

definition of the estimated item matches the CES definition. The mean square error is defined as:

$$MSE = CV^2 \cdot \bar{X}^2 + (\bar{X} - \bar{X}(PCE))^2 \cdot CL$$

where CV = coefficient of variation for the item in the Diary or Interview

\bar{X} = item mean of either the Diary or Interview

$\bar{X}(PCE)$ = item mean calculated from PCE

CL = PCE means are multiplied by a factor of 1, .5 or 0, based on the completeness of information used in comparing CES and PCE item definitions.

In general, the estimate with the lower mean square error will be selected for the integrated publication, though exceptions may occur. For example, an exception might be made in categories where expenditure estimates commonly reflect outlays less reimbursements, such as medical care. The time frame and structure of the Interview survey make it superior to the Diary in collecting reimbursements, thereby making Interview survey expenditure estimates somewhat 'truer.' Of course, either the Diary or Interview can be automatically selected as the data source if information for the item is collected uniquely in only one of the surveys.

Once sources have been identified for each item, programs will be written to compile and aggregate the proper estimates from each component into publication tables. Such tables containing 1984-85 data will form the basis of a proposed integrated bulletin with a tentative release date of August 1987.

DATA ON DISKETTE

The usual media for dissemination of CES data have historically been printed materiel--news releases, reports, bulletins, Monthly Labor Review articles and research summaries--and reels of magnetic tape--the sets of public use tapes. In the foreseeable future, diskettes will join these media as a source of data.

In September 1985, the BLS moved to make time-series data contained in LABSTAT, available to the public on diskette. The LABSTAT computer system incorporates data storage, retrieval, and analysis capabilities for over 150,000 time-series created from BLS statistical surveys. Among the major price, employment, and productivity programs represented in LABSTAT are: Consumer Price Indexes for the Urban Population; Producer Price Indexes; National Industry, Employment, Hours, Earnings Statistics; Industry Productivity Statistics; and Productivity and Cost Measures.

Six methods are available for retrieving data from the LABSTAT data base. Each provides a different type of service, ranging from data presentation for statistical analysis to special formatting, tabulation, and charting for data display. Analytical capabilities encompassed in LABSTAT feature the Statistical Analysis System (SAS) and the Statistical Package for the Social Sciences (SPSS). Two seasonal adjustment programs are also available. All data are considered available to the public, regardless of whether they have actually been released in print or any other form. Such diskettes, available on a subscription basis, would be packaged with documentation outlining data uses, limitations, and future revisions. The files would be formatted for compatibility with the more common PC applications programs.

As the CES is now continuous and is producing time-series information, CES data have become eligible for transfer to diskette. In fact, as a special project, CES tables requested by the Internal Revenue Service have been successfully downloaded to diskette for their internal use. It should be emphasized that this is a long-term development. No dates have been scheduled as yet for the general production of diskettes. Preliminary plans involve putting Diary and Interview bulletin publication tables on diskette. Purchasers of the diskette would also receive a printed copy of the bulletin which would act as documentation.

CONCLUSION

The CES projects described in this paper show that the work of the Division of Consumer Expenditure Survey extends beyond the regular publication/public use tape program with which most data users are familiar. Developments on data comparisons, integration of Diary and Interview data, the durables public use tape, and data on diskette promise users both additional new CES data and alternative views and formats of previously released data. These projects, like the CES itself, will be operating on a continuing basis. With the Forms Design Team insuring that the Diary and Interview questionnaires remain clear and current, users can look forward to a future of timely expenditure data for different types of households in a wide variety of forms from the CES.

CONSUMER PRICE INDEX: INCORPORATING RESULTS FROM THE CONSUMER EXPENDITURE SURVEY

Charles C. Mason, Bureau of Labor Statistics¹

-----ABSTRACT-----

The first major activity in a Consumer Price Index (CPI) Revision is the implementation of a Consumer Expenditure Survey as a basis for selecting and weighting a new market basket of goods and services to be priced. Until these data are in hand, it is impossible to complete a revision of the CPI. It is essential to update the market basket periodically so that the CPI reflects price changes of items currently purchased by consumers. The Bureau of Labor Statistics is in the midst of a five year program to update and improve the CPI. This revision will use the CES data from 1982-84 and population distribution from the 1980 census to update the CPI market basket. The updated market baskets will be introduced in the January 1987 index.

in wage rates based on increases in the CPI. In addition to private sector workers whose wages or pensions are adjusted according to changes in the CPI, the index affects the income of more than 60 million persons largely as a result of statutory action: 38 million social security beneficiaries, about 3 1/2 million retired military and Federal Civil Service retirees and survivors, and about 20 million food stamp recipients. Changes in the CPI also affect 23 million children who eat lunch at school. In addition, since 1985 the CPI Urban Index has been used to adjust the Federal Income Tax structure in order to prevent inflation induced tax increases.

The uses of the CPI as a measure of inflation and economic policy, as a deflator of other series, and as an income or benefits escalator requires that it be a current accurate index. Buying patterns change over time as a result of price change, demographic changes in the population, or changing tastes or habits, market basket of goods and services priced by the CPI needs to be periodically updated. If the CPI measures an outdated market basket, the significance of any inaccuracy is heightened and errors can result in misleading economic signals at crucial phases of the business cycle. Historically the Bureau has updated the CPI market basket approximately once every ten years. The last revision of the CPI took place in 1978 and incorporated a market basket of goods and services based on 1972-73 spending patterns. The Bureau is currently in the process of revising the CPI again. The new index will be available in January 1987. The revised CPI will utilize a market basket of goods and services based on 1982-84 buying patterns. The first step in updating the market basket of goods and services is to survey the index population to determine what they are purchasing. The Consumer Expenditure Survey of the Bureau of Labor Statistics fulfills this need and forms the basis for designing the new market basket. The remainder of this paper discusses the formulation of the CPI and how the new CPI market basket will change as a result of the data collected in the Consumer Expenditure Survey.

INTRODUCTION

The first major activity in a Consumer Price Index (CPI) Revision is the implementation of a Consumer Expenditure Survey as a basis for selecting and weighting a new market basket of goods and services to be priced. The Consumer Price Index is a measure of the average change in the prices paid by urban consumers for a fixed market basket of consumption goods and services. It is calculated for two population groups. One group, the CPI Wage Earner and Clerical Worker Population, is composed of families and individuals where the majority of total income is derived from wage and salary occupations. The other group is composed of all urban families and individuals and is referred to as the CPI Urban Index. The CPI has three primary uses: first, as a measure of inflation and an indicator for Government economic policy; second, as a deflator of other economic series to adjust for price change and to translate these series into inflation free dollars; and third, as a means of adjusting income payments. More than 4.3 million workers are covered by collective bargaining agreements which provide for increases

¹Chief, Revision Cost Weights Staff

CPI FORMULATION AND STRUCTURE

The CPI is defined as a fixed quantity price index, and is the ratio of the cost of purchasing a set of items of constant quality and constant quantity in two different time periods. The index can be denoted by $I_{t,0}$ where t is the comparison period for which a new index number is to be

$$I_{t,0} = \frac{\sum_i P_{it} Q_{ir}}{\sum_i P_{i0} Q_{ir}} \times 100$$

P_{it} is the price for the ith item in period t
 P_{i0} is the price for the ith item in the reference period 0
 Q_{ir} is the quantity of the ith item consumed in the expenditure base period r

For sampling and computational purposes the set of all consumer expenditures is divided into progressively smaller and more narrowly defined categories. At the top is the aggregation of all items. The index for this aggregation is the number commonly used in the news coverage of the CPI. The "all items" CPI is derived by combining all items in all areas priced for the CPI. The next level of aggregation is the major group level. The CPI has seven major groups: Food and Beverages, Housing, Apparel and Upkeep, Transportation, Medical Care, Entertainment, and Other Goods and Services. Each major group is divided into two or more expenditure classes (EC). The current CPI has 68 EC's. The EC's have been historically defined to represent classes of closely related expenditures. Examples include Bakery products, Household appliances and Boys' apparel. To select specific items for pricing the EC's are partitioned into item strata composed of one or more entry level items (ELI). ELI are the first stage sampling units in the selection of an item for pricing. Table 1 shows an example of the CPI coding structure.

TABLE 1. AN EXAMPLE OF CPI CODING

EC	STRATA	ELI	TITLE
01			Cereal and Cereal Products
	01		Flour and Prepared Flour Mixes
		1	Flour
		2	Prepared Flour Mixes
	02		Cereal
	03		Rice, Pasta and Cornmeal
		1	Rice
		2	Macaroni, Other Pasta Products and Cornmeal

Item strata and ELI are designed to be relatively narrow homogeneous groupings of similar items which can be priced in the market place. Prices are collected for each item strata in the CPI. The number of prices collected for a particular item strata is based on a variance cost model which takes into account the relative importance of an item, the variance of the item strata index, and the cost of collecting the price data. For example, food items have a lower marginal cost of collection than other items since grocery stores stock a large selection of items and, therefore, the cost of pricing an additional item once the data collector is in the outlet is low. Given the choice to price an additional food item or some other item with similar relative importance and variance characteristics, the more efficient use of resources would be to price the food item.

For item strata including a broad range of product types such as girls' apparel, entry level items are defined to facilitate selection of a particular item for pricing. Girls' apparel is composed of seven ELIs--1) coats and jackets, 2) dresses, jumpers and suits, 3) sweaters, blouses and tops, 4) skirts and pants, 5) active sportswear, 6) underwear and sleepwear and 7) hosiery and other accessories. One or more of these ELIs are selected for pricing in each of the 91 local areas in which CPI prices will be collected. The selection is a statistical sampling procedure with the probability of selection equal to the relative importance of the ELI expenditure within the strata. Table 2 shows the average expenditure and the probability of selection for the ELIs in girls' apparel.

TABLE 2. AVERAGE EXPENDITURES AND PROBABILITY OF SELECTION FOR THE ELIS IN ITEM STRATA GIRLS' APPAREL

ELI	Title	Average Expenditure	Probability of Selection
1	Coats and Jackets	\$ 6.33	.099
2	Dresses, Jumpers and Suits	11.01	.172
3	Sweaters, Blouses and Tops	21.82	.341
4	Skirts and Pants	22.46	.351
5	Active Sportswear	5.31	.083
6	Underwear and Sleepwear	5.50	.086
7	Hosiery and Other Accessories	6.99	.109

NOTE: Average expenditures based on 1982-83 CES Interview Survey data published in bulletin 2246 Consumer Expenditure Survey: Interview Survey 1982-83 U. S. Dept. of Labor, Bureau of Labor Statistics, February 1986

In the current CPI the probability of selection for an ELI within each of the local areas was based on 1972-73 Consumer Expenditure Survey data tabulated at the regional level. In the revised CPI a rotating sample of areas (one fifth of the geographic areas each year), will have new probabilities of selection calculated annually at the regional level. The expenditure data used to calculate the proportions will be based on the two most recent years of Consumer Expenditure Survey data available.

CES DATA AND THE REVISED CPI MARKET BASKET

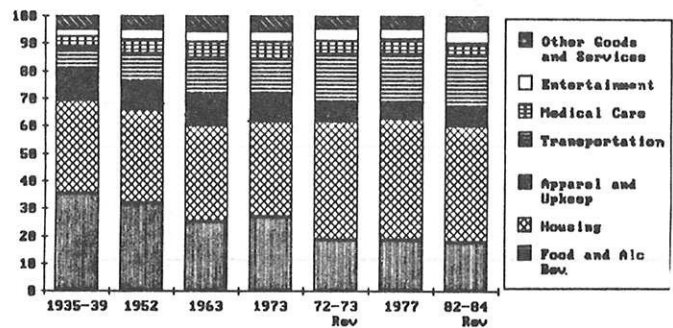
Analysis of the Consumer Expenditure Survey data in the construction of the revised CPI market basket confirms the continuation of many of the trends that were noticed in earlier revisions. The relative importance (i.e., the share of the total expenditure represented by the particular item strata or set of strata) for services continues to grow in relation to the 1972-73 data. Food at home continues to decline in relative importance and the relative importance of airline transportation continues to grow.

Relative importance can change as a result of changes in (1) the physical amounts consumed, (2) relative prices of goods and services, and (3) definition. An example of physical consumption change is the rapid growth of video cassette players and tape purchases. This growth has resulted in a larger relative importance for TVs video and audio equipment. The impact of relative price change on relative importance can be seen using an example of a consumer initially purchasing an equal dollar amount of steak and poultry. If the price of the steak rises twice as fast as the price of poultry and the consumer continues to purchase the same quantity of each, then the relative importance of steak in the consumer's budget may approach twice that of poultry. The change in relative prices may, however, induce a partially offsetting change in relative physical volumes of consumption. The third cause of changes in relative importance, definitional, can be demonstrated by the CPI decision to change the definition of used cars to the sum of all used car purchases less the value of vehicles sold or traded in on new car purchases. This change resulted in a reduction in the used car weight. The remainder of the paper examines the changes in the CPI relative importance within the major groups and looks at some of the underlying reasons.

