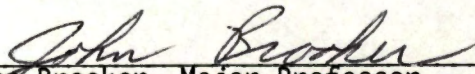
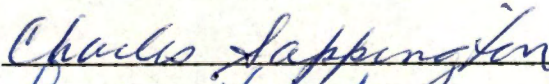
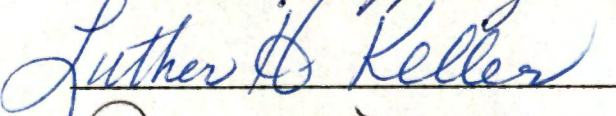
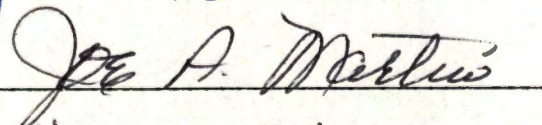
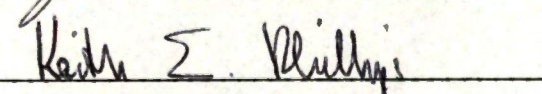


To the Graduate Council:

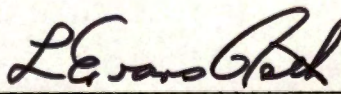
I am submitting herewith a dissertation written by Roger A. Hinson entitled "Household Budget Allocation: The Interrelationship Between Fixed Expenses and the Purchase of Food." I recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Agricultural Economics.


John Brooker, Major Professor

We have read this dissertation
and recommend its acceptance:

Accepted for the Council:


Vice Chancellor
Graduate Studies and Research

HOUSEHOLD BUDGET ALLOCATION: THE INTERRELATIONSHIP BETWEEN FIXED
EXPENSES AND THE PURCHASE OF FOOD

A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Roger A. Hinson

June 1980

3041700

ACKNOWLEDGMENTS

I would like to thank all the people who have made my graduate work possible. For financial assistance I am indebted to both the Tennessee Agricultural Extension Service and The Department of Agricultural Economics and Rural Sociology. For constructive and helpful suggestions on this research effort, committee members Dr. Joe Martin, Dr. Charles Sappington, Dr. Luther Keller, and Dr. Keith Phillips have my appreciation. Dr. James Snell was responsible for my interest in the area of demand for food, and he provided valuable guidance in formulating the project. Dr. Clark Garland has been a source of information and inspiration on matters of all sorts. Morgan Gray helped immeasurably with computer programming. The secretarial and statistical staff have helped more than I could say, especially Pat, Gail, Melitta, and Marcia. My cohorts, the other graduate students present and past, have helped make the stay here an experience. I have learned much from them.

Finally, Dr. John Brooker has served as my major professor. To him, I owe many thanks for the time and effort he devoted to supervision, criticism, and support. May he prosper!

ABSTRACT

This project was designed to evaluate the hypothesis that the consumption function for food changes at alternative levels of income due to the effect of a group of expenses over which the consumer unit (CU) has little control in the short run. These "fixed" expenses include shelter, utilities, debt repayment in each time period, and other expenses which must be met. The analysis assumes that food expenditure may be influenced by the fixed expenses, since the latter have a first claim on income. Total food expenditure, on the other hand, may be varied, through quality and possibly even quantity adjustments, to permit additional income to be allocated to fixed.

The changes in the marginal utilities were expected to occur at \$8,000 and again at \$16,000. Below \$8,000, food and basic necessities are emphasized. From \$8,000 to \$16,000, the CU may incur proportionately more fixed expenses than the additional income should precipitate. Above that level, a more equitable relationship between food and fixed may be revealed, more similar to the case of the low income group.

The fixed expenses were expected to be responsive to the take-home pay of the CU, while food was responsive to take-home pay less fixed expenses, referred to hereafter as residual income. The objectives were to compare alternative forms of the Engel curve using residual income and food expenditures, to evaluate the effects of income and a set of socio-economic variables between income levels, to compare distributional effects on different food categories between income levels, and to estimate elasticities with respect to both food and fixed expenses.

Data were obtained from two sources. The aggregate relationship between food and fixed expenses was estimated from the 1972-73 Bureau of Labor Statistics survey. Distributional effects were examined from data collected in a 1978 survey in Knox County, Tennessee.

Preliminary investigations revealed that only two age groups could be identified with respect to food expenditures, that heteroscedasticity was a problem, that some groups of dummy variable slope coefficients could be eliminated, and that specified subgroups did not seem to be different in terms of vectors of estimated coefficients.

Engel curves were estimated with OLS. Elasticity estimates were somewhat higher than most earlier estimates, an occurrence that was anticipated because residual income was used in the calculation.

The expense categories fixed and food were estimated by a system of simultaneous equations. Two stage least squares was employed to obtain consistent estimates.

The basic assertion of the research was that fixed and food are interrelated, and that the relationship changes between income levels as specified above. There appeared to be no evidence to support the assertion that changes in food induced changes in fixed expenses that were consistent with this hypothesis. On the other hand, fixed caused no change in food at low incomes, and had a negative effect on food at moderate and high incomes. These effects are in line with expectations at low and moderate incomes. Other variables included in the model were consistent with the idea that various factors are more important at some income levels than at others. Consumer unit size influenced food expenditure strongly.

Elasticities from this BLS data indicated that, with respect to income, fixed expense was elastic and inferior at low income, and elastic and superior at moderate incomes. The other values were inelastic, and some were normal and some inferior. Food was positive and inelastic with respect to residual income, and inelastic and inferior with respect to fixed expenses.

The distributional effects among meats, produce, dairy, grocery, and meals out were estimated. With respect to residual income in the lower incomes, meats decreased and meals out increased. The response to changes in fixed expense was negative, except for meats. A variable designed to reflect different stages in the life cycle seemed to have some explanatory power. In general, there were few dummy variables significantly different from the base groups.

The elasticity estimates were generally less than 1. With respect to residual income, there seemed to be quite wide fluctuations between income groups, and there were indications that some variables were inferior at particular income levels.

From these results, policy implications may be drawn with respect to income assistance designed to accomplish nutritional goals. It may be the case that cash assistance would not be used to purchase food, but would instead be diverted to consumption of items which contribute to the fixed expenses.

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION	1
II. REVIEW OF LITERATURE	7
III. APPLICATION OF ECONOMIC THEORY	20
IV. THE REGRESSION MODEL	29
V. DATA SOURCES AND PROCEDURES	37
VI. MODEL SPECIFICATION.	57
VII. RESULTS.	73
VIII. CONCLUSION	123
REFERENCES	132
APPENDIX	138
VITA	142

LIST OF TABLES

TABLE		PAGE
1.	Error Sums of Squares From High and Low Income Groups and Number of Observations	31
2.	Total Number of Census Tracts and Number Selected, by Income Group	38
3.	Increased Dollar Expenditure on Food Per Additional Subgroup Member for Eleven Age-Sex Subgroups, and Ratio of Each to the Adult Male (20-54) Subgroup	50
4.	Test of Differences Between Regression Coefficients for Selected Age-Sex Subgroups	51
5.	Increased Dollar Expenditure on Food Per Additional Subgroup Member for Two Age-Sex Subgroups, Ratio of Each to Adult Male (20-54) Subgroup, and Test of Difference Between the Two Subgroups	53
6.	Elasticity of Food Expenditure with Respect to the Natural Log of RESINC, Estimated From Semilog Engel Function, by Income Groups	76
7.	Estimated Coefficients, t-Ratios, and R^2 for Linear Engel Curves by Income Group.	81
8.	Elasticity of Food Expenditure with Respect to RESINC by Income Groups, Estimated From Linear Engel Function	82
9.	Mean Values for Food and Fixed Expense by Income Group, and Percent Change Between Groups in Each Expense Category.	84
10.	Parameter Estimates and t-Ratios for Explanatory Variables in the Fixed Expenditure Equation, BLS Model, by Income Groups	87
11.	Parameter Estimates and t-Ratios for Explanatory Variables in the Food Expense Equation, BLS Model, by Income Groups.	94
12.	Elasticities ^a for Fixed Expenditure and Food Expense with Respect to Selected Variables, by Income Groups	100

TABLE

PAGE

13.	Parameter Estimates and t-Ratios for Explanatory Variables in the Fixed Expenditure Equation, Knox Model, by Income Groups	104
14.	Parameter Estimates and t-Ratios for Explanatory Variables in the Meats Expense Equation, by Income Groups	106
15.	Parameter Estimates and t-Ratios for Explanatory Variables in the Dairy Equation, by Income Groups	109
16.	Parameter Estimates and t-Ratios for Explanatory Variables in the Produce Equation, by Income Groups	110
17.	Parameter Estimates and t-Ratios for Explanatory Variables for the Grocery Equation, by Income Group	112
18.	Parameter Estimates and t-Ratios for Explanatory Variables in the Meals out Equation, by Income Group	114
19.	Elasticities for Fixed Expenditure and Food Expense with Respect to Selected Variables, Knox Model, by Income Group	118
A-1.	Data Means, Number of Observations and Standard Deviations for Selected Variables by Income Group.	139
A-2.	Data Means, Number of Observations and Standard Deviations for Selected Variables by Income Group, From Knox Data	141

CHAPTER I

INTRODUCTION

Each year, the government of the United States must make decisions concerning allocation of the revenue it generates in the form of taxes and loans. Allocations are based on goals established by those people in policy-making positions, and one of these goals is to try to enable each person in the country to have a nutritionally adequate diet. For the most part, people acquire purchasing power, go to the marketplace, and purchase nutrients. Their purchasing power is in the form of cash and/or grants from the government, such as Food Stamps, which are designed to increase food consumption by lower income families. If the consumer's budget is viewed as a choice between food and nonfood items, and if an increasing share of available resources (the total budget) is allocated to nonfood items, then the share remaining for food purchases must necessarily decline. Since food expenditure is positively correlated with quantity consumed, the consumer's level of nutrition may be less adequate if the food budget is reduced (1, pp. 31-32).

The allocation of consumer income to various consumption categories has been studied extensively for both private commercial use (40, p. 116) and for government policy purposes (40, pp. 116-139). In the former case, information is generated to assist in the formulation of marketing strategies. In the latter, accurate information is needed to try to specify actions which give Pareto-better programs (programs in which the net utility change is positive) in cases in which government policy redistributes income. The way consumers allocate income to food

and to other types of goods and services is important. It has relevance to farm policy because greater demand for agricultural products tends to increase prices and to reduce the need for government price supports. The categories of goods and services purchased by low income consumers would likely have impacts on the effectiveness of government programs designed to provide them assistance. And, other government activities might be affected as well, an example of which might be income tax withholding policies. In short, information about consumer actions with respect to income and budget changes should be of considerable use to individuals and agencies concerned with government programs regarding food consumption.

A. PROBLEM

Because of the humanitarian and economic goals of government policy and the policy implications of the division of the consumer budget between food and nonfood categories, economists have been interested in trying to understand the decision-making process used by consumers and to identify and evaluate the important explanatory variables in the consumption function for any good. The paradigm of neoclassical economic theory has evolved a concept of demand which assumes that a few variables explain most of the shifts in the function. These variables are income, tastes and preferences, number of consumers, price of the good, and the price of related goods.

These factors underly the consumer's preference structure. Some ideas about the shape of that structure are given by estimated income elasticities. Researchers have evaluated these estimates, in terms of mean quantities and expenditures, for complete sets of data and for data

sets subdivided according to income levels. The former case is rather limited since estimates are collected at only one point. The latter gives information about the elasticities at various points on the curve.

Results when data have been subdivided by income level have resulted in conflicting estimates of income elasticities for food. It has been presumed that these elasticities, although relatively low, are positive. However, through some ranges of estimated functions for both the food category "meats" and for food as an aggregate, negative income elasticities have appeared. Tests using alternative data sets have not rejected the results. Because of the inherent nature of goods and changing preferences at different income levels, there may be income ranges at which some goods will not be purchased, or that purchases will either increase or decline. Thus, there is no theoretical obstacle to the phenomenon of different elasticities for food at different levels of income, and even for negative elasticities in some ranges.

Now, if income elasticities for food are negative in some specified income levels, it must be because consumers feel they gain more utility from consumption of items other than food when income increases through these ranges. It seems reasonable that consumers might have reached a point at which some minimal level of necessities (food, shelter, clothing, etc.) has been attained and it becomes preferable to consume more of, or to begin to consume, other goods and services. Those items which the consumer has just begun to consume may be items which require a larger proportion of the family budget, and it might also be the case that such purchases require the use of credit. The items purchased may be either durable goods or some other type of good

which requires a large budget share. The effect of the purchase on the consumer's budget, and on food purchases in particular, is the important consideration.

The size of this group of "fixed" expenses is influenced when the household decides on a level of living, or on some level of quality of goods and services it wants to attain. Some of the items may be selected on the basis of the permanent income of the household, and the household may be forced to use credit if current consumption is desired. This credit arrangement usually implies that some regular payment per period of time must be made.

In addition, there may be other parts of the budget which are regularly occurring events in the situation faced by the consumer, and these may be "fixed" in the short run. Expenses such as rent or mortgage payments, utilities, and even regular medical bills may be included in the group of expenses presumed to have an effect on food expenditures.

The concept of permanent income, which is assumed to be the basis for household purchases, is not directly observable. An attainable measure which may have an effect on fixed expenses is gross income less all payroll deductions. The logical measure on which food and other expenses besides fixed may be based is the income remaining after both payroll deductions and fixed expenses are paid, or the "residual" income. This seems reasonable, since these expenses are made more frequently and are smaller parts of the budget, and can be varied if the residual income is reduced by an increase in the fixed expenses.

B. OBJECTIVES

The general purpose of this study is to examine the relationship between household expenditures for "food" and a set of expenses over which the household has little control, referred to as "fixed" expense items, at selected income levels. The basic contention is that, with respect to these three categories of expenses, there are three distinct income groups. In the first income group, food and other basic necessities predominate. In the second, the household will commit a larger proportion of available capital to "fixed" expenses (those which must be paid on a regular basis), and proportionately less will be available for "food." In the third income group, another turning point is reached where food expenditures again exhibit the characteristics of a normal good. The specific objectives are:

1. To estimate alternative forms of the "residual" income-food consumption relationship.
2. To estimate the effects of income and selected socio-economic variables on the interrelated variables fixed expenses and food expenditures, at three income levels.
3. To estimate the effect of income on the distribution of the food budget among specific food categories.
4. To estimate aggregate and individual elasticities for food expenditures and fixed expenses with respect to income and other selected variables.

The method of least squares regression, extended through the dummy variable technique to deal with the effects of qualitative variables, is the tool used to evaluate the research objectives. In the

analysis of various forms of the Engel curve in objective 1, ordinary least squares will be used. The hypothesized relationship between the consumer unit's fixed expenses and its food purchases (objective 2) will utilize a two equation simultaneous system, and will employ the two stage least squares (2SLS) variant to handle the expected correlation between the endogeneous variables and the error terms. The third objective, the distributional effects among food categories, will also utilize the 2SLS procedure and a larger, six equation system of equations. And finally, the elasticities will be calculated from the estimated regression coefficients and the appropriate variable values.

CHAPTER II

REVIEW OF LITERATURE

Research hypotheses normally evolve through time and through reflection on past research. Ideas are developed and tested, and their strengths and weaknesses evaluated. New hypotheses may be formed in the process and they may also be tested. Thus, one of the critical points in the process of research is the review of literature in the problem area to benefit from earlier efforts. Failure in this area may lead to repeated mistakes and inefficiencies.

This review of literature will be systematized by looking at each of the factors which have been assumed to influence the consumption function. The factors are (1) income, (2) consumer unit size, (3) tastes and preferences, and (4) other influences.

According to Stigler (52), early economists had established themselves and their work by the 1870's. Engel, on the basis of his studies of family budgets, published a paper in 1857 which contained the result that the proportion of expenditure devoted to food decreases as the standard of living of the household rises. Stigler also commented on the evolution of the theory of demand, in particular the influence of factors such as income, tastes, organization of sellers, and prices of substitutes and complements on the quantity of a particular good or service sold.

A. The Relevance of Income

An important issue in the literature on consumption has been the exact definition of the term income. Several theories were spawned,

more or less sequentially. Keynes (31, p.90) implied an absolute income hypothesis (AIH) -- that consumers react to the actual level of income they have at any one time. On theoretical grounds, the AIH was rejected because it considered only current income and because it was contradicted by proportionality, the proposition that the consumption function should go through the origin (32).

The relative income hypothesis (RIH) was developed, and it incorporated the idea that consumption is influenced by peer groups. It met with objections similar to those associated with the AIH, and with the objection that sociological variables were being used to explain economic phenomena. In response to these deficiencies, Friedman developed the permanent income hypothesis (PIH), which follows the "currently accepted pure theory of consumption behavior, seems consistent with existing empirical evidence, and has implications capable of being contradicted by additional evidence" (19, p. 6). The PIH defines total income and consumption as composed of two parts, the permanent and transitory components, as follows:

$$Y = Y_p + Y_t, \text{ and } C = C_p + C_t$$

where p is the permanent component and t is a transitory component.

Each of the types of income has a separate effect on consumption. Permanent consumption, C_p , is based on permanent income, y_p , on the basis of the formula

$$C_p = k(i, w, u)y_p,$$

where k depends on the interest rate i , wealth w , and a tastes or disturbance term u . Transitory income affects permanent consumption only

to the extent that part of it becomes viewed as permanent income. Friedman says that "the transitory components of income show up primarily in changes in the consumer unit's assets and liabilities, that is, in his measured savings" (19, p. 220).

Because permanent income is unobservable (19, p. 222) studies have generally used a measure of current income adjusted for certain types of taxes. From the theory, one expects that the income variable is the basic explanatory variable, but attempts to verify this have resulted in low explanatory power in terms of measured R^2 (a common result with cross-sectional data). It has been argued (44, pp. 80-81) that a possible alternative would be to use total expenditure in any given period as a better estimate of permanent income than measured total income in any one period, because the CU has more control over expenditure than over income. The sensitivity of expenditures to purchases of durables in particular time periods is a drawback of that method.

With Friedman's suggestion that the transitory portion of income may appear as measured savings, Holbrook and Stafford (25, pp. 15-20) evaluated five sources of income with regard to the marginal propensity to consume each source (MPC). Their result was a lower MPC for some types of income, particularly transfer income. These differences are well documented.¹ In addition, Friedman's ideas of the uses to which transitory income may be put (consumption of durables, savings) are not

¹Other authors have evaluated various types of income with respect to some idea of their level of permanence. For examples, see West and Price, or Benus Kmenta and Shapiro. These are illustrative examples, not exhaustive in any sense.

contradicted by the implication that transitory income is not spent for durables, because of the definition of consumption to include the flow of services from durables but not their purchases. This suggestion led to interest in the role played by expenditures on durables in the CU budget. The expanding use of credit as a purchase instrument and the influence of available credit on durable goods expenditure also played a role in the increased interest.

Fisher (16) was concerned with consumers' use of credit. She concentrated on durable goods expenditure, the use of cash or credit to finance purchases, and used multiple regression procedures to explain durable goods expenditure. The procedure was to try to evaluate factors enabling purchase, such as (1) purchasing and nonpurchasing households, (2) size of outlays, and (3) the use of cash or credit, as well as the motivating (psychological) factors and needs which influence purchases. These latter factors correspond in a sense to the biogenic and psychogenic needs discussed by Currie (11, p. 21).

Fisher's study used the month as the time frame. She felt that figure was probably more accurately reported, since repayment schedules are usually referenced in months.

In Fisher's formulation, durable goods expenditure was defined as monthly mortgage payments and installment payments. Current disposable income was used as the explanatory variable; other specifications looked at liquid resources and at an aggregation to all assets. Results were termed good in the first two cases, but rather less good than more in the latter case.

More recently, Hamilton (21) used least squares to estimate consumption of beef (fed and nonfed). The specification used a method of

classifying beef into two qualities and included an index of durable goods' prices as an explanatory variable. One premise used was "that the short-run elasticity of demand for durable goods is elastic." The conclusion was that "when durable goods' prices rise total expenditures for durable goods decline, leaving a larger amount of income to be spent on nondurable items. . ." This result seems to support the hypothesis of this study, that food expenditures may be influenced by the proportion of income devoted to fixed expenses. Durables are included in "fixed" expense if purchase requires the use of credit.

Hamilton also supports Williams, who evaluated the effect of income, household size, and a series of socio-economic variables on both quantity of and expenditures on total meat consumption and on five meat categories. He found net income (gross income minus taxes) to be a significant explanatory variable only for the higher income groups for total meat for both the quantity and expenditure models. The R^2 was low, generally less than 0.10, except for the highest income group. The interesting result of that study was that observed means and elasticities may indicate that meat is an inferior good through a particular range of income. Further investigation of this phenomenon by Craun (10) revealed that, from statistical evaluation on the basis of R^2 , the evaluated mathematical form which gave the best fit showed that at some income levels elasticities for both quantity and expenditure were negative.

Dorminy (14, pp. 40-43) compared the data from Williams' survey on meats with the 1965 Household Survey conducted by Consumer and Food Research Division, ARS, USDA (56). After evaluating some alternative models, he concluded that the samples seemed to be drawn from different populations. Each data set indicated that through a range of income, a

decrease in the MPC was evident, although in some cases it was not pronounced enough to produce negative elasticities.

Williams, Craun, and Dorminy indicated that these changes were unexpected for consumption of meats, and said that further investigation would need to be done before one could accept the conclusion that elasticity is in fact negative for meats at particular income levels. It has been noted earlier that theorists (5, p. 15, p. 98-106) have indicated that marginal utilities may change with income level, and that they reflect the degree of aggregation of the product concerned.

The conclusion may be reached that the measures of elasticity of meats observed were not particularly surprising or unexpected from the theoretical standpoint, but might be more so for an aggregate such as total food expenditure. It would seem more unlikely that negative elasticities would be observed with expenditure than with quantity because of the possible quality trade-off. These observations mean that Williams' study may be used to develop hypotheses about more general, policy-oriented questions. An example would be whether the meat expenditure result could be generalized to include all food, and the precise meaning if one could find such evidence about MU's for aggregate food expenditure? And, if there were a trade-off between food and other consumption, which food items were involved?

In summary, the effect of durables on expenditures has become a subject of concern to economists. Fisher used various measures of income to estimate durable good expenditures. Hamilton used an index of durables prices as an explanatory variable for beef consumption. Finally, Williams noticed patterns of means and elasticities which made him suspect that some factors not accounted for might be responsible,

and this was supported by work by Craun and by Dorminy. It seems to be established that there is a relationship between durables prices and the consumption of at least some foods. If the effect of "fixed" expenses, of which durables is a component, does exist, it may be the case that incorporating some measure of that effect in the model specification will show that the appropriately defined income variable is an appropriate explanatory variable for food expenditure.

