DESIGN ISSUES FOR THE RETAIL TRADE SAMPLE SURVEYS OF THE U.S. BUREAU OF THE CENSUS[1/]

Ruth E. Detlefsen and Carol S. Veum, U.S. Bureau of the Census
Ruth E. Detlefsen, Room 2653-3, Washington, D.C. 20233

KEY WORDS: Stratum bounds, allocation, sample size, multiple constraints

I. Introduction

The Business Division of the U.S. Bureau of the Census provides several measures of retail activity in the U.S. Retail establishments are classified in major groups 52 through 59 of the 1987 Standard Industrial Classification Manual. They are primarily engaged in selling merchandise for personal or household use and servicing goods they sell. Exceptions to this definition accommodate trade practices.

Measures are produced from censuses and sample surveys. Censuses were taken irregularly during the years 1929 through 1963 and beginning with 1967, have been taken every five years to cover years ending in "2" or "7". Sample surveys providing monthly and annual statistics began in 1952.

The Census of Retail Trade provides statistics on annual sales, annual and first quarter payroll, employment and number of establishments. Statistics are given for the U.S. and many other geographic areas for detailed kinds of business (KBs) as defined in the 1987 Industry and Product Classification Manual. These statistics are also summarized by sales and employment size categories. The census gives statistics on sales by merchandise line and presents various tabulations for KB specific data items.

Sample surveys conducted for the retail trade area include the Monthly Retail Trade Survey (MRTS), the Monthly Advance Retail Trade Survey (MARTS), the Monthly Retail Inventory Survey (MRIS) and the Annual Retail Trade Survey (ARTS). The MRTS, MARTS and MRIS provide monthly estimates of levels and of changes in monthly level estimates from a month ago and from a year ago. Monthly sales are estimated from the MRTS and the MARTS, while end-of-month inventories are estimated from the MRIS. Estimates for the MRTS, MARTS and MRIS are produced at the U.S. level for various KBs. In addition, the MRTS estimates are produced at more detailed KBs and for some geographic areas. The ARTS provides estimates of annual sales and purchases, end-of-year inventory and accounts receivable and gross margin at the U.S. level for various KBs. The ARTS estimates include levels and changes in levels from a year ago.

Every five years the Bureau reselects samples of employers to include in the MRTS, the ARTS and the MRIS. Samples are reselected to take advantage of updated KB and sales and other information available from the latest census, and to reduce respondent burden. This paper will describe these surveys and steps taken to reselect samples to be introduced in early 1992. Included is a discussion of a study of a stratification-allocation scheme proposed by Lavallée and Hideroglou. In addition, the sample for the MARTS has historically been reselected about every two years. This paper will describe the current MARTS design and consideration given to changing the design to parallel the design used for the MRIS.

II. Overview of the Sample Surveys

II.A. Monthly Retail Trade Survey

MRTS estimates are available about 40 days after the end of a month. They are derived from a list sample representing most retail employer businesses and an area sample representing all retail nonemployers and employer births not represented by the list sample. Since the main concern in this paper is to address list sample issues, little more will be said about the area sample.

The MRTS sample to be introduced in 1992 was selected using a stratified random sample design with stratification on KB (and sometimes geography) and annual sales. Sample sizes were determined to meet set variability constraints for each KB (and geographic area, if applicable) and for various KB aggregates.

The list sample for the MRTS survey to be introduced in 1992 was initially selected from sampling units constructed from establishments extracted from the 1989 Standard Statistical Establishment List (SSEL). The list sample will be updated with a sample of births prior to being introduced, and quarterly after that. For the MRTS, companies contributing the most sales to each KB were taken into the survey with certainty and will be required to report sales monthly. All other employer companies were reformed into Employer Identification (EI) number sampling units and subjected to sampling. An EI is a number issued by the Internal Revenue Service (IRS) to employer businesses for reporting payroll. Selected noncertainty sampling units were assigned to one of three rotating panels. EIs in each panel will be required to report current (CM) and previous month (PM) sales once each quarter.