Food and Beverages

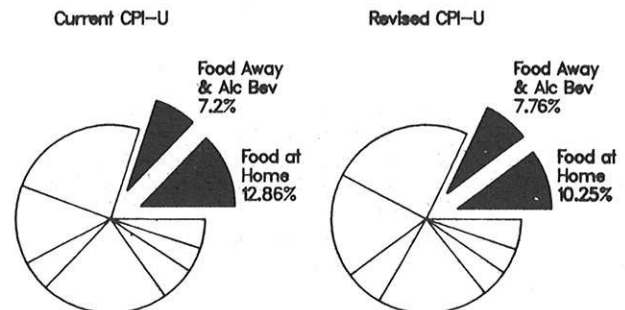
The food and beverages group continues the long established trend of declining relative importance in the CPI market basket. As Figure 1 shows, in the 1935-39 period food and beverages accounted for over one third of the CPI market basket. The preliminary cost weights for the revised CPI indicate that the new relative importance will be just over one sixth, nearly a 50% decline. All of the decline since the 1978 revision has taken place in the food at home category as shown in Figure 2. In fact, food away from home and alcoholic beverages have grown somewhat since the last revision. The trend towards declining food at home and increasing food away from home importance is probably related to the growing phenomena of two earner households.

Figure 1. Changing Relative Importances (1935/39 - 1982/84)



The decline in food and beverages reported since the last revision has resulted in publication changes for the CPI. The CPI currently publishes at the U.S. level 71 food at home item strata indexes, after the revision 52 will be published. The reduction represents the smaller pricing sample for food at home. Most of the old indexes will be combined with others so that a new higher level of aggregation will be available.

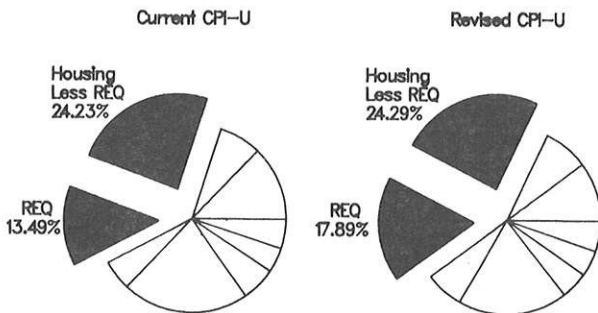
Figure 2. Food and Beverages



Housing

The relative importance for housing continues to grow. Even with the introduction of rental equivalence in 1983 the relative importance for housing reported in the CES has grown faster than the CPI relative importance. The change to rental equivalence converted the homeownership component from a method that included investment as well as consumption elements, to a flow-of-services approach that measures the cost of shelter services consumed by homeowners. The flow-of-services approach was implemented using the rental equivalence technique (REQ), which estimates the change in costs of renting housing services equivalent to those provided by owner occupied homes. As Figure 3 indicates, the largest change in housing is in homeowner's equivalent rent. The change in the cost weight from 13.49 percent of total expenditures to 17.89 percent reflects the differential growth in the value of owned homes and rents charged. The CPI estimates the REQ cost weight by asking CES households what they think their home would rent for and updates the index for price change by using the change in rent for housing similar to the owner stock as a proxy for the change in REQ. An analysis of CES and CPI rent survey data showed that during the past 15 years the quality of owner occupied housing has improved. Evidence of this quality improvement ranged from an increased number of rooms and bathrooms to an increased proportion of air conditioned homes. The proportion of the population living in owner-occupied dwellings has also increased.

Figure 3. Housing



The growth of certain other housing related expenditures has also resulted in changes in the CPI housing major group. New indexes will be published in the revised CPI when a sufficient sample is available for video cassette recorders and other video equipment, indoor plants and fresh cut flowers, care of invalids in the home, and home computers and other information processing equipment.

Apparel and Upkeep

Figure 4. Apparel and Upkeep

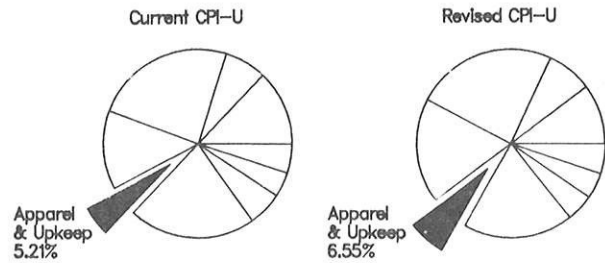
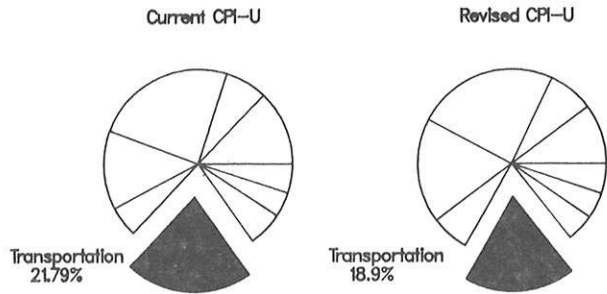


Figure 4 shows the change in relative importance for apparel and upkeep between the current CPI and the revised CPI. When the relative importance from the 1972-73 Consumer Expenditure Survey is compared with these two, its relative importance, 7.4 percent, is very similar to that of the revised CPI, 6.6 percent. Both of these relative importances are substantially higher than the current CPI's 5.2 percent. The similarity of the relative importances between the CES data in 1972-73 and 1982-84 suggests that expenditure proportions have remained fairly stable over the past decade. The difference between the CES data and the CPI suggests that average prices of purchased apparel have risen faster than the CPI index. These observations suggest either that consumers are buying a more expensive grade of clothing or that the CPI understates price change or some combination of the two. There is some prima facie evidence that the latter is at least partially the cause. For example, the index for women's suits has declined 14 percent since December 1977. Certain methodological problems with the large number of sales and with substituting new fashion year designs for the old contribute to this effect and are under careful review.

Transportation

Figure 5 shows that the relative importance of transportation has declined relative to the current CPI, although in this major group there are nearly as many increases as decreases at the EC or item strata level. One area which shows an increase is new cars. This differential reflects basic CPI concepts. The CPI prices a constant quality automobile and when new cars become available price change which are the result of quality changes and not true "price change" are factored out. The consumer however has not bought a "constant quality" automobile but has purchased an automobile with more and more features. This results in an increase in the relative importance for new cars in the revised index compared to the current CPI.

Figure 5. Transportation



Another contributing factor to the change in the transportation weight is the decline in the relative importance for gasoline. This change is probably due to consumers' conservation efforts, including the purchase of more fuel efficient vehicles in response to the rapid increase in gasoline prices experienced in the late 70's. These factors tend to raise the relative importance for gasoline much more than consumers increased their consumption.

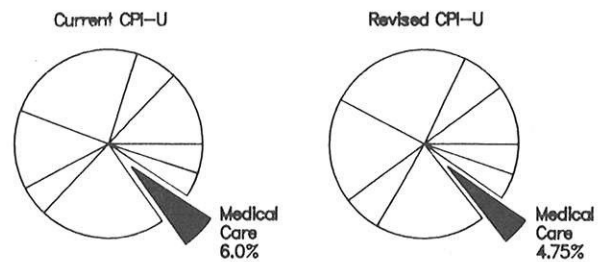
The relative importance for transportation has also changed for definitional reasons as evidenced by the decline for used cars from 3.6 percent to 1.3 percent. The change does not necessarily represent a shift from the purchase of used cars. It does however represent a change in the way the CPI defines the cost weight. In the past all purchases of used cars were valued at the transaction price, i.e., the negotiated price less any trade-in value. New cars were treated similarly. In the revised CPI the treatment of trade-ins and of outright sales by consumers will change. Trade-ins will continue to be netted out of the price of used cars, but, in addition, trade-ins on new cars will also be netted from used car purchases. A trade-in on a new vehicle is, in effect, a sale to the business sector. Ultimately another consumer will purchase the traded-in vehicle at a higher price since it will include the dealer margin (profit and preparation costs). Outright sales of vehicles from one consumer to another will be netted against the corresponding purchase since the sale/purchase is really an intra-population transfer of consumption, with one consumer giving up some other form of consumption in exchange for the car and the other consumer trading the car for the income to finance current consumption. The sale of a used car is construed as a negative expenditure providing the consumer with income with which to make other purchases.

The remaining used car sales are comprised of the sale of vehicles from the business sector to consumers. This used car purchase along with the dealer markup on used cars purchased from dealers constitutes the new CPI weight for used cars.

Medical Care

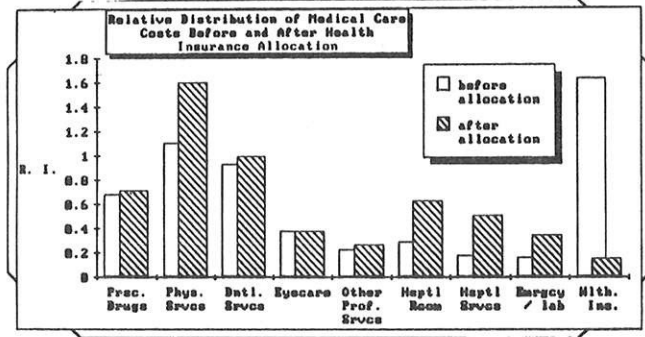
The relative importance for the revised CPI's medical care major group is about 25 percent less than the current CPI medical care component. The observed difference is at least in part affected by the way in which we pay for medical care. Major medical expenses are almost always partially paid for by health insurance. So while medical care prices have risen at a fairly high rate over the last decade, (156 percent vs 129 percent for all items), average family health insurance expenditures have risen only 19 percent. That health insurance expenditures have risen less rapidly than has medical care costs may be a result of reduced utilization of the policies or more probably reflects an increasing proportion of medical insurance paid by employers or government transfer payments.

Figure 6. Medical Care



Two changes in the medical care component are worthy of special note. In the current CPI health insurance relative importance is based on premiums paid by consumers and is allocated into benefits and retained earnings. In the revised CPI the health insurance relative importance will contain only the retained earnings. The benefits portion of health insurance has been allocated to the medical care service or commodity for which it is used. This change results in, *ceteris paribus*, a smaller relative importance for health insurance. However this allocation more accurately reflects the manner in which health insurance is priced and the true cost of medical care to the consumer. Figure 7 shows the relative distribution of each medical care item strata before and after the allocation from health insurance.

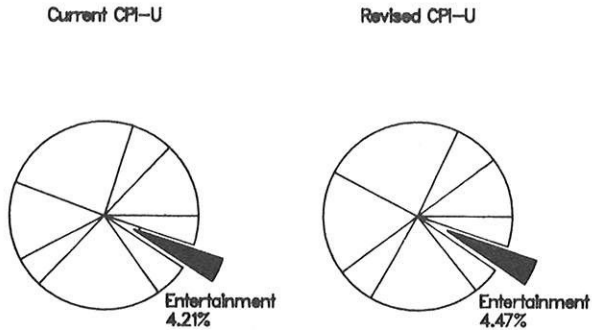
Figure 7. Health Insurance Allocation



The other change is a restructuring of the professional services index and calculation of an inpatient/outpatient index. The new construction results in a crossing of the line between commodities and services in that eyeglasses and contact lenses will be moved from medical care commodities into medical services and priced along with other eyecare services. This change facilitates the collection of prices for contact lenses and eye examinations. Certain outpatient services such as lab fees and x-rays will be moved from "Professional Services" to "Hospital and Other Medical Care Services".

Entertainment

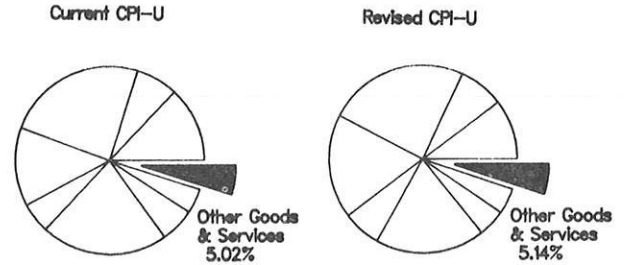
Figure 8. Entertainment



This major group shows almost no change between the current CPI and the revised CPI. However, the development of pricing methodologies for pet services, veterinary services and other entertainment services has led to a published index. In addition, separate indexes will be published for club membership and fees and for fees for participant sports.