B. Consumer Unit Size

The literature on effects of income on household expenditure patterns indicated that decisions were not based solely on income. While income coefficients were generally significant, R^2 's indicated that a large proportion of the variance in food expenditure was left unexplained. Demand theory has indicated that some other variables are important, and in terms of food expenditure the physical requirements of the CU (Consumer Unit) are probably next in importance (55, p. 11). Inclusion of CU size increases R^2 considerably. In addition, Currie (11, p. 27) pointed out that failure to include the CU size will cause specification bias if there is any correlation between it and income. It is necessary to include the effect of number, age, and sex of household members if food expenditure is to be explained.

The effect of CU size has seemed to focus on two basic issues. The traditional approach has been to try first to select some groups of individuals who seem to be reasonably homogeneous in terms of consumption of the commodity of interest. This is obviously fraught with difficulties, because it has to do with the differing physical and mental demands by the individuals, and whether they are more similar within

groups or between groups. Then, the first of the two issues is whether the expected effects should enter as a quantitative variable (a scale to represent the age/sex composition of any particular household as a single number), or as some number of discrete categories assumed to be reasonably homogenous in terms of requirements (the qualitative variable).

The second issue is a subset of the quantitative approach. If this approach is taken, each household member may be said to have a given effect on the food budget. This effect may be evaluated either as a "total" effect or as the sum of two individual effects. These are the "specific" effect (the actual physical requirements) and the "income" effect, which reflects that at a given income level, another family member makes the CU relatively poorer. These two effects sum to a total effect.

Assuming the problem of selection of groups is solved, in the qualitative approach dummy variables are inserted in the regression for each of the groups specified. This approach offers flexibility if the objective is to test hypotheses about the effect of CU size on food expenditure (11, p. 28).

The alternative to the qualitative approach involved an attempt to convert the composition of the household to a continuous scale. In the 1920's, Sydenstricker and King (54) recognized the importance of CU size and attempted to devise a scale that would allow the effects to be represented continuously. Building from that base, Prais identified the specific and the income effects mentioned earlier and said that the two effects implied separate scales for specific commodities and a general scale for income. Estimation of these scales required an iterative solution; a standard of living was assumed, the specific scale

was estimated, the income scale was remeasured, etc. until satisfactory estimates were obtained. The procedure was recommended subsequently in a book by Prais and Houthakker, but was not actually used because they did not have a complete commodity set (11, p. 30).

Forsyth (17) attempted to use the Prais and Houthakker method but reached the conclusion that the separate scales could not be estimated by that or any other method because of restrictions on the parameters of the Engel functions (7, pp. 460-461). He said that one should be content to estimate the total effect. Barten refuted Forsythe's conclusion, and separate schools of thought formed.

Price (46) continued the Prais-Houthakker approach by estimating the effect of CU size with a double log expenditure system and obtained specific scales by assuming an income scale. Tomek observed that this entire approach seems to leave the authors mentioned in his survey with the feeling that no completely satisfactory method of capturing the effect of CU size is available.

A second line of thought evolved, initiated by Friedman in an article which suggests that the scale should be specified as continuous functions of age with differences in effect on food expenditure between sexes from birth. This seems consistent with the later development by Forsythe that the separate scales were not important if the objective was the estimation of the total effect. Blokland and Somermeyer later specified utility in terms of households, and included an adult equivalent scale in the utility function; the adult equivalent scale was separate by sex and was a continuous cubic function of age. Peterson later expanded this concept by introducing welfare concepts and a statistical evaluation of household composition. Buse and Salathe

continued with the ideas of continuous functions and sex differences, and expressed household expenditures as a function of the number of adult equivalents, income, and a set of socio-economic characteristics (summary from 42). Buse and Salathe (7, p. 468) noted that the theoretical and statistical implications of their model had not yet been worked through.

In summary, the main thrust of empirical and applied work on the question of the effect of household size and composition has been devoted to the development of scales to form a continuous variable. The focus initially was to find some measure to complement the literature on income, so the first work was empirical in nature rather than theoretical. The effort was toward scales rather than toward use of qualitative variables to test particular hypotheses about effects of CU size. Despite the efforts of many researchers, most seem less than optimistic about the development of scales which are satisfactory. New developments have focused on the estimation of scales which are continuous functions of age, rather than stepwise functions.

C. Tastes and Preferences

General equilibrium analysis suggests that all things are interdependent. The goods and services selected by the CU are influenced by income and household size, but many other factors are also important. Factors common to identifiable groups may have influences which affect food expenditure. Some examples are race, education, and occupation of the household head.

Social class is one of the factors considered. This is often viewed as some combination of education and occupation (44, pp. 153-160). Benus, Kmenta, and Shapiro (6, p. 131) used an average of the past five

years income as an indicator of socio-economic status, but this formulation was dropped due to insignificance.

Race and religion have been evaluated in studies summarized by Tomek (55, p. 13) and in Williams (60, p. 21-22). It has been assumed that race has an influence on food consumption and this has been verified (59, pp. 725-727). Religion seems to have been generally dismissed as a major factor. The information was not reported in the 1972-73 BLS survey.

Region or geographic location has been identified as a factor which has an effect on consumption expenditures. The effect seems to be associated with the customs and values held by the area's population. A variable for region is a proxy variable for those customs, etc., which cannot be easily identified and quantified in surveys. Regional differences have been seen in data from surveys of the English people (44, pp. 162-164) and in U.S. data (55, p. 13).

The degree of urbanization is often assumed to have some influence on consumption. Lee and Phillips (35) and Herrman (24), have evaluated the effect of population levels. This factor probably reflects the different consumption opportunities available in areas as diverse as rural areas and urban environments. Certainly, it would seem that the opportunity to spend on a variety of consumption items would have some effect on other expenditure areas, such as food. This distinction is equally valid between the larger and smaller urban areas, although it is probably a less distinct, more subtle difference.

In addition to these individual influences, Herrman (24, pp. 826-830) found some significant interactions between explanatory variables.

These were between household size and income, between household size and urbanization, and between household size, income, and urbanization.

These socio-economic factors have been evaluated by many other authors in addition to those noted. The results seem to have been spotty, without evidence of any of the factors being significant regressors in every case. Other factors have been evaluated (60, pp. 21-23) but they seem in general to be less influential with respect to food expenditure than those included in this review. The results have even been so inconclusive in some cases that McDonald and Mukherjee (37) found no socio-economic variables significant for meats.

D. Other Influences

Finally, some variables have been considered in the specification of models designed to explain consumption expenditures. Prais and Houthakker (44, Ch. 8 and Ch. 10) mention and deal with both quality variations and economies of scale in consumption. Quality variations have an effect on whether one uses quantity or expenditure as the dependent variable. Economies of scale in consumption deals with the possibility that larger households may be able to use the goods they purchase more efficiently. Examples would be less wastage of "leftovers" from meals, or the handing down process in clothing, from one child to the next.

The relevance of prices to a cross-sectional study has been reviewed by Tomek (55, p. 12). He affirms the reasonability of the practice of treating price as a given factor, one which is faced by all consumers uniformly and is constant. He also notes Mincer's work, which took account of the costs and benefits of the search for lower

prices, and that implied a direct relationship between income and the price of necessities. Therefore, "omitting prices paid may bias the income coefficient downward," but "the omitted variable effect could be very small" (38).

CHAPTER III

APPLICATION OF ECONOMIC THEORY

A. Indifference Curves, Engel Curves, and Elasticities

Economics is the allocation of scarce resources among competing ends. Choices must be made because neither small nor large economic units can satisfy all the demands they would like to make with respect to available resources. Usually, some sort of constraint becomes effective and the unit must rank those things it might consume. This ranking is on the basis of utility (or satisfaction) derived from the consumption of goods and services. Utility is couched in terms of satisfaction, but it is not designed to convey ideas of pleasure or pain. It simply demonstrates that a need or desire exists, so utility is derived. An item has no innate goodness; the utility is entirely in the perception by the economic unit that the good or service will serve some purpose.

Economic units are said to have a preference structure which describes the way choices are made between alternative categories of goods (the utility function), and between particular items within categories. The Classical economists, whose work was most completely laid out by Marshall (35), assumed that utility could be measured with sufficient accuracy to make statements that one good or bundle of goods was preferred to another, and about the strength of that preference. Utility was functional, and the derivative of the total utility function was a measure of the utility change from the consumption of an additional increment of the good (marginal utility). The principle of diminishing

marginal utility was deduced; if additional units of the good were consumed, a point would be reached at which the last unit consumed would not give as much satisfaction as the one before. Because of the restrictiveness of the cardinal utility assumption, economists investigated other possibilities and found that the demand curve properties desired could be derived from ordinally measured scales. It was no longer necessary to try to justify the distance properties of the cardinal measure, but the idea of marginal utility was still valid.

Consumer preference in terms of the Neoclassical ordinal demand curve is usually illustrated with the indifference curve map, a graphic representation of the general utility function $U = (X_1, X_2, \dots, X_n)$. The X 's are goods purchased. Restrictions are placed on the function so that the rank ordering and convexity conditions are satisfied. Any curve on the indifference map represents some given level of utility $U = U_0$, and so long as the consumer stays on the same curve, the utility derived from the different bundles of goods represented by the indifference curve is constant. This implies that the same level of utility may be derived from different combinations of goods.

Consumers are assumed to have some method for choosing among the different combinations of goods they might consume. Specifically, they are attempting the process of constrained optimization, a process of maximizing the utility function subject to the constraint which describes the limit of purchasing power. The consumer moves (on his utility map) as far away from the origin as possible by starting at the top of his list of preferences, where marginal utilities are highest, and by moving down the list until available resources are gone. This constraint means that an infinite amount cannot be demanded. In an

example in which the consumer may purchase only two goods, it would be possible (Figure 1, panel a) to purchase all of either good, or to purchase some combination of the two. Because utility gains are represented by indifference curves further away from the origin, movement will be away from one point until a constraint becomes effective. At the limit, consumption would be at the point where the ratio of marginal utility to price is equal ($MU/P_A = MU/P_B$) for the two goods. Graphically, the budget constraint is tangent to an indifference curve.

The indifference curves in Figure 1, panel b, show the possible combinations of food expenditure and expenditure on all other goods and services at utility levels U_1 and U_2 . The budget constraints AA' and BB' represent the relative prices, and the point of tangency between the indifference curve and the budget constraint is the maximum utility

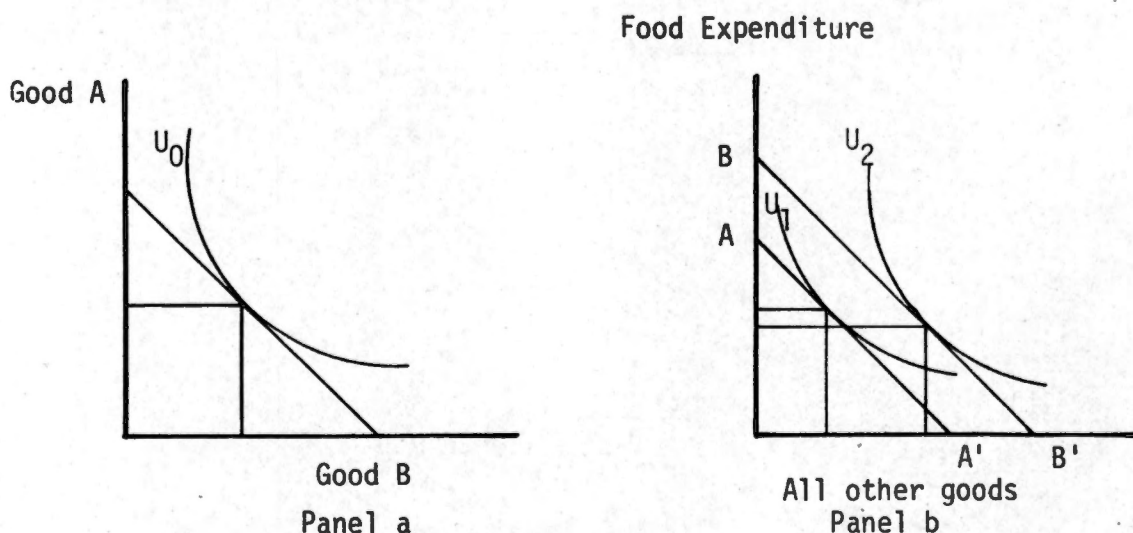


Figure 1. Indifference curves illustrating tradeoffs between bundles of goods.

to be gained from that amount of resource. If the amount of resource is increased from AA' to BB' expenditure on food decreases, but expenditure on all other goods is increased. Thus, at least through the range of resource constraints AA' to BB', the indifference map shows a decrease in consumption of a good as the level of available resource is increased. A negative marginal utility is implied. There is no theoretical restriction on the shape of the curve preventing that occurrence.

In cases such as panel b, the inference may be drawn that if this were the actual consumer preference structure consumption of "all other goods" must have gone up because their marginal utilities were higher, and conversely, the marginal utility of food must have been lower through this range of income. It may be hypothesized that food and other necessities of life are preferred at low levels of income. As income increases consumption of necessities becomes less urgent. For a variety of reasons, the consumption of other goods becomes more desirable. A relative sense of food satiation in terms of quantities might be deduced here.

A measure of this change in preferences or marginal utilities may be derived from the market decisions of consumers. This measure is called elasticity, and it measures the percentage change in quantity associated with a given percentage change in income. Elasticities are by construction independent of units of measurement, and they may be positive, zero, or negative.

The sign associated with the elasticity shows the type of good. A positive sign indicates a normal or superior good, where expenditures increase with income. An inferior good is indicated by a negative sign to show that consumption decreases with an increase in income. If the

absolute value of the estimate of elasticity is greater than 1, then the good is said to be elastic, and a value of less than 1 is inelastic. Both normal and inferior goods result from consumer preference patterns rather than from innate qualities of the goods. As a consumer's available resources increase, he may purchase more of the same product, vary product quality, or he may abandon consumption of the good entirely and choose something else. Preference for a good is not independent of the preference for other goods, and Becker (5, pp. 31-34) states that a good may have different elasticities, and that many goods may be inferior at some expenditure levels and normal at others.

Elasticities may be stated in terms of physical quantities or expenditures. Quantity measurements are straight forward -- they are the changed percentage of the good consumed associated with the percentage change in income. Expenditure elasticities, on the other hand, reflect consumer preferences with respect to several dimensions of a good which include quantity, quality, conveniences, psychological satisfactions, etc. Expenditure elasticities will be used in this study because quantity of food, which contains many diverse components, cannot be satisfactorily aggregated.

A concept often used in budget expenditure studies is the Engel curve, which relates expenditure on a good to the level of income. The expenditure on good X is plotted with the income level, and the various sets of points are mapped. The shape of the Engel curve shows the income elasticity, the effect of a change in income on consumption of the good. Figure 2 show examples of low and high income on consumption of the good. by Engel curves.

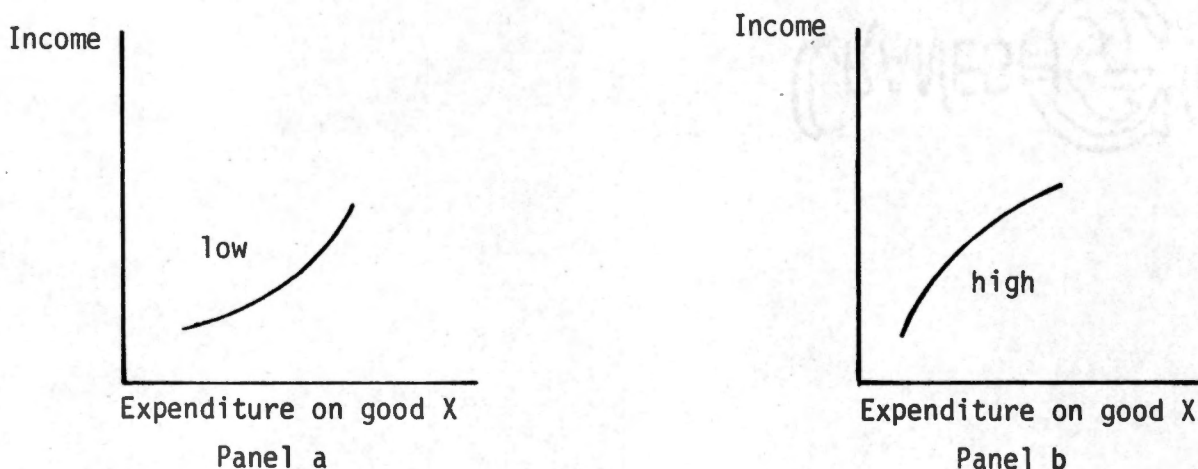


Figure 2. Engel Curve Elasticities

B. The Income-Consumption Relationship

In the analysis of family budgets, income is normally considered to be the basic explanatory relationship, with household size and other sociological variables also playing significant roles. In view of the theoretical notions, it would seem that income would have more influence than is indicated by empirical research results -- those results have generally been significant, but the explanatory power is rather low. This result has been noted in a previous section.

Several approaches to the measurement of income have been suggested. Among those reviewed (pages 7-9), from the theoretical standpoint the PIH seems more relevant in explaining consumption particularly with respect to the larger purchases which are expected to influence the "fixed" portion of CU budgets. It is, however, unobservable because of the limitations of data collection and because CUs may not consciously recognize all the factors which affect permanent income.

The observable variables for model specification are those which come from measures such as total household income or income of the head

of the household, and these are usually adjusted for taxes paid. This is nothing more than absolute income in some time period, and it will be influenced by transitory income in the period. Another possibility is that total expenditure in the period may offer an alternative estimate of permanent income, because the CU has more control in a period of time over expenditure than over income. The CU would be expected to know permanent income and to adjust expenditure accordingly. Expenditure is, on the other hand, subject to variation because of the timing of large purchases, and total expenditure may be over or under-estimated.

The concept of income used in this study involves adjustments to the measured or absolute income that is usually collected. With the objective in mind that an evaluation of the negative elasticities phenomenon observed by Williams would be useful, the conclusion that the proper specification of income may not have been made in previous studies seems appropriate. It is implied that consumers may react to some variable besides the traditional measure of income. Income levels at which the negative elasticities occurred suggest that there may be some connection between food expenditure and "big-ticket" or durable goods. If that relationship can be specified, food expenditures may respond to income increases as a normal good.

The position of the CU with respect to the proportion of income that is in the category of fixed expenses may be examined in terms of the long and short runs. The long run is thought of as the position at which all resources are variable. In the short run, at least one factor is fixed. This notion may be extended to the CU, where food expenditures are thought of as occurring in the short run because the decision is related to the level of fixed expenditures, things which cannot be

adjusted immediately. An auto or home commitment is in this category -- it is a legal commitment and if not met, some action will be taken by another party. The payment must continue. Since something is fixed, one can say that a short-run position exists. If, on the other hand, the CU has no commitments that cannot be easily broken, and can take its income and essentially allocate where ever it prefers, it would be in a long-run position.

The argument is that the discrepancy might be resolved if income were evaluated with respect to two types of expenses. The first type will be called "fixed," and is defined as a set of expenses which cannot be changed by the CU in the short run because the expenses are either contractual in nature or because the CU would lose either monetary or psychological value if the expense is not met. The second type of expense is that over which the CU has some control in the short run, such as entertainment, food, etc. Here, by a combination of quantity and quality adjustments, the proportion of the budget can be changed in response to an increase in the fixed portion. If the marginal utility of consuming additional items in the variable category is less than that of consuming additional items in the fixed category, then the latter category will be selected as more resources become available. The assertion here is that if the CU must borrow to be able to consume these "big-ticket" items, then more consumption implies more debt and the fixed portion must increase by enough to cover the additional payments. A preference to purchase with cash will not have the same effect but will be reflected in some sort of capital accumulation until the purchase price is attained.

The consumer unit is expected to make another adjustment in consumption after it has reached another income plateau. At some point, after CU's have purchased enough durables to become "satiated," they are expected to begin to purchase items in the variable expense class. This implies that the relative positions of marginal utilities have changed again. Hence, expenditure on food, which is in the variable expense category, should increase, although the proportion will probably continue to decline.

In summary, consumer units would meet expenses for necessities at low income levels. As incomes increase and purchasing power and access to credit grow, the proportion of expenses devoted to necessities would not increase as fast as the proportion devoted to other categories of goods. After the CU acquires some desired level of stocks of large ticket goods and services, and as income continues to grow, consumption of both necessities and other goods and services may increase.

CHAPTER IV

THE REGRESSION MODEL

A. Violation of Assumptions

The least squares regression model is a method of estimating population parameters from sample data. The statistical properties of these estimators have been shown to have certain desirable properties. If the usual set of assumptions about the residual or disturbance term are not met, the estimators may lose some of these properties. Given the expected characteristics of the samples used in the estimation procedure, two of these assumptions (homoscedasticity and no exact multicollinearity) may be violated, and therefore require comment. The remainder of the assumptions (i.e., linearity, zero expected value of the disturbance term, uncorrelated disturbances, and non-stochastic explanatory variables) are either met or can be shown to produce estimators which retain acceptable properties in their absence.²

The first of the two assumptions mentioned above is constant variance over the range of observations, or homoscedasticity ($E(u_i^2) = \sigma^2$). The classic violation of this assumption is the relationship between income and consumption expenditures. If cross-sectional data are used, one would expect that consumers with low income would not vary expenditure much, because most of the expenditure goes for such necessities as food and shelter. With little income remaining, it is difficult for expenditures to vary greatly.

²See discussion in Kmenta or Johnston.

On the other hand, consumers generally spend proportionally less on necessities as incomes increase. Hence, they have more discretion about when to spend, and how much to spend. This higher proportion of income not devoted to necessities makes it probable that the variance is different between low and high income consumers.

Since the possibility exists that the error term is heteroscedastic (nonconstant variance), tests have been developed to evaluate the relationship (43, p. 104). The Goldfeld-Quandt test is one such test, and it involves the following steps:

1. The ordering of the data set by the values of the variable expected to be associated with the size of the error variance.
2. Deleting some proportion (d) of observations from the center of the data set (20 percent suggested).
3. Obtaining estimates of the error sums of squares (ESS) associated with the observations at each end of the data set.
4. Testing the hypothesis of homoscedasticity with the statistic $F = ESS_2/ESS_1$, with $(N - d - 4)/2$ degrees of freedom in both the numerator and denominator.

Following this procedure, the data was arranged to set d at 350, and to have at least as many observations in the lower income groups as in the higher income group. This last condition was to insure that ESS for that group is not larger simply due to the larger number of observations. The ESS for each group is given in Table 1. Then, $F = 3.387$, and the hypothesis of homoscedasticity was rejected at the 0.01 level.³

³ $F_{120,120,0.01} = 1.53.$

TABLE 1. ERROR SUMS OF SQUARES FROM HIGH AND LOW INCOME GROUPS AND NUMBER OF OBSERVATIONS

Income Group	ESS	n
High	853930033	624
Low	254629074	627

Data Source: 1972-73 Bureau of Labor Statistics Survey (58).