Each month Horvitz-Thompson estimates of CM

and PM sales (X'_t and X''_{t-1}) will be generated as input to calculate CM and PM composite sales estimates. For any three digit KB, the CM preliminary and PM final composite estimators will, respectively, have the forms:

$$X'''_t = (1-b)\, X'_t + b \frac{X'_t}{X''_{t-1}} X'''_{t-1}$$

$$X''''_{t-1} = (1-a)\, X''_{t-1} + a X'''_{t-1}$$

where the Horvitz-Thompson estimators are for months t and t-1 from the panel for t, and a and b are constants between 0 and 1.

The month-to-month ratio estimator will be the ratio between the CM preliminary and PM final. Through a study of the covariance structures of the Horvitz-Thompson estimators, b was set to .75 and a to .8. These values reduce the variances of the level and ratio estimators.

Monthly estimates are benchmarked to the latest ARTS estimates. Both seasonally adjusted and non-seasonally adjusted estimates are provided.

A fuller description of the MRTS is documented in Isaki, et. al., 1976; Konschnik, et. al., 1985, and Wolter, et. al., 1976.

II.B. Monthly Retail Inventory Survey

MRIS estimates are published along with the MRTS estimates and are derived in a manner similar to that for the MRTS. The MRIS sample is a subsample of the MRTS sample. A paper by Garrett, et. al., 1986 gives more details.

II.C. Annual Retail Trade Survey

The ARTS estimates are derived by obtaining data from all MRTS certainty companies, all EIs in two of the three MRTS panels and about one fourth of the area sample establishments in the MRTS. Level estimates are made using a Horvitz-Thompson estimator and are benchmarked to the latest census.

II.D. Monthly Advance Retail Trade Survey

MARTS estimates are available 10 to 15 working days after the end of a month and thus provide the earliest indication of retail sales. Estimates are derived from a sample selected from all MRTS certainties, about one fourth of the MRTS list sample noncertainty sampling units and one fourth of the MRTS area sample nonemployers. Currently, the MARTS subsample is reselected about every two years and is not updated to reflect births.

The MARTS is selected using a stratified random sample design, where stratification is by KB, size of MRTS sampling unit and annual sales. Sample sizes are determined using Neyman allocation restricted by total sample size. Since only one panel of sampling units is selected, each sampling unit reports one month of data each month.

Level estimates are calculated by forming the ratio of weighted CM sales to weighted PM sales for each detailed KB. The ratio is formed using only sampling units reporting for both months. This ratio is multiplied by the preliminary composite estimate from the MRTS that corresponds to the PM for the MARTS. Level estimates for the detailed KBs are then summed to give aggregate level estimates. Ratios of change in CM sales from the PM are obtained by dividing the MARTS CM estimate by the preliminary composite estimate from the MRTS that corresponds to the PM for the MARTS.

Both seasonally adjusted and non-seasonally adjusted estimates are provided.

III. Sample Design for the Monthly Retail Trade Survey

Before selecting the new samples, studies of the universe were made to set stratum bounds, sample sizes and other parameters required for selection and estimation.

III.A. Lavallée-Hideroglou Method

As a first step in the parameter studies, a stratification-allocation scheme proposed by Lavallée and Hideroglou (L-H) was investigated. Methods used in the past required considerable amounts of human intervention and did not guarantee optimal stratification. If adaptable to the MRTS, the L-H procedure would have cured both ills.

The method is an iterative procedure to determine certainty and noncertainty stratum bounds for highly skewed populations. The bounds are set to minimize sample size for a fixed level of reliability, simple random sampling without replacement from the noncertainty strata and use of a power allocation. The algorithm was easy to program and, for the data sets studied by L-H, converged rapidly to optimum bounds to give substantial reductions in sample size for desired levels of reliability.

