{\sigma}_s^2}{c}\right] + \lambda(cKs + cwK_w)\right\} = 0$$

(12) $$\frac{\partial L}{\partial w} = -\left\{\frac{1}{w^2}\left[\frac{\hat{\sigma}_\varepsilon^2}{cs} + \hat{\sigma}_w^2\right] + \lambda csK_w\right\} = 0$$

(13) $$\frac{\partial L}{\partial \lambda} = B - cK_c - csK_s - cswK_w = 0$$

These four equations collectively contain the three variance components estimated from the ANOVA model and the cost of sampling chains, stores, and weeks.

The optimization procedure requires an estimate of the costs involved in adding chains to a sample (K_c), adding stores to a chain (K_s) and adding weeks to a store (K_w). As an estimate of the respective costs, K_c was assumed to be \$500, K_s was assumed to be \$25, and K_w was assumed to be \$5. Of course, these costs could be changed to suit different situations and a new solution would be obtained if relative costs of sampling changed. A special computer program was written for the solution of these equations.

Results

Sample Size--Weekly Deliveries

Dallas and Toledo data were analyzed using the MINVAR analytic procedures. Using weekly deliveries, results for specific combinations of accuracy and data transformation are reported for each city. For each given level of accuracy in terms of C.I. on mean deliveries, three data transformations were utilized; raw or untransformed data, a square root transformation, and a logarithmic ($\log_e$) transformation. The data transformations are used to reduce variability.

The results of these analyses are summarized in Tables E-2 and E-3. The indicated sample size required is the value of N from equation (6) using the degree of accuracy specified at the top of each column. As previously indicated, the values in these tables must be squared in order to obtain store-week requirements. For example an N of 90 results from the log transformation and selection of a 95 percent C.I. within 3 percent of mean cheese deliveries in Dallas, Table E-1. This suggests that as many as 8,100 store-weeks of data are required to obtain this accuracy for cheese in Dallas.

The universe of food chain stores in Dallas is 6 chains with an average of 36.8 stores per chain. Thus 221 store units are available per week. A total of 36 weeks of data from all stores would be required to estimate average weekly deliveries within an accuracy deviation of 3 percent. If one is willing to accept 5 percent accuracy, store-week requirements are 1089 (N=33) and, therefore, 5 weeks of data would suffice, using all 221 stores. Reducing the confidence level to 90 percent and using the 5 percent accuracy level calls for only 400 store weeks (N=20). Thus, two weeks of data would suffice, utilizing all 221 stores.

For Toledo, cheese deliveries, converted to a log basis, can be measured within 3 percent accuracy at the 95 percent confidence level by using only 100 store-weeks (N=10), Table 3. In general, sample requirement in Toledo is less demanding than Dallas requirements.

Clearly, many of the required N in these tables exceed practical limits of data collection. In fact, for data collection in only one metropolitan area, a practical upper limit on N for weekly data may be regarded as 25 or 26. This N was derived regarding 5 as the upper limit on chains per city, 10 stores per chain as the upper limit of stores, and 13 weeks as the upper limit on weeks per store (13 weeks is one quarter of a year's data).

Sample Size--Biweekly Deliveries

Sample size was also estimated using biweekly data (combining adjacent weeks) in order to investigate the amount of reduction in total variation which might be obtained. The results of using biweekly data are reported in Tables E-4 and E-5. Interpretation of the numbers in these tables is similar to those in Tables E-2 and E-3.

Table E-2. N required for specified confidence interval on mean deliveries per store per week, Dallas

Product and Transformation of Data	Measurement error 95% C.I. on $\overline{X}$ within			Measurement error 90% C.I. on $\overline{X}$ within		
	3%	5%	7%	3%	5%	7%
	N value for store weeks of data					
Cheese						
Raw Data	1021	368	188	617	223	114
Square Root	231	83	43	139	50	26
Log	90	33	17	54	20	10
Butter						
Raw Data	5243	1888	963	3166	1140	582
Square Root	878	317	162	532	192	98
Log	826	298	152	499	180	92
Margarine						
Raw Data	619	223	114	374	135	69
Square Root	179	65	33	108	39	20
Log	71	26	13	43	15	8

Source: Computed

Table E-3. N required for specified confidence interval on mean deliveries per store per week, Toledo

Product and Transformation of Data	95% C.I. on $\overline{X}$ within			90% C.I. on $\overline{X}$ within		
	3%	5%	7%	3%	5%	7%
Cheese						
Raw Data	303	109	56	183	66	34
Square Root	77	28	15	47	17	9
Log	10	4	2	6	3	2
Butter						
Raw Data	1187	428	218	717	258	132
Square Root	273	98	50	164	59	30
Log	107	39	20	65	24	12
Margarine						
Raw Data	806	290	148	487	175	90
Square Root	189	68	35	114	41	21
Log	22	8	4	13	5	3

Source: Computed

Table E-4. N required for specified confidence interval on mean deliveries per store, biweekly, Dallas