A procedure for correcting for heteroscedasticity is weighted least squares. This procedure involves the adjustment of the values of each variable observation by some factor. The selection of the factor is the key. The factor may be estimated from the data by calculating s^2 for some discrete groups of the data. This procedure has been used on many sets of data similar to the BLS consumption data. It has been shown that the degree of heteroscedasticity may be assumed to be more or less proportional to the square of income. If this assumption is accepted, then the variance is $E(u_i)^2 = \sigma^2 x_i^2$, and may be corrected by multiplying by $1/\sqrt{x_i^2}$ (25, p. 85).

A second assumption made with respect to the heteroscedastic error term is about the source of the problem. Here, the problem is asserted to lie with the income variable. This seems reasonable in light of the previous statement that additional income certainly seems to provide more discretion to the CU in terms of its level of consumption.

Some of the other variables may be related to the error term, but these possibilities are not as well grounded theoretically or empirically.

On the other hand, the other variables may well be related to income. If this is true, then an adjustment for income alone will at least partially correct for heteroscedasticity associated with other variables.

Thus, the correction factor applied to the BLS data set was of the form:

$$Y_i / X_i = a / X_i + b_1 X_1 / X_i + \dots + u_i / X_i$$

where X_i = residual income.

The application of least squares procedures faces another problem when the correction for heteroscedasticity is applied because of the definition $RESINC = INCOME - FIXED$. If the correction is based on $RESINC$, then the inclusion of $RESINC$ as an explanatory variable in the $FOOD$ equation creates a relationship where $INCOME = FIXED - 1$ for each observation. If $INCOME$ were used, then an analogous situation would be created in the $FIXED$ equation. The estimating procedure would break down because of the perfect relationship. This problem may be addressed by correcting the $FIXED$ equation by $RESINC$ and the $FOOD$ equation by $INCOME$. This is a legitimate procedure since the exact source of heteroscedasticity is never known and because $RESINC$ and $INCOME$ are probably quite highly correlated.

The second problem has to do with the degree of linear relationship between the explanatory variables in the model. The assumption specifies no exact linear relationship, but Kmenta (32, p. 380) says that "multicollinearity is a question of degree and not of kind." In

this context, one is concerned with the actual level and the point at which it might become a problem.

The effect of a high degree of multicollinearity may be seen in the variances and covariances of the estimated regression coefficients (32, p. 388). These values increase because of the effect on the determinant of $X'X$. The precision of the estimators becomes low and it may become difficult to reject the hypothesis that estimates are 0. Variables may be incorrectly dropped from the model because calculated t values are small.

In addition to the problem related to acceptance and rejection of the null hypothesis, difficulties in interpretation of regression coefficients arise. The reference to partial correlation coefficients emphasizes that they are measures with other independent variables constant. However, these other variables are not constant when the multicollinearity exists, and precise meaning is, consequently, difficult to assign.

Multicollinearity appeared to be a problem with respect to the models specified in this study. It has been noted that income and size of the consumer unit are correlated (almost always positively). This effect does not seem to be a major problem, because the variables are generally found to be significantly different from zero. Thus, even if the estimates are imprecise, the effects are still apparent.

A second source of multicollinearity is introduced into the model through the dummy variables required for analysis of the qualitative variables. These factors can and do move together over households. Scarce attributes can cause problems, since they differ from zero in few elements. Finally, the "possibility of approximate linear

relationships between intercept dummies and slope dummies . . ."

(51, p. 62) is not precluded.

The focus of the problems presented by multicollinearity is on the lack of precision in the statistical tests performed. The adjustments lie in the possibility of finding additional sample information, since the problem is one which lies inherently in the sample. The linear relationship between variables may be investigated by regressing each independent variable on all of the other independent variables. If a large value of R^2 appears, multicollinearity is indicated and one of the variables may have to be omitted from the model.

The result of this procedure was a set of consistently high R^2 s (0.80-0.90), an indication that there may be problems with sorting out effects. The effects may be strong enough to be statistically significant even with the existence of this data problem.

B. Two Stage Least Squares

The error term associated with regression estimates is often a source of loss of desirable properties of least squares estimators. One source of such problems for the estimating technique is the case of correlation between the error term and an endogenous variable. This condition causes problems with the property of consistency. The 2SLS procedure produces consistent estimates by regressing each endogenous variable in the system on all the exogenous variables, producing a set of estimated values of the endogenous variables. These estimates are by construction independent of the error term. In the second stage, these predicted values replace the original observations on the endogenous variables, and the least squares procedure is applied to this new set of data (28, pp. 382-383).

In the interpretation of the 2SLS coefficients, the usual statistic considered is called the t-ratio. The use of this statistic is not strictly valid. The asymptotic properties of the statistic do not exactly follow the t distribution, but the properties of the estimators and of the hypothesis-testing mechanism indicate that the procedures seem to be acceptable, although some errors may be introduced. ". . . at least from the point of view of hypothesis testing or interval estimation, the OLS method is inferior to the consistent methods of estimation" (32, p. 583). The knowledge of the small sample properties "is far from complete" (32, p. 584), but the consistent estimators are not worse. These comments indicate that the t-ratios may serve as reasonable approximations to the t distribution, so the usual tests are applied. The actual meaning of the statements as to which variables are significant and which are not is somewhat ambiguous, and the statements regarding significance levels are made with this warning in mind.

C. Dummy Variables

The regression model may be extended to include variables which are not easily measured. Due to the nature of these variables, an observation either possesses a particular characteristic or it does not. The variables are handled through the assignment of 0,1 values to indicate whether the trait is present in any particular observation. These variables are commonly called dummy variables. The effect is that essentially two models are estimated, one for the excluded class and a second which is measured as the difference of the included class from the excluded class.

The formulation allows only the intercept of the regression equation to change. It is possible to allow changes in the estimated slope of the line through the interaction of the variable which allows the intercept change and the variable which is presumed to influence the slope of the line.

Both intercept and slope dummy variables may be specified for more than two characteristics for each dummy variable. An example would be the comparison of several different occupational groups or levels of education. Multiple sets of dummy variables may be included in the same model.

An additional change in the format has to do with the relationship between the dummy variables before and after the correction for heteroscedasticity. Before the correction, an intercept change was defined by a 0 or a 1, depending on whether the observation has the particular characteristic in question. If the slopes are to be allowed to vary also, that may be accomplished by the multiplication of the intercept change by the variable assumed to be responsible for the changing slopes. When the data are transformed, all variables are divided by the variable assumed to be responsible for the nonconstant variance. In this particular case, the same variable (RESINC) is assumed to be responsible for both the changing slopes of the dummy variable and the problem with the error term. The result of the application of both procedures is that the variable which appeared to be allowing the intercept to change actually allows the slope to change, and the new intercept is the original intercept change divided by the variable which corrects for heteroscedasticity.

CHAPTER V

DATA SOURCES AND PROCEDURES

In this chapter, several procedural steps will be considered. The sources of data are described, first for a survey conducted in Knox County Tennessee, and secondly, for the 1972-73 Bureau of Labor Statistics survey. The manner in which relevant information was sorted from the mass of information in the BLS survey is described. Information to calculate a consumer unit size scale was obtained from the BLS data, and a test to evaluate whether the SMSA sizes and the two years included in the survey were different was conducted.

A. THE KNOX SURVEY

For maximum usefulness, research should be based on up-to-date information, either primary or secondary. A primary data set can be expected to reflect current trends and can be tailored to complement research interests in a particular geographic area. The evaluation of the research hypothesis was to be based on a survey in Knox County, Tennessee. The scope of the study included all food expenditure, including meals away from home. Information was requested which would allow the generation of data on income (by particular types), socio-economic variables, payroll deductions, change in savings account for the year, and the category of "fixed" expenses. A general question attempted to generate some idea of the consumer unit's income expectations over the next year.

To insure that an adequate sample proportion was obtained in each of the income groups, information from the Knoxville Metropolitan

Planning Commission on location of poverty by census tract was utilized. These data identified 65 tracts and gave the proportion of low, moderate and other CUs by percentage of median income in the Knoxville SMSA. To attain confidence levels of 90%, Yamane's (61) procedure required 100 observations in each of the income groups. The number of tracts in each group was rather low, so the selection of particular tracts is shown in Table 2.

TABLE 2. TOTAL NUMBER OF CENSUS TRACTS AND NUMBER
SELECTED, BY INCOME GROUP

Income	Tracts	
	Total Number	Number Selected
High		
\$16000 and above	6	6
Moderate		
\$8000 - \$16000	39	20
Low		
\$0 - \$8000	20	10

The tracts to be sampled in the moderate and low income groups were selected by reference to a random number table. A decision to over-sample by 10% was reached, a figure which seemed reasonable with respect to previous research efforts. The total number of observations required in each income group was divided by the number of tracts in that income group to obtain the number of interviews to be completed in

each tract. Random entry points into the tract were selected as starting points, every fourth house thereafter was interviewed, and the interviews continued on the same side of the street until a CU declined to participate, was not home, etc. In that case, the home across the street was interviewed, and the procedure continued in the same manner.

At its inception, the survey objective was to obtain 330 completed CU observations. Time was a constraint -- in order to avoid both vacation periods (summer) and holiday food purchases (Thanksgiving, Christmas), the survey had to be completed by a date early enough to avoid the problems associated with data collection in those presumably nonrepresentative periods. An initial interview was conducted with each CU to explain the purpose and procedures of the survey, to get a commitment to cooperate, and to collect income and other socio-economic information. By the target date, approximately 225 CUs had agreed to cooperate. Information to complete the survey was mailed to each of the participants.

Participants were asked to mail information on food expenses on a weekly basis for three weeks. On Thursday of the second week, those CUs who had not responded were telephoned. The same procedure was used during the third week and on the Thursday after the survey period ended. Despite this effort results from the survey were discouraging. Of the original 225 CUs, only about one-third (85) reported for the full three week period, with 14 additional CUs reporting for one or two weeks.

It seemed likely that one reason for the disappointing participation rate was that the respondents were asked for considerable detail about the arrangement of their financial affairs, and then were asked to keep grocery store cash register receipts and a record of food expenditures outside the home for the three-week period. In addition,

the responses were not anonymous. Several interviewees raised the question of confidentiality, and many others probably had unspoken reservations. Although this problem was addressed in the interview, it, in addition to the other detail requested, must have had an effect. Finally, no financial incentive was offered for completion of the survey. Previous research (40, p. 128) had indicated that such incentives had not always contributed significantly to better participation rates.

The completed observations from the Knoxville survey were used insofar as the limited sample size permitted. The data were deflated by the Consumer Price Index to an average of the 1972 and 1973 levels. It contained information about some selected food categories. Specifically, these categories were meats, dairy products, produce, and groceries. A fifth category was collected because it completed the expenditures on food in stores. The category pertained to nonfood purchases in grocery stores. The categories were selected because they are required by law on cash register receipts from grocery stores. Since it seemed that very personal financial information was being requested, this method of reporting food expenditure was selected because it seemed to minimize the effort required to complete the survey, thereby increasing the likelihood of cooperation. This trade-off between detailed financial information and accurate and detailed food expense reporting was to be accomplished with the simple process of collecting grocery store cash register receipts and reporting a single aggregate for expenses on meals outside the home.

The detail associated with the Knoxville survey was used to derive estimates of the effect of changing income on the composition

of the food bundle purchased. It seemed that this type of information would be useful to policymakers.

B. THE BLS SURVEY

The possibility of an unsuccessful survey had been considered during the planning stages of the research, and an alternative data set, the 1972-73 Bureau of Labor Statistics Consumer Expenditure Survey was available (58). This set was collected by the Bureau of the Census with several objectives in mind (8, 33), foremost of which was the revision of the Consumer Price Index. It was no surprise that the data was not ideally suited to the objectives of this project, and that considerable data transformation would have to be done. The form of the data was such that the major portion of the items that have been termed "fixed" expenses could be sorted and aggregated on an annual basis (principal and interest payments on homes, rent, auto and recreational vehicle payments, etc.), but there is a missing element compared to the Knoxville survey. There, obligated payments were collected, and no questions were asked about classification of particular items. The only concern was the effect on the family budget. Some smaller obligations were picked up in the Knoxville survey, such as payments on smaller purchases and medical payments if they were a regular, ongoing part of the family budget. These were the sort of things which would be expected to have the same effect on the budget as larger items. In the BLS survey, there was no way to evaluate whether such items were a continuing part of the budget, or whether they were one-time expenses.

The data set provided by BLS was the interview portion of the survey, and consisted of information gathered in a series of quarterly

visits. The approximately 2.2 million individual records pertained to about 22,000 consumer unit observations. These were sorted by location (region of the country) and by size (population of the SMSA and associated rural areas) in an attempt to get a data set generally corresponding to the Knoxville data.

BLS boundaries for SMSA size were 50,000 - 399,999 and 400,000 - 1,000,000. Because Knoxville's population was just greater than the boundary between these groups, the appropriate group for comparison was unclear. Knoxville may have retained more of the attributes of smaller cities than it had gained the atmosphere of a large urban center. Hence, both groups were included in the sample.

The sample size in the former SMSA size group was 453 in 1972 and 439 in 1973. For the latter size group, the analogous numbers were 356 and 560.

In summary, there were two distinct data sets, the Knox data and the BLS survey. The former was used to estimate the distributional effects, if any, of the interrelationship between "fixed" expenses and food expenses. The latter was the basis for the estimation of these two variables as aggregates.

A. Derivation of Variables

The BLS data contained an exceptional amount of detail about the expenses incurred by the consumer units. The relevant items were collected and the result was a data set which contained socioeconomic information about consumer units and variable values, some of which existed in the original data set and some generated from the data. The variables were:

Food expense. This group included all food expenses reported by the consumer unit, except catered food at weddings, confirmations, or Bar Mitzvahs. These excepted items were not thought to be affected by income in a period as short as a year. They were also likely to be occasional, rather large expenses, not related to the usual food consumption habits of the unit. An imputed value for food received without direct expense was included in the total. The total food expense was about 16% of income.

Fixed expense. Several categories of expenses were included in this value.

Rent expense. Rent was included for principal residence, rent received as pay (since that amount would be freed for other expenses), and garage and parking fees charged to renters.

Ownership expense. If the consumer unit owned its principal residence, the various costs associated with that ownership were collected. The first case (the usual one) was payment of mortgage installments on a home purchased prior to the survey year. These payments were decomposed into interest, principal reduction, and tax, insurance, and other associated expenses. The principal reduction was recorded as a negative number, and had to be subtracted to get the correct effect on consumer unit expenses. A second case was involved when the unit purchased a home or sold and bought a home in the survey year. Then, the logical procedure was to take the purchase price, subtract the mortgage balance at the end of the year, and add the associated costs of closing, points, penalties, and other sale expenses. This figure was assumed to be a reflection of the housing cost for the year. When a house was sold and another purchased, the trade-in or sale price was subtracted from the

mortgage balance at the end of year (EOY), in addition to the other factors mentioned. This brought some additional factors into the calculation, since in some cases it appeared that the CU had reported incompletely. These cases were those in which the result of the data evaluation was a large positive number (the possible result of omission of the EOY balance) or negative number (EOY balance but no purchase price, etc.). The possibility existed that these large numbers were correct, because the unit might sell and not repurchase a home. The more likely cause, however, seemed to be reporting error. Some observations were deleted. The criteria for deletion was whether the deletion made a large reduction in the standard error of the variable, that is, whether it seemed that the observation was one of a few outliers. Finally, a third case was identified, where the consumer unit owned the residence outright. Then, the expense was simply a summation of property taxes, property insurance, maintenance, etc., associated with owning the home.

Transportation services. Another contributor to the consumer unit's fixed expense is transportation. These contribute in the sense that an auto purchase using payments over time requires a continuous payment until the obligation is discharged. The first attempt to obtain a value for transportation was formulated from the logic that if a finance charge was reported, the likelihood was great that the auto had been purchased with credit. In that case, the algorithm summed the entries under purchases and finance charges and subtracted an item titled "amount owed on vehicles purchased in survey year."

This produced negative entries in almost half the observations with entries in this category. As in the case of home ownership, the

implication was that reporting was incomplete -- to get negatives, either the end of year entry was greater than purchase price or there was no entry for purchase price. To check whether this series of negative numbers was the result of a simple omission of the finance charge, the opposite approach was taken, by first looking for an end of year value, subtracting that if one was found, and then adding purchases and finance charges. This evaluation had essentially the same result as the first formulation, so the evidence of missing information seemed confirmed. The possibility that some CUs "trade down" for a less expensive auto is admitted, a factor which could explain some of the negatives, but which seemed unlikely to explain so many.

The procedure adopted was to assume that if a CU had a number or variable value resulting from the algorithm, then it had expenses which should be included in the fixed portion of the budget. The further assumption was made that the CUs who had reported correctly and, as a result, had a positive entry, came from the same population as those who seemed not to have reported correctly. Therefore, the average of the seemingly correct responses was substituted for the negative results. The average was chosen because it is the best estimate of a population parameter if only sample information is available.

Additional expenses such as auto lease expenses and payments on autos purchased prior to the survey year were included in this category.

Recreational vehicles, selected gift expenses, and vacation home expenses. Motor homes, motorcycles, boats, etc., seemed to be the sort of purchases which would have an effect on fixed expenses. They were deemed likely to generate credit payments. The selected gifts were exclusively autos given as gifts. The vacation home contains essentially

the same information as the earlier "home ownership" discussion. These expenses were summed.

A sort of catch-all category was included in the analysis. This was the "net change in amount owed to other creditors at end of survey year," and its value was included in the fixed expenses.

Income. Income was defined to include wages and salaries, as well as items such as net income from business or farm, veterans' compensations, public assistance, and various sorts of refunds. It was adjusted for payroll deductions such as union dues, occupational expenses, life insurance, social security, retirement fund contributions, and income taxes.

Other sources of funds. These included net changes in amount in saving accounts and in checking accounts.

With respect to some of the included variables, there was no way to determine how much of the debt had been repaid in the form of installment payments, so the total of purchase price and interest expense was subtracted, on the assumption that this would cause the bias introduced to have the effect of causing residual income to be lower than it actually should be. Hence, the estimating procedure should have produced a coefficient for fixed with some upward bias, and one for residual income with some downward bias. The statistical tests may be somewhat biased, but the effect is not expected to be large.

B. Derivation of Consumer Unit Size

This procedure involved the classification of each CU member into one of the 11 categories defined by the Consumer and Food Economics Institute, Agricultural Research Service, USDA, in publications concerned

with the "Cost of Food at Home for a Week" (39, pp. 47-49). Calculation of the variable designed to represent the number and age/sex composition of the CU followed a procedure developed in Currie (12, pp. 67-69). The sample was separated into groups assumed to be homogeneous in terms of the expected effect on food expenditure. By comparing all included groups to one of the groups chosen to be the standard, each family member was assigned a weight factor. The weights for each family member were summed, creating a kind of index value to represent the CU.

The assumed homogeneity within groups and the more or less arbitrary selection of the number of groups has received criticism (7, pp. 460-462). The critics contend that such groupings are unnecessarily restrictive in view of the fact that food requirements, hence expenditures, change as part of a continuous biological process and should be represented by a continuous variable. The establishment of discrete groups does allow changes in food expenditure associated with changing biological requirements, but has boundaries which sometimes seem rather large.

Despite this criticism, the empirical development has primarily taken this stepwise approach. Additionally, use of these groupings may be justified because many policy decisions are based on these standards.⁴ For example, "USDA's family food plans have served for more than 40 years as guides" for both public and private agencies. The U. S. Department of Agriculture uses the information to prepare nutritional guidance materials and to establish coupon allotments for the Food Stamp program;

⁴These are based on actual consumption by families from 1972-73 BLS Consumer Expenditure Survey.

the Internal Revenue Service determines rates of care for dependents based on this series; and various other agencies establish food cost standards on this basis. In the private sector, the consumption information is used by camps and institutions to estimate food needs, and by charitable groups to establish rates of care for foster children. It would appear that standards already extensively used should be more useful to policymakers.

Interest centers on the estimation of the total effect of income (Y) and the age/sex groups (N) on total food expenditure (E). In the equation, m groups are assumed. The form is

$$E = a + b \log Y + c_1 N_1 + \dots + c_j N_j + \dots + c_m N_m$$

The c_j 's are regression coefficients and measure the additional expenditure by the CU due to the addition of a member of a group to a household, other things constant. The scale can be estimated by adjusting the equation to show the coefficient as a proportion of the standard group (m).

$$E = a + b \log Y + c'(w_1 N_1 + \dots + w_j N_j + \dots + w_m N_m)$$

where $w_j = c_j / c_m$.

Both income and CU size must be included in this formulation because of the correlation between the two. If CU size were excluded, the income coefficients would be biased to the extent of the correlation.

It has been conventional to use the adult male as the base group. Weights were calculated and summed for each CU to get the value for the variable CUSIZE.

For each of the age/sex subgroups, the effect on food expenditure of the addition of another family member was estimated. Because of the criticism of the discrete approach, it seemed appropriate to test whether the subgroup coefficients were statistically different. If not, then fewer groups would have to be specified to obtain an estimate of consumer unit size.

From the parameter estimates and the ratio of each subgroup to the base (male 20-54), there appeared to be major groupings at less than 12 years and 12 years and greater (Table 3).

To test whether these groupings were relevant, a t-test of the form

$$\hat{c}_j - \hat{c}_k / s(\hat{c}_j - \hat{c}_k), \text{ where}$$

$\hat{c}_j - \hat{c}_k$ = the difference between two coefficients j and k, and

$s(\hat{c}_j - \hat{c}_k)$ = standard error of the difference between c values

for j and k, may be used.

Because of the similarities in terms of size of coefficients and the large number of possible combinations of coefficients, some representative differences were evaluated (Table 4). At the $\alpha = 0.10$ significance level, the only significant difference was between the 3-5 years subgroup and the 12-14 years male subgroup.

Because of the expected biological differences between the age subgroups and between sexes, the distinction between adults and teenage males and females was retained, along with the children subgroup (<12). The coefficients were reestimated, and it again was not possible to distinguish the difference between adults and teenagers with respect to

TABLE 3. INCREASED DOLLAR EXPENDITURE ON FOOD PER
ADDITIONAL SUBGROUP MEMBER FOR ELEVEN AGE-SEX
SUBGROUPS, AND RATIO OF EACH TO THE ADULT
MALE (20-54) SUBGROUP

Age/Sex Group	Estimated Coefficient	Ratio to Adult Male
20-54 female	257.6*	1.16
>54 female	191.2*	0.86
20-54 male	222.4*	1.00
>54 male	270.5*	1.21
1-2 years	127.0*	0.57
3-5 years	104.4	0.46
6-8 years	140.3*	0.63
9-11 years	158.1*	0.71
12-19 female	243.4*	1.09
12-14 male	278.3*	1.25
15-19 male	285.6*	1.28

* Statistically significant at $\alpha = 0.10$ level.