Because populations within retail KBs are highly skewed with regard to sales and because sample selection goals match those of L-H, a study of the L-H method applied to the Census Bureau's retail sales survey was appropriate. Modifications to the L-H method were made to accommodate Neyman allocation.

To apply the procedure, it is assumed that the frequency distribution of the population may be represented by a continuous density.

To find optimum bounds, derivatives of an equation expressing total sample size in terms of desired sampling variability and the allocation method are taken with respect to each of the bounds and set to 0. The

resulting equations are expressed as quadratic equations.

Since the populations under consideration are finite, continuous density expressions for the weights, means and variances are replaced by corresponding expressions for finite populations. The equations can then be solved using the L-H iterative procedure:

1. Sort the population in ascending order.
2. Start with arbitrary bounds.
3. Calculate stratum weights, means and variances for the starting bounds.
4. Using the coefficients corresponding to the quadratic equations resulting from these bounds, calculate another set of bounds by solving the quadratic equations again.
5. Repeat 3 and 4 until consecutive solutions are identical or differ by a small amount.

Before implementing the L-H method, the procedure was applied to various KBs to determine if it was practical for the MRTS. This investigation showed convergence to be slow (often 50 to 100 or more iterations) or nonexistent. When convergence occurred it was faster with fewer strata and fewer cases being stratified. Convergence was fastest when starting bounds were set equal to values set for the last sample selection and when sampling units were grouped prior to the application of the procedure. Different starting values gave different ending bounds. Many times the bounds differed substantially.

Because of these findings, the L-H procedure was abandoned. Had the procedure proved beneficial, other modifications would have been required. In particular, the procedure would need to accommodate two kinds of sampling units and allow stratification by sales and allocation by payroll.

III.B. Parameter Studies

After the L-H procedure investigation, studies were conducted to determine regression coefficients, create sampling units, establish stratum bounds and determine sample sizes and allocations for the MRTS.

III.B.1 Universe Creation and Analysis

A computer file containing data records for all employer establishments tabulated in the 1987 Census of Retail Trade was constructed.

To verify that the establishment universe contained no systematic data errors and to present a general picture of the universe, a preliminary analysis of the universe was performed. The analysis included general descriptive statistics, frequency distributions and edit failure listings.

III.B.2. Regressions and Scattergrams

After reviewing frequency distributions and edit failures, regressions were run to model relationships between establishment sales and annual payroll. Regression coefficients were used to compute measures of size for allocation and will be used to set measures for new establishments and for imputation purposes throughout the life of the sample.

The following two models were considered:

$$Y = \beta_1 X + \varepsilon \quad \text{and} \quad \sqrt{Y} = \beta_2 \sqrt{X} + \varepsilon$$

where Y = sales and X = annual payroll. Only no-intercept models were considered because resulting coefficients are simpler to apply and the sales-to-payroll relationships are adequately modeled.

Using data for all retail establishments, coefficients for each model were estimated for the 5-, 4-, 3- and 2-digit KBs. For each KB, one of the two models was selected to represent the sales-to-payroll relationship. The following two R-squared statistics were calculated:

$$R_1^2 = \frac{\sum_i \hat{y}_i^2}{\sum_i y_i^2} \quad \text{and} \quad R_2^2 = 1 - \left[\frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \right]$$

For linear models without intercept, R_1^2 is the ratio of the regression sum of squares to the total sum of squares. R_2^2 represents the variation of y about its mean that is accounted for by the fitted model. For the square root model, both statistics were calculated using data in the original units.

The effect of outliers on the coefficients was considered. At the 5-digit KB level, regression parameters for both models were recomputed excluding establishments with absolute residuals over three times the standard error of the estimate. Outliers at the 5-digit KB level were eliminated from subsequent 4-, 3- and 2-digit computations.