Other Goods and Services

Figure 9. Other Goods and Services



The final major group, other goods and services is comprised of personal care, tobacco products and personal services. Figure 9 shows that a slight growth in relative importance has taken place in this category. All of the growth observed is in the personal services area. In response to this growth the revised CPI will publish separate indexes for legal fees, banking and accounting fees, and cemetery lots and funeral expenses.

CONCLUSION

The Consumer Expenditure Survey data have been tabulated and used to update the Consumer Price Index market basket of goods and services. The updated market basket will ensure that the CPI remains a current and accurate measure of the prices that consumers pay for well into the next decade.

WEIGHTING THE CONSUMER EXPENDITURE SURVEY DATA: THE GLS STUDY

Thesia I. Garner¹ and Kimberly D. Zieschang,² Bureau of Labor Statistics
U.S. Department of Labor

ABSTRACT

Weighting is the method by which a sample of interviewed households is made to represent the population. This paper describes the generalized least squares (GLS) weighting procedure. The procedure, developed at the Bureau of Labor Statistics to improve the accuracy of Consumer Expenditure (CE) Survey estimates of counts of consumer units, will soon replace the principal person method currently used in the CE Survey. The GLS procedure is the result of a major theoretical and applied research effort at the Bureau devoted to developing a new method that would address weighting problems resulting in part from the use of principal person weighting. A description of the GLS weighting adjustment method and results from applying the method to several quarters of CE data covering the early 1980's are presented.

INTRODUCTION

The Bureau of Labor Statistics, through a contract with the Bureau of the Census, collects information on consumer expenditures, income, and demographic characteristics from a two-component Consumer Expenditure (CE) Survey. The components are distinguished by survey instrument, the first being a three hour Interview conducted with four rotating panels of consumer units (CU's) for five consecutive quarters, and the second being a two-week Diary of expenditures to be completed in large part by the sampled consumer units. Each component in effect constitutes a separate survey, since there is no overlap between the respective samples, the only relationship between the two being that the samples for each are chosen from the same frame at the same time. There is considerable overlap in the kinds of information collected between the surveys, primarily though not exclusively demographic in nature. Some overlap also occurs in information collected on expenditures, though the surveys are oriented toward obtaining different kinds of expenditure information. The Interview component is designed to provide information on major, infrequent purchases, "infrequent" meaning for practical purposes with a frequency of less than two weeks to a month. The Diary component is oriented toward purchases with more than weekly or biweekly

frequency, food purchases being one of the major categories of interest. However, the Diary component is not explicitly limited to certain categories at present, in part to minimize confusion for the households completing the Diary. This is the primary source of overlap in expenditure information collected.

In this paper the emphasis is on improving CE Survey estimates of counts of consumer units partitioned into categories of demographic and economic interest, such as family composition and tenure status. The mode of accomplishing this improvement is to develop a technique of weighting adjustment that uses ancillary information on the age, race, and sex composition of persons as part of a uniform, mean square error (MSE) minimizing estimation procedure for CU totals. The exposition first briefly reviews the current practice for weighting adjustment and estimation in the CE Surveys. We then introduce a method of weighting adjustment first proposed by Don Luery and considered by Anthony Roman [1] of the Bureau of the Census. This generalized least squares (GLS) method is extended to define a procedure that integrates overlapping demographic information from the surveys in a way that should reduce the MSE and improve the accuracy of key CU counts. We then describe application of the GLS procedure to thirteen quarters of data from the Consumer Expenditure Survey covering 1980 IV - 1983 IV and its evaluation against the current principal person weighting procedure.

CURRENT PROCEDURES

The basic sampling weights for the CE Surveys are determined according to the size measured in number of housing unit addresses of each of seven subframes from which sample addresses are chosen, relative to the sample size allocated to the subframe. The basic weight assigned to an address is the inverse of the probability of selection of the housing unit. These basic weights are adjusted by application of a "weighting control factor" to the basic weight, accounting among other things for unexpectedly clustered addresses, as for example, student housing in the "area segment" subframe. A further "noninterview factor" is applied to adjust for inability to obtain an interview for some sample units, for example, those in occupied housing units who refuse to participate or units where no one is home. The adjustment is performed within groups of consumer units defined by geographic area,

¹Research Economist, Division of Price and Index Numer Research

²Research Economist, Chief, Division of Price and Index Number Research

tenure, household size, and race. All consumer units within each household at an address are assigned this adjusted weight. Each member then receives a "second stage" adjustment to force the population of members classified into 48 age, race, and sex categories, as estimated by the CE Surveys in a given month, to equal the U.S. population for that month in those age, race, and sex categories. These latter quantities are called "control totals" and we will refer to this as a weighting control procedure. The sources of control totals for the CE Surveys are the U.S. Census and the Current Population Survey (CPS). A single "principal person" is chosen to represent the CU. The CU receives this person's weight. The "principal person" is the female of a reference person and spouse pair, or the reference person when no spouse is present. If the "principal person" is male, his weight and hence his CU's weight is multiplied by a "principal person factor" to adjust for a historical tendency for males to be underrepresented compared with females. The current stages of weighting beyond noninterview adjustment have two objectives: weighting control to add additional information to the estimation process, and the assignment of a weight to the CU, a sampling unit unique to the CE surveys.

The resulting "principal person" CU weights are then used for computing estimates of totals and averages for various quantities of interest from each of the surveys. To make "integrated" estimates of average expenditures for various demographic, tenure, and income categories, a selection is made from one or the other survey when information is available from both at a very detailed level of expenditure classification, by Uniform Classification Code (UCC).

Although no attempt is made at present to make "integrated" estimates of counts of persons or CU's within the demographic and economic categories for which average expenditures are currently computed, interest has arisen in making this "integration". There is the need to know which survey, the CE Diary or Interview, is producing the best estimate for each of a number of counts. More generally, there is a need to combine information from both surveys in arriving at a single estimate more efficiently than merely "integrating" them. The alternative generalized least squares procedure described next is designed not only to perform the weighting "control" function of the current method, but this weighting "composition" function as well.

THE GENERALIZED LEAST SQUARES PROCEDURE

The Adjustment Equation and Computational Layout

The GLS procedure adjusts the sample weights after noninterview adjustment by minimizing the weighted squared adjustments subject to a set of linear "control" constraints the adjusted weights must satisfy. The problem can be expressed as

$$\min_W (\Omega - W)' \Lambda^{-1} (\Omega - W) \text{ subject to } MW = N \quad (1)$$

where Ω = the pre-adjustment sample weight vector of dimension $n \times 1$, where n is the sample size;

W = the adjusted sample weight vector of dimension $n \times 1$;

Λ = the GLS weighting matrix of dimension $n \times n$;

M = a $k \times n$ matrix whose columns provide the counts of k items for each of n CU's;

N = a $k \times 1$ vector of known "control" counts for each of the items corresponding to the rows of M .

In the usual single sample situation, the rows of M and N correspond to persons in each of a number of age, race, and sex categories, or any other item on which there is known control information that is also collected for each sample unit. The CE context is characterized by two samples. This implies dual control constraints, one set for each sample. In addition, it suggests the possibility of equating the survey counts of items on which no control information exists, but on which comparable information is collected in each sample. Both the "control" and "composition" aspects of weighting adjustment are straightforwardly accommodated in the constraint system $MW = N$ by a suitable partitioning of matrices. Thus, let

$$\Omega = \begin{bmatrix} \Omega_1 \\ \Omega_2 \end{bmatrix}; \quad \Lambda = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ & \\ \Lambda_{12} & \Lambda_{22} \end{bmatrix};$$

$$M = \begin{bmatrix} 0 & 0 \\ M_1 & 0 \\ 0 & M_2 \\ M_1^c & -M_2^c \end{bmatrix}; \quad N = \begin{bmatrix} N^0 \\ N^0 \\ 0 \end{bmatrix};$$

where subscripts index the samples, superscript "0" refers to control items, and superscript "c" refers to composite items.

The solution of the minimization (1) is given by the following generic formula, with appropriate substitutions for the matrices Ω , W , Λ , M , and N as above for the two sample control/composition case:

$$\hat{W} = \Omega + \Lambda M' (M \Lambda M')^{-1} (N - M \Omega). \quad (2)$$

The weight MSE matrix is Λ ; Λ can be determined from the sample data under certain assumptions

about the sampling process. For details, see Zieschang [2,3]. Expression (2) for \bar{W} was computed for each test quarter of the CE Surveys for the full sample and the twenty sample replicates.

Bounding Extreme Weight Adjustments

Experience with weights generated by the GLS procedure indicated little propensity to generate extremely large weights or proportional weight adjustments relative to those produced by the current principal person procedure. However, occasional extreme downward adjustments did occur in either the replicate or full files of both surveys that resulted in negative adjusted weights. These negative values preclude interpretation of the weight as the number of consumer units represented by a given CU, and are conceivably indicative of high variance estimates of derivative statistics for some set of small subdomains of the CE universe.

The method used to deal with this problem simply recodes the adjusted weight when it falls outside a tolerance interval containing the unadjusted weight. In this study the tolerance interval was set, with lower bound only, at one fourth of the unadjusted weight. For example, if the unadjusted weight was 16,000, the lower bound would be set at 4,000. Upper bounds were not enforced because the "pre stage two" weight, Ω , becomes progressively biased downward as population growth occurs between sample selections. Secularly rising proportional adjustments are therefore reasonable and upper limits on tolerance regions of weights potentially risky. Setting a lower bound only will bias up estimates of totals produced with the recoded weights; however, evidence from the empirical trial below clearly demonstrates that this bias is extremely small for the twenty-five percent lower bound used here in the CE context.

Another consideration involves the computation of variances and the appropriate tolerance interval lower bound for proportional adjustment in the replicates, given that this is set at one fourth of the unadjusted weight in the full file. The replicate proportional tolerance should be "looser" than that of the full file to produce accurate variance estimates, though how much "looser" is not straightforwardly determined. However, comparison of recoded coefficients of variation (CV's) of estimates of CU subdomain size and mean family income before tax with CV's generated by the unrecoded GLS weights indicated very little effect here either. The proposed method therefore stands as an attractive and viable near term alternative to the current procedures.

Time Interval Selection and Control Subdomains

To evaluate the performance of the GLS procedure in the context of the Consumer Expenditure

Surveys, the time interval for weighting batches of consumer units was set at one quarter. Current principal person procedures are implemented on a monthly basis; however, this results in weighting batches of consumer units of between 300 and 400 for the Diary Survey, or for rotation groups in the Interview. The number of CU's drops to 150-200 for the sample replicates (used for variance calculation) that are weighted in parallel with the full sample. Batches of this size tend to have patchy coverage of CU-member age/race/sex characteristics that are to be controlled to known Census/CPS counts. This is not surprising, since there are 48 control categories within which these at most 300-400 CUs' weights are being adjusted. Current procedure deals with this by a large amount of collapsing of control cells into one another, sacrificing control detail. The primary purpose for monthly weighting adjustment is to insure that the Diary months and Interview rotation group/months are correctly scaled relative to one another in the aggregate. This purpose should be served well enough by controlling the CU member weights in each month or rotation/month so that they add to the total population for that month.

To provide better coverage of the control cells, quarterly weighting is preferable to monthly because the weighting batch is three times as large. Under current principal person procedures, going to a quarterly time interval would require sacrificing the aggregate population controls on Diary months and Interview rotation/months within quarter. On the other hand, a quarterly GLS procedure for CU member control cells can accommodate these aggregate monthly and rotation controls simply as additional linear constraints to be satisfied. It is therefore able to take advantage of the better control cell coverage of large quarterly batches of data while retaining essential controls at the month and rotation/month level.

Even for quarterly batches of data, these coverage considerations lead to an aggregated set of controls for the Diary Survey via a reduction in the number of member age groups from the available twelve to the six used in CE publications. Certain of the 48 detailed age/race/sex cells were empty for the Diary in either the full file or replicate samples in certain quarters. The 48 controls were applied to the aggregate of all rotations and months for each quarter in the Interview Survey, resulting in weighting batches of between 4200 and 4600, a size sufficient to preclude empty control cells in the full or replicate samples over the thirteen quarter period of data used for this study. The quarterly age/race/sex cells used in this study for the Diary and Interview Surveys are detailed in Table 1. In addition, monthly total population controls were added for the Diary Survey and rotation group/month total population controls were added for the Interview.