Data Source: 1972-73 Bureau of Labor Statistics survey (58).

TABLE 4. TEST OF DIFFERENCES BETWEEN REGRESSION
COEFFICIENTS FOR SELECTED AGE-SEX
SUBGROUPS

Age-Sex Group	t-Value Between the Two Coefficients
2-54 female to > 54 female	0.71
1-2 years to 3-5 years	0.16
12-19 female to 12-14 male	-0.29
3-5 years to 12-14 male	1.50*

* Statistically significant at $\alpha = 0.10$ level.

Data Source: 1972-73 Bureau of Labor
Statistics survey (58).

food expenditure. A final estimation was done, with two subgroups, designated adults and children. The tests indicate that the coefficients are different from 0 and from each other (Table 5).

This relationship was also estimated with the semi-logarithmic form, in addition to the linear form actually used. The distinction seemed to make little difference to the analysis.

This statistical evidence led to the conclusion that additional information could not be gained from partitioning the data set into a large number of age/sex groups. It may be questioned whether these results are reasonable in view of the theoretical and empirical work on the biological and psychological effect on expenditures from differences in the age/sex composition of the household. However, much literature has dealt with other goods and services in addition to food, and the differences may be more pronounced in these other products than in the more stable and predictable demand for food.

C. Test of Equality Between Survey Years and SMSA Sizes

After the data was processed and the variable used to represent the age/sex composition of the family created, a test was conducted to determine if the subgroups defined were similar enough to be used as a single set or if they could be viewed as distinctive subgroups. An F test on residual sums of squares may be used to determine whether observations came from the same population (28, pp. 136-137). Here,

$$E_1 = X_1 \hat{C}_1 + u_1$$

$$E_2 = X_2 \hat{C}_2 + u_2,$$

where

TABLE 5. INCREASED DOLLAR EXPENDITURE ON FOOD PER
 ADDITIONAL SUBGROUP MEMBER FOR TWO AGE-SEX
 SUBGROUPS, RATIO OF EACH TO ADULT MALE
 (20-54) SUBGROUP, AND TEST OF
 DIFFERENCE BETWEEN THE TWO
 SUBGROUPS

Age/Sex Group	Estimated Coefficient	Ratio to Adult	Calculated t-Value Between Coefficients
Adults (≥ 12)	313.7	1.0	5.73*
Children (< 12)	118.0	0.376	

*Statistically significant at $\alpha = 0.10$ level.

Data Source: 1972-73 Bureau of Labor Statistics survey (58).

E_1, E_2 = vectors of observations on the dependent variable

X_1, X_2 = matrix of observations on independent variables

$\hat{c}_1, \hat{c}_2$ = vectors of estimated coefficients

u_1, u_2 = vectors of residuals, or disturbance term,

represented the two sets. The test procedure involved the relationship between the error sum of squares (ESS) from the pooled data and from the individual regressions. Now,

Q_1 = ESS from pooled data

Q_2 = ESS from individual regressions, summed

$Q_3 = Q_1 - Q_2$

m, n = number of observations in the respective groups

k = number of explanatory variables,

and the test was given as

$$F = Q_3/k / Q_2/m + n - 2k$$

This procedure gave a calculated value for F which may be compared with critical values of the F distribution, to test the null hypothesis that $\hat{c}_1 = \hat{c}_2 = c$, or that the observations come from the same population. This hypothesis may be rejected if the calculated value of F exceeds the critical value. Using ordinary least squares, income, household size, and a group of socio-economic variables were regressed on total food expenditure. Residual sums of squares were evaluated in accordance with the F test described earlier.

It was expected that the four subgroups (page 37) from the BLS survey were different. Firstly, sampling was done in consecutive years, calendar 1972 and 1973. In this case, there might have been events in

one year which caused the observations to deviate from their "normal" values. The survey design recognized this possibility, and the two-year procedure was adopted to try to eliminate that source of bias. In fact, part or all of 1972 and 1973 were influenced by wage and price controls which dampened potential price changes, which are one significant source of consumption differences, as well as other adjustments. In addition, the first Arab oil embargo impacted the economy in late 1973, but the economic brunt of that embargo may have been taken later because of inherent lags in the system. Additionally, other year to year differences were expected, though they were likely to be more subtle.

The degree of urbanization may have affected the consumer unit's food expenditure. The consumption patterns were expected to differ because of the alternative consumption possibilities (increased level of available activities of a cultural or recreational nature, etc.) associated with the size of the metropolitan area.

The test of the null hypothesis that the matrix of $\hat{c}$'s was equal indicated that the individual years were not statistically different. The test was two tailed -- the influence of prices, which edged upward about 3%, did not seem likely to be dominant enough to completely offset the possibility that other factors could have caused consumption in the second year to be greater than the first.

Despite the theoretical notion that differing degrees of urbanization would affect consumption patterns, the evidence available from the sample did not support that conclusion. The critical value of the F distribution at $\alpha = 0.10$ is 1.44, while the calculated F ratio between the two SMSA sizes is 0.95. The analogous ratio between the two survey years was 1.17. The hypothesis that the difference between the B

vectors is 0 could not be rejected for either of the subsets of the data, so all observations were pooled to form one data set for further analysis.

CHAPTER VI

MODEL SPECIFICATION

A particular economic model is selected on the basis of economic logic; mathematic forms and statistical procedures are chosen to represent the relationship on the basis of whether or not they reflect the economic requirements. To the greatest extent possible, the researcher must formulate and translate into correct model specification the three elements mentioned, or be resigned to the fact that parameter estimates are most likely biased. If any one of the three elements of specification contains error, it will affect the research regardless of the correctness of the other two. Thus, care must be taken when decisions about model specification are made, and the effect that each element has on the other two must be considered throughout the entire process. If it seems that a particular procedure in one element may affect other elements adversely, adjustments should be made which will allow all elements to represent the research hypothesis correctly.

The interrelationships between the categories of expenditures in the consumer's budget are complex. If one looks at food expenditures, expenditures on the category of "fixed" expenses, and all other expenses, the categories are clearly related. Each part of the budget is not determined by itself, but in conjunction with the other parts. Since the consumer has limited resources, the decisions about allocations to budget categories must consider the entire preference structure. Income and the socio-economic characteristics of the consumer are also involved because of the effects on tastes and preferences. Model specification

must allow these expected effects to be revealed if they exist. Since the variables are expected to affect each other, that effect may be permitted through a simultaneous system of equations, in which the estimated values satisfy both equations.

The model must include (and exclude) the correct variables. Errors here can lead to bias in estimated parameters, just as can the choice of equation system. Exclusion of relevant variables will cause bias unless the correlation with other variables is zero, because the effect of the excluded variable still exists but is shifted to the error term and to the included explanatory variables to the extent of their correlation with the excluded variable. Inclusion of irrelevant variables may also cause problems, but these seem not as serious as the former type (43, pp. 187-190).

The variables included in the model were selected either on the basis of significance in previous research or on relevance to the logic of the model. They were of two basic types -- those that comprised the various budget categories, and those that were expected to reflect the tastes and preferences of the CU.

Since the study utilized two data sets and some differences in the form of the variables between the two data sets could not be resolved, a separate model was specified for each set. The differences in the variables were in the ability to obtain equivalent data on the CUs fixed expenses and in the detail of reported food expenditures. The former case had to do with the ability to identify those expenses which should be added to the fixed portion of expenses; the BLS survey offered no way to determine if they were regularly occurring expenses. In the latter, the Knox survey picked up information about particular groups of

food expense which may be of interest to policymakers. This detail was not contained in the BLS data.

A. THE BLS MODEL

The central interest of this study was an analysis of the allocation of income to certain categories in the CU budget. Specifically, it was hypothesized that the CU evaluates budget shares on the basis of the share of income going into the obligated portion of the budget. It then considers the residual to decide whether to purchase more items which would increase the fixed portion of the budget. The general form of the equation system was:

$$\text{FIXED} = f(\text{FOOD}, \text{INCOME}, \text{CUSIZE}, \text{AGEHD}, \text{LC}, \text{RACE}, \text{EDUC}, \text{OCCUP}, \\ \text{RENTOWN}, \text{OCCSPO})$$

$$\text{FOOD} = f(\text{FIXED}, \text{RESINC}, \text{CUSIZE}, \text{AGEHD}, \text{LC}, \text{RACE}, \text{EDUC}, \text{OCCUP}, \\ \text{RENTOWN}, \text{OCCSPO})$$

where

FIXED = necessary expenses plus debt service

FOOD = all food expenses and value imputed for food acquired without cost

INCOME = total income minus payroll deductions

RESINC = INCOME minus PAYMENTS

CUSIZE = number and age/sex composition of the family

RACE = race of CU head, white or nonwhite

OCCUP = occupation of CU head

EDUC = education of CU head

LC = interaction between AGEHD and CUSIZE, a life cycle variable

RENTOWN = renter or owner of principal dwelling

OCCSPO = spouse occupation

AGEHD = age of CU head

Specification error exists when the conceptual model does not properly represent the real situation. The simultaneous system written above contains error; it would seem that the PAYMENTS variable (the decision about how much of the budget will be "fixed" each period of time) is made in one time period, and FOOD would be affected in subsequent periods, with the effect based on the income available after the first decision is made. This would require the use of lagged values of RESINC. An effect is implied in subsequent periods on FOOD. It would reflect the decision of the CU regarding the amount of assets held as cash, as savings, and as other forms of equities. Data needed evaluate that hypothesis are not available in this data set.

B. THE KNOX MODEL

The Knox model was designed to permit estimation of the variables collected in the Knoxville survey. The first difference between the Knoxville survey data and the BLS data was in the collection of the fixed portion of CU expenses -- respondents were asked to report all their regular monthly payments -- those things which would cause the loss of some good or service if the payment were not made. With this in mind, more expenses were drawn into the category, because the respondents were asked specifically about such expenses. The BLS data offered no way to separate these items, which were expected to have the same effect as debt repayments, from the one-time kind of payments. The expected effect was that the fixed expense would increase as a proportion of income in the Knoxville model.

One of the objectives of the study was to determine which categories of food were most affected if the hypothesis that total food expenditure is affected by size of fixed expenses was not rejected. This extra information was available from the Knoxville survey and was used to evaluate distributional effects among food groups. If, as supposed, food expenditures declined and other expenditures increased, just what foods were responsive to the change? A possibility was that all food categories might change proportionally, but it was more likely, given the differences between individual food products, that the effect would be felt in particular categories. It may be of interest for policy-makers to have an idea of the effect on diet as residual income levels change. Identification of these differences was not possible with the BLS survey information.

The Knox model was a six equation system designed to estimate the values of FIXED and of the individual food categories. The first equation is similar to the first equation from the BLS model. The exceptions are two. First, information about the occupation of the spouse was not collected in the Knox survey. Second, information about the number of incomes in the CU was collected in the Knox survey, and was included in this model.

The individual food categories were specified as functions with both FIXED and FOOD as explanatory variables, along with the socio-economic variables. The logic for the inclusion of FIXED in the Knox model was the same as for its inclusion in the BLS model. FOOD was included in order to investigate whether a relationship exists between total food expenditure and the share devoted to the individual food categories.

The general forms were:

FIXED = f(FOOD, INCOME, CUSIZE, AGEHD, LC, RACE, EDUC, OCCUP,
RENTOWN, NUMINC)

MEAT = f(FIXED, FOOD, RESINC, CUSIZE, AGEHD, LC, RACE, EDUC,
OCCUP, RENTOWN, BULKSTOR, NUMINC)

DAIRY = f(FIXED, FOOD, RESINC, CUSIZE, AGEHD, LC, RACE, EDUC,
OCCUP, RENTOWN, BULKSTOR, NUMINC)

PRODUCE = f(FIXED, FOOD, RESINC, CUSIZE, AGEHD, LC, RACE, EDUC,
OCCUP, RENTOWN, BULKSTOR, NUMINC)

GROCERY = f(FIXED, FOOD, RESINC, CUSIZE, AGEHD, LC, RACE, EDUC,
OCCUP, RENTOWN, BULKSTOR, NUMINC)

MEALS OUT = f(FIXED, FOOD, RESINC, CUSIZE, AGEHD, LC, RACE, EDUC,
OCCUP, RENTOWN, BULKSTOR, NUMINC)

where the variable names are the same in the BLS model, with the addition of

MEAT = expenditure on meats

DAIRY = expenditure on dairy products

PRODUCE = expenditure on produce

GROCERY = expenditure on grocery

MEALS OUT = expenditure on meals outside the home

NUMINC = number of incomes in the CU

BULKSTOR = bulk food storage capacity

C. ALGEBRAIC FORM OF ENGEL CURVES

The two models above and the specific variables included gave a general form for the model. There must also be specified a particular algebraic form, one which allows the expected relationship to be revealed if it exists in the data set. It should maintain the economic logic

behind the hypothesis. The expected relationship is that food expenditures will increase, then grow at a slower rate or even decline, then again increase through particular income ranges.

The hypothesized behavior of food expenditure requires a form which will allow those characteristics to be represented. A cubic function of the general form

$$y = a + b_1X + b_2X^2 + b_3X^3$$

will produce a curve with the general form shown in figure 3a, but the specific shape is determined by size and sign of coefficients. Fitting such a function would be ideal because it permits the construction of two turning points. It is linear in the parameters and can be estimated with least squares. The cubic function should give an estimate about the location of the points of change in the curve (Figure 3a).

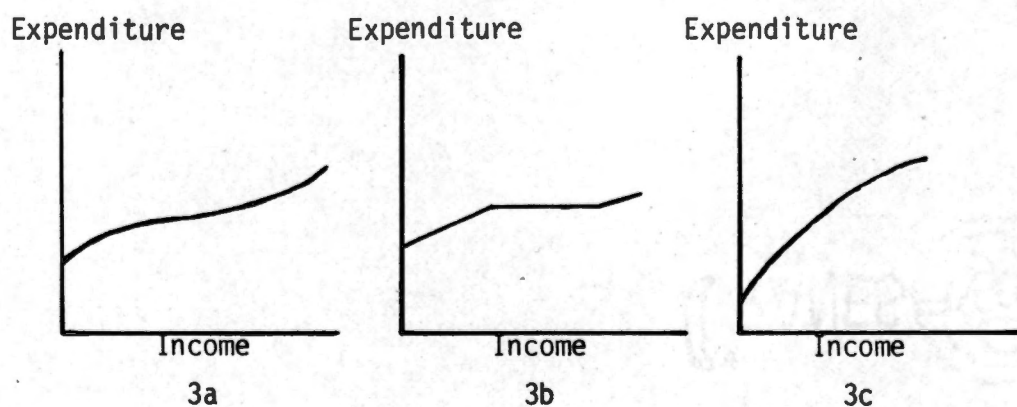


Figure 3. Graph of Alternative Forms of Engel Functions

The second possibility is depicted in Figure 3b. Here, the model contains three linear segments which may be estimated with dummy variables or with separate regressions. It may be argued that this procedure produces a sufficiently good approximation, because over a short distance a straight line approximates a curved line.

Selection of the transitional points was a rather arbitrary process and the real points of change were unlikely to be those actually selected. Earlier research had indicated that three income groups might appropriately be specified. The values of the points were \$3000 and \$9000. The values were inflated to 1978 levels. Since other researchers had presumed that behavioral changes might occur somewhere above the poverty level designated by the government,⁵ the lower point was moved from its inflated level of about \$5500 to \$8000, or about 140% of the poverty level. A similar evaluation process was followed to obtain an estimate of the upper boundary, which was set at \$16000.

Finally, the classic Engel curve may be observed and food expenditures may grow but not relatively as fast as income. That situation would imply that food is a normal good through all income ranges. This case may be represented algebraically with either the inverse form or the semilog form. The inverse approaches some level of income asymptotically. The semilog increases at a decreasing rate, but does not reach an asymptote. Prais and Houthakker (44, p. 96) conclude that the semilog form seems to provide the best fit for food expenses of those forms they evaluated. The general form is:

$$Y = a + b \log X$$

and is depicted graphically in Figure 3c.

⁵West and Price imply that 125% of the poverty level may be a threshold.

D. FORM OF THE VARIABLES

Income. Income is the variable of central interest in this model. The basic objective is to identify the alternative measure of the income variable considered by the CU when it decides how much to spend on food. If that variable can be specified, it should be possible to explain more adequately the variation in food expenditure.

For this analysis, the income variables will be calculated according to the following logic. It is commonly assumed that the CU reacts to the amount of income remaining after adjustment for income taxes. As an alternative, the CU may allocate between fixed and non-fixed expenses in a first step, based on take-home pay. This figure would be gross pay minus all deductions, such as taxes (income and FICA), insurance premiums, regular savings contributions, and union dues. A second step would involve the allocation among the nonfixed expenses. In the first step the CU decides whether purchases will be made currently or whether income will be saved so that cash purchases may be made later. If the CU does not have sufficient resources to pay for items in the current time period, then a decision to consume now will require the use of credit, and the credit arrangement will affect the fixed budget expenses.

Race. The variable race will be included as a dummy for white and nonwhite groups of households -- the decision as to group assignment will be on the basis of race of the head of the CU.

CUSIZE. The justification for inclusion of the variable representing consumer unit size was given in Chapter IV.

Education. The level of educational attainment is one indicator of the social status of the CU. It is to be expected that consumption

will be higher as values of this variable increase. The groupings specified in the BLS data are reported below; the Knox data was coded to be consistent with the BLS.

GROUP 1 - some grade school complete

GROUP 2 - some high school complete

GROUP 3 - high school graduate

GROUP 4 - some college complete

GROUP 5 - college graduate, graduate work

GROUP 6 - none, not reported

Occupation. Occupation of CU head is another variable which may have connotations of social status, and consumption may differ between the groups because there are different motivating factors and peer groups. The variable was broken into 12 groups in the BLS survey, but was recoded into 6 groups. The groups seemed to be similar enough that the gain from having fewer dummy variable groups more than offset whatever loss of information resulted.

GROUP 1 - not working, not retired, other

GROUP 2 - clerical, sales, service workers

GROUP 3 - self-employed, salaried professional and technical,
and salaried managers and administrators

GROUP 4 - craftsmen, operatives

GROUP 5 - retired

GROUP 6 - unskilled laborers

LC. The stage in the life cycle of the consumer unit probably has an effect on the consumption of both FOOD and FIXED expenses. For example, a CU with only two adults probably has different patterns of consumption than does a CU with two adults and teenage children. In

order to have some measure of this effect, a variable has been created which is the interaction between AGEHD and CUSIZE. This distinction ignores any differences which might occur because of variables such as sex of household head.

In this analysis, LC may be viewed as some estimate of the stage in the life cycle of the CU because it is a measure of the combined effects of the two variables. The biological changes expected to occur and to cause differences in consumption of different ages should be captured through inclusion of the age of the CU head. CUSIZE evaluates increases in size in terms of the effect on consumption. When these two interact, one would expect that the values would be low for lower values of AGEHD and would increase as both AGEHD and CUSIZE increase. The rate of increase would begin to slow as CUSIZE begins to decline due to the departure of children and would probably eventually show an actual decrease in the value of the variable.

The form of the variable actually used in the model deviates somewhat from the earlier discussion. The correction for heteroscedasticity, discussed in a previous section, produced a series of observations for both variables that is small (<1). Multiplication of these decimals produces an even smaller set of numbers, a set that is always declining. Hence, the signs associated with the results of the regression model would be negative at low income, still negative but smaller in an absolute sense in the next income group, and probably positive in the highest income group.

RENTOWN. The groups were:

GROUP 1 - owners

GROUP 2 - both owner and renter

GROUP 3 - renters

The assumption here is that consumption patterns differ between CUs who own their principal residence and those who rent. The dummy variable groups are included to test that assertion.

OCCSPO. Occupation of the CU spouse is based on the same groupings and justification for inclusion as occupation of CU head.

NUMINC. The number of income sources may have some effect on consumption expenditure. Multiple sources may be viewed as inherently more stable than a single source. A problem here is the definition of income sources. More specifically, should incomes which are only a supplementary source of funds be equated with more basic sources, such as income from an occupation. The formulation used as the basis for determining the value was to include only income sources accruing to the CU head and spouse. Other sources may be small and may provide only pocket money to individual CU members.

AGEHD. This variable should reflect biological changes in consumption patterns associated with age differences.

BULKSTOR. The CUs food expenditure may be lower because of the economics resulting from purchase and home storage of food in bulk quantities.

E. TESTS OF DUMMY VARIABLE GROUPS

The model specification attempted to impose as few restrictions as possible on the estimating procedure. One source of restrictions

was the knowledge of the researcher about the subject. This knowledge permitted the statement that some variables had no discernable effect on the dependent variable, and inclusion was not necessary. This action implied that the variable coefficient was 0. On the other hand, no consensus had developed with respect to whether some socio-economic variables should be included, so a test to determine whether they should be retained was conducted.

All of this group of variables were qualitative, so it was possible to impose 0 restrictions on either slope or intercept dummies. The usual approach had been the imposition of the restriction on slopes, while intercepts were allowed to shift. From the theoretical viewpoint, slopes may also vary, and it is not difficult to allow them to do so.

The test evaluated the hypothesis that the deletion of the group of dummy variables (either slopes or intercepts) did not significantly increase the error sum of squares from the regression. If the hypothesis was not rejected, then this group was deleted from the model.

These tests of restrictions on RACE, OCCUP, EDUC, OCCSPO, and RENTOWN were conducted as deletions from a completely specified model to avoid specification bias. Preliminary results indicated that all intercept groups were significant, so they were included in the model. The tests of slope coefficients revealed that RACE and OCCSPO were not significant, so they were deleted from the model. Slopes for EDUC, OCCUP, and RENTOWN were included.

F. IDENTIFICATION

The proposed equation system shown earlier in general form is a simultaneous one, which means that the disturbance terms are not

independent and that some procedure besides OLS must be used to estimate the parameters. Another reason for an alternative estimation procedure is that in order for the equations to be estimated, each must meet a condition which specifies that it excludes information contained in the other equations. This excluded information will allow other equations to shift while the first remains constant. This procedure defines the curve for the particular equation, or the endogenous variable it seeks to explain. This is the essence of the problem of identification -- that in order for an equation to be identified, it must contain information not contained in other equations. One form, a commonly accepted and necessary condition, of the identification rule is that the number of explanatory variables excluded from the equation must be at least as great as the number of endogenous variables included in the equation minus one (28, p. 274; 40, pp. 271-276). The variables are not to be deleted at random, but the process must be based on the internal logic of the system, or the a priori information which makes it possible to identify specific curves.

The equation system defined for the BLS data set contains the endogenous variables FIXED and FOOD. Inspection of the exogenous variables suggests that RESINC may be deleted from the first equation because the CU is assumed to look at the INCOME variable rather than at the RESINC variable when it makes the decision about how much of available resource will go into the fixed category. RESINC is that amount left over after the fixed expenses are paid, so it would be reasonable that since the residual income can be changed, if required, to make additional funds available for the payment of fixed expenses, RESINC could be deleted. In addition, RESINC plus FIXED are equal to INCOME, so a

statistical problem exists if both are included. The other variables are expected to have some influence on the FIXED variable, so they should be included in the equation. With the deletion of RESINC, the equation is over identified.