Scattergrams of sales vs. annual payroll and $\sqrt{\text{sales}}$ vs. $\sqrt{\text{annual payroll}}$ were also generated for each 5-digit KB to assist in detecting potential problems due to outliers, heteroscedasticity and nonlinearity in the sales to payroll relationships.

The model selected for a KB was the one that best fit the data in terms of R_2^2 (recommended by Kvålseth, 1985). Coefficients calculated with outliers excluded were chosen over the original coefficients if excluding outliers provided a better fit and changed the coefficient by more than ten percent. The coefficient, ß, to be used in subsequent parameter study, sample selection and

estimation operations for any KB was set to β_2^2 if the square root model was selected and to β_1 if the other was selected.

III.B.3. Creation of Company Summaries

To determine certainty boundaries, company summaries were formed from establishment records. Prior to forming the summaries, establishments were assigned sampling KBs, annualized sales, annualized annual payroll, pseudo sales and metropolitan codes for general merchandise, apparel or furniture (GAF) establishments. Annualized sales and annualized annual payroll were set by multiplying the census sales and annual payroll by 12 divided by the months-in-business. Pseudo-sales were set equal to ß times the annualized annual payroll. The sampling KB was set based on the census KB and the metropolitan code was set based on the census state and county code for each GAF establishment.

Establishment data was summarized within each company to provide total annualized sales, annualized annual payroll and pseudo sales across all establishments by sampling KB, by geographic code within each GAF KB, by geographic code across all GAF KBs and across all GAF KB and geographic codes for GAF totals.

The sampling KB for the company was the one contributing most to total annualized sales. If the sampling KB was GAF, the sampling geographic code for the company was the geographic area with the largest total annualized sales.

III.B.4. Certainty Determination

Initial certainty boundaries were set by visually inspecting sales size frequency distributions for each KB to determine where the companies began to thin out in number to form the "tail" of the distribution. A single cutoff was determined for all geographic areas within each GAF sampling KB.

Certainty bounds were adjusted based on number of certainties and percent of total sales represented by the certainties in the current sample.

A company became certainty if its annualized sales exceeded the cutoff for the sampling KB of the company or its total annualized sales for any associated KB exceeded the cutoff for that KB.

In addition, survey analysts were allowed to add any company as certainty if the company had known unique characteristics that could potentially lead to problems with the estimates if not included as certainty.

Companies determined to be certainties during the parameter studies were taken as certainties in the sample selection operations. Thus, the sampling unit for certainty company summaries was the company.

III.B.5. Creation of Sampling Units

Establishments of all noncertainty companies were resummarized into EI summaries. EI summaries were formed in the same way as the company summaries. The EI summaries were used to determine noncertainty strata, sample sizes and allocations for each sampling KB.

III.B.6. CV Constraints

Prior to the parameter studies, survey analysts specified target coefficients of variation (CV) for monthly level and ratio estimates. Before these CVs were used as constraints to determine sample sizes, adjustments were made to ensure that both level and ratio CVs could be met simultaneously and to account for variability from the area sample. Adjustments were based on application of theoretical relationships to data observed over several months and on actual relationships observed during that time.

These design CVs were used as input to determine sample sizes and allocations.

III.B.7. Initial Sample Allocation

The cumulative $\sqrt{f}$ rule was applied to the EI sampling units for each sampling KB/geographic area to determine noncertainty stratum bounds. Neyman allocation was used to generate sample sizes for a number of target CVs and strata. Stratification was by annualized sales, while allocation was by pseudo sales. Target CVs included the adjusted CVs along with slight variations of the CVs. The number of strata were 3, 6, 9 and 12. One allocation was chosen for each sampling KB/geographic area to achieve minimum sample size while preventing stratum sizes from falling below three (two if GAF) and stratum weights from exceeding 1000. Where possible, weights in smaller sales strata were forced to be higher than in larger ones.