Table 1. Member Control Categories^a

AGE	BLACK MALE		BLACK FEMALE		NON-BLACK MALE		NON-BLACK FEMALE	
	I	D	I	D	I	D	I	D
14 - 17	1		13		25		37	
18 - 21	2	1	14	7	26	13	38	19
22 - 24	3		15		27		39	
25 - 29	4	2	16	8	28	14	40	20
30 - 34	5		17		29		41	
35 - 39	6	3	18	9	30	15	42	21
40 - 44	7		19		31		43	
45 - 49	8	4	20	10	32	16	44	22
50 - 54	9		21		33		45	
55 - 59	10	5	22	11	34	17	46	23
60 - 64	11		23		35		47	
65*	12	6	24	12	36	18	48	24

^aI = Interview
D = Diary

Composite Subdomains

Subdomains chosen for the rows of the M^C matrices of equation (2) include region of CU residence, sampling frame from which CU was drawn, tenure status of the CU, and four family types. These groupings are described in Table 2.

Table 2. Consumer Unit Composite Subdomains

REGION	
NEAST	= NORTHEAST REGION
NCENTRAL	= NORTH-CENTRAL REGION
SOUTH	= SOUTHERN REGION
WEST	= WESTERN REGION
SAMPLING FRAME	
CEN70	= 1970 CENSUS FRAME
SPECPLAC	= SPECIAL PLACES FRAME
ARSEG	= AREA SEGMENTS FRAME
NEWCON	= NEW CONSTRUCTION FRAME
TENURE	
OWNER	= OWNER CONSUMER UNITS
RENTER & STUDENT	= RENTER CONSUMER UNITS INCLUDING THOSE IN STUDENT HOUSING
FAMILY TYPE	
ALL_HM	= ALL HUSBAND/WIFE CONSUMER UNITS
SPT1+<18	= SINGLE PARENT CONSUMER UNITS
SINGLE	= SINGLE PERSON CONSUMER UNITS
OTHER	= ALL OTHER CONSUMER UNITS

Region, tenure, and family type were chosen as composite subdomains because of their use in CE publications. Additional family type detail plus a set of CU size categories would exhaust the set of demographic subdomains on which tables are published for the CE. Augmenting the composite classification chosen above to include these additional groups would be a useful extension of the current study. The frame classifications were included because of prior knowledge that the frames from which CU's are drawn are the same size (since they are identical) between surveys. Also, improved

composite estimates of numbers of CU's in each subframe should aid in planning future CE samples, and in assessing ongoing sampling performance and identifying problems in the management of the sampling process.

SOME EMPIRICAL RESULTS

Data

The GLS weighting procedure described in the previous section was applied to the CE Diary and Interview data for thirteen quarters covering 1980 IV through 1983 IV. The urban samples only were weighted, since rural CU's were eliminated from the universe from 1981 IV through 1983 IV. Urban controls included age/race/sex (A/R/S) population totals computed by the Bureau of the Census from (a) updated 1970 census data and (b) urban/rural population distributions obtained from the Current Population Survey (CPS). The controls for the 48 A/R/S cells were available monthly over the period covered. These were summed across months for each quarter and across A/R/S types for each month to obtain the quarterly A/R/S and monthly total population controls to be used by the quarterly GLS procedure. The summed monthly total population controls were divided by three for the Diary and by twelve for the Interview so that the controlled weights would sum to average total population for the quarter: in the first case across the three months within the quarter, and in the second across the three months and four rotations. The quarterly sums of the monthly A/R/S controls were divided by three to obtain average quarterly population in each of the 48 cells.

Other data files originated from the BLS CE database. The "pre-stage two" weight Ω to be adjusted was computed as the product of the base weight (BASEWT), weighting control number (WTCONTNO), and monthly noninterview factor (NONINTAD). The M matrix of counts of persons or indicators of subdomain membership was generated straightforwardly from the AGE, RACE, and SEX, and REGION, SERIAL, CUTENURE, and FAMTYPE variables in the data base.

Control Errors

To provide an indication of the success of the GLS and GLS-recoded procedures in meeting the control objective of weighting (see Bounding Extreme Weight Adjustments in previous section), the percentage deviation of the adjusted weights from the control counts were computed. For comparison purposes, these deviations were computed for the currently used monthly principal person weights in the database as well. Table 3 displays the computed percentage control errors for 1980 IV, using BLS publication aggregates of age.

The principal person errors are under the headings suffixed 'FINLWT', while the GLS-recode

Table 3. Percent Control Errors 1980 Quarter Four

MEM_RACE	MEM_SEX	AGG_AGE	DFINLWT	DRECODE	DM_HAT	IFINLWT	IRECODE	IM_HAT	CONTROL
			0.77	0.00	-0.00	2.58	0.01	-0.00	139835053
BLACK			3.74	0.01	-0.00	1.58	0.07	0.00	16472402
NONBLACK			0.38	-0.00	-0.00	2.72	-0.00	-0.00	123362651
BLACK	FEMALE		1.04	0.01	-0.00	2.41	0.13	0.00	9108624
BLACK	MALE		7.08	0.01	-0.00	0.56	0.00	0.00	7363778
NONBLACK	FEMALE		-0.42	-0.00	-0.00	0.97	-0.00	-0.00	64812014
NONBLACK	MALE		1.26	-0.00	-0.00	4.65	-0.00	-0.00	58550637
BLACK	FEMALE	14-24	0.14	-0.00	-0.00	15.55	0.33	-0.00	2728482
BLACK	MALE	14-24	25.01	-0.00	-0.00	0.11	0.00	0.00	2403580
NONBLACK	FEMALE	14-24	-3.28	-0.00	-0.00	2.64	-0.00	-0.00	15439254
NONBLACK	MALE	14-24	-3.89	-0.00	-0.00	6.11	-0.00	-0.00	15119474
BLACK	FEMALE	25-34	-1.68	-0.00	-0.00	-8.84	0.00	0.00	2046992
BLACK	MALE	25-34	-13.94	-0.00	-0.00	16.47	-0.00	-0.00	1628839
NONBLACK	FEMALE	25-34	-0.26	-0.00	-0.00	0.46	-0.00	-0.00	1345380
NONBLACK	MALE	25-34	4.16	-0.00	-0.00	4.81	-0.00	-0.00	13008970
BLACK	FEMALE	35-44	-9.99	-0.00	-0.00	6.78	0.00	0.00	1319357
BLACK	MALE	35-44	-13.73	-0.00	-0.00	-17.33	0.00	0.00	1054166
NONBLACK	FEMALE	35-44	0.32	0.00	0.00	0.55	-0.00	-0.00	9069243
NONBLACK	MALE	35-44	-3.28	-0.00	-0.00	-4.02	-0.00	-0.00	8738016
BLACK	FEMALE	45-54	9.63	-0.00	-0.00	3.89	0.00	0.00	1103269
BLACK	MALE	45-54	34.27	-0.00	-0.00	-10.88	-0.00	-0.00	873999
NONBLACK	FEMALE	45-54	0.52	-0.00	-0.00	-0.22	-0.00	-0.00	8209192
NONBLACK	MALE	45-54	4.02	-0.00	-0.00	4.07	-0.00	-0.00	7635931
BLACK	FEMALE	55-64	-0.97	-0.00	-0.00	-21.60	0.25	-0.00	914057
BLACK	MALE	55-64	16.22	0.09	-0.00	10.53	-0.00	-0.00	734995
NONBLACK	FEMALE	55-64	0.73	-0.00	-0.00	0.45	-0.00	-0.00	8200210
NONBLACK	MALE	55-64	-5.11	0.00	0.00	11.06	-0.00	-0.00	2195534
BLACK	FEMALE	65*	16.01	0.07	-0.00	14.97	0.00	0.00	996468
BLACK	MALE	65*	-18.94	0.00	0.00	-4.42	0.00	0.00	668200
NONBLACK	FEMALE	65*	1.33	0.00	0.00	0.85	-0.00	-0.00	10439735
NONBLACK	MALE	65*	16.53	0.00	0.00	6.12	-0.00	-0.00	6852711

and GLS results are under those suffixed, respectively, 'RECODE' and 'W HAT'. Prefixes 'D' and 'I' refer to the Diary and Interview Surveys, respectively. It can be seen in this table that the GLS weights hit the controls exactly, that the errors of GLS-recode are negligible, and that the errors of the existing principal person weights are often substantial.

Composition Errors

The success with which the composition objective was met on the composite subdomains was measured by the 'arc-discrepancy' between weighted CU totals from the Diary and Interview Surveys. This is computed as twice the difference between survey estimates divided by their sum, yielding an average measure of the proportional difference. Results for the region composite subdomains are presented in Table 4 for five test quarters. Principal person weights are indicated as 'FINLWT', while GLS-recode and GLS weights are indicated respectively as 'RECODE' and 'W HAT'. The composite subdomains are labelled as in Table 2, with the blank subdomain referring to all CU's. As with meeting the control objective, GLS achieved the composition objective, equating the surveys exactly, GLS-recode generated negligible discrepancies, and principal person generated often substantial discrepancies between surveys.

Effects on Survey Discrepancies in Non-Composite Subdomains

Arc-discrepancies were also computed for subdomains related to BLS publication categories but not composited during weighting adjustment. These included six age of CU head categories, a tenure category for non-student housing renter status, and five CU size categories. Also included were five husband and wife family/age

Table 4. Arc Percentage Discrepancies, Composite Subdomain Region

	SUBDOMAIN=				
	_B04	_B11	_B12	_B13	_B14
WEIGHT					
FINLWT	4.08	3.43	0.64	1.33	3.96
RECODE	-0.00	-0.00	0.22	0.04	0.07
GLSMT	0.00	0.00	0.00	0.00	0.00

	SUBDOMAIN=WEST				
	_B04	_B11	_B12	_B13	_B14
WEIGHT					
FINLWT	11.20	20.83	12.52	16.67	4.89
RECODE	0.00	0.00	0.01	0.09	0.01
GLSMT	0.00	0.00	0.00	0.00	0.00

	SUBDOMAIN=CENTRAL				
	_B04	_B11	_B12	_B13	_B14
WEIGHT					
FINLWT	1.68	-1.54	-4.51	-8.49	5.64
RECODE	0.00	0.00	0.17	0.07	0.01
GLSMT	0.00	0.00	0.00	0.00	0.00

	SUBDOMAIN=SOUTH				
	_B04	_B11	_B12	_B13	_B14
WEIGHT					
FINLWT	-7.28	-8.79	-10.90	-8.63	-4.64
RECODE	-0.01	-0.01	0.57	-0.02	0.20
GLSMT	0.00	0.00	0.00	0.00	0.00

	SUBDOMAIN=WEST				
	_B04	_B11	_B12	_B13	_B14
WEIGHT					
FINLWT	15.32	7.53	10.37	10.84	14.55
RECODE	0.00	0.00	0.01	0.01	0.01
GLSMT	0.00	0.00	0.00	0.00	0.00

of oldest child classes, two single person CU/employment status classes, three multi-person CU/employment status classes, and an incomplete income response class. In all subdomains, GLS and GLS-recode usually were better than principal person on a quarter by quarter basis. Discrepancies in the incomplete income response category were large because income response is markedly different between the surveys, the most likely explanation involving the difference in survey instruments.

Effects on the Magnitude of Estimates of Subdomain Size and Mean Family Income Before Tax

Because GLS is proposed as an alternative to the current principal person methods, it is of some interest to know if there is a systematic or pervasive difference in estimates produced with the adjusted weights. To evaluate this, two types of variable were considered. The first was subdomain size, the estimated total for assorted subgroups of consumer units. The second was mean family income before tax (FIBT) for the same assortment of CU subgroups. The behavior of mean FIBT should be indicative of the behavior of other important survey variables, notably detailed expenditures, because mean FIBT and mean expenditure by product classification are highly correlated for the most part.

Ratios of estimates of CU population generated by GLS-recode and principal person weights, for both the composite and a set of noncomposite subdomains, were calculated. There was a tendency for GLS to estimate lower CU counts than principal person across time and across subpopulations for Diary and Interview.

Ratios of mean FIBT produced by GLS and principal person weights for complete reporters on the income question, in general, indicated that the relationships between the estimates derived from GLS and principal person weighting were opposite to those for population size. There was a small but pervasive tendency for GLS estimates of mean FIBT to be higher than those for principal person, an effect slightly more pronounced for the Diary Survey compared with the Interview.

Precision of Estimates of Subdomain Size and Mean Family Income Before Tax

Ratios of GLS with principal person estimates of the CV's of estimates of subdomain size and FIBT were computed to evaluate the precision of the estimates. The CV is the ratio of the standard deviation of an estimate with the estimate itself; the CV provides a 'scale independent' measure of the variability of the estimate. The standard deviations used in computing the CV's were computed using twenty independently (GLS or principal person) weighted replicate samples to generate twenty replicate estimates for each cell considered. The standard deviations were computed as the square root of the average squared difference between the replicate estimates and that of the full sample. The improvements in precision obtained with the GLS weights are rather striking for subdomain size, particularly for the composite subdomains in the Diary Survey. Results demonstrate that overtime, within the rather lengthy test period, GLS weighting improved the precision of estimates of CU population in each of the subdomains considered.