Product and Transformation of Data	Measurement error 95% C.I. on $\bar{X}$ within			Measurement error 90% C.I. on $\bar{X}$ within		
	3%	5%	7%	3%	5%	7%
	N value for store biweekly periods of data					
Cheese						
Raw Data	449	162	83	271	98	50
Square Root	91	33	17	55	20	10
Log	10	4	2	6	3	2
Butter						
Raw Data	2791	1005	513	1684	607	310
Square Root	456	165	84	276	100	51
Log	265	96	49	160	58	30
Margarine						
Raw Data	234	85	43	142	51	26
Square Root	59	22	11	36	13	7
Log	6	2	1	4	2	1

Source: Computed

Table E-5. N required for specified confidence interval on mean deliveries per store, biweekly, Toledo

Product and Transformation of Data	95% C.I. on $\bar{X}$ within			90% C.I. on $\bar{X}$ within		
	3%	5%	7%	3%	5%	7%
Cheese						
Raw Data	130	47	24	79	29	15
Square Root	151	55	28	91	33	17
Log	4	2	1	2	1	1
Butter						
Raw Data	329	119	61	199	72	37
Square Root	64	23	12	39	14	7
Log	24	9	5	15	6	3
Margarine						
Raw Data	526	190	97	317	114	59
Square Root	113	41	21	68	25	13
Log	10	4	2	6	2	1

Source: Computed

The practical upper limit on N in the tables when biweekly data are considered is about 19. This was derived regarding 5 as the upper limit on chains per city, 10 stores per chain as the upper limit on stores, and 7 biweekly periods as the upper limit on weeks (14 weeks necessary to collect 7 biweekly periods).

Constrained Optimization

Constrained optimization involves optimizing sample size given a fixed budget. Results from the constrained optimization technique, equations (10) through (13), are illustrated by some examples which follow. Using weekly data from Dallas and an arbitrary $9,000 budget allocated for data collection, results indicate that 11 chains, 7 stores per chain, and 4 weeks is the optimum combination to obtain the greatest accuracy of mean cheese deliveries per store per week. This would yield a 95 percent C.I. on mean cheese deliveries of approximately $\pm$ 4.4 percent.

For butter in Dallas using weekly data and the same $9,000 budget allocation, the optimum combination would be 12 chains, 4 stores per chain, and 9 weeks. This combination would yield a 95 percent C.I. on mean butter deliveries of approximately $\pm$ 17.3 percent. For margarine in Dallas using weekly data and a $9,000 budget, the optimum combination would be 12 chains, 5 stores, and 4 weeks. This would yield a 95 percent C.I. on $\overline{X}$ of approximately $\pm$ 4.0 percent.

Note that budget allocations for data collection could be changed which would result in differing accuracy and combinations of stores, weeks, and chains. The relationship between cost and accuracy can be investigated with the constrained optimization technique.

The disadvantage of this approach is that accuracy cannot be initially stipulated but rather is a result of the budget allocated; and, therefore, the relative cost of securing various segments of data such as chains versus stores. Since it is preferable to constrain sampling error for ADA purposes, methodology which allows accuracy to be stipulated and budget to be minimized is preferable to this approach. Another difficulty with this approach as applied here is that the ANOVA model is based upon the assumption of an infinite universe. This, of course, is not the situation for a metropolitan market.

Summary and Conclusions

For this approach, delivery data were transformed alternatively to square root and logarithms in order to reduce variance arising from differences in delivery volume arising from variations in store size.

Summary of the sample size of store-weeks of data required for specified levels of accuracy for the logarithm of deliveries are presented in the top portion of Table E-6. Requirements for biweekly data are noted in the bottom section of Table E-6. Sample requirements in this table assume an infinite universe. Therefore, no population correction factor is applied.

Comparison of the number of store-biweeks of data required for 3 percent accuracy when compared with the number of stores in Dallas and

Toledo, provides one view of the possible ability to use those metropolitan areas as market test locations, given a limited budget for data collection. Clearly, Dallas is a more feasible test market than Toledo for 3 percent measurement accuracy, given the pre-stated analysis assumptions, Table E-6. Butter deliveries are too variable in either city to permit measurement at the 3 percent accuracy of mean deliveries per store.

Selection of a 5 percent permissible measurement error alters the acceptability of the two cities. Measurement is achievable on 3 products in Toledo, whereas butter measurement is still not feasible in Dallas, so long as no sample size to universe size correction factor is utilized. Emphasized from this result is the difficulty of generalizing about the suitability of any specific city as a test market judging from its mere physical size. Prior knowledge of the particular product's retail store delivery behavior in the individual market is absolutely essential.

This approach permits a decision as to the attainable measurement accuracy, within a given budget limitation, for any market once the variance in mean deliveries of chains, and stores within chains, is known. For example, a $9,000 budget for the collection of weekly delivery data for margarine in Dallas would yield ± 4 percent accuracy in determining the mean level of deliveries per store, with a 95 percent confidence level. In contrast, a ± 17.3 percent accuracy results for butter deliveries in Dallas. Such a large error is usually not acceptable. Therefore, Dallas would be dismissed as a possible butter test market.