III.B.8. Final Sample Allocation

Information about the selected allocations was fed into a multiple CV constraint routine (developed by Dr. Bev Causey) to determine optimum sample sizes and allocations to simultaneously meet all specified CV constraints for sampling KBs and geographic areas and for aggregate KBs. The information used included sampling KB/geographic area, CV, stratum upper bounds, universe counts, sample weights and within-stratum sums and variances for the allocation variable.

Following the multiple CV constraint analysis, sample counts were adjusted to prevent stratum sample sizes from exceeding the universe counts, keep stratum sizes above three (two if GAF), and where possible, force weights of smaller strata to exceed those of larger strata. Sample sizes were increased for KBs in which potential problems in meeting target CVs were anticipated.

III.C. Sample Selection

Stratum bounds, sample sizes and sampling rates determined during the parameter studies were used to select samples for the MRTS as described earlier in the paper.

IV. Design for the Monthly Retail Inventory and Advance Retail Trade Surveys

In addition to the MRTS, Business Division conducts the MRIS and the MARTS on a monthly basis. Beginning with the last sample reselection, the MRIS design has paralleled that of the MRTS. Because of the success of the MRIS design change and because other benefits might be derived from a similar design for the MARTS, consideration was given to designing the MARTS similar to the MRIS. In considering a design change for the MARTS, the following benefits were expected:

- The design would parallel the MRTS and MRIS designs so that one system for selecting and updating samples and producing estimates could be maintained for all monthly retail surveys.
- Differences between the MARTS and MRTS sales estimates would be reduced.
- Time spent to select new MARTS samples would be halved since the sample would be selected every five years instead of every two years.
- The MARTS sample could be updated regularly to include births.
- The MARTS estimates would always be from a subset of sampling units in the current MRTS. Presently, the MARTS sampling units are a subset of the prior MRTS sample for the first several months after introducing a new MRTS sample.

Initially, the plan was to design the MRIS and the MARTS samples to be nonoverlapping subsamples of the MRTS. The sample sizes required to simultaneously meet CV constraints for these two surveys were such that MRTS sample sizes would have increased greatly to accommodate this. As a result, parameter studies were conducted to determine stratum boundaries, sample sizes and allocations for one sample that would accommodate both MRIS inventory and MARTS sales estimates.

IV.A. Parameter Studies

For these two surveys, all phases of the parameter studies up to the initial allocation were done as a part of the MRTS studies or were done in a manner similar to MRTS.

Because inventory data was not collected in the census, the studies relied on census sales data and the strong correlation between sales and inventories for most retail KBs. Sales and inventory data from the 1987 ARTS were used in determining the sample allocation for MRIS and MARTS.

IV.A.1. Initial Allocation

Prior to determining the initial sample allocations for MRIS and MARTS, the list sample tabulating units in the 1987 ARTS were assigned to MRIS/MARTS sampling KBs based on the ARTS sampling KB and assigned to MRIS/MARTS strata based on their ARTS sales. Tabulating units that would have been MRTS certainties, but not MRIS/MARTS certainties based on their ARTS sales size were assigned to a special stratum within each sampling KB.

Using 1987 ARTS data, initial sample allocations were derived separately for each of the MRIS and MARTS. Allocations were based on 1987 end-of-year inventory for MRIS and on 1987 annual sales for MARTS. For both, CV constraints were applied individually to each MRIS/MARTS KB and several aggregate KBs. Initial conservative sample sizes for each of MRIS and MARTS were determined by selecting for each stratum the largest sample size required to satisfy the individual CV constraints. This analysis was done several times using various numbers of noncertainty strata for each KB. To assist in determining the final set of strata, an estimate was made of the number of MRTS sampling units that would fall into each MRIS/MARTS stratum. For each MRIS/MARTS KB, the number of noncertainty strata that minimized the total sample size and that minimized the number of MRIS/MARTS strata that did not contain sufficient MRTS sampling units to draw the required number of MRIS/MARTS sampling units was chosen.