Results for mean FIBT reveal that over time GLS estimates are more precise than principal person for a majority of the composite subdomains in either survey. For the selection of noncomposite subdomains, Diary GLS mean FIBT estimates are about even with principal person in precision, while the Interview estimates are generally noticeably better using GLS.

As a final note to this section, the measurement of the precision of GLS versus principal person estimates of totals and means is only one of two components of the accuracy of the estimates as measured by mean square error. The relative CV's just discussed provide information on the variance component of MSE, but not the squared bias component. GLS weighting as specified here is unbiased if the sample design and execution are unbiased. A great deal of effort is expended by BLS and the Census Bureau to collect unbiased samples, and to the extent this effort is successful, the GLS estimates are probably more accurate than the principal person, even in the infrequent event they are slightly less precise. Of course to the extent that these efforts to achieve unbiasedness fail, this assertion loses force. However, even in this case, GLS provides a well-ordered tableau within which issues of survey bias can be examined.

CONCLUDING REMARKS

The results of this empirical evaluation of a quarterly GLS weighting procedure for the CE Surveys have been for the most part highly favorable to GLS when compared with the existing principal person procedures. Not only does the operational, recode variant of GLS display a high degree of numerical consistency with the control totals and in equating composite CU totals, it also demonstrably improves the precision with which the sizes of both the composite subdomains and a broad selection of other, noncomposite subdomains are estimated. For the analytical variable family income before tax, improvements in precision are minor at the subdomain level, but are of notable magnitude at the all CU level. The CV's of the Diary and Interview GLS estimates of mean FIBT average, respectively, eight and twelve percent below the same statistics computed for the corresponding principal person estimates. To the extent that these gains are inherited by estimates of mean expenditures, it is reasonable to expect that GLS weighting will improve the quality of the data published by the Division of Consumer Expenditure Studies. GLS weighting might also be used in future revisions of the Consumer Price Index, providing possibly more precise expenditure weights for producing that series.

REFERENCES

1. Roman, Anthony. "Proposal for Weighting Research in the Consumer Expenditure Surveys," interval memorandum, Demographic Surveys Division, U.S. Bureau of the Census, 1982.
2. Zieschang, Kimberly D. "Consumer Expenditure Survey GLS Weighting Study; I. Weighting and Estimation in the Consumer Expenditure Surveys," Division of Price and Index Number Research, Office of Prices and Living Conditions, Bureau of Labor Statistics, 1985a.
3. Zieschang, Kimberly D. "Consumer Expenditure Survey GLS Weighting Study; II. Empirical Evaluation of GLS and Principal Person Procedures," Division of Price and Index Number Research, Bureau of Labor Statistics, 1985b.

METHODOLOGICAL RESEARCH IN THE CONSUMER
EXPENDITURE SURVEY PROGRAM

Clyde Tucker, Bureau of Labor Statistics¹

ABSTRACT

This paper reports on two research projects involving the CE Diary Survey. One project, almost completed, examines the attitudes and record-keeping behavior of diary respondents. The other project, still in the data-collection phase, is an experiment to evaluate alternative diary formats. The paper discusses the reasoning which led to these projects and various aspects of their designs.

INTRODUCTION

Survey methods can always be improved. The Consumer Expenditure Survey Program supports ongoing methodological research designed to improve the quality of the CE data. This paper reports on two such research projects involving the CE Diary Survey. One project, almost completed, examines the attitudes and record-keeping behavior of diary respondents. The other project, still in the data-collection phase, is an experiment to evaluate diary formats.

The CE Diary Survey was designed to collect all of the daily expenditures of sampled consumer units over two consecutive weeks. A consumer unit is defined as one of the following: (1) the collection of all members of a household who are related by blood, marriage, adoption, or other legal arrangement; (2) a person living alone or sharing a household with others or living as a roomer in a private home or lodging house or in a permanent living quarters in a hotel or motel, but who is financially independent; or (3) two or more persons who live together and pool their income to make joint expenditure decisions. To be considered financially independent, at least two of the three major expense categories (housing, food, and other living expenses) have to be provided by the respondent.

The diary is especially effective for gathering information about small, frequently purchased items which are normally difficult to recall. These expenditures include grocery items, meals eaten out, household supplies, inexpensive articles of clothing, and personal care products and services. In addition to expenditures, data are also collected on the income, work experience and demographic characteristics of consumer unit members.

Each year data are collected from approximately five thousand consumer units in more than one hundred primary sampling units (PSU's) throughout the country. The sample is designed to be

representative of the national, noninstitutional population. Besides the population residing in regular housing, persons residing in selected group quarters, such as college dormitories, are also represented. Weights are assigned to each consumer unit in the survey in order to provide estimates for the U.S. population.

RESPONSE ERROR IN THE DIARY

The program of research described here is intended to reduce response error, in particular response bias. We believe that most response errors in the diary come in the form of underreports of expenditures. It is difficult to imagine an individual recording more items than were purchased or consistently overreporting the price of items, but the failure to report all items is quite likely given the time and effort required to keep the diary. There is substantial evidence to support this assertion. Sudman and Ferber [3] found that the estimates from consumer diaries recorded by hand were less than those obtained using tape recordings, and estimates of food expenditures from the CE diary were below those from the National Accounts in both the 1972-73 and 1980-81 periods [2, 6].

Response errors can contribute either to the bias or variance in estimates. There are other nonsampling errors which also can increase the total error in a survey and are not included in the sampling variance given in published reports.

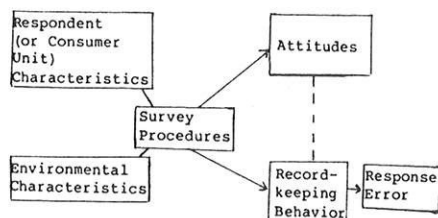
These errors include sampling bias, clerical errors and processing errors. We have concentrated on response error because we believe this form of nonsampling error represents the greatest threat to the integrity of our expenditure estimates. We are already using supplemental sampling frames, weighting and imputation to correct for sampling bias; and clerical and processing errors are more likely to be random than are response errors.

APPROACH TO THE STUDY OF RESPONSE ERROR

Response errors occur at the micro level and can be viewed as outcomes of the survey process. Figure 1 provides a picture of this process for the CE diary. The elements found here are similar to those found in other survey processes. One element which is not the same as in the typical survey situation is the respondent's record-keeping behavior. Usually the respondent's behavior is confined to thought processes and verbal responses. In the diary survey, however, the respondent has a more active role.

¹Mathematical Statistician, Office of Prices and Living Conditions.

FIGURE 1. Diagram of the Survey Process Leading to Response Error



We can see from this diagram that respondent (or consumer unit) characteristics, in combination with environmental circumstances and the intervening survey procedures (including interviewer characteristics), influence both the respondent's attitudes toward the survey and his or her record-keeping behavior. The attitudes and record-keeping behavior (both thought processes and physical actions) are collateral in that they occur at approximately the same point in time. Attitudes are thought processes, but they are not a part of record-keeping behavior. To make this point more clearly, attitudes are not responsible for the outcome of the survey process; the record-keeping behavior is. Obviously, in many cases, attitudes and record-keeping behavior coincide; but there may be a number of instances where this is not true.

As a first step toward the reduction of response error in the diary survey, we have to determine the causes of response errors. Once the causal factors are identified, we can develop new survey methods for overcoming their effects. In order to identify these factors, we have to understand how the antecedent elements in the survey situation interact to affect both attitudes and record-keeping behavior. We then can complete our understanding of the causal sequence leading to response errors by examining the relationships between particular behaviors and response errors. Since response errors occur at the micro level, we would like to carry out our studies at that level. To this end, we are engaged in research to define measures of an individual respondent's response error (5); but we also are interested in conducting aggregate analyses of the expenditures themselves by the demographic, attitudinal and behavioral variables. The remainder of this paper discusses two research projects which will provide us with our first information about the causal sequence outlined above.

THE SUPPLEMENTAL DIARY SURVEY: UNIVARIATE DISTRIBUTIONS

The causal sequence leading to response error is extremely complicated. We have designed our studies to build on one another until we are able to analyze the survey process in its entirety. I have already mentioned the work we are doing to develop measures of response error. Another study focuses on the role of attitudes and record-keeping behavior. The data for this project came from a supplemental survey administered to the diary respondents and interviewers in the second quarter of 1984 at the conclusion of the second diary week. This

questionnaire, found in the appendix, was designed to measure the feelings and behaviors associated with keeping the diary. Some of the information also can be used to establish the environmental context.

Before proceeding with a discussion of the supplemental survey, let me mention the major drawback to a survey of this kind--social desirability effects. Most of the respondents are aware of the recommended method for keeping the diary. They are also aware of the kinds of attitudes which will be viewed in a favorable light. Therefore, there may be a tendency for respondents to give false answers in the socially desirable direction.

With this in mind, I want to turn to a review of the initial findings from the supplemental survey, some of which are presented in Table 1. One of the most important findings is that almost twenty percent (215) of the respondents completed at least one of the diaries by total recall. This means that no entries had been made in the diary during the week so the interviewer had to record as many of the previous week's expenditures that respondent could recall. The extent of recall poses an obvious threat to the quality of the data received from the diary survey. The fact that so many respondents completed the diary by recall complicates our analysis because they could not be asked most of the questions on the supplemental survey.

A potential problem with this type of survey is that all family members may not report their expenditures to the person keeping the diary. However, most (84%) of the CU's with more than one person indicated expenditures by members other than the record keeper were included; and slightly more than 70% of these respondents said that the expenditures were reported daily. Another 15% received reports every 2 to 3 days. Unfortunately, we do not know how complete the reports were.

In order to adequately evaluate the respondent's reports, we asked whether the expenditures for the two-week diary period were about what the CU usually spent. Forty percent of the respondents said that these two weeks did not give an accurate indication of their usual expenditures. Almost 75% of these respondents said their reported expenditures were lower than usual while 25% said they were higher. The large portion falling in the "lower" category may indicate an attempt by some respondents to misrepresent underreporting.

When we place a diary, we recommend that the respondent record expenditures on a daily basis to insure that none are forgotten. As a check on how closely respondents follow these instructions, they were asked how often expenditures were usually recorded. About 70% said they recorded expenditures immediately after they occurred or at the end of each day. Two-thirds of these respondents said they never entered an expenditure after the day it occurred. As for those respondents reporting a delay in

recording expenditures, the reasons most often given were that they forgot, they were busy, or they were out of town.

When asked how they felt about keeping the diary, a number of respondents thought that the size of the diary was awkward, that the task required too much time and effort, and/or that interest in the diary declined over the two weeks. However, somewhat larger proportions felt that keeping the diary was enjoyable and/or beneficial. A smaller group of respondents found it difficult to obtain expenditures from other members of the CU.

Only about a quarter of the sample had any suggestions for improving the diary. While there was a variety of these suggestions, the most common involved changing the diary format. For instance, several respondents felt that the weight and quantity fields should be eliminated.

We also asked the interviewers to give their opinions of the attitudes and behaviors of the respondents. Almost a third of the time, the interviewer reported the respondent was reluctant to agree to keep the diary. According to the interviewer's opinion, almost a quarter of the respondents did not make a conscientious effort. These findings appear to coincide with the fact that nearly 20% of the diaries required a recall interview. Impressions about the respondents conscientiousness usually were gained from observing the state of the diary at pickup or from statements made by the respondent.

TABLE 1. Unweighted Frequency Distribution for Selected Variables from the Supplemental Diary

<u>Method of</u>	<u>How Often Expenditures</u>
<u>Diary Completion</u> (N=1184)	<u>Recorded?</u> (N=961)
By Total Recall 18%	End of Each Day 57%
By Respondent 81	End of Week 6
No Answer 1	Every Few Days 16
	Immediately After 15
	Expenditure Occurred
	Other 6
<u>Other CU</u>	<u>Feelings About</u>
<u>Members Reporting</u> (N=661)	<u>Diary</u> (N=969)
Yes 84%	Size Awkward 20%
No 15	Too Much Effort 22
Don't Know 1	Enjoyable 29
	Beneficial 27
	Boring 11
<u>Frequency of Other</u>	<u>Lost Interest</u> 15
<u>Members Reports</u> (N=550)	<u>Hard To Get Others'</u> 8
Daily 71%	<u>Reports</u>
At End of Week 6	Not Enough Space 2
Never 7	Other 14
Other 1	
<u>Were Expenditures</u>	<u>Interviewer's Impression</u>
<u>Representative?</u> (N=1153)	<u>of Respondent</u> (N=1149)
Yes 50%	Eager (or Cooperative) 50%
No 40	Reluctant 33
Don't Know 4	Don't Know 17
<u>How Expenditures</u>	<u>Did Respondent</u>
<u>Unrepresentative?</u> (N=399)	<u>Seem Conscientious?</u> (N=1151)
Higher 27%	Yes 67%
Lower 73	No 25
	Don't Know 8

*The total N is 1184, but percentages are based on the number of respondents for each question.