This analysis can be useful in assisting in selection of test markets, especially when research budgets impose a limitation on the research design.

Table E-6. Sample requirements for indicated accuracy of measuring the logarithm of changes in product deliveries to chain stores, selected cities and products.

City, Product, and Time Period	Measurement Error With 95% Confidence Level Of a/			Measurement Error With 90% Confidence Level Of a/		
	3%	5%	7%	3%	5%	7%
	---store weeks of data---					
WEEKLY MEASUREMENT						
Dallas						
Cheese	8,100	1,089	289	2,916	400	100
Butter	682,276	88,804	23,104	249,001	32,400	8,464
Margarine	5,041	676	169	1,849	225	64
Toledo						
Cheese	100	16	4	36	9	4
Butter	11,449	1,521	400	4,225	576	144
Margarine	484	64	16	169	25	9
	---store biweeks of data---					
BIWEEKLY MEASUREMENT						
Dallas						
Cheese	100	16	4	36	9	4
Butter	70,225	8,216	2,401	25,600	3,364	900
Margarine	36	4	1	16	4	1
Toledo						
Cheese	16	4	1	4	1	1
Butter	576	81	25	225	36	9
Margarine	100	16	4	36	4	1

a/ The required sample size is chains times stores per chain times weeks or biweeks per store. For example, N of 8,100 could be obtained by sampling 6 chains with an average of 36.8 stores per chain for a 36 week period.

Source: Computed from survey data.

APPENDIX F

ANALYSIS OF THE INDIVIDUAL CHAIN STORE INVENTORIES

The adequacy of deliveries to retail stores as a measure of product movement during some time period depends partly on the behavior of retail store inventories. In order to evaluate the influence of inventory changes, two inventories were taken at randomly selected times in Terre Haute, Indiana (Dec. 15 and Jan. 10) and Toledo, Ohio, (Jan. 5 and Jan. 15).

Inventories were taken randomly for two reasons. First, store delivery data was collected weekly for eight weeks from each participating food chain store. No a priori basis was available to predetermine the optimum times that store level inventories should be taken. Therefore, a random selection of inventory dates was as good as any other alternative. The second reason for randomly selected inventory dates was that the research project budget was insufficient to permit taking weekly store-level inventories for eight weeks.

Inventories have two dimensions, absolute level and the relative change in inventory level from one time period to another. The latter is important to the measurement of product movement during a specific time period. If large inventory changes occur relative to store deliveries, then deliveries alone inadequately monitor retail sales.

The relationship in Terre Haute and Toledo between product deliveries per store for four weeks and changes in the two sample inventory levels are noted in Table F-1. Within Terre Haute inventory changes for cheese and butter were of about the same relative magnitude to deliveries. The same was true in Toledo. However, the difference between the two cities was

Table F-1. Relationship between chain store inventory changes and product deliveries, Terre Haute, Indiana and Toledo, Ohio

City and Product	Average Quantity Delivered per Store per 4-Week Period	Average Inventory Change per Store	Inventory Change As Percent of 4-Week Deliveries
	pounds	pounds	pounds
Terre Haute			
Cheese	4,472	-415	-9.2
Butter	735	+ 60	+8.1
Margarine	6,444	- 8	-0.1
Toledo			
Cheese	2,786	+ 71	+2.5
Butter	990	- 25	-2.5

Source: Survey data

appreciable. In Terre Haute average inventory changes per food chain store for cheese and butter were equivalent to 9.2 and 8.1 percent, respectively, of four-week deliveries. The corresponding percentages in Toledo were 2.5 and 2.5, respectively. These figures suggest that even if delivery data for four weeks are measured within 3 percent accuracy, at the 95 percent confidence level, inventory changes at times could result in deliveries misrepresenting sales by perhaps as much as 10 percent, given a 4 week sampling unit. Especially is this a possibility in a city such as Terre Haute. At the time of the survey it had only 14 chain stores among 7 chains, and these stores were served by distribution centers located in other cities. Toledo by comparison, had 93 chain stores representing 5 chains. The large number of stores in Toledo possibly contributes to more balanced inventory changes from one time period to another among the stores.

Also, for Terre Haute, the time during the year inventories were taken must be considered. December 15 and January 10 inventories probably represent a very severe comparison. The first date reflects possible inventory build-ups for per-Christmas holiday sales. January 10, by comparison, is likely to reflect a more normal stock condition. This implies that any research design used to measure product movement should have the capacity to handle such divergencies in the market data.

Estimation of inventory changes by a store sampling procedure as opposed to a total census was not explored in this study. However, application of sampling may well be considered for very large markets. Sampling appears to be inadvisable in markets the size of Terre Haute or even for markets in the 500 thousand population category. This follows since the standard error associated with the mean inventory level per store was so large that no significant difference in the two inventory levels measured could be detected. Yet the inventory changes amounted to from 2.5 percent in Toledo to 8 percent in Terre Haute of total four-week period deliveries. Where measurement needs to be accurate within 3 percent or even 5 percent, a latin square or other type research design must be employed which allows causal relationships to be statistically tested.