With regard to the second equation in the system, the objective is to estimate the amount spent on food. In this case, we assume that the CU looks at the RESINC to decide how much to allocate to food, rather than at INCOME. This is because the logic of the research hypothesis involves the actual measure of income the CU evaluates to make decisions about the various parts of its budget, and the ideas of the fixed and residual portions of income are very important to that decision. This should allow the variable INCOME to be excluded from the equation. Some of the other variables may be excluded as well. The distinction between renters and owners seems to offer no additional information because either source of expense enters in RESINC in the same way. Food will be purchased from this fund regardless of whether it is a residual with respect to rent or mortgage payments.

Occupation and education of the CU spouse need not be included in the equation. The number of incomes in the CU may be deleted as well because a family must purchase food, and total income seems to be more important than the expected stability of income.

With the deletions, the equations are over identified for the BLS model:

$$\text{FIXED} = f(\text{FOOD}, \text{INCOME}, \text{CUSIZE}, \text{AGEHD}, \text{LC}, \text{RACE}, \text{EDUC}, \text{OCCUP}, \\ \text{RENTOWN}, \text{OCCSPO})$$

$$\text{FOOD} = f(\text{FIXED}, \text{RESINC}, \text{CUSIZE}, \text{AGEHD}, \text{LC}, \text{RACE}, \text{EDUC}, \text{OCCUP})$$

The identification procedure for the Knoxville model proceeds along the same lines, with analogous arguments for the inclusion and exclusion of variables.

FIXED = f(FOOD, INCOME, CUSIZE, AGEHD, LC, RACE, EDUC, OCCUP,
RENTOWN, NUMINC)

MEATS = f(FIXED, FOOD, RESINC, CUSIZE, AGEHD, LC, RACE, EDUC,
OCCUP, BULKSTOR)

DAIRY = f(FIXED, FOOD, RESINC, CUSIZE, AGEHD, LC, RACE, EDUC,
OCCUP, BULKSTOR)

PRODUCE = f(FIXED, FOOD, RESINC, CUSIZE, AGEHD, LC, RACE, EDUC,
OCCUP, BULKSTOR)

GROCERY = f(FIXED, FOOD, RESINC, CUSIZE, AGEHD, LC, RACE, EDUC,
OCCUP, BULKSTOR)

MEALS OUT = f(FIXED, FOOD, RESINC, CUSIZE, AGEHD, LC, RACE, EDUC,
OCCUP, BULKSTOR)

CHAPTER VII

RESULTS

Previous chapters have been devoted to a detailed specification of the research objectives and research hypothesis, the procedures followed in the manipulation on the data file, and the confrontation of certain problems associated with the estimation procedure. In this chapter the results from the BLS and the Knox models will be presented. These will be:

1. An analysis of the three algebraic forms of the Engel Curve.
2. The analysis of the relationship between FOOD and FIXED on the basis of:
 - a. the percentage change from one group mean to the next, and
 - b. the regression results and elasticity estimates from the Bureau of Labor Statistics (BLS) and the Knox models.

A. ALTERNATIVE FORMS OF ENGEL CURVE

Functions which describe the relationship between income and consumption are called Engel curves. Here, the two variables are defined as follows: 1) income is taken as a residual, the difference between take-home pay and the fixed or committed expenses of the CU (described earlier), and 2) since different food items cannot, in an overall sense, be expressed as a single quantity, expenditure on food is used as an aggregative measure.

Three functional forms were selected for use in evaluating the behavior of Engel curves. These were discussed at some length in an earlier chapter. They were the semilog, the cubic (an attempt to generate an estimate of the income levels at which the preference structure changes), and the linear (by income groups to allow slope changes between groups).

The functional forms were evaluated basically on the a priori expectations from the research hypothesis. Some considerations for Engel curves in general are "that there is an initial income below which a commodity is not purchased, and, that there is a satiety level, that is, a maximum to the quantity of the commodity consumed which is not exceeded no matter how high income may rise" (44, p. 82).

The first of these criteria is probably not applicable, since food is a necessity. The second is applicable, since there is a maximum in terms of quantity of food, that an individual will consume. However, if one thinks in terms of a substitution of quality for quantity, a satiety level for food expenditure would be quite different.

It has been hypothesized that the linear and cubic forms will prove useful in the explanation of food expense and the CUs fixed expense. The linear form has been corrected for heteroscedasticity and may not be tested directly against the semilog and cubic. The semilog was not corrected because the studies referenced did not apply the transformation, so it would not be possible to compare them to the current study if the semilog estimation had used the corrected data. The cubic form was not estimated with the corrected data because of the desire to evaluate it with respect to the semilog. And, the basic objective with that form is to get an idea of the points of change in the data. The

main interest is in the distinction between IG1 and IG2 (income groups 1 and 2, respectively), and it might not be an extremely unrealistic assumption to say that the variances are fairly homogenous at the lower incomes.

Another problem with testing the transformed data against the untransformed data lies with the R^2 statistic. This statistic, the commonly used measure of goodness of fit, must utilize the same dependent variable. This is because the unexplained sum of squares, a component of the R^2 , cannot change if there is to be consistency of measurement.

The semilog form. Here, the natural logarithm of residual income (RESINC) was taken to form the variable XLOG. It has been suggested earlier that Engel curve estimation must account for any correlation between income and the size of the consumer unit. Hence, CUsiz was included in the regression. The model is:

$$\text{FOOD} = -3587 + 527 \text{ XLOG} + 291 \text{ CUSIZE} \quad R^2 = 0.38$$

(-15) (29) (17)

TABLE 6. ELASTICITY OF FOOD EXPENDITURE WITH RESPECT TO
THE NATURAL LOG OF RESINC, ESTIMATED FROM SEMILOG
ENGEL FUNCTION, BY INCOME GROUPS

Income Group	Elasticity ^a
1	0.535
2	0.341
3	0.217

^aThe elasticity estimating formula is b/Y ,
where:

b = value of the estimated regression coefficient
for the natural log of RESINC.

Y = mean value of the dependent variable FOOD in
income group i , $i = 1, \dots, 3$ (25, p. 62).

Data Source: 1972-73 Bureau of Labor
Statistics Survey (58).

Tomek reported some representative elasticities from studies at different periods in time, as early as 1936 and as late as 1972. The reported range was from 0.08 to 0.52 (55, p. 18). He included both expenditure and income elasticities, and a range of geographical areas and family types. These differ from the results reported on page 94 in that the value of residual income should be smaller than the other measures of income used. That difference should make the estimates reported here larger. The results indicate that this study was not out of line with those earlier efforts with respect to magnitudes of elasticities, and they were in the higher range of those estimates as expected.

The cubic function. The estimated model is

$$\text{ALLFOOD} = 384.2 + 62.5 A + 0.789 A^2 - 0.01 A^3 + 306.6 \text{ CUSIZE}, R^2 = 0.39$$

(7.4) (11.1) (7.5) (-4.6) (18.6)

() = t scores

A is RESINC/1000. In this case, the R^2 was marginally better than the semilog form. Although all the estimates were significant at the 0.001 level, the results were not encouraging on an a priori basis. The graph in Figure 4 shows that, as expected, the food budget initially increased at an increasing rate. It continued to behave in the same manner until residual income reached about \$26,000, where an inflection point was reached. The only local extremum at positive income was somewhere in the range of \$80,000. This result was illogical, and was most likely the result of the few observations in the very high income range. On the basis of these results, it seemed logical to disregard this functional form as a rational explanation of the consumer's allocation of resources to food.

The linear function. This group of estimates was estimated from data corrected for heteroscedasticity. The procedure was to apply the Goldfeld-Quandt test to determine if the problem was indeed present, and then carry out the data transformation. The transformation itself was accomplished by simply dividing each observation's values by the variable RESINC, which was assumed to cause the variance problem. The data were changed and interpretation of the estimated coefficients must be adjusted accordingly. In the former case, the representation was in terms of dollars or some other physically identifiable unit. Now, the data has undergone a sort of deflation process, as each becomes a proportion of

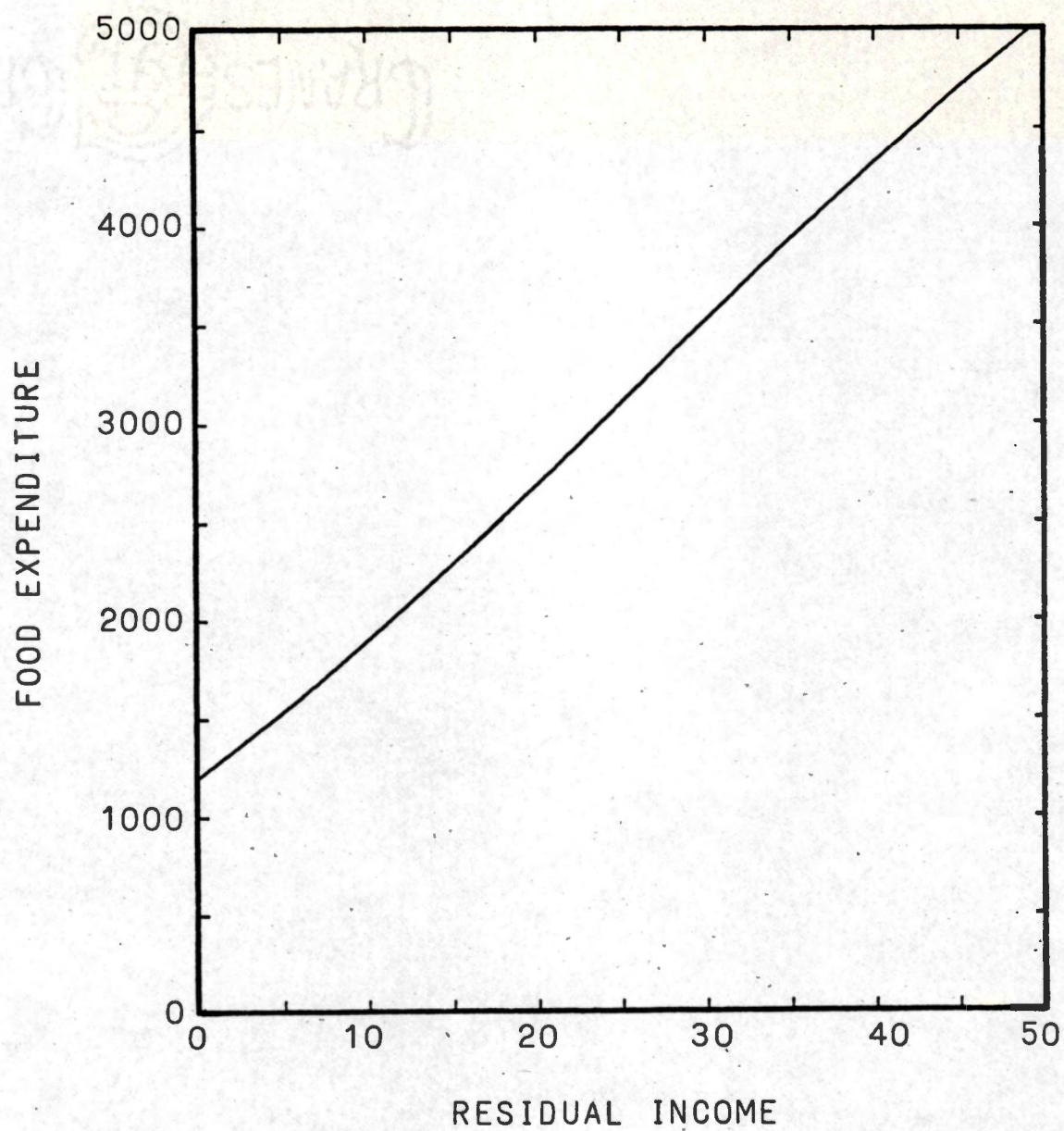


Figure 4. Cubic Engel Curve Calculated with Residual Income.

residual income. They have the same meaning in terms of the effect of the explanatory variable on the dependent variable.

One further comment on the transformation is in order. Now, when the variable RESINC (the assumed source of the heteroscedasticity) is divided by itself, the result is a column of 1's. The usual intercept term, which is a column of 1's, has been divided by the income series, just as have all the other variables. If an attempt were made to retain RESINC in the equation, a second column of 1's would be inserted by the estimation procedure, a perfectly collinear set would be formed, and the procedure would break down. To resolve this problem, RESINC has been omitted and the variable created from the intercept term (INTA) has been included. The estimation now appears in the following form:

$$Y = a + b_1 \text{ INTA} + B_2 \text{ CUSIZE} + u$$

However, the INTA term may still be interpreted as the true intercept, and the a will be interpreted as a marginal propensity to consume FOOD from RESINC (MPC_{food}).

In terms of the new units which have been introduced, it is possible to get an idea of the behavior of FOOD in response to changes in RESINC. The MPC_{food} is the vehicle used in the evaluation. Because of the fact that previous research efforts have been in terms of some other specification of income, the results of those efforts have little to say about the behavior which might be expected from RESINC. From knowledge of the expected behavior of consumers and the preference for (or against) food, some expectations may be derived.

In lower income groups, the FOOD budget share constitutes a relatively larger portion of the budget, and the FIXED share would be

expected to be a relatively smaller portion. Other goods and services still exert an influence on the budget as well, and this must be considered. But, because FOOD is necessary consumption, and is a larger portion of the budget at lower incomes, it would seem that FOOD would have a relatively strong effect at those low incomes.

The forces which cause FOOD to have a stronger influence on the budget at lower incomes may be relaxed in higher income groups. As income grows, the residual may grow as well, hence the response of FOOD to RESINC may be less. This is somewhat the same reason that heteroscedasticity is a problem with studies involving income.

In the high income group, the relationship is expected to be weakest because FOOD becomes a progressively smaller proportion of income and because FIXED is assumed to grow less as a percentage of income than the middle group. The expected effect of changes in INCOME is that consumers respond by maintaining some level of RESINC in the low income groups. The next income level should see RESINC increase less than FIXED, and there might be even a decrease in that value. In the high income group, RESINC and FIXED should both increase, with the former increasing faster than it did in the intermediate income range.

The estimated coefficients (MPC_{food}) showed this progressively less responsive reaction of FOOD to RESINC. In terms of relative positions, the low income group had the largest effect at 0.39, IG2 had an intermediate effect at 0.24, and the high income group had the low effect at 0.04 (Table 7). This pattern conformed to the effects stated above, and it appeared to agree with the semilog form of the Engel function.

TABLE 7. ESTIMATED COEFFICIENTS, t-RATIOS, AND R^2 FOR LINEAR ENGEL CURVES BY INCOME GROUP

Income Group	Intercept	Variable	CUSIZE	R^2
		<u>MPC_{food}</u>		
1	154 (1.6) ^a	0.39 (0.6)	-45 (-1)	0.75
2	617 (4.0)	0.24 (6)	99 (4.7)	0.3
3	367 (2)	0.04 (2)	350 (11)	0.37

^aNumbers in parentheses are t-ratios.

Data Source: 1972-73 Bureau of Labor Statistics Survey (58).

Other evidence about the FOOD-FIXED relationship may be obtained from the mean values of the income groups, and from the percentage change between mean values of the groups. This discussion will be presented in the next section, where that relationship will be the major topic.

Estimates of elasticities from the linear models were calculated. These estimates may be criticized on the grounds that they were rather large, most obviously at the lowest income group. However, due to the data transformation the values represented elasticities of proportion. A very strong preference for the consumption of FOOD was in evidence at the low income level, where the value was 1.8 (Table 8). The values moved lower in the next two income groups, and seemed to be generally in the range reported earlier. Another comment is that the MPC_{food} used to calculate the value at IGI was not statistically significant. The other

TABLE 8. ELASTICITY OF FOOD EXPENDITURE WITH RESPECT TO RESINC BY INCOME GROUPS, ESTIMATED FROM LINEAR ENGEL FUNCTION

Income Group	Elasticity ^a
1	1.8
2	0.8
3	0.2

^aThe elasticity estimating formula is $MPC_{\text{food}} \cdot \frac{\bar{X}}{\bar{Y}}$,
where:

MPC_{food} = the marginal propensity to consume food from RESINC

$\bar{X}$ = mean of RESINC in income group i , $i = 1, \dots, 3$

$\bar{Y}$ = mean of FOOD in income group i , $i = 1, \dots, 3$ (26, p. 62).

Data Source: 1972-73 Bureau of Labor Statistics Survey (58).

MPC_{food} 's were significant. Therefore, there was evidence that the low income estimate may not be as accurate as the other two.

The values moved downward as expected. If the elasticities move as great a distance as the estimates indicate, then the demand for expenditure on additional FOOD is much less at IG3 than at IG1.

In summary, the specified forms of Engel curves varied widely with respect to results. The cubic form was disappointing because it did not reflect a declining proportion of expenditure allocated to food as the residual income increases. The linear and the semilog forms produced estimates which were in general accord with earlier empirical results.

The semilog has become the standard, because previous research had, in more than one case, given it the distinction of being termed "best" of the alternative forms evaluated. Because results summarized in Table 8 conformed to earlier efforts, confidence in this income specification and in the data set is gained.

The linear form provided support for the research hypothesis. The coefficients changed in a way consistent with the semilog form, and the elasticities for the two forms became more nearly similar as income increased. Both the forms seemed to indicate that substantial changes took place in the consumption function as consumers moved between income ranges. Whether this change in FOOD was related to FIXED as well as RESINC was the objective to be evaluated in the next section of the report.

B. RELATIONSHIP BETWEEN THE FOOD AND FIXED EXPENSES

The evaluation of the relationship between the two categories of expenses, FOOD and FIXED, and the appropriate measures of income, INCOME and RESINC, was an objective of this study. That evaluation was treated in a comprehensive manner through the use of linear regression techniques and the results will be presented in this section. First, some findings on this relationship but derived from the linear Engel curve will be presented.

A. Comparison of Differences Between Means

It was observed (Table 9 and associated discussion) that MPC_{food} did fit the expected pattern of change between income groups, and that this linear form appeared to agree reasonably well with the

TABLE 9. MEAN VALUES FOR FOOD AND FIXED EXPENSE BY INCOME GROUP, AND PERCENT CHANGE BETWEEN GROUPS IN EACH EXPENSE CATEGORY

Income Group	Variable				
	<u>Food</u>		<u>Fixed</u>		
	<u>Mean Value</u>		<u>Mean Value</u>	<u>Percent Change Between Groups</u>	
1	0.216	---	0.233	---	91%
2	0.291	---	0.444	---	38%
3	0.210		0.272		

Data Source: 1972-73 Bureau of Labor Statistics Survey (58).

results from the semilog form. That pattern was based on the premise that CU adjusts the budget share devoted to FOOD and FIXED, with the level of each type of expense dependent on the income group. The CU was expected to be forced to hold down FIXED in IG1, because a larger share went to areas such as FOOD, which could not be decreased very much. As income increased into IG2, and into IG3, this effect was expected to change because of some level of FOOD satiation and because more capital was available from additional income and credit.

Even though the regression coefficients from both the semilog and the linear forms of Engel curves progressed as expected as income changed, the Engel curve is not the only way to view the relationship.

More information might be gained if the change from one income group mean to the next was considered. The group means for both FOOD and FIXED were reported along with the percentage change from one group to the next (Table 9). This change supported the hypothesis that FOOD increased less slowly in the second income group than did FIXED.

The mean values of both FOOD and FIXED increased between the first two income groups and decreased between the last two. One aspect of the table was the change between groups expressed as a percentage of the group from which one measured, i.e., from IG1 to IG2 and from IG2 to IG3. In the first case, FOOD increased by 34% as opposed to an increase of 91% in FIXED. It appears that consumers prefer the additional FIXED expenses by a substantial margin over additional FOOD.

This provides some evidence of support for the research hypothesis. Although there was not an actual decline in food expenditure, the percentage increase in FOOD was much less than the increase in FIXED.

A comparison of the move between the higher income groups showed that the proportion decreased but that the drop was smaller for FOOD than for FIXED. There may have been support for the assertion that the proportion devoted to FOOD did not drop as fast as FIXED, just as it did not increase as fast. There may be some "stickiness" to FOOD expense.

It may also be noted that the proportion of residual income devoted to FOOD is about the same at low and high incomes, whereas FIXED is a larger proportion in the high income group than in the low income group.

There appeared to be some discrepancy between these results and those from the Engel curve analysis. As stated earlier, there were differences brought about by the heteroscedasticity adjustment and by

the fact that elasticities are regression coefficients adjusted by mean expenditures, and not just the means themselves.

B. 2SLS Results From the BLS Model

Coefficients. The estimated coefficients using 2SLS, with the dependent variable FIXED and the transformed data, are presented in Table 10. The number following the coefficient is the t-score.

The endogeneous variables FIXED and FOOD were presumed to come from separate budget allocation steps, and to follow a pattern determined by the income group to which the CU belongs. The expected effect here is that the endogeneous variables might increase together at low incomes but, due to a change in the consumer preference structure, FIXED could increase while FOOD might increase by a smaller proportion or even decrease at the next income level. As available resources increase from the second to the third group, the CU may have reached a point where the consumption of additional FIXED produces a decreased level of gain in utility and resources are again devoted more strongly toward food expenditure. Hence, one would expect to see a positive response from FIXED if FOOD increases in IG1, a smaller (and possibly negative) response in IG2, and a positive response in IG3.