THE SUPPLEMENTAL DIARY SURVEY: SUMMARY MEASURES

The Attitude and Behavior Scales

Earlier I spoke of the possible inconsistencies

between a respondent's attitudes toward the diary and his or her record-keeping behavior. As I pointed out then, behavior is what we must change in order to improve our estimates even though changes in attitudes and behavior will often coincide. I want to focus now on the relationship between attitudes and behavior found in the supplemental survey. This analysis will be potentially very useful for differentiating respondents with respect to the accuracy of their weekly expenditure reports. Complicating this effort is, again, the fact that respondents completing the diary through total recall could be asked only a few of the questions on the survey. I will deal with these respondents both separately and in conjunction with those who did record expenditures in the diary.

Because a number of items on the supplemental questionnaire were used to gather information about attitudes and behavior, a method was needed for summarizing these data so that concise statements about the relationship between attitudes and behavior could be made. Seven questions were chosen for use in forming an attitude scale. These questions are listed in Table 2 along with an explanation of how they were recoded to create the scale. The five behavioral items used to develop the behavior scale are also listed in Table 2. Both respondent and interviewer questions are used in the scales. Some inconsistencies exist between respondent and interviewer reports; but, by and large, they are in agreement. The recoded values for each set of items were added together to produce scale values. The higher the value on either scale, the more likely it is that the respondent gave an accurate expenditure report.

TABLE 2. Items Used in the Formation of the Attitude and Behavior Scales*

<u>Attitude Scale</u>	<u>Behavior Scale</u>
Respondent Q. 10	Respondent Q. 8
Yes = 1	1 or 4 = 2
No = -1	2 = -2
Don't Know or	3 = -1
Missing = 0	5, 6 or Missing = 0
Respondent Q. 11-2	Respondent Q. 11-7
Checked = -1	checked = -1
not checked = 0	not checked = 0
Respondent Q. 11-3	Respondent Q. 11-8
checked = 1	checked = 1
not checked = 0	not checked = 0
Respondent Q. 11-4	Interviewer Q. 2
checked = 1	Yes = 2
not checked = 0	No = -2
Respondent Q. 11-5	Don't Know or Missing = 0
checked = -1	Interviewer Q. 4a
not checked = 0	Total Recall = -2
Respondent Q. 11-6	Partial Recall = 0
checked = -1	No Recall = 2
not checked = 0	
Interviewer Q. 1	
Eager or Cooperative = 2	
Reluctant = -2	
Don't Know or	
Missing = 0	

* Refer to the supplemental questionnaire in the appendix for the text of the questions.

Respondents who completed the diary by total recall were assigned the most negative value on each scale. This seemed entirely appropriate in the case of the behavior scale. As for the

attitude scale, I assumed their behavior indicated extremely negative attitudes toward the diary. We will see later, when these respondents are considered separately, that the assignment of these scale values may not have always been justified.

The Attitude/Behavior Typology

To further simplify the analysis, the scales were combined to form a four-category attitude/behavior typology. Greater weight was given to the behavior scale since, ultimately, the respondents must be judged by their behavior. Small positive values on the scale were considered to be negative as a way of compensating for social desirability effects. The method used for classifying respondents is given below:

- Category 1: Behavior ≤ 2 and Attitude ≤ 1
- Category 2: Behavior ≤ 2 and Attitude > 1
- Category 3: Behavior > 2 and Attitude ≤ 1
- Category 4: Behavior > 2 and Attitude > 1

This typology offers a contrast in respondent styles. The first category is composed of respondents with poor attitudes and poor record-keeping behavior. At the other extreme, in category four, are respondents who have the attitudes and behavior we desire. The middle categories are actually the most interesting. Category two includes respondents who express positive attitudes, but their behaviors do not coincide with these attitudes. In category three are respondents who have poor attitudes but desirable behavior. To emphasize the differences between these groups of respondents, I have given them descriptive names. These names along with the percentage of the sample of respondents falling in each category are found in Table 3.

TABLE 3. The Attitude/Behavior Typology (N = 1184)

Category 1 (Resisters)*	34%
Category 2 (Misleaders)	7
Category 3 (Complainers)	21
Category 4 (Accommodaters)	38

*Over half of the respondents in this category completed the diary through recall.

The members in the first category are labeled "resisters" because their attitudes and behavior clearly show that they resisted keeping the diary. Respondents in category two gave a misleading picture of themselves. They display positive attitudes toward the diary, but their behavior indicates otherwise. Respondents in category three, on the other hand, were quite the opposite. They disliked keeping the diary although they kept it correctly. Category four is made up of what might be called "model" respondents. They seem to be extremely accommodating and often enthusiastic.

This typology was created in the belief that it would differentiate respondents in terms of how well they reported their expenditures. A simple test of this hypothesis using information contained in the supplemental survey is offered in Table 4. Remember that respondents were asked how their expenditures for the diary period compared to their usual expenditures. Clearly, the "accommodaters" were the ones most likely to report their expenditures were about the same as

usual or higher. The fact that the "complainers" said their reports were the same as usual more often than did the "resisters" and "misleaders" leads us to believe that, other than the "accommodaters," they provided the most accurate expenditure reports. Again, it is the behaviors and not the attitudes which determine the nature of the expenditure report.

TABLE 4. Attitude/Behavior Typology and the Respondent's Comparison of Reported Expenditures to Usual Expenditures

Reported vs. Usual	Resisters (N = 400)		Misleaders (N = 84)	Complainers (N = 245)	Accommodaters (N = 455)
	Total* Recall	Others			
Same	46%	45%	49%	53%	63%
Higher	1	12	12	12	9
Lower	7	32	32	31	25
DK/NA	46	11	7	4	3

* Many of these respondents were asked only if their reported expenditures were the same as usual and not the follow-up question ascertaining the direction of the difference.

Other differences in the table are less dramatic (except for the total-recall "resisters"); but if we combine the "lower" category with the "don't know/not ascertained" category, a pattern does emerge. These categories can be combined if we assume that a respondent gives a "don't know" answer or no answer at all in order to avoid the question and that the probable answer would have been "lower". When the percentages from these two categories are added, we get results which, as expected, are the opposite of those from the "same" category.

I have already stated that the assignment to category one of respondents who completed the diary through total recall may not always have been justified. As it turns out, a number of these respondents were elderly, so they may have been physically unable to keep the diary. We do not know what their attitudes were or whether they would have kept the diary properly if they could have. However, an indication of this possibility is gained by examining the opinions given by the interviewers about the attitudes and behaviors of these respondents. This information is found in Table 5. In almost twenty percent of the cases, the interviewer said the respondent had a positive attitude or made a conscientious effort to report expenditures. While these data should be viewed with some skepticism since the interviewer's perception may be in error, more than forty percent of these respondents did say their expenditure reports reflected their usual expenditure.

TABLE 5. Interviewers' Opinions of Respondents Completing the Diary Through Total Recall (N = 215).

	Opinion of Attitudes		Opinion of Behavior	
Eager or Cooperative	18%	Conscientious	17%	
Reluctant	65	Not Conscientious	67	
DK/NA	17	DK/NA	16	

The typology developed here provides us with a better understanding of our respondents, and we should gain further insights when we examine the expenditures of the four groups. The attitude/behavior typology also can be used in the comprehensive analysis of the survey process

pictured in Figure 1. The most striking finding from the present analysis is that attitudes and behavior are often inconsistent, especially in the case where negative attitudes are associated with positive behavior. Again, it is the respondent's behavior which we must change. However, in the process, we undoubtedly will affect attitudes; and this certainly will be a desirable outcome. According to the earlier description of the survey situation, the same variables influence both attitudes and behavior. We have no control over consumer unit characteristics, but we can change the environmental context and survey procedures. The next section describes one attempt to improve survey procedures.

THE DIARY OPERATIONAL TEST

The Rationale

The supplemental survey was designed to give initial information about the attitudes and record-keeping behavior of diary respondents and also provide some data on environmental characteristics. As indicated above, we can use this information in conjunction with consumer unit characteristics, other environmental characteristics, and expenditure reports to examine the survey process. However, this examination will be incomplete because we have left the survey procedures largely unconsidered. Since each respondent receives the same diary, we cannot know how different survey procedures would affect expenditure reporting. The purpose of another project we have recently undertaken, the Diary Operational Test, is the evaluation of the effectiveness of different survey procedures.

The Design

The effects of two variables will be examined in a design which uses data from both a special research sample and the regular diary sample to make comparisons. One variable measures the effect of the current practice of having the diary and the quarterly interview survey conducted by the same interviewers. Since the quarterly interview involves so much more of the interviewer's time than the diary, it is suspected that less emphasis has been placed on the quality of data gathered from the diary survey. To evaluate the extent to which this is true, two interview conditions have been used. In the regular sample, interviewers have continued to conduct both surveys. In the research sample, interviewers worked only on the diary.

The other variable concerns the physical layout of the diary itself. The diary's format can either hinder or facilitate the reporting of expenditures. Recent research into reporting rates from both the 1972/73 and 1980/81 diaries indicated that explicit references to particular products in the diary increases the likelihood that these items will be reported, especially if the reporting rates are low to begin with [1, 4]. To evaluate diary formats which provide more explicit instructions as to the commodities to be

reported, two experimental diaries were developed. Portions of these two diaries and the diary currently in use can be found in the appendix.

The experimental diaries have fewer expenditure categories than the current diary. They are also smaller and have more attractive covers. Respondents are not required to specify the quantity and weight of items they purchase when completing the two experimental diaries as they must in the current diary. What distinguishes the two experimental diaries from one another is the specificity of the item descriptions within each section. In experimental diary A, only blank lines for recording purchases are provided under each of the section headings just like the current diary; but, in contrast to the latter, the section headings contain more complete descriptions of the items to be reported. Experimental diary B has only category titles; however, the lines beneath each heading have specific items printed on them. Respondents need only check whether an item was purchased and give the price.

These two formats were chosen as the most promising alternatives for increasing the specificity of item descriptions. Diary A still gives the respondent the freedom to describe purchases, but it also requires a significant amount of writing. In the other experimental diary, much less writing is necessary; but the respondent must make classification choices which will take more thought than the blank-line condition and may be prone to error. Respondents must also add expenditures for all items appearing on the same line and record the total.

Table 6 depicts the features of a design employing the variables described above. In addition to the two experimental diaries, the current diary will be used as a control. These three diaries were administered at random to consumer units in the research sample, and the interviewers assigned to this sample did not work on the quarterly interview survey. The fourth cell represents the on-going "production" diary where interviewers conducted both surveys.

TABLE 6. Diary Operational Test Design

	<u>Format</u>	<u>Interview Condition</u>
Diary A	Blank-Line	Diary Only
Diary B	Specified-Line	Diary Only
Control	Current Diary	Diary Only
Production	Current Diary	Both Surveys

Two other features of this experiment are worth mentioning. The first is that an additional section was added to the household characteristics questionnaire which is very similar to the supplemental questionnaire discussed previously. We hope to be able to distinguish between the attitudes and record-keeping behavior of respondents exposed to the different experimental conditions. The other feature is a new method of collecting recalled expenditures. Currently, the interviewer records these expenditures directly into the diary using unscripted procedures and also asks a series of follow-up questions about specific commodities

which the respondent may have forgotten to report (See appendix.). The new procedure, used with the two experimental diaries, involves a scripted recall section contained in the household characteristics questionnaire. One part of this recall section can be found in the appendix. The new recall section was designed to improve our ability to collect recall information; and, by keeping this data separate from expenditures recorded by the respondent, we will be able to determine the extent of recall in the diary.

The Sample

The sample cases for the experiment were drawn from 22 of our largest (self-representing) PSU's, and the random assignment of diaries to members of the research sample was carried out within each PSU so that approximately a third of the respondents in every PSU would receive one of the three format conditions. We expect about 1000 consumer units per research cell. These consumer units were interviewed between May and November of 1985. The production sample should have yielded about 1300 interviews during that period in the same 22 PSU's.