Information similar to that for the MRTS was fed into the multiple CV constraint routine to determine optimum sample sizes and allocations to meet CV constraints for all sampling KBs and aggregates. This was done independently for MRIS and MARTS.

IV.A.2. Final Allocation

Finally, a single allocation was selected to meet all variability and budget constraints for both surveys. In deriving this allocation, all previous allocations, the previous MRIS sample size, the previous MARTS

sample size and CVs achieved by the current MRIS and MARTS samples were considered.

IV.B. Sample Selection for the Monthly Advance Retail Trade Survey

When the parameter studies were underway, no decision to change the MARTS design had been made. Later a decision was made to continue the MARTS under the current design with slight modifications to be implemented for the next MARTS reselection. Reasons for not moving to the new design now include:

- There is no evidence that a design change without improvements in other areas would result in better estimates.
- There is evidence that nonsampling problems (like early reporting) must be resolved before real improvements can be made in the estimates.
- Under the new design, sampling units in the MARTS and MRIS would be identical so that respondent burden may be a negative factor. Further study is required.
- The proposed design would mean major operational changes with few resources available to make those changes.

Since the major source of problems is from nonsampling factors, first priority will be to identify and resolve problems from those sources rather than to make major design changes.

In the short term, modifications to the existing design will be explored. Changes to the existing design will be implemented if they provide better estimates and do not require major resources. An investigation into stratifying area sample establishments separately from other sampling units will be undertaken. Consideration will also be given to imposing CV constraints.

As resources allow, more extensive studies will be conducted to determine major design changes that will improve the MARTS estimates. Depending on results of these studies, impact of improvements from nonsampling factor research and on available processing and operational resources, consideration will be given to moving to a design similar to the one used for the MRIS.

IV.C. Sample Selection for the Monthly Retail Inventory Survey

For MRIS, the parameter study results were used to select a subsample from MRTS.

References

Causey, B. (1972), "Optimal Allocation in Stratified Sampling with Multiple Variance Constraints." Technical Notes No. 5, U.S. Bureau of the Census, Washington, D.C., 8-13.

Garrett, J., Detlefsen, R., and Veum, C. (1987). "Recent Sample Revisions and Related Enhancements for Business Surveys of the U.S. Bureau of the Census." Proceedings of the Business and Economic Section, American Statistical Association, 141-149.

Isaki, C. T., Wolter, K., Sturdevant, T., Monsour, N., Trager, M. (1976). "Monthly Business Surveys." Proceedings of the Business and Economic Statistics Section, American Statistical Association.

Konschnik, C., Monsour, N., and Detlefsen, R. (1985). "Constructing and Maintaining Frames and Samples for Business Surveys." Proceedings of the Section on Survey Research Methods, American Statistical Association.

Kvålseth, T., (Nov. 1985). "Cautionary Note About R^2." The American Statistician, Vol. 39, No. 4 (Pt. 1), 279-285.

Lavallée, P. and Hideroglou, M. (June 1988). "On the Stratification of Skewed Populations." Survey Methodology, Vol. 14, No. 1, 33-43.

Office of Management and Budget (1987), Standard Industrial Classification Manual 1987. Superintendent of Documents, U.S. Government Printing Office, Washington, D.C.

U.S. Bureau of the Census (1979), Standard Statistical Establishment List Program, Technical Paper 44, Washington, D.C.

U.S. Bureau of the Census (1987), 1987 Industry and Product Classification Manual, Washington, D.C., Department of Commerce.

Wolter, K., Isaki, C., Sturdevant, T., Monsour, N., and Mayes, F. (1976). "Sample Selection and Estimation Aspects of the Census Bureau's Monthly Business Surveys." Proceedings of the Business and Economic Section, American Statistical Association, 99-109.

[1/] This paper reports the general results of research undertaken by Census Bureau staff. The views expressed are attributable to the authors and do not necessarily reflect those of the Census Bureau.