The estimate for MPC_{fixed} was negative (-1.03) and significantly different from 0 in IG1. CUs spent less on FIXED when INCOME increased, so the additional income was available to purchase other goods and services. In IG2, the coefficient was positive and significant, an indication that FIXED increased when INCOME increased, as hypothesized. Less income was available for other expenditure. There was no significant change in IG3. The results in IG1 and IG2 did not seem to be

TABLE 10. PARAMETER ESTIMATES AND t-RATIOS FOR EXPLANATORY VARIABLES IN THE FIXED EXPENDITURE EQUATION, BLS MODEL, BY INCOME GROUPS

Variable	Income Group					
	1		2		3	
	Parameter Estimate	t-Ratio ^a	Parameter Estimate	t-Ratio ^a	Parameter Estimate	t-Ratio ^a
MPC _{fixed} ^b	*-1.03	-6.8	*0.45	2.9	0.002	0.05
FOOD	- 0.005	-1.1	*0.37	5.6	*0.92	6.4
CUSIZE ^c	-48.0	-1.0	-230.0	1.6	*-432.0	-2.4
AGEHD ^d	*-10.0	-3.5	*-15.0	-1.8	*-35.0	-3.0
LC ^e	*217.0	3.8	5611.0	0.3	22125.0	0.5
RACE ^f	*183.0	2.2	*387.0	2.2	-58.0	-0.2
Education of CU head ^g						
I1 (Intercept)	*-350.0	-1.8	1802.0	1.3	-1332.0	-1.1
S1 (Slope)	-0.02	-0.1	-0.25	-1.3	0.07	0.9
I2	*-333.0	-1.9	*2852.0	2.1	-799.0	-0.7
S2	0.09	0.7	*-0.36	-1.9	0.06	0.7
I3	*-546.0	-2.4	1506.0	1.3	*-1497.0	-1.8
S3	0.13	0.9	-0.21	-1.2	0.07	1.2
I4	642.0	2.2	2899.0	1.8	*-1468.0	-1.7
S4	-0.2	-1.1	*-0.36	-1.7	0.07	1.2
I6	*-576.0	-2.7	1764.0	0.9	405.0	0.2
S6	-0.005	-0.03	-0.25	-0.9	-0.02	-0.1

Table 10 (continued)

Variable	Income Group					
	1		2		3	
	Parameter Estimate	t-Ratio ^a	Parameter Estimate	t-Ratio ^a	Parameter Estimate	t-Ratio ^a
Occupation of CU Head ^h						
I1	*-4066.0	-24.0	-217.0	-0.2	-771.0	-0.6
S1	*1.25	11.5	0.03	0.2	0.07	0.8
I2	*-3526.0	-16.1	-434.0	0.4	-1100.0	-1.4
S2	*1.11	9.7	0.03	0.2	0.06	1.05
I4	*-3415.0	-10.4	*1733.0	1.7	-1038.0	-1.1
S4	*1.05	7.5	*-0.27	-1.9	0.07	1.0
I5	*-3614.0	-13.9	679.0	0.5	435.0	0.2
S5	*1.14	9.2	-0.16	-0.9	-0.01	-0.1
I6	*-3731.0	-10.4	-1849.0	-1.0	-30.0	-0.0
S6	*1.1	6.6	0.26	1.0	0.03	0.14
Renter or owner ⁱ						
I2	-1151.0	-0.6	*10257.0	7.2	675.0	0.6
S2	0.5	0.9	*-1.11	-5.6	0.1	1.1
I3	-228.0	1.2	14.0	0.0	3.0	0.0
S3	0.01	0.1	0.06	0.1	0.04	0.6
Occupation of CU Spouse ^j						
I1	70.0	0.6	*308.0	1.9	136.0	0.7
I2	*582.0	1.9	80.0	0.4	24.0	0.1
I4	255.0	0.5	-99.0	-0.3	-245.0	-0.7

Table 10 (continued)

^a*indicates statistical significance at the 0.10 level.

^bMarginal propensity to consume.

^cConsumer unit size.

^dAge of head of consumer unit.

^eStage in life cycle.

^fRace: Whites and other nonblacks = 0
Blacks = 1

^gEducation of CU head:

- Group 1 = some grade school complete.
- Group 2 = some high school complete.
- Group 3 = high school graduate.
- Group 4 = some college complete.
- Group 5 = college graduate, graduate work.
- Group 6 = none, not reported.

^hOccupation of CU head:

- Group 1 = not working, not retired, other
- Group 2 = clerical, sales, service workers
- Group 3 = self-employed, salaried professional and technical,
and salaried managers and administrators
- Group 4 = craftsmen, operatives
- Group 5 = retired
- Group 6 = unskilled laborers

ⁱRent-own:

- Group 1 = owners
- Group 2 = both owners and renters
- Group 3 = renters

^jOccupation of CU spouse:

- Group 1 = not working, not retired, other
- Group 2 = clerical, sales, service workers
- Group 3 = self-employed, salaried professional and technical,
and salaried managers and administrators
- Group 4 = craftsmen, operatives

Data Source: 1972-73 Bureau of Labor Statistics Survey (58).

inconsistent with the consumer behavior described earlier. However, it must be recalled that the data were grouped by INCOME, so the interpretation is valid only over those limited ranges.

The effect of changes in the other endogeneous variable FOOD was not significant in IG1, but was significant and positive, at 0.37 and 0.92 respectively, in IG2 and IG3. These estimates indicated that increases in FOOD have a positive relationship to FIXED expenses.

Consumer unit size (CUSIZE) is significant in IG3, and the sign is negative (-432). The coefficient for the variable is negative in both the IG1 and IG2 income groups; however, the coefficients were not significant. One possible explanation of this outcome is that the real income of each family member decreases with the increase in CUSIZE. Thus, it would not be unreasonable to anticipate the negative sign.

AGEHD caused a significant decrease in the value of FIXED expenditure at all three income levels. The change was largest in IG3 at -35. There are two likely causes. First, a more conservative attitude toward additional debt is likely as age increases. Secondly, commitments made some years in the past do not reflect the inflated price level associated with today's terms of trade.

The variable reflecting stage in the consumer unit's life cycle (LC) was significant only in IG1. It appeared likely that multicollinearity might have been a factor here, because the estimates in IG2 and IG3 were very large but not significant.

Five qualitative variables concerned with socio-economic characteristics were specified. They were race, education and occupation of the CU head, occupation of CU spouse, and a distinction between renting CUs and those who owned a principal residence.

The excluded or base group in the case of the RACE variable was the "whites and other nonblacks." A rather surprising result was that the included group, "blacks," appeared to have higher expenditures on FIXED than the base, with positive and significant parameter estimates in IG1 and IG2 of 183 and 387, respectively.

The coefficients for educational groups showed that the slope coefficients, or the rates of change of the included groups, as measured from the excluded group "college graduates'," were not significant in IG1; on the other hand, all the included intercept changes were significant. That exception was not different from the base at any IG. In IG2, the "some high school" group and the "some college" group were significant for both slope and intercept. IG3 revealed only intercepts for "high school graduate" and "some college" groups to be different.

It was observed that the educational group coefficients were largest, in absolute terms, in IG2. An implication about behavioral changes between levels of income might be associated with the base group, "college graduate." In IG1, the intercepts were negative except for the "some college" group. In IG2, all intercepts were positive. Such a result might have been anticipated if few CUs were both "college graduate" and in the low income group. For those CUs, the lifestyle could deviate from those CUs with similar education but higher incomes.

The "salaried professional and technical" group was selected as the base for the occupational groups. With respect these occupational groups, a distinction between IG1 and IG2 was evident. Every parameter estimate in IG1 was significantly different. In IG2 only "craftsmen" was different from the base group. The various groups in this category

appeared to have less impact on FIXED as income increases. This was accentuated by the results in IG3, where none of the groups was significantly different from the base.

The distinction between renters and owners of a principal residence, in terms of effect on various consumption expenditures, had not been extensively pursued in previous research efforts. This distinction reflected the permanent income of the CU to the extent that the asset level was affected. The base group was the "owner" group. Somewhat contrary to the expected result was the finding that only "renters" was significantly different, with an intercept value of 10257 and a slope value of -1.11, and this difference was only in IG2.

The occupation of CU spouse contained no observations in the categories of "unskilled laborers" or "retired," so those groups were omitted. The "sales, service" group in IG1 and the "not working" group in IG2 were the only significantly different coefficients.

The second equation in the simultaneous system, with FOOD expense specified as the dependent variable, involved a different logic than the FIXED expenditure equation. Food was a necessary regular, and continuing expense. This regularity should have enhanced estimation, in a statistical sense. In this equation the other endogeneous variable, FIXED, was included as an explanatory variable. After an accounting was done for FIXED, the CU was expected to make a sort of second step allocation of resources. It was in this step that the influence of residual income on FOOD should have been evident. The estimates of the parameter were reflected in the variable MPC_{food} (marginal propensity to consume food from residual income).

The estimates for MPC_{food} were significant in IG1 at 0.2 and in IG3 at 0.12 (Table 11). The finding that the estimate was highest at IG2 did not agree entirely with the results of previous sections dealing with Engel curves, where the response of FOOD was highest at IG1. This might have resulted from the simultaneous solution of the equations, or it might have been the case that it would be possible for the CU to increase spending for FIXED relatively more than FOOD, since the expenditure system was not complete.

The pattern of the response of FOOD to changes in FIXED across the income groups appeared to support the assertion that there was a relationship at least somewhat similar to the one specified in the problem statement. The coefficient showed the expected sign in IG2, but was ambiguous in IG1 and "wrong" in IG3.

In IG1, there was essentially no change in FOOD in response to changes in FIXED (-0.02); FOOD was essentially stable regardless of changes in FIXED. In IG2, the relationship was inverse, and the coefficient was significant (-0.4). The result was the same in IG3, but the estimate was not as large (-0.33).

Changes in CUSIZE, which represented the added expense to the CU of the addition of an individual in a particular category, seemed likely to produce changes in the same direction in FOOD. Additional expense could have been generated through induced consumption on the part of other CU members, and on the other hand, some economics of scale could have been introduced. The estimates show that CUSIZE was positive and significant at all three income levels, with values of 195, 345, and 397 respectively. Hence, the expected direction was the one revealed by the coefficients, and the effect seemed to be strong.

TABLE 11. PARAMETER ESTIMATES AND t-RATIOS FOR EXPLANATORY VARIABLES IN THE FOOD EXPENSE EQUATION, BLS MODEL, BY INCOME GROUPS

Variable	Income Group					
	1		2		3	
	Parameter Estimate	t-Ratio ^b	Parameter Estimate	t-Ratio ^b	Parameter Estimate	t-Ratio ^b
MPC _{food}	0.09	0.4	*0.2	3.6	*0.12	4.2
FIXED	-0.02	-1.6	*-0.4	-4.6	*-0.33	-2.3
CUSIZE	*195.0	5.9	*345.0	10.7	*397.0	6.3
AGEHD	*-5.0	-2.5	3.1	1.1	2.9	0.5
LC	*-335.0	-36.0	*-3441.0	-5.7	-5852.0	-0.5
RACE	*710.0	13.8	36.0	0.3	*383.0	2.3
Education of CU Head						
I1 (Intercept)	96.0	0.6	-262.0	-0.9	139.0	0.2
S1 (Slope)	0.2	0.8	-0.01	-0.2	-0.05	-0.8
I2	*850.0	5.3	*-693.0	-3.2	182.0	-0.4
S2	-0.35	-1.3	*0.12	2.3	-0.03	-0.6
I3	*425.0	2.9	-277.0	-1.4	359.0	1.1
S3	0.04	0.2	0.006	0.1	-0.03	-0.9
I4	178.0	1.2	*-1119.0	-3.3	*862.0	-2.4
S4	-0.02	-0.1	*0.2	2.6	*-0.07	-1.9
I6	-197.0	-1.1	-463.0	-1.2	-14.0	-0.03
S6	0.4	1.4	0.03	0.4	-0.02	-0.4

Table 11. (continued)

Variable ^a	Income Group					
	1		2		3	
	Parameter Estimate	t-Ratio ^b	Parameter Estimate	t-Ratio ^b	Parameter Estimate	t-Ratio ^b
Occupation of CU-Head						
I1	209.0	1.8	252.0	0.7	-398.0	0.7
S1	0.2	1.0	-0.04	-0.6	0.01	0.3
I2	203.0	1.5	*805.0	4.3	157.0	0.5
S2	0.003	0.0	*-0.17	-3.5	-0.01	-0.3
I4	*-810.0	-6.8	34.0	0.2	-448.0	-1.1
S4	*0.5	2.6	-0.02	-0.4	0.04	0.9
I5	*443.0	3.3	*919.0	2.1	1143.0	0.9
S5	-0.09	-0.5	*-0.21	-2.4	*-0.17	-1.6
I6	*-246.0	02.4	*664.0	2.6	-1078.0	-0.8
S6	*0.5	2.7	*-0.2	-2.5	0.07	0.5

^aSee footnotes to Table 10, page 89, for variable definitions.

^b*indicates statistical significance at $\alpha = 0.10$ level.

Data Source: 1972-73 Bureau of Labor Statistics Survey (58).

FOOD expense decreased significantly in response to an increase in AGEHD in IG1, but there was no significant difference in IG2 and IG3. Increases in income appeared to cause AGEHD to become a neutral influence as income increases. The life cycle variable (LC) was significant in IG1 and IG2 and was negative in all three income groups. This conformed at least to some extent with a priori expectations with respect to the change in FOOD associated with increased income. In other words, it appeared that as LC increased, the influences which cause FOOD to decrease, such as the decrease in real income and other forces, dominated.

The effect of the intercept change between races reveals that "blacks" had higher food expenditures regardless of income group. The coefficient was not significant, however, in IG2.

Education of the head of the CU might indicate some changes in behavior of expenditures between income groups. In IG1, two intercepts were significant. Except for the "no education" group, however, the intercepts were positive. IG2, on the other hand, revealed that all intercepts were negative, although only two were significant. It might be inferred that the former result was due to the behavior pattern described earlier, where highly educated CUs could have been at low income levels because they subscribed to a different set of values. In IG2, all intercepts were negative, although only two were significant. This might be interpreted as a change on the part of the base group, or on the part of all other CUs, but it seemed to signal a behavior change from IG1 to IG2. In IG3, the patterns reverted to essentially the same pattern as IG1.

The slope changes associated with different educational groups were seldom significant and showed little pattern with respect to signs.

As was the case with the FIXED expenditure equation, there were more significant differences from the base in the occupation groups. In IG1, all the intercepts were significant except "sales, clerical" and "not working," but the signs varied. In IG2, the "not working," "retired," and "unskilled" groups were significant and positive. In IG3, the only significant occupational group was "retired," and the signs were mixed.

For slope changes in the occupational groups, in IG1 only "craftsmen" and "unskilled" were significant, both at 0.5. In IG2, the estimates were all negative and three were statistically significant. In IG3, there was only one statistically significant estimate and the direction of change moved at random. Based on the change in slope signs, from generally positive in IG1 to negative in IG2, there appeared to be evidence that the consumption function of the base group changed, relatively, when the income level changed.

Some overall comments may be made here. Both the FOOD and the FIXED equations seemed to explain more of the variation in the endogeneous variables in IG1 and IG2 than in the higher income group. The focus of national nutritional objectives has been on lower income consumers rather than higher. In terms of policy implications, then, the relatively less fruitful results should not be a major concern.

With regard to the FIXED equation, the socio-economic variables (defined as all except the endogeneous variables, MPC_{fixed} , and MPC_{food}), had more significant differences in IG1 and IG2. The race of the CU also seemed to have more influence in this equation. It may also be noted, in regard to race, that it was rather surprising that "blacks" had higher expenditure on both FOOD and FIXED whenever the differences

were significant. The lower income and asset levels generally attained by blacks would lead to the conclusion that their expenditure levels should be lower. The variable definition, in which whites were classified with all other nonblacks, may have had an effect on this result.

The most important results, however, were those relating to the measures of income on the endogenous variables and their effect on each other. In the FIXED equation, INCOME (reported as MPC_{fixed}) seemed to be diverted from FIXED expenditure in IG1, but toward FIXED in IG2. This was consistent with the notion of an increase in available funds being used for FIXED when some level of FOOD and certain other goods had been attained. The response of FOOD to RESINC (MPC_{food}) was really an unknown quantity, but it appeared to be rather well behaved, in terms of the expectations from the Engel curve analysis. The exception is the coefficient in IG1, which was the lowest of the three groups. The research hypothesis focused most specifically on the responses of FOOD and FIXED to changes in the value of each other. In IG1, neither seemed to have an effect on the other. In IG2, the results were encouraging. There, a one unit change in FIXED caused a -0.4 unit change in FOOD. As FIXED moved upward, FOOD was adversely affected, i.e., FIXED was substituted for FOOD in IG2. In IG3, the effect of a one unit change in FIXED on FOOD was still -0.33 units -- the expected effect was that FIXED and FOOD might both increase. Thus, it appeared that in IG2 a change in the pattern of expenditures similar to the one predicted from the research hypothesis was evident.

Elasticities. As was the case with the regression coefficients, the concern with elasticities was in the pattern of change across the spectrum of income. Regression coefficients are defined as the change in the dependent variable as some independent variable is changed by one unit, with the other variables held at some level. The elasticities, on the other hand, remove the effect of the units in which variables are measured by changing to measures which are based on percentages. Thus, they are directly comparable. Some implications regarding the relationship between FOOD and FIXED may be more evident with elasticities than with regression coefficients.

The elasticity of FIXED with respect to INCOME showed a negative percentage change (-1.54) in IG1 as income increased 1%, an indication that the good was inferior (Table 12). The analogous value in IG2 was positive (1.64), indicating that it was a normal good. Both these values were greater than 1 in absolute value, so they were elastic. The designation of goods as normal or inferior and elastic or inelastic said nothing directly about the relationship between the variables FIXED and FOOD. The type of good and the degree of elasticity did indicate that these estimates were consistent with the notion that FIXED may be adjusted to allow different levels of FOOD expense.

The elasticity of FIXED with respect to FOOD, on the other hand, did not follow the pattern expected if changes in FIXED were influenced by changes in FOOD. In this case, a positive progression through the income ranges existed. It appeared that FOOD was not competitive with FIXED in any income group (the -0.01 value in IG1 seemed too small to alter this conclusion).

TABLE 12. ELASTICITIES^a FOR FIXED EXPENDITURE AND FOOD EXPENSE
WITH RESPECT TO SELECTED VARIABLES, BY INCOME GROUPS

Variable	Income Group		
	1	2	3
		<u>Fixed</u>	
MPC _{fixed}	-1.54	1.64	0.01
FOOD	-0.01	0.28	0.75
CUSIZE	-0.07	-0.30	-0.44
AGEHD	-0.53	-0.34	-0.53
LC	0.16	0.04	0.09
		<u>Food</u>	
MPC _{food}	0.41	0.68	0.59
FIXED	-0.02	-0.60	-0.44
CUSIZE	0.80	0.59	0.58
AGEHD	-0.64	0.09	0.06
LC	-1.37	-0.47	-0.03

^aSee footnote to Table 8, page 82, for elasticity formula.

Data Source: 1972-73 Bureau of Labor Statistics Survey (58).

The elasticity with respect to both CUSIZE and AGEHD indicated that FIXED was inferior with respect to either. For CUSIZE, the effect of the addition of a family member on real income caused FIXED to decrease by larger percentages as income went up. AGEHD and LC seemed rather consistent from group to group.

The estimates of FOOD elasticities with respect to the independent variables revealed some interesting implications. With respect to RESINC, FOOD was a normal good; the highest value was 0.68 in IG2. There appeared to be a relationship between FOOD and the CUs residual income.

The effect of FIXED on FOOD, particularly in IG2, supported the basic research hypothesis. In IG2, there was a decrease of -0.6% in FOOD when FIXED increased by 1%. This supported the contention that consumption behavior did change regarding the distribution of income between FOOD expenses and FIXED expenditure. Note that the effect on FOOD was stronger in IG2 than in the other income groups. The effect of FIXED on FOOD in IG1 was negligible; in IG3, that effect was -0.44, an indication that FIXED may have had a less pronounced effect at higher incomes.

Food was a normal good with respect to CUSIZE at all income levels. It was normal in IG2 and IG3 with respect to AGEHD, and inferior in IG1. The stage of the life cycle variable was consistently negative, but approached 0 in IG3. From the elasticities for both FIXED and FOOD, there was general support for the following assertions: 1) FIXED had an effect on FOOD, the relative effect changed between income groups, and was strongest in IG2, 2) MPC_{fixed} changed in a

pattern consistent with the idea that a weak inverse relationship may have existed between FIXED and FOOD in IG1, but was direct in IG2, 3 that RESINC was directly related to FOOD, and that relationship was strongest in IG2, and 4) that CUSIZE, through changes in real income, prompted inverse changes with respect to FIXED and direct changes with respect to FOOD. On the other hand, FOOD did not seem to have an effect on FIXED which was consistent with the research hypothesis.

C. 2SLS Results From the Knox Model

Coefficients. The Knox model utilized data which were different from the data used in the BLS model. The Knox model, it was noted, contained detail about individual food categories.

The Knox data were collected in current dollars, which were deflated to 1972 levels so they would be comparable to the BLS data. Values were annualized because they were collected in terms of months and weeks. Nonconstant variance was again considered a problem, so the same correction technique was applied to the Knox data as was used on the BLS data.

The lower income group yielded the poorest response to the survey, and degrees of freedom would have been a problem if the model had been applied to that group. Due to this problem, the first income group was combined with the second group for purposes of analysis. In the text, the combined group was referred to as IG1, and the high income group was called IG3.

Some changes were also made in the qualitative variables. Zero restrictions were imposed on slope coefficients to retain more degrees of freedom. In the higher income group, empty cells were found for

education (some high school complete), occupation (sales, service, and clerical workers and retired), and race.

The 2SLS estimation of the FIXED equation (Table 13) revealed that the included explanatory variables did not seem to do a very good job of explaining the variation in the dependent variable. This might have been due to statistical problems of evaluating significance with 2SLS procedures, or inadequate model specification. In IG1, none of the quantitative variables was significantly different from 0. The only significant difference in the group of qualitative variables was the "renter-owner" distinction.

In IG3, the results were not much better. At the higher income level, increases in the size of the FOOD variable caused FIXED to actually decrease (-0.18), and the change was significant. However, the other continuous variables were not significant.

Among the dummy variables, there were two significant estimates. The "high school graduates" had significantly lower FIXED expenditure. The "craftsmen, operatives" group was higher than the base group.

The effect of the number of incomes on FIXED was not significant in either income group.

The remainder of this set of equations was the individual food groups. These represented an attempt to evaluate the changing composition of the food basket as income increased, and whether the other explanatory variables were related to the changing composition of that basket.

The equation for MEATS implied that there may be differences in consumption between the income groups (Table 14). The higher income group had no significant coefficients. This latter case may have been

TABLE 13. PARAMETER ESTIMATES AND t-RATIOS FOR EXPLANATORY
VARIABLES IN THE FIXED EXPENDITURE EQUATION,
KNOX MODEL, BY INCOME GROUPS

Variable ^b	Income Group ^a			
	1		3	
	Parameter Estimate	t-Ratio ^c	Parameter Estimate	t-Ratio ^c
MPC _{fixed}	0.14	1.4	0.03	0.2
FOOD	-0.07	-0.7	* -0.18	-2.1
CUSIZE	270.0	1.4	625.0	0.5
AGEHD	-22.0	-1.3	-77.0	-1.0
LC	-3712.0	-0.4	-157419.0	-0.5
RACE	153.0	1.6	--- ^d	
Education of CU Head				
I1 (Intercept)	278.0	0.5	-4804.0	-1.6
I2	-746.0	-1.2	--- ^d	
I3	-223.0	-0.4	*-4154.0	-2.1
I4	308.0	0.5	1089.0	0.7
Occupation of CU Head				
I1	-853.0	-0.9	-1937.0	-0.8
I2	-212.0	-0.3	--- ^d	
I4	-157.0	-0.2	*4493.0	2.1
I5	314.0	0.4	--- ^d	
I6	-269.0	-0.3	476.0	0.3
RENTOWN ^e	*987.0	2.4	-2688.0	-1.7
NUMINC ^f	118.0	0.6	-671.0	-0.8

Table 13 (continued)

^aGroup 1 is equivalent to Groups 1 and 2 in BLS model.