The Analytical Model

The simplest statistical model which will be used to estimate the effects of the experimental conditions is the following:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk} \quad (1)$$

where Y_{ijk} is the amount of the expenditures reported by a respondent or, alternatively a measure of the response error in that report; α_i ($i = 1, 2$) represents the effect of the interviewer variable, β_j ($j = 1, 2, 3$) represents the effect of the format variable, $(\alpha\beta)_{ij}$ is the interaction term and ϵ_{ijk} is the error. Estimation of the parameters of this model is complicated by the fact that we do not have a complete factorial design. The two new format conditions did not appear in cells where interviewers worked on both surveys. Therefore, the interaction term in (1) cannot be estimated. The two treatment effects can be estimated in a roundabout manner where each is considered individually. The results from the production cell are compared to those from the control cell to determine the effect of the interviewer variable. Results from the three research cells can be used to gauge the effects of the different diary formats. Assuming no interactions, these separate estimates can be added to arrive at the total effects of conditions in each of the experimental diary cells when compared to the production cell. That is,

$$\mu_{RA} - \mu_p = (\mu_{RA} - \mu_{RC}) + (\mu_{RC} - \mu_p) \quad (2)$$

$$\mu_{RB} - \mu_p = (\mu_{RB} - \mu_{RC}) + (\mu_{RC} - \mu_p) \quad (3)$$

Data from the experiment can be used in more complicated analyses than that described in (1) in order to examine the complete survey process pictured in Figure 1. Of particular interest will be the effects of the different diary forms on the attitude/behavior typology.

APPENDIX

1. Did the respondent seem eager or reluctant about keeping the Diary?
 - 1 Eager
 - 2 Reluctant
 - 3 Don't know

- 2a. Do you think the respondent made a conscientious effort to complete the Diary every day?
 - 1 Yes
 - 2 No
 - 3 Don't know - Go to item 3a
- b. What gave you this impression?

- 3a. Did you contact the consumer unit at any time between the 1st week placement and pickup?
 - 1 Yes
 - 2 No - Go to item 4a
- b. How was the contact made?
 - 1 By personal visit - WHEN _____
 - 2 By telephone - WHEN _____
- c. Was the CU having problems completing the Diary?
 - 1 Yes - Specify _____
 - 2 No

- 4a. Was any portion of the week 2 Diary completed by recall at the time of pickup?
 - 1 Entire Diary by recall
 - 2 Part of the Diary by recall
 - 3 None completed by recall - STOP
- b. Was the respondent able to recall the day each expenditure occurred?
 - 1 Always - STOP
 - 2 Sometimes
 - 3 Never
- c. Where were entries obtained by recall recorded?
 - 1 First Diary day
 - 2 Second Diary day
 - 3 Third Diary day
 - 4 Fourth Diary day
 - 5 Fifth Diary day
 - 6 Sixth Diary day
 - 7 Seventh Diary day
 - 8 Additional page

- 5a. INTERVIEWER CHECK ITEM
 - 1 Diary completed by total recall by interviewer - Complete only items 7a and 12
 - 2 Diary completed by respondent (including partial recalls) - Complete items 5b and 6-12, as appropriate
- b. Is this a one person CU?
 - 1 Yes - Go to item 7a
 - 2 No

- 6a. Who was responsible for recording expenditures for this consumer unit (CU)? Enter line number(s) and name(s) of each person(s) mentioned.

ASA IF APPROPRIATE, OTHERWISE GO TO ITEM 7a
- b. Besides (Person(s) in 6a), were expenditures of other persons in this CU entered in the Diary?
 - 1 Yes
 - 2 No
 - 3 Don't know } Go to item 7a

- c. How frequently did other CU members report their expenditures to (Person(s) in 6a)? Read all categories and mark (X) only one.
 - 1 Daily
 - 2 Every 2 to 3 days
 - 3 At the end of the week
 - 4 Never

- 7a. Were the expenditures recorded for the past two weeks about what this CU usually spends in a 2 week period?
 - 1 Yes - Go to item 8
 - 2 No
 - 3 Don't know - Go to item 8
- b. Were the expenditures for these 2 weeks higher or lower than usual?
 - 1 Higher - PROBE Specify any specific reason _____
 - 2 Lower

8. When were entries usually recorded on the Diary?

Read all categories and mark (X) all that apply.

- 1. At the end of each day
- 2. At the end of each week } Go to item 9f
- 3. Every few days
- 4. Immediately after an expenditure occurred
- 5. Other - Specify _____

9a. Was it ever necessary to make entries in the Diary for expenditures made earlier during the week?

- 1. Yes
- 2. No
- 3. Don't know } Go to item 10

b. Which of the following reasons explain why the Diary was not completed on a daily basis? Mark (X) all that apply.

- 1. Forgot to record expenditures for one or more days!
- 2. Was too busy to record expenditures on the Diary
- 3. Misplaced the Diary
- 4. Other - Specify _____

c. Were the expenditures remembered later in the week recorded -

- 1. On the page for the day the expenditure occurred?
- 2. On the page for the day the expenditure was remembered?
- 3. On another page?
- 4. Don't know.

10. Were the instructions and examples contained on the Diary form helpful?

- 1. Yes
- 2. No
- 3. Don't know

11. Which of the following describe your feeling about keeping the Diary? Read all categories and mark (X) all that apply.

- 1. Diary size was awkward
- 2. Keeping the Diary took too much time and effort
- 3. Keeping the Diary was enjoyable
- 4. Keeping the Diary was beneficial for my CU
- 5. Keeping the Diary was boring
- 6. Interest in keeping the Diary declined
- 7. It was difficult to obtain expenditures from other CU members
- 8. Not enough lines on the Diary form to record expenditures
- 9. Other - Specify _____

10. No opinion

12. Would you like to make any comments about the Diary, or suggestions for improving the way we collect the data?

1. Yes (Exhaustive)

2. No

Example of a follow-up question currently in use:

INTERVIEWER: Are expenditures reported in Diary? (a) Yes No

• Ask questions a, b, and c if the "NO" box in column (a) is marked. We have found that certain types of expenses are often forgotten. I would like to ask a few questions concerning these items.

1a. Does anyone in the CU eat out - lunches, dinners, snacks, etc. - even occasionally? Yes No - Go to next "No" box marked in column (a)

b. Did anyone in the CU eat out at all during the past week? Yes No - Go to next "No" box marked in column (a)

c. How much was spent for meals eaten out during this past week? \$ _____ 00

Part from the new recall section:

2. Were there any meals and snacks purchased at a restaurant or carry-out which you may have forgotten to enter in the Diary?		1 <input type="checkbox"/> Yes		2 <input type="checkbox"/> No - Go to part C		3 <input type="checkbox"/> Don't know - Go to part C		
PGM 3B Line No.	PROCESSING USE	Describe item purchased	Total cost Excluding sales tax		Were alcoholic beverages included in total cost? Mark (X) one		If "Yes," How much?	
			Dollars	Cents	Yes	No	Dollars	Cents
001								
002								
003								
004								
005								

Section from Current Diary:

Item description	LINE NO.	PROCESSING USE	Describe item purchased	Number of cans, bottles, etc.	Net weight or volume per bottle, can, etc.	Is this item - Mark (X) one				Total cost Excluding sales tax	
						Fresh	Frozen	Bottled or Canned	Other	Dollars	Cents
Part 1 FOOD AND BEVERAGES	001										
Dairy and Bakery Products	002										
Pastry, brownies, bread, milk, cream, eggs, etc.	003										
	004										
	005										
	006										
	007										
	008										
Meat, Fish, and Poultry	009										
Beef brisket, trout, chicken parts, etc.	010										
	011										

Section from Experimental Diary B:

Line No.	Item purchased	Mark (X) if Item purchased	Total cost Excluding sales tax	
			Dollars	Cents
	Part 1 - Food for Home Consumption (Continued)			
	Fish and Seafood			
032	Canned fish and seafood			
034	Fresh and frozen fish			
036	Fresh and frozen shell fish			
	Eggs and Dairy Products			
036	Eggs			
037	Fresh whole milk			
038	Other fresh milk and cream			
039	Butter			
040	Cheese			
041	Ice cream, frozen yogurt, and related products			
042	Non dairy cream substitutes			
043	Margarine			
044	Other dairy products including powdered milk and fresh yogurt			

Section from Experimental Diary A:

Line No.	PROCESSING USE	Describe item purchased	Total cost Excluding sales tax		Is this item - Mark (X) one			
			Dollars	Cents	Fresh	Frozen	Bottle or can	Other
		Part 1 - Food for Home Consumption (Continued)						
		Eggs and Dairy Products (Eggs, milk, cream, butter, margarine, cheese, ice cream, etc.)						
040								
041								
042								
043								
044								
		Fish and Seafood (Canned, fresh or frozen - fish or seafood, etc.)						
035								
036								
037								
038								
039								

REFERENCES

- Jacobs, C.A. "Diary Reporting Levels-1972/73 and 1980/81." BLS memorandum to E.Jacobs, December 16, 1983.
- Pearl, R.B. "Revaluation of the 1972-73 U.S. Consumer Expenditure Survey." Tech. Paper No. 46, U.S. Department of Commerce, Bureau of the Census, July, 1979.
- Sudman, S. and Ferber, R. "Experiments in Obtaining Consumer Expenditures by Diary Methods." Journal of the American Statistical Association 66: 725-735.

- Tucker, C. "Decision-Making in Diary Research Planned for 1985." BLS memorandum to C.A.Jacobs, March 20, 1984.

- _____. "Identifying the Measurement Error in a Consumer Unit's Report of Expenditures." Paper presented at the annual meeting of the American Statistical Association, Las Vegas, August, 1985.

- U.S. Department of Labor. Consumer Expenditure Survey: Diary Survey, 1980-81. Bull. 2173, Bureau of Labor Statistics, September, 1983.

ABSTRACT

Data from the 1980-81 Consumer Expenditure Survey were used to determine whether wife's education has an influence on family food expenditures once the effects of income and other factors are taken into account. The double-log function was used in the multiple regression analysis of 22 food expenditure categories at four different levels of aggregation. Income and education elasticities were also compared. All income elasticities were positive and all but three were statistically significant. Ten of the 22 education elasticities were statistically significant; eight of these coefficients were positive while two were negative.

Considerable attention has been focused on the factors influencing food consumption patterns. Although economists have often emphasized consumption-income relationships, the influence of other sociodemographic variables such as age, race, family size, employment status, urbanization and region have also been studied [2, 4, 9, 10]. Less consideration has been given to the relationship between wife's education and food expenditures. One exception is research by Abdel-Ghany and Schrimper [1] which used data from the 1965-66 USDA Food Consumption Survey. Even after controlling for the effects of income and other factors, wife's education was found to have a significant positive effect on 10 of the 22 expenditure categories studied. Additional research, however, is needed in order to ascertain whether the increased levels of education attained by women in recent years has changed the composition of family food expenditures.

The objectives of this study are: (1) to examine, after taking into account other factors determining family consumption behavior, the influence of wife's education on family food expenditures, and (2) to compare education and income elasticities for selected food expenditure categories. Findings from this research could assist in understanding possible changes in food consumption patterns resulting from future changes in the educational level of the general population.

RATIONALE

Despite the increase in the number of dual-earner families, wives still perform the majority of work in the household. Since food shopping, preparation and cleanup account for the greatest proportion of household production, it may be safe to assume that in most families the wife has the

¹ Extension Consumer and Family Economics Specialist

² This research was partially supported by a grant from the Auburn University Research Grant-in-Aid Program.

major influence on food consumption patterns.

One reason why a wife's education should influence food consumption patterns is that increased education may enhance her household production skills. The rationale is provided by Michael's neutrality model [8] which assumes that education increases efficiency in all nonmarket activities and that the influence of education is neutral (has the same effect on each activity). This greater efficiency increases a household's "real" income at any given level of money income, thus increasing a household's ability to purchase goods and services. Michael hypothesized that the increase in real income would be associated with proportionately higher outlays on commodities, such as food away from home and other services which are considered luxuries, but proportionately lower outlays on necessities, such as food at home.

To empirically test his model Michael [8] analyzed the spending patterns of all households participating in the 1960-61 Consumer Expenditure Survey (CES). Education of the household head was used to measure education. Food for home use and food away from home were the only dependent variables measuring food spending. Findings for these expenditure categories were consistent with Michael's model, since a negative education elasticity was found for food for home use, while a positive education elasticity was found for food away from home.

Other reasons have also been used to explain why a wife's education might influence family food consumption. One is that knowledge of nutrition and concern for health may vary with a wife's level of education, resulting in differing food consumption patterns. Another is that tastes or preferences associated with different levels of education could affect consumption patterns.

METHODOLOGY

The diary component of the 1980-81 Consumer Expenditure Survey (CES)³ was the data base for this research. The specific data analyzed were from the 1981 survey year. The study sample consisted of urban husband-wife families with pretax incomes of \$75,000 or less. Families not reporting information on income, wife's education and other sociodemographic characteristics of interest were deleted, resulting in a study sample of 4,299 families.

Multiple regression analysis was used to investigate the influence of wife's education and family income on food expenditures at four different levels of aggregation. The first level was expenditure on total food while the second consisted

³ For a description of CES methodology see [11].

of spending on food at home and food away from home. Expenditures were also broken down into eight major food groups and 11 individual food products. These major food groups and individual food products will be discussed in more detail in the findings section.

Wife's education and family income were treated as continuous variables. Wife's education was the number of years of schooling completed, while family income was the total of 1980 pretax income of all family members 14 years old or over. Other independent variables included as control variables were: (1) wife's age, (2) wife's race, (3) number of hours worked by the wife in 1980, (4) family size, (5) region and (6) SMSA status. The double-log function was used in the multiple regression analysis of the 22 food expenditure categories.

Finding the most appropriate form of Engel curve has always been a problem in economics. The choice of consumption-income relation is influenced by both economic and statistical considerations. Ferber [3] found that the functional form used in the majority of studies was the double-log, essentially the same as that used by Ernst Engel. The main reason for the popularity of the double-log form is that it results in a constant income elasticity. It passes through the origin and has an upward rather than downward curvature when the elasticity is greater than one.⁴

FINDINGS

Income and education elasticities for the 22 expenditure categories analyzed are shown in Table 1. All income elasticities were positive and all, except those for poultry, eggs and fresh milk and cream, were statistically different from zero. As may be seen, income elasticity estimates indicate that among sample families demand for the food products studied was inelastic. In other words, a given percentage change in income led to a less than proportional change in expenditure, which is consistent with Engel's law.

Only 10 of the 22 education elasticities were statistically significant; eight of these coefficients were positive, while two were negative. Although wife's education had no influence on spending on food at home it was positively related to outlay on food away from home which accounted for 30 percent of total food spending.

Findings for the food groups show that wife's education had no influence on spending on grain products, sugars and sweets and fats and oils, but was positively related to spending on fruits. These findings are consistent with those of Abdel-Ghany and Schrimper [1]. Expenditure on dairy products and vegetables was not influenced by the level of wife's education. These findings are contrary to those of Abdel-Ghany and Schrimper who found that wife's education had a positive influence on these outlays. It is uncertain whether these differences are due to changes in the influence of wife's

⁴For additional information see [5-7].

education between 1965 and 1966 and 1981 or differences in the research methodology used in each study.

Spending on meat, poultry, fish and eggs was not related to wife's education. Miscellaneous foods, which included frozen, canned and packaged meals, meats and other products, was positively influenced by wife's education. Since items in this category contain a preponderance of convenience foods, this finding may reflect a preference for built-in marketing services among more educated wives. For six of the 11 individual food products analyzed, education elasticities were found to be statistically different from zero. The positive elasticity for bakery products also indicates a preference among the more educated for commodities with built-in services. Wife's education had no effect on spending on beef, pork or poultry but was positively related to outlay on fish. As may be seen, wife's education was associated with lower expenditure on eggs and fresh milk and cream. These findings suggest that more highly educated wives may be attempting to lower the serum cholesterol levels of family members, thus reducing the chances of cardiovascular disease.

Wife's education increased outlay on both fresh and processed fruits but had no influence on either fresh or processed vegetables. The positive elasticity for fresh fruits could be interpreted as a concern for nutrition among more highly educated wives, but this argument is weakened by the fact that spending on fresh vegetables was not affected by education. The positive education elasticity for processed fruits, often considered less nutritious than fresh, further weakens this argument. Although the finding for processed fruits could reflect a preference for built-in service, the fact that this desire was not reflected in increased outlay on processed vegetables is curious. These findings could be attributed to a preference for fruit in all forms among families of more highly educated women.

CONCLUSIONS

Study findings show that wife's education had a significant influence on family food consumption even after the effects of income and other factors were taken into account. The reasons for these findings, however, are not clear-cut. The positive elasticities for food away from home, miscellaneous foods and bakery products could be explained by Michael's hypothesis in which, other factors being equal, increased education is expected to shift spending to services.

A concern for health among more highly educated wives could explain the positive elasticity for fish and negative elasticities for eggs and fresh milk and cream. The fact that wife's education was positively associated with outlay on fresh and processed fruits, but neither fresh nor processed vegetables, is more difficult to interpret on the basis of the previous two explanations. The most plausible reason appears to be a preference among the more highly educated for fruit in any form. It should be noted that the nature of the Consumer

TABLE 1. Income and Education Elasticities of Food Expenditure Categories.

Categories	Elasticities		
	Income	Education	\bar{X} Weekly Expenditure
Total Food	.25***	.19***	69.70
Food at Home	.15***	.06	48.82
Food Away from Home	.51***	.55***	20.88
Food Groups			
Grain Products	.09***	.05	6.32
Meat, Poultry, Fish, Eggs	.13***	-.01	14.12
Dairy Products	.10***	.06	6.57
Fruits	.10***	.27***	3.87
Vegetables	.12***	-.04	4.60
Sugars and Sweets	.06***	.08	1.86
Fats and Oils	.05***	.03	1.42
Miscellaneous Foods	.10***	.15**	4.76
Individual Products			
Bakery Products	.10***	.10*	4.25
Beef	.08**	.02	6.42
Pork	.08***	-.01	3.30
Poultry	.03	-.03	2.06
Fish	.10***	.10**	1.44
Eggs	.01	-.11***	.90
Fresh Milk and Cream	.01	-.15***	3.40
Fresh Fruits	.08***	.14***	2.17
Fresh Vegetables	.11***	-.07	2.24
Processed Fruits	.06***	.21***	1.70
Processed Vegetables	.07***	.04	1.36

*Significant at the .05 level.

**Significant at the .01 level.

***Significant at the .001 level.

Expenditure Survey data does not allow determination of how close to reality these explanations are.

Study findings differ somewhat from those of Abdel-Ghany and Schrimper [1]. It is uncertain whether this reflects changes in the relationship between wife's education and food consumption that have occurred in the 15 years separating the two studies or whether this is the result of methodological differences.

Consumer units participating in the 1980-81 CES Diary Survey were asked to record expenses for two consecutive one-week periods. Because of the random timing of many consumer purchases and the relatively short survey period, it is possible that the results may not have been typical of the long-run spending patterns of the families

studied. While this problem is more acute for such infrequently purchased items as durable goods, it is possible that a consumer unit reporting no expenditure on a food item may have been depleting its stock of the item during the survey period.

At the present time there is no "ideal" data base which collects such weekly information over a longer time period, such as a year. For this reason, using data from the 1982-83 CES to clarify the influence of wife's education and other factors on family food consumption should not be precluded.

REFERENCES

1. Abdel-Ghany, Mohamed and Schrimper, Ronald A. "Food Consumption Expenditures and Education of the Homemaker," Home Economics Research Journal, 6, June 1978, pp. 283-292.
2. Crockett, Jean. "Demand Relationships for Food," Consumption and Saving, Volume 1. Edited by J. Crockett and R. Jones. Philadelphia: University of Pennsylvania, 1960, pp. 293-310.
3. Ferber, Robert. "Research on Household Behavior," Surveys of Economic Theory, Volume 3. New York: St. Martin's Press, 1966.
4. Foster, Ann C. "Wife's Employment and Food Expenditures," Conference Proceedings, Southeastern Regional Association of Family Economics - Home Management, 14th Annual Conference, February 14-16, 1985, pp. 89-95.
5. Hu, Teh-wei. Econometrics: An Introductory Analysis. Baltimore: University Park Press, 1973.
6. Madala, G. S. Econometrics. New York: McGraw-Hill, 1977.
7. Mayes, David G. Applications of Econometrics. Englewood Cliffs, New Jersey: Prentice-Hall International, 1981.
8. Michael, Robert T. The Effect of Education on Efficiency in Consumption, National Bureau of Economic Research Occasional Paper No. 116. New York: Columbia University Press, 1972.
9. Salathe, Larry E. Household Expenditure Patterns in the United States, U.S. Department of Agriculture, Economics Statistics and Cooperative Service, Technical Bulletin No. 1603. Washington, D.C.: U.S. Government Printing Office, April 1979.
10. Salathe, Larry E. and Buse, Reuben C. Food Consumption Patterns in the United States, U.S. Department of Agriculture, Economics Statistics and Cooperative Service, Technical Bulletin No. 1587. Washington, D.C.: U.S. Government Printing Office, January 1979.
11. U.S. Department of Labor, Bureau of Labor Statistics. Consumer Expenditure Survey: Diary Survey, 1980-81, Bulletin 2173. Washington, D.C.: U.S. Government Printing Office, September 1983.

Sherman Hanna, Kansas State University*
 Lanita Carter, Kansas State University**

ABSTRACT

Regression models of food away from home spending on demographic variables were tested with U.S. BLS Consumer Expenditure data for married couple households with husbands employed fulltime. Use of interaction variables demonstrated patterns not apparent in a regression without interaction variables. The number of weeks worked per year by the wife had a significant effect on food away from home spending for white and other nonblack households, but not for black households.

INTRODUCTION

An important trend in food consumption during the past 30 years has been the increase in consumption of food away from home (FAFH). U.S. Department of Commerce [22] data on aggregate expenditures on food away from home show an increase in food away from home as a percentage of total food expenditures between 1955 and 1985. When both are measured in current dollars, FAFH increased from 18.8 to 29.6 percent of total food expenditures. When both are measured in constant dollars, the increase was only from 22.3 to 28.3 percent. Perhaps the major factor contributing to the change in the consumption of food away from home is the effect of demographic shifts that have taken place in recent years. The demographic shifts that have contributed to the increase in the consumption of FAFH have been changes in household composition and size, and the increase in the labor force participation of women [12].

The labor force participation rate of married women with spouses present increased from 28.5 percent in 1955 [23] to 53.8 percent in 1985 [24]. The increase during that time period was steady, with an increase in the rate every year. Though these increasing rates are often cited as one of the major factors contributing to the increase in the consumption of FAFH, many cross-sectional studies have not found wife's employment status to be positively related to expenditures for FAFH [6;11;17]. The objective of this paper is to evaluate the usefulness of using interaction terms from a set of demographic and socioeconomic variables in order to better understand patterns of food away from home spending.

FACTORS RELATED TO EATING OUT

Previous Research

Household Composition. Household size decreased from 3.6 in 1955 to 2.7 in 1983 [21]. Pro-

*Professor of Consumer and Family Economics

**Ph.D. student, Consumer and Family Economics

chaska and Schrimper [15] found that the presence of preschool children in the household tended to decrease the number of meals eaten away from home. Family size may have more impact on food at home expenditures than on FAFH expenditures since a larger household may experience a greater income constraint and may only have one adult in the labor force [19].

Wife's Labor Force Participation. Another major factor influencing the increase in the demand for FAFH is the increase in the number of two-earner households [18]. As more and more women enter the labor force, production activities within the home are changing. The major shift in household production may occur in meal preparation, so that households with an employed wife may eat out more often [12]. A result of the increase in the number of two-earner households is not only a changing lifestyle, but also a higher income [12].

Foster, Abdel-Ghany, and Ferguson [7] found that wife's employment did not have a significant influence on expenditures. Foster [6] found that working wife households spent less on FAFH than non-working wife households. A study of husband-wife two child families in 11 states [8] found that the wife's employment status did not have an independent effect on expenditures for meals away from home. These findings are in contrast to those of Lazear and Michael [13] who reported that two-earner households spent 55 percent more than one-earner households on FAFH. In a study of Wisconsin husband-wife families with two children, Ortiz, et al. [14] found that fulltime employment of the wife was significantly related to higher proportions of meals purchased away from home. Bellante and Foster [3] found that the wife's weeks worked and fulltime employment status of the wife were positively related to FAFH.

Income. Most cross-sectional studies cite family income as having the greatest impact on the consumption of food away from home. Kinsey [11] concluded that since income elasticities for FAFH were often less than 1.0, FAFH was not a luxury good for all households. Salathe [16] estimated the income elasticity for FAFH of .843 and a household size elasticity of -.057 indicating that larger households spend less on FAFH than small families at the same income level.

Economic Models

Traditional Demand Theory. The simplest economic model for eating out expenditures is considering eating out as a normal good, possibly as a luxury. The economic definition of a luxury is a good with an income elasticity greater than one. If all types of households share a strong preference for eating out versus preparing food at home, income