^bSee footnotes to Table 10, page 89, for variable definitions.

^c*indicates statistical significance at 0.10 level.

^dNo respondents in the survey identified in this group.

^eRenter or owner of principal dwelling.

^fNumber of incomes in the CU.

Data Source: 1978 Knox County Survey.

TABLE 14. PARAMETER ESTIMATES AND t-RATIOS FOR EXPLANATORY VARIABLES IN THE MEATS EXPENSE EQUATION, BY INCOME GROUPS

Variable ^b	Income Group ^a			
	1		3	
	Parameter Estimate	t-Ratio ^c	Parameter Estimate	t-Ratio ^c
MPC _{meats}	*-0.11	-1.9	0.01	1.1
FIXED	*0.27	2.3	-0.003	-0.1
FOOD	*0.22	2.1	0.13	1.4
CUSIZE	-25.0	-0.4	79.0	1.5
AGEHD	*8.0	2.7	0.6	0.1
LC	*356.0	2.4	-1506.0	-1.7
RACE	-123.0	-1.3	--- ^d	
Education of CU Head				
I1	159.0	1.2	261.0	1.1
I2	149.0	0.9	--- ^d	
I3	*234.0	1.8	118.0	0.9
I4	66.0	0.5	79.0	0.7
Occupation of CU Head				
I1	*-395.0	-2.1	-81.0	-0.3
I2	163.0	1.1	--- ^d	
I4	*313.0	1.8	-85.0	0.7
I5	-175.0	-1.1	--- ^d	
I6	*-283.0	-1.8	3.0	0.0
BULKSTOR ^e	6.0	0.1	58.0	0.8

^aGroup 1 is equivalent to Groups 1 and 2 in BLS model.

^bSee footnotes to Table 10, page 89, for variable definitions.

^c*indicates statistical significance at 0.10 level.

^dNo respondents in the survey identified in this group.

^eBulk food storage capacity.

the result of statistical or specification problems, as mentioned earlier. Another possibility is that little variation existed among high income consumers with respect to these consumption categories.

These equations showed differences between IG1 and IG3 which may have policy implications. The change in MEATS associated with an increase in RESINC (MPC_{meats}) was negative and significant at -0.11 in IG1, but was not significant in IG3.

The next two variables were the values of FIXED and FOOD predicted from the procedure's first step. The response of MEATS was positive (0.27 per unit change in FIXED) and significant in IG1, and was not significant in IG3. That result seemed to indicate, particularly in IG1, that MEATS was a product strongly demanded by consumers, since they chose to increase MEATS expenditures along with increases in FIXED expenditures. The response with respect to FOOD seemed reasonable, because MEATS had traditionally been viewed as a normal good. Expenditures should have increased as total food expenditure increased. In IG3, the change was in the same direction, but was not significant.

CUSIZE was not significant at either IG1 or IG3. AGEHD was positive and significant in IG1 (8), as was the case with the life cycle variable (356).

The level of education attained was significantly different from the base for the "high school graduates" in IG1. All the coefficients, however, were positive, a possible indication that there may have been some resistance to consumption of additional quantities of meats as the level of educational attainment increased.

Again, in IG1, the occupation of the CU head caused differences in terms of consumption of meats, for "clerical, etc." (-395), for

"craftsmen, operatives" (313), and for "unskilled laborers" (-283). The differences between the income groups seemed to be smaller in IG3.

The dummy variable for race was not significant, not was the variable designed to distinguish differences between CUs who had some method of storing food in bulk quantities and those who did not.

The equation for DAIRY contained the same variables as MEATS. In this equation, however, there was no evidence of significant differences from the base groups for any of the qualitative variables (Table 15). As was the case with MEATS, the educational groups were all positive, a hint that there might have been some connection between high levels of education and lower expenditures on DAIRY, the same result as in the MEATS equation.

The level of residual income, represented by the marginal propensity to consume dairy products, showed no evidence of being significantly different from 0 at either income group.

The levels of FIXED and FOOD did have an effect on DAIRY expenditure. In the former case, an increase in the level of FIXED caused a decrease in DAIRY in IG1 (-0.06), but no significant change in IG3. FOOD increases caused increases in DAIRY in both income groups, in roughly the same amounts (0.11), and these increases were significant.

Neither CUSIZE nor AGEHD was significant at IG1, but CUSIZE was significant at IG3 (44). The interaction between these two, or the stage in the life cycle, was significant at both levels of income and was positive.

For PRODUCE, the $MPC_{produce}$ was again not significantly different from 0 in either income group (Table 16). All the other continuous

TABLE 15. PARAMETER ESTIMATES AND t-RATIOS FOR EXPLANATORY VARIABLES IN THE DAIRY EQUATION, BY INCOME GROUP

Variable ^b	Income Group ^a			
	1		3	
	Parameter Estimate	t-Ratio ^c	Parameter Estimate	t-Ratio ^c
MPC _{dairy}	0.02	1.2	-0.003	-0.8
FIXED	*-0.06	-1.9	-0.003	-0.2
FOOD	*0.11	3.9	*0.11	3.1
CUSIZE	8.2	0.5	*44.0	2.1
AGEHD	-1.1	-1.3	1.9	1.1
LC	*132.0	3.2	*1962.0	5.5
RACE	37.0	1.5	--- ^d	
Education of CU Head				
I1 (Intercept)	12.0	0.4	119.0	1.3
I2	23.0	0.5	--- ^d	
I3	26.0	0.7	13.0	0.2
I4	34.0	0.9	49.0	1.0
Occupation of CU Head				
I1	22.0	0.4	75.0	-0.8
I2	52.0	1.3	--- ^d	
I4	15.0	0.3	-68.0	-1.4
I5	48.0	1.1	--- ^d	
I6	46.0	1.1	-16.0	-0.3
BULKSTOR ^e	-22.0	-1.1	5.0	0.2

^aGroup 1 is equivalent to Groups 1 and 2 in BLS model.

^bSee footnotes to Table 10, page 89, for variable definitions.

^c*indicates statistical significance at 0.10 level.

^dNo respondents in the survey identified in this group.

^eBulk food storage capacity.

TABLE 16. PARAMETER ESTIMATES AND t-RATIOS FOR EXPLANATORY VARIABLES IN THE PRODUCE EQUATION, BY INCOME GROUPS

Variable ^b	Income Group ^a			
	1		3	
	Parameter Estimate	t-Ratio ^c	Parameter Estimate	t-Ratio ^c
MPC _{produce}	0.01	1.4	-0.002	-0.3
FIXED	*-0.05	-2.6	0.02	0.6
FOOD	*0.09	5.0	0.07	1.1
CUSIZE	*-39.0	-4.1	25.0	0.6
AGEHD	*1.1	2.1	0.3	0.1
LC	*183.0	7.2	-28.0	0.1
RACE	-3.0	-0.2	--- ^d	
Education of CU Head				
I1	*-37.0	-1.7	97.0	0.6
I2	*-57.0	-2.0	--- ^d	
I3	24.0	1.1	144.0	1.5
I4	-35.0	-1.5	28.0	0.3
Occupation of CU Head				
I1	*-107.0	-3.4	182.0	1.0
I2	*-51.0	-2.1	--- ^d	
I4	-7.0	-0.2	-39.0	-0.4
I5	*-86.0	-3.2	--- ^d	
I6	*-70.0	-2.6	80.0	0.7
BULKSTOR ^e	*-57.0	-4.4	67.0	1.3

^aGroup 1 is equivalent to Groups 1 and 2 in BLS model.

^bSee footnotes to Table 10, page 89, for variable definitions.

^c*indicates statistical significance at 0.10 level.

^dNo respondents in the survey identified in this group.

^eBulk food storage capacity.

variables were significant in IG1 but not in IG3. In IG1, the signs were consistent with expectations, with the possible exception of CUSIZE which was negative.

The same pattern of no significance at IG3 prevailed for the dummy variables. In IG1, however, the educational groups became mostly negative, indicating that the "college graduate, graduate school" may have consumed relatively more PRODUCE than the other educational groups. The lower two groups, with respect to formal education, were significantly different. Essentially the same pattern was the case for occupation, as well, with the "salaried professional, etc." significantly lower in four of the five cases.

In this equation, the bulk food storage dummy was significant. In this case, the CUs who had such storage capacity had lower expenditure on PRODUCE, a reasonable result.

The estimated coefficients for the GROCERY equation revealed the same results (Table 17), with respect to significance in IG3, as those for PRODUCE. There were no significant variables.

In IG1, the significant variables were FOOD at 0.45, AGEHD at -9, and LC at -676. It was interesting to note that GROCERY showed a larger change than did any of the other categories in response to changes in FOOD, and it was positive as expected. Age and LC were both negative.

For occupation, the three significant estimates were higher than the excluded group, the "salaried professional, etc." The bulk food variable was significant and indicated that those who had storage capacity had higher expenditures.

TABLE 17. PARAMETER ESTIMATES AND t-RATIOS FOR EXPLANATORY
VARIABLES FOR THE GROCERY EQUATION, BY INCOME GROUP

Variable ^b	Income Group ^a			
	1		3	
	Parameter Estimate	t-Ratio ^c	Parameter Estimate	t-Ratio ^c
MPC _{grocery}	-0.02	-0.5	0.02	1.1
FIXED	-0.002	-0.0	-0.13	-1.6
FOOD	*0.45	7.6	0.22	1.2
CUSIZE	53.0	1.6	138.0	1.4
AGEHD	*-9.0	-5.2	-14.0	-1.7
LC	*-676.0	-7.8	2607.0	1.5
RACE	61.0	1.1	--- ^d	
Education of CU Head				
I1	113.0	1.5	115.0	0.3
I2	136.0	1.4	--- ^d	
I3	6.0	0.1	-216.0	-0.9
I4	-12.0	0.2	15.0	0.1
Occupation of CU Head				
I1	*543.0	4.9	58.0	0.1
I2	43.0	0.5	--- ^d	
I4	-97.0	-0.9	88.0	0.4
I5	*253.0	2.7	--- ^d	
I6	*367.0	4.0	-64.0	-0.2
BULKSTOR ^e	*89.0	2.0	27.0	0.8

^aGroup 1 is equivalent to Groups 1 and 2 in BLS model.

^bSee footnotes to Table 10, page 89, for variable definitions.

^c*indicates statistical significance at 0.10 level.

^dNo respondents in the survey identified in this group.

^eBulk food storage capacity.

The estimation of the MEALS OUT equation (Table 18) was the last in the system of simultaneous equations comprising the Knox model. It appeared that different factors may have been influential in this decision, as opposed to the other food categories. For instance, in IG1, residual income ($MPC_{\text{meals out}}$ at 0.13) had a significant and positive effect on expenditures. Also significant in IG1 was the coefficient for FIXED, and the effect was negative and much larger than the analogous effect of this variable in other equations in the system.

The FOOD coefficient was significant in IG3, and was positive. It seemed that as FOOD expenditure grew, the growth of expenditure on meals outside the home was even greater than other foods within this income group.

The effect of CUSIZE was negative and significant in IG3. AGEHD and LC were not significant at either income level.

The dummy variables revealed very little difference from the excluded base groups. In terms of the direction of differences, some statements may be made tentatively. For education of CU head, the base group was higher in all cases. The same statement may be made with respect to occupation, with one exception. It may be implied from these coefficients that MEALS OUT was different, in a sense, from the food categories of MEATS, DAIRY, and GROCERY, because CUs with higher levels of education seemed to prefer higher expenditure on MEALS OUT and PRODUCE, and less on the MEATS, DAIRY, and GROCERY categories. The negative signs in occupation may have implied something about the preferences of the "salaried professional, etc." with respect to the other occupation groups. The excluded group was probably highest in terms of a social status scale, and would probably have been highest in

TABLE 18. PARAMETER ESTIMATES AND t-RATIOS FOR EXPLANATORY VARIABLES IN THE MEALS OUT EQUATION, BY INCOME GROUP

Variable ^b	Income Group ^a			
	1		3	
	Parameter Estimate	t-Ratio ^c	Parameter Estimate	t-Ratio ^c
MPC _{meals out}	*0.13	2.1	-0.004	-0.3
FIXED	*-0.24	-1.8	-0.009	-0.1
FOOD	0.15	1.3	*0.37	2.3
CUSIZE	7.0	0.1	*-226.0	-2.4
AGEHD	0.8	0.2	4.0	0.5
LC	-3.0	-0.1	-1454.0	-0.9
RACE	28.0	0.3	--- ^d	
Education of CU Head				
I1	-231.0	-1.6	*-709.0	-1.8
I2	-250.0	-1.4	--- ^d	
I3	*-286.0	-2.0	-102.0	-0.4
I4	-54.0	-0.4	-103.0	-0.5
Occupation of CU Head				
I1	-62.0	-0.3	-487.0	-1.2
I2	-217.0	-1.4	--- ^d	
I4	-243.0	-1.3	159.0	0.7
I5	-39.0	-0.2	--- ^d	
I6	-65.0	-0.4	-101.0	-0.4
BULKSTOR ^e	-22.0	-0.3	-144.0	-1.1

^aGroup 1 is equivalent to Groups 1 and 2 in BLS model.

^bSee footnotes to Table 10, page 89, for variable definitions.

^c*indicates statistical significance at 0.10 level.

^dNo respondents in the survey identified in this group.

^eBulk food storage capacity.

terms of income and education. These factors might have combined to produce the result that the CU spent less time in the kitchen and would choose to eat outside the home more often than other occupation groups.

This simultaneous equation system was an attempt to investigate the interdependencies between the FIXED and the FOOD expenses, to try to determine relevant socio-economic variables and their effect on both types of expenses, to analyze the effect of INCOME and RESINC on the types of expenses, and to take into account the effect of level of income.

In terms of the effect of RESINC and INCOME, it was not clear that these specifications were useful in a general sense. INCOME was not a significant regressor in the FIXED equation, and RESINC was significant in only two food categories, both in IG1. A possible reason for these results might be that a good deal of the variance to be explained by the model may have been removed in the correction for heteroscedasticity.

The interrelationship between FOOD and FIXED may be explored by an examination of the estimated coefficients. In the food categories in IG1, there should logically have been a negative effect between the direction of change for FIXED and the food categories, and the individual food categories should have increased as total food expenditure did. The agreement with a priori expectations and the strength of the effect should be revealing.

With regard to the first effect, between FIXED and the individual categories, the effect was significant and negative for PRODUCE and DAIRY, and not different from 0 for GROCERY. This was expected, as was the general size of the changes.

The effect on MEATS and MEALS OUT seemed to indicate that they might have been somewhat stronger than expected. For the case of MEATS, the coefficient was positive and significant -- the sign here was not expected. It indicated that the preference for meats was very strong, enough so that even in the face of increases in FIXED, CUs continued to purchase. This was consistent with the apparent preference of U. S. citizens for meat as a primary source of protein.

In the case of MEALS OUT, the coefficient was negative and significant, but far greater in magnitude than the other categories. It suggested that when the consumer felt that something had to be sacrificed, this category was most likely to be selected. It was a logical choice, because these expenditures were thought to be more elastic because they include the consumption of services besides food.

The various categories of food may be summed to get the value of FOOD. Hence, the estimates given in the tables were simply the proportions of the food budget spent on each food category. These did not sum to 1.0 because the values for food were those predicted from the first stage of the procedure.

Some general comments may be made with respect to the conformity of the data from the Knox survey to the general hypothesis. In the first place, it was not possible, for the statistical reasons given earlier, to have a distinction between poverty level CUs and those expected to be a step above poverty (moderate incomes, IG2 in the BLS model) on the Knox model. A valuable distinction was lost. There were implications to be drawn from the two income groups specified even with this limitation.

The coefficient estimates for the FIXED equation were significant in very few instances. It seemed unusual to experience such diversion between this result and the result for the BLS model, which seemed much more fruitful.

The results from the food category estimation were more useful because they indicated that differences exist between the levels of income.

Elasticities. As was the case with the BLS model, elasticities were calculated for each of the endogenous variables with respect to other selected variables. The elasticities are interpreted in the manner described in that earlier section.

The relationship between the FIXED and FOOD expense, and the interface with income measures, have been the central concern of this report. With the Knox model, the elasticity of FIXED with respect to INCOME and of FIXED with respect to FOOD showed that, in the first case, there was a rather large difference between income groups, and in the second case there appeared to be little. A change of 1% in income was associated with a 0.35 change in FIXED in the lower group, while the analogous figure in IG3 was only 0.1 (Table 19). FIXED became more inelastic at higher levels of income. The elasticity of FIXED with respect to FOOD was also inelastic at both income levels (-0.05 and -0.08). One can say that FOOD expense was slightly negatively associated with FIXED, or that FOOD was an inferior good there.

As in the case of the BLS model, the other elasticities calculated were not central to the evaluation of the research hypothesis.

TABLE 19. ELASTICITIES FOR FIXED EXPENDITURE AND FOOD EXPENSE
WITH RESPECT TO SELECTED VARIABLES, KNOX MODEL,
BY INCOME GROUP

Variable	Income Group	
	1	3
<u>FIXED</u>		
INCOME ^a	0.35	0.1
FOOD	-0.05	-0.08
CUSIZE ^c	0.34	0.40
AGEHD ^c	-0.55	-0.75
LC ^c	-0.08	-0.31
<u>MEATS</u>		
RESINC ^b	-0.43	0.43
FOOD	0.70	0.67
CUSIZE ^c	-0.1	0.69
AGEHD ^c	0.86	0.08
LC ^c	0.34	-0.33
FIXED	1.51	-0.03
<u>DAIRY</u>		
RESINC ^b	0.19	-0.18
FOOD	0.89	0.81
CUSIZE ^c	0.08	0.55
AGEHD ^c	-0.30	0.36
LC ^c	0.32	0.61
FIXED	-0.84	-0.04
<u>PRODUCE</u>		
RESINC ^b	0.01	-0.18
FOOD	0.77	0.75
CUSIZE ^c	-0.40	0.45
AGEHD ^c	0.33	0.08
LC ^c	0.47	-0.01
FIXED	-0.75	0.40

Table 19 (continued)

Variable	Income Group	
	1	3
<u>GROCERY</u>		
RESINC ^b	-0.08	0.42
FOOD	-1.58	0.54
CUSIZE ^c	0.22	0.58
AGEHD ^c	-1.06	-0.88
LC ^c	-0.71	0.27
FIXED	-0.01	-0.59
<u>MEALS OUT</u>		
RESINC	0.97	-0.21
FOOD	0.93	2.30
CUSIZE ^c	0.05	-2.26
AGEHD ^c	0.17	0.63
LC ^c	-0.01	-0.38
FIXED	-2.59	-0.10

^aINCOME is equivalent to MPC_{fixed} .

^bRESINC is equivalent to MPC_{food} .

^cSee footnotes to Table 10, page 89, for variable definition.

Data Source: 1978 Knox County Survey.

Because of this, comments will be made only when the figures appear to indicate either a particularly expected or unexpected event. In the present case, both CUSIZE and AGEHD were inelastic, and it appeared that CUs tended toward positive changes in FIXED while AGEHD indicated negative changes. Both were consistent with rational behavior in terms of the research hypothesis.

The elasticities of various expenditure categories of food were also presented (Table 19). As expected, both a priori and with regard to the regression coefficients from which they were partially derived, some rather large differences existed between these elasticities. MEATS was the first category considered. The elasticity with respect to RESINC was negative in IG1 and positive in IG3, as expected. From this, it may be said that MEAT was an inferior good with regard to this measure of income in IG1.

The change associated with a 1% change in FOOD showed that about 0.7% increase could be expected regardless of income group. An interesting point was the elasticity of MEAT with respect to FIXED. In IG1, strong evidence of preference for meat was seen. Meat was actually elastic with a value of 1.51, in terms of percentage changes, with a budget share with which it was expected to be competitive. In IG3, the change was very low. It might be postulated that some other variable, such as educational level, might have played a role here, since higher education was probably positively correlated with increasing income. This could have translated into nutritional awareness of the possible health related problems of additional consumption of meat.

CUSIZE, AGEHD and LC all had rather large changes between the income groups.

With respect to RESINC, DAIRY was inelastic but positive in IG1 and negative in IG3. Since this was a substitute for meat as a source of protein and might be considered cheaper, it may be that rational CU behavior was seen here.

The percentage change in DAIRY with respect to food was in the 0.8 to 0.9 range, and showed that an increased percentage in the food budget resulted in an increase in DAIRY expenditure.

The elasticity of DAIRY with respect to FIXED also showed what appeared to be a competitiveness when compared with MEAT. Now, in IG1, DAIRY was negative -- an increased percentage of FIXED implied an almost corresponding negative change in DAIRY, a result just opposite the one for MEAT. This seemed to be further evidence of a competitive relationship between MEAT and DAIRY.

For PRODUCE, changes in RESINC seemed to generate percentage changes which were rather small in comparison to the other categories. It seemed rather unexpected that the change would be negative in IG3, as it was. It can also be seen that PRODUCE increased with changes in FOOD. With respect to FIXED, in IG1, a change of -0.75 occurred. If FIXED and FOOD were competing for additional budget dollars, some of the increase in percentage change in FIXED might have come from PRODUCE. This changed in IG3, however, where the elasticity was positive.

GROCERY as defined in this study included anything not defined in one of the previous three categories, and excluded nonfoods. As such, it decreased with respect to all variables except CUSIZE in IG1, and was elastic for FOOD and AGEHD. In IG3, only AGEHD and FIXED were negative, and the others were normal goods.

The category of food expense with the most volatility was MEALS OUT. Such a result may be expected, because MEALS OUT included other services consumed at the same time. These other services were expected to be more elastic than food-stuffs themselves.

With respect to RESINC, the elasticity of MEALS OUT was 0.97 in IG1 and -0.21 in IG3. This seemed to be a very large change, and indicated that meals outside the home may be of greater value to CUs in lower income range, an unexpected finding. Some consistent explanation might be provided by the value of elasticity of MEALS OUT with respect to FOOD. In IG3, that elasticity value was 2.3, an indication that MEALS OUT was a very large component of additional food purchases. And, a 1% change in CUSIZE was associated with a large negative (-2.26) value of MEALS OUT. It may be the case that the FIXED budget share had become large enough that expenditures on MEALS OUT were squeezed down.

The elasticity of MEALS OUT with respect to FIXED revealed that it was an inferior good at both IG1 and IG3. In IG1, however, the change was large. Since this food category was expected to show the most movement in response to percentage changes in other variables, it seemed logical that changes in FIXED expense would produce the largest change in elasticity. In IG3, although the elasticity was negative, there was not much change associated with changes in FIXED.

CHAPTER VII

CONCLUSION

A. SUMMARY

The problem defined in Chapter I delves into the factors underlying the process in which consumers collect information about the goods and services available to them, process that information, and form a preference structure which guides the purchases they make. It postulates that the level of income of the household, the per capita level of income (a function of consumer unit size), a set of socio-economic variables, and other goods and services affect the level of consumption of a good. The goods considered in the study are aggregate categories in the budget. They are FOOD, (including food purchased and consumed at home, outside the home, and obtained without cost), and FIXED, a category which includes expenses that cannot be stopped in the short run without the loss of vested interest. It is hypothesized that because of the tendency toward current consumption and the use of credit, the consumer may reach some critical level of these FIXED expenses, and that level may have an effect on FOOD. This effect is expected in the moderate range of income. An objective of the study is to evaluate whether the changes described above exist as consumers move from the low to the moderate and into the higher incomes. Further objectives deal with Engel curves derived using RESINC, with a set of socio-economic variables and their effect on both FOOD and FIXED, and with the effect of this model specification on five individual categories of food.

Initially, the primary source of data was to be a Knox County, Tennessee survey designed to collect data specifically to evaluate the objectives mentioned above. The completion rate from that survey was low, however, so emphasis was shifted to the 1972-73 Bureau of Labor Statistics survey. This survey was designed to collect comprehensive information on the flow of income to various budget shares, to be used primarily to adjust the Consumer Price Index. Formulating a data set consistent with the hypothesis being tested required the collection, aggregation, and adjustment of the original data set. To the extent permitted by the limited nature of the sample, the data from the Knox survey were utilized.

A primary analytical tool used was the least squares regression model, with dummy variables to evaluate qualitative variables. The ordinary least squares model was used to estimate the Engel curves. Since the hypothesis stated that the share devoted to FIXED might affect FOOD, and the opposite effect might also occur, a two equation simultaneous system was specified. And, since the endogeneous variables and error term were expected to be correlated, consistent estimates were obtained by the application of two stage least squares.

The results of the application of OLS to estimate Engel curves showed that, with respect to the cubic form, the expected changes in direction did not occur. Given the hypothesis, the function should have curved somewhere in the vicinity of the low to moderate and the moderate to high income boundaries. This change would have indicated that the preference for FOOD was not as great, relatively, in the moderate range as the preference for FIXED. The opposite case was expected to

hold in the low income group. With respect to this a priori notion, the cubic form seemed less than adequate.

Both the semilog and the linear forms seemed to be better specifications. The semilog elasticity estimates with respect to RESINC conformed closely to the expected difference from previous studies; that is, they were somewhat higher than the previous estimates.

The linear case was estimated separately for each income group so that evidence of the changed relationship could be reflected. The parameter estimates here seemed to reflect much the same pattern as the semilog. The indication in both cases was that the elasticity of FOOD with respect to RESINC declined as income level increased.

The estimation of the interrelationship between FOOD and FIXED from the BLS model was evaluated through comparison of mean values and with the simultaneous system. The comparison of means with respect to the percentage change between groups revealed that the change in FOOD was not as large as the change in FIXED between either set of groups. This appeared to support the idea of a changing relationship between the variables.

The simultaneous system revealed further evidence about the FOOD-FIXED relationship. In general, it may be said that evidence exists to support the assertion that FIXED expenses had an effect on FOOD, particularly in the moderate income range, but that FOOD had no effect on FIXED consistent with the expected inverse relationship. The marginal propensity to consume FOOD in relation to FIXED changed in a manner which would allow more income to be devoted to FOOD at moderate incomes. The response of FOOD to RESINC seemed to say that a larger proportion of RESINC was devoted to FOOD in the moderate group than in

the low group. This quite likely happened because FIXED became larger, hence RESINC was smaller, in the moderate incomes.

In the FIXED equation, it appeared that the various socio-economic groups differed little from the base groups except at low income. The consumption function seemed to be little changed in dollar amounts at the high and moderate levels, regardless of occupation, education, etc. With respect to the FOOD equation, there seemed to be some distinction between the various qualitative variable groups in both the low and moderate income groups. At the high income level, the differences with respect to FOOD expenditure were less prominent, which was the same general result found with the FIXED equation.

The elasticities calculated from this group of parameter estimates showed some interesting patterns. MPC_{fixed} appeared to change in a manner which was consistent with the idea that FOOD could be increased at low incomes, and that in the moderate income range there seemed to be strong preference for additional expenditure on FIXED, as hypothesized. The elasticities of FOOD with respect to the different levels of income showed no evidence that additional expenditure on FOOD caused a decrease in FIXED expense.

The elasticities for FOOD showed that FOOD increased with RESINC increases. With the measures of income used in some earlier research, this had not always been the case. The elasticity with respect to FIXED contained a pattern which indicated that FOOD was responsive in the general pattern set out in the problem statement, except in high incomes. At low incomes, FIXED appeared to produce a very small negative change in FOOD. The moderate range contained a change which is negative, an indication that the hypothesized effect existed in this income range.

The estimation of the Knox model was an attempt to gain more specific information about the effect of FIXED on individual food expense categories. Generally, the dummy variable groups showed little ability to differentiate, except in the lower income group for PRODUCE.

The MPC_{fixed} was not significant at either income group. There did appear to be an effect on FIXED from changes in FOOD in the high income group, but the sign was not consistent with its counterpart in the BLS model.

The interesting results from the individual food categories were the MEATS and MEALS OUT groups. There was a strong preference for MEATS; in the lower income group, this estimate increased significantly even when FIXED increased. On the other hand, it decreased significantly at that level as RESINC increased. Neither coefficient was significant at the high income level.

For the dependent variable MEALS OUT, $MPC_{meals\ out}$ was positive and significant in the lower group. The response to FIXED was significant and negative. This category seems most likely to be contracted if the CU must decrease food expense.

The elasticities followed the same patterns and had the same signs. They did tend to get larger as the level of income increased because of the relative size of the dependent and explanatory variables.

B. LIMITATIONS

It has already been noted that problems existed with respect to the data sources. The Knox County, Tennessee survey response rate was much lower than anticipated. This limited response effectively deleted a proposed investigation of the effect of transitory income on food

expenditure which might be brought about by differences in the level of "fixed" expenses. In addition, distributional effects among the food categories might have been clearer with a more adequate sample size. This same problem with sample size forced the union of the low and moderate income groups, resulting in information loss. The high income group had some empty cells in the dummy variable groups.

The BLS data set contained both known and suspected limitations. The data was collected specifically to update the Consumer Price Index, and generally to provide information about how household income was spent. A problem was to garner the required specific information from the general data source. It appeared, from the estimates of Engel curves that the data were in the same general range as data sets used in other studies.

C. POLICY IMPLICATIONS

Earlier sections of this report have dealt with the evaluations of the consumer preference structure through the Engel curve and through analysis of the interrelationship between two categories of expense in the consumer's budget. Another purpose of this report is to relate the results of the statistical model to the realm of public policy. Beyond the collection of facts and inferences for strictly academic purposes, it is an attempt to use the results of the application of research methodology to assist the public decision-maker to make informed choices.

The objective of public policy must be kept in mind when one attempts to apply research findings. In the area of human nutrition, policy instruments are designed to assist consumers in the determination

of the components of nutritious meals, of economical preparation of meals, and other programs which, it is hoped, will positively affect the adequacy of the diet. The areas just mentioned are basically educational in nature. Beyond that, the thrust of nutritional programs involve income supplements to try to insure the consumer sufficient purchasing power to permit consumption of a nutritionally adequate diet.

One of the income supplements is the Bonus Food Stamp program. Through this program, low income citizens are given an amount of income which should allow them to consume an adequate diet if they follow the "thrifty plan" developed by the Consumer and Food Economics Institute. The supplement is graduated so that fewer dollars are given as income from other sources increases. It is designated that the supplement be used for food, and the supplement is allocated in the form of coupons which may be redeemed for food. The coupons may not legally be converted to cash or used to purchase nonfood items.

Some proposals in the past have called for the granting of cash payments rather than coupons to purchase food. Arguments supporting this change are usually based on the premise that resource allocation is efficiently accomplished when individuals have the freedom to purchase whatever combination of goods and services gives them the greatest utility.

If the results from the semilog form of the Engel curve are accepted, then it would seem to make little difference whether the supplement is in coupon or cash form. Increases in income would lead to changes in food expenditure, since the elasticity is always positive because of the constraint imposed by the functional form.

If, on the other hand, the simultaneous system specified in this report is the model used for analysis, a quite different conclusion can be reached. Referring to Table 11, page 94, it appears that FOOD expense is stable in IG1. In IG2, in contrast, the effect of FIXED on FOOD expense is -0.4, a significant difference from 0. The conclusion which may be drawn is that if the income of the CU is supplemented with cash, then the actual expenditure on FOOD will decrease, a result which in no way achieves the policy objective. It would seem that other industries are subsidized, when the real goal is to allow low income consumers to achieve an adequate diet. The allocation of coupons which are limited in use to the purchase of food seems to be a more effective means of accomplishing the objective.

This effect is limited in IG2, since most Food Stamp subsidies go to consumers in IG1. It can be seen even in IG1, however, that an income subsidy in cash might not provide the additional consumption of food intended by the policy. The change in expenditure in IG1, noted above, was not significant in the BLS model. In the Knox model, where the coefficient is an aggregate for both IG1 and IG2, one can see that the effect of FIXED on FOOD is negative for DAIRY, PRODUCE, GROCERY (not significantly), and MEALS OUT. In short, there seems to be little evidence that the nutritional objectives of public policy will be better accomplished by a switch to cash supplements.

A further implication of such a switch (from coupons to cash supplements), given the validity of the results concerning elasticity, is the effect on demand for both raw agricultural commodities and the services produced by food processors. One goal of policy in the agricultural sector is to try to influence supply so that market mechanisms

will keep price at an acceptable level. In the absence of that balance, some government action may be triggered to maintain price. If government supplements for food were diverted, a rather large market for agricultural products would be gone, and additional measures would be required to restore the market to equilibrium. Further action might be in the area of more restrictions on output. It might be that action to support price would not be sufficient to maintain current output, a move which would force additional resources out to the agricultural sector. Whatever the choices, the loss of part of the demand for food, as a consequence of a change in the method of supplementing food expenditures (or for any other reason), could have far-reaching implications.

CRANES CREST

REFERENCES

CRANES & CREST

REFERENCES

1. Adrain, J., and R. Daniel. "Impact of Socioeconomic Factors on Consumption of Selected Food Nutrients in the United States." American Journal of Agricultural Economics, Vol. 58, No. 1, pp. 31-38, February 1976.
2. Ando, R., and F. Modigliani. "The 'Life Cycle' Hypothesis of Saving: Aggregate Implications and Tests." American Economic Review, Vol. 58, No. 1, pp. 55-84.
3. Baird, C. W. Macro-economics: An Integration of Monetary, Search, and Income Theories. Chicago: Science Research Associates, Chapter 7, 1973.
4. Barten, A. P. "Estimating Demand Equations." Econometrica, 36:213-251.
5. Becker, G. Economic Theory. Random House, New York, 1st edition.
6. Benus, J., J. Kmenta, and H. Shapiro. "The Dynamics of Household Budget Allocation to Food Expenditures." Review of Economics and Statistics, 58 (1976):129-138.
7. Buse, R. C., and L. E. Salathe. "Adult Equivalent Scales: An Alternative Approach." American Journal of Agricultural Economics, Vol. 60, No. 3, pp. 460-468, August 1978.
8. Carlson, M. D. "The 1972-73 Consumer Expenditure Survey." Monthly Labor Review, Dec. 1974.
9. Christ, C. F. Econometric Models and Methods. John Wiley and Sons, Inc., New York, 1966.
10. Craun, J. W. "Analysis of Meat Consumption and Expenditure in Knoxville, Tennessee." Unpublished M.S. thesis, The University of Tennessee, Knoxville, 1970.
11. Currie, J. M. "The Analysis of Family Budgets." The Demand for Food, ed. W. J. Thomas, pp. 15-36. Manchester: University of Manchester Press, 1972.
12. Currie, J. M., A. J. Rayner, and J. Stewart. "Postscript." The Demand for Food: An Exercise in Household Budget Analysis, Manchester University Press, William Clowe and Sons, Limited, 1972.
13. Deusenberry, J. S. Income, Savings, and the Theory of Consumer Behavior. Cambridge: Harvard University Press, 1949.

14. Dorminy, H. G., Jr. "A Comparison of the Estimated Consumption Function for Meat from Two Different Samples." Unpublished M.S. thesis, The University of Tennessee, Knoxville, 1972.
15. Ferguson, C. E., and S. C. Maurice. Economic Analysis. Richard D. Irwin, Inc., Homewood, Illinois, 1974.
16. Fisher, J. "An Analysis of Consumer Goods Expenditures." Review of Economics and Statistics, 44:64-71.
17. Forsyth, F. G. "The Relationship Between Family Size and Family Expenditure." Journal Royal Stat. Soc., A123 (1960):367-397.
18. Friedman, M. "A Method of Comparing Incomes of Families of Differing Compositions." Studies of Income and Wealth, Vol. 15, New York: National Bureau of Economic Research, 1952.
19. Friedman, M. A Theory of the Consumption Function. (Princeton: Princeton University Press), Summary Chapter, 1957.
20. Goreux, L. M. "Income and Food Consumption." Monthly Bulletin of Agricultural Economics and Statistics, 10, (1960):1-13.
21. Hamilton, Philip C. "Fed and Non-Fed Retail Beef Demand for the 1970's." Unpublished M.S. thesis, Oklahoma State University, 1978.
22. Hassan, Z. A., and S. R. Johnson. "Price Elasticity Estimates from Cross-Sectional and Panel Data: A Survey." Food Demand and Consumption Behavior: Selected Research Topics, ed. Robert Raunikaar, pp. 116-139. Workshop Papers printed by the University of Georgia AES, 1977.
23. Herrmann, R. O. "Income-Expenditure Elasticities for Food in Households of Differing Size and Composition." Michigan AES Quarterly Bulletin, 47:43-50, August 1964.
24. Herrmann, R. O. "Interaction Effects and the Analysis of Household Food Expenditures." Journal of Farm Economics, 49 (1967):821-832.
25. Holbrook, R., and F. Stafford. "The Propensity to Consume Separate Types of Income: A Generalized Permanent Income Hypothesis." Econometrica 39, (1971):1-21.
26. Hu, Teh-Wei. Econometrics: An Introductory Analysis, University Park Press, Baltimore, Maryland, 1973.
27. Hymans, S. and H. Shapiro. "The Allocation of Household Income to Food Consumption." Five Thousand American Families-Patterns of Economic Progress, Volume II, ed. J. N. Morgan, pp. 253-278. Ann-Arbor: Institute for Social Research, The University of Michigan, 1974.

28. Johnston, J. Econometric Models. McGraw-Hill Book Company, Inc., New York, 1963.
29. Johnston, J. Econometric Models. McGraw-Hill Book Company, Inc., New York, 2nd edition, 1972.
30. Judge, G. G. "On Estimating the Parameters of Economic Relationships." Surveys of Ag. Economics Literature, Vol. II. Minneapolis: University of Minnesota Press.
31. Keynes, J. M. The General Theory of Employment, Interest, and Money. Harcourt, Brace, and World, Inc., New York, 1935.
32. Kmenta, J. Elements of Econometrics. The Macmillan Company, New York, 1971.
33. Kuznets, Simon. "National Income, A Summary of Findings." New York, National Bureau of Economic Research, 1946.
34. Layng, W. J. "The Revision of the Consumer Price Index." Statistical Reporter, No. 78-5, United States Department of Commerce, February 1978.
35. Lee, F. Y., and K. E. Phillips. "Differences in Consumption Patterns of Farm and Nonfarm Households in the United States." American Journal of Agricultural Economics, 53(1971):573-582.
36. Marshall, A. Principles of Economics. London-Macmillan, 8th edition, 1920.
37. McDonald, R., and T. K. Mukherjee. Selected Socio-Economic Factors Influence on Red Meat Consumption in the Baltimore-Washington Metropolitan Area. Maryland Agricultural Experiment Station Bulletin 476, April 1966.
38. Mincer, J. "Market Prices, Opportunity Costs, and Income Effects." Measurement in Economics, pp. 67-82. Stanford: Stanford University Press, 1963.
39. Modigliani, F. "The Life-Cycle Hypothesis of Saving, the Demand for Wealth, and the Supply of Capital." Social Research, Vol. 33, pp. 160-216, Summer 1966.
40. Pearl, R. B. "Data Systems on Food Demand and Consumption: Properties, Uses, Future Prospects." Food Demand and Consumption Behavior-Selected Research Topics, ed. Robert Raunika, pp. 116-139. Workshop Papers printed by the University of Georgia AES, 1977.
41. Peterkin, B. "USDA Family Food Plans and Their Costs." National Food Review, NEAD, ESC, United States Department of Agriculture.

42. Peterson, H. P., and R. Buse. "A Bibliography on the Theory and Research on Household Expenditures." (In cooperation with the College of Agriculture and Life Sciences, University of Wisconsin.) Agricultural Economics Reports-293, 44 pp., July 1975.
43. Pindyck, R. S., and D. L. Rubinfeld. Econometric Models and Economic Forecasts. McGraw-Hill Book Company, New York, 1976.
44. Prais, S. J., and Houthakker. The Analysis of Family Budgets. Second edition, abridged. Cambridge: At the University Press.
45. Prato, A. A. "Estimation of Food Demand and Consumption Relationships." Food Demand and Consumption Behavior--Selected Research Topics, ed. Robert Raunika, pp. 94-115. Workshop Papers printed by the University of Georgia AES, 1977.
46. Price, David W. "Unit Equivalent Scales for Specific Food Commodities." American Journal of Agricultural Economics, 52(1970):228-233.
47. Pyatt, F. G. Priority Patterns and the Demand for Household Durable Goods. Cambridge (England) University Press, 1964.
48. Rockwell, G. R., Jr. Income and Household Size: Their Effects on Food Consumption. USDA Marketing Research Report 340, June 1959.
49. Samuelson, P. "A Note on the Pure Theory of Consumer's Behavior," Economica (NS), 5:61-71, 1938.
50. Smithies, A. Economics and Public Policy. Washington, Brookings Institute, 1955.
51. Stewart, J. "Econometric Techniques." The Demand for Food: An Exercise in Household Budget Analysis, Manchester University Press, William Clowes and Sons, Limited, 1972.
52. Stigler, G. J. "The Early History of Empirical Studies of Consumer Behavior." Journal of Political Economy, 62:95-113, April 1963.
53. Suits, D. B. "The Determinants of Consumer Expenditure--A Review of Present Knowledge." In Impacts of Monetary Policy--Commission Money and Credit, Prentice-Hall, Englewood Cliffs, New Jersey, pp. 1-53, 1963.
54. Sydenstricker, E., and W. I. King. "The Measurement of the Relative Economic Status of Families." Quarterly Publication of American Statistical Association, 17-842, 1921.
55. Tomek, W. G. "Empirical Analysis of the Demand for Food: A Review." Food Demand and Consumption Behavior, ed. Robert Raunika, pp. 1-30. Agricultural Experiment Station, University of Georgia, 1977.

56. United States Department of Agriculture. "Food Consumption of Households in the South, Spring 1965." Agricultural Research Service Household Food Consumption Report No. 4, 1965-66.
57. United States Department of Commerce. Survey of Current Business (Washington, D. C., U. S. Government Printing Office, 1978).
58. United States Department of Labor, Bureau of Labor Statistics. 1972-73 Consumer Expenditure Survey Interview Survey Detailed Public Use Tape.
59. West, D. A., and David W. Price. "The Effects of Income, Assets, Food Programs, and Household Size on Food Consumption." American Journal of Agricultural Economics, 58(1976 Part 1):725-730.
60. Williams, E. R. "The Influence of Selected Economic and Social Factors on Meat Consumption." Unpublished Ph.D. dissertation, The University of Tennessee, Knoxville, 1968.
61. Yamane, T. Statistics, An Introductory Analysis. Second edition, New York, Harper and Row, 1967.



APPENDIX



TABLE A-1. DATA MEANS, NUMBER OF OBSERVATIONS AND STANDARD DEVIATIONS FOR SELECTED VARIABLES BY INCOME GROUP

Variable	Number	Mean	Standard Deviation
Income Group 1			
Raw Data			
FIXED	449	1057.72	969.80
FOOD	444	985.20	678.56
INCOME	451	2774.63	2371.73
RESINC	449	1803.75	2587.71
CUSIZE	451	1.89	1.15
AGEHD	451	55.65	19.60
Income Group 2			
Raw Data			
FIXED	572	1981.10	1477.23
FOOD	568	1544.42	934.11
INCOME	572	7389.64	1431.61
RESINC	572	5356.43	3914.65
CUSIZE	572	2.65	1.45
AGEHD	572	44.80	16.45

Table A-1 (continued)

Variable	Number	Mean	Standard Deviation
Income Group 3 Raw Data			
FIXED	624	2923.61	1812.62
FOOD	621	2418.47	1393.20
INCOME	624	15706.80	6611.53
RESINC	624	12905.38	6764.72
CUSIZE	624	3.28	1.59
AGEHD	624	45.30	12.33

^aIncome boundaries were specified in 1978 dollars, and deflated to 1972 dollars.

Data Source: 1972-73 Bureau of Labor Statistics Survey (58).

TABLE A-2. DATA MEANS, NUMBER OF OBSERVATIONS AND STANDARD
DEVIATIONS FOR SELECTED VARIABLES BY INCOME GROUP,
FROM KNOX DATA

Variable	Number	Mean	Standard Deviation
Income Group 1 Raw Data			
FIXED	65	2248.87	1523.82
FOOD	65	1430.07	754.57
INCOME	65	5438.97	3211.21
RESINC	65	3190.10	2863.87
CUSIZE	65	2.36	0.94
AGEHD	65	44.66	18.77
Income Group 3 Raw Data			
FIXED	30	4732.14	3198.33
FOOD	30	1992.74	927.85
INCOME	30	18138.00	9123.83
RESINC	30	13405.86	9619.25
CUSIZE	30	3.05	1.31
AGEHD	30	46.23	14.26

Data Source: 1978 Knox County, Tennessee Survey.

VITA

The author was born near Moss, Tennessee in 1948. He graduated from Celina High School, Celina, Tennessee, in 1965, and attended The University of Tennessee, Knoxville. He graduated with the B.S. degree in 1969, with a major in Agricultural Business. After being drafted for service in the Army in 1969, he was married, released from active duty, and accepted into graduate school in 1971. In 1972, he graduated with the Master of Science degree in Agricultural Economics from The University of Tennessee, Knoxville, and accepted an appointment with the Resource Management Farms Program in the Tennessee Agricultural Extension Service.

In 1976, he returned to graduate work as a graduate assistant with the Agricultural Extension Service, and later with the Agricultural Experiment Station. He was divorced in 1977. He completed requirements for the Doctor of Philosophy degree from The University of Tennessee, Knoxville, in Agricultural Economics in 1980.