AN ECONOMETRIC STUDY OF EXPENDITURE PATTERNS OF CONSUMERS IN SELECTED URBAN AREAS OF NIGERIA

by

SAMUEL OSENI ADAMU, M.Sc.(ECON)

A Thesis Presented for the Degree of Ph.D.

in the University of London,

February 1972

ABSTRACT

Optimum utilisation of available statistical data is an objective which a country, particularly a developing one, should pursue. Adequate information about the structure of an economy is necessary for purposeful and meaningful planning, but a developing country lacks long statistical series that can make such information available. The objective of this study is to construct a simultaneous equation model of expenditure patterns of consumers based on the theory of consumer's behaviour; estimate parameters of the model; examine statistically various other factors, in addition to income, which affect consumption; and subsequently, with the help of certain assumptions on consumer preference orderings, derive income and price elasticities using Nigerian family budget data.

The achievement of this objective depends on sound theoretical basis and appropriate data. These are covered in Chapters 2, 3 and 4. The findings are discussed in Chapters 5 and 6. These findings show that different levels of factors like area and occupation affect average group expenditures (tests between means) but the slopes of group expenditures with respect to total expenditures (the β_i's) are not significantly affected by these factors; that area effects are reasonably explained by occupational composition of areas; that the way estimates of β_i vary with family size confirms common indication of economies of scale particularly for commodities like food and transport; that because of broad classification of commodities,

family composition cannot be meaningfully incorporated into the analysis; that two-way classification by income and family size is superior to one-way classification by income alone and, that with the aid of the Leser-Frisch approach to assumptions concerning substitutability or complementarity between broadly classified commodities, some kind of estimates of price elasticities is possible in addition to the usual income elasticities from family budget data.

ACKNOWLEDGEMENTS

I thank Mr. E. A. Adewole, Chief Statistician, Federal Office of Statistics, Lagos and his able assistant, Mrs. Sodipo, who made available summary files where appendix B was compiled. University of Ibadan provided research grants for data collection and free computing time at the University Computer Centre.

I am grateful to the authorities of the London School of Economics and Political Science for allowing me to use the facilities of the School during my current study leave. In particular, I owe much to Professor Sir Roy Allen who had been involved in my statistical training since my first admission to the School as an undergraduate student in October 1961 and became my supervisor at my second visit in October 1971. That this work is finally submitted, in good time as a Ph.D. thesis, owes much to his experience, efficient supervision and advice.

I have benefitted from discussions and correspondences with Professors J. Durbin and J. D. Sargan of the London School of Economics and Political Science. Professor O. Aboyade, as head of my department at the University of Ibadan was very helpful in making the six-month study leave possible.

I thank all those who took part in typing this work, within a short time, for the systematic and efficient manner they carried out the assignment.

Last but by no means least, I thank my wife for her moral support and for the good care of the young members of the family during my leave of absence.

With all these, the full responsibility for the content of this work is mine.

S.O.A.

CONTENTS

LIST OF TABLES IN THE TEXT

Page

Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Consumers' behaviour is a meeting point of many disciplines in the social sciences and studies of it can be viewed from different angles. The core of economic theory is based on it. The relation between economics and other social sciences is more felt in it than in any other sections of economic analysis. It is impossible to go very far without a knowledge of the behaviour of the household in typical circumstances. The study of family budgets seems of great importance from a number of points of view, particularly for economists and businessmen. The businessman produces ultimately to satisfy the consumers, although the concept of demand in economic theory is different from that in business. In economics, demand analysis is chiefly concerned with the demand for a commodity irrespective of brand,[1] or a group of commodities like meat or dairy products; or on a higher level of aggregation, for all food, housing or clothing.

This thesis is directed towards economic aspects of consumers' behaviour with particular attention to the expenditure patterns of consumers in selected urban areas of Nigeria. The study will derive demand functions for commodity groups making up the total expenditure

---

1 An important exception is the demand schedule under imperfect competition as covered by Robinson (1969). (All references in this form are listed in alphabetical order with regards to author's name at the end of the thesis, under the heading, 'References'.)

of the consumers in the areas. In so doing, estimates on consumption patterns will be made available which will be useful in demand analysis and in allocation problems concerning social and economic planning of Nigeria. The study takes into consideration some non-economic variables like occupational, geographic and demographic factors and examines their effects with the aid of statistical techniques. Even though emphasis is on the results rather than the techniques, this cannot be done without comparing estimating procedures. In this respect, some contribution will be made to methodology, even if it only amounts to restating old solutions in some other ways. The computational approach is reasonably simple but various estimates of the sets of parameters are made and compared for the purpose of choosing the best, statistically, as far as the available data allow. The final results show the importance of cross-classified data (if grouped data are used) in family budget studies. It means that, however small the sample of observations used, first priority must always be given to using this kind of data as against using data grouped only by income.

Empirical studies in consumers' behaviour using family budget data are usually involved with economic theory as shown by Allen and Bowley (1935), Stone et. al. (1954), Prais and Houthakker (1955), Wold and Jureen (1956) and Schultz (1937) to mention a few of the detailed and standard studies since the 1930's. They also provide useful information on growth and development problems. There is also the related allocation problem particularly for those countries getting out of crises, in which case there is need for some rationing. Studies of this kind are usually carried out by individuals, or governments sponsoring individuals, because of the interest in standards of living or general welfare. Traces of this motive can be found in some of the major works cited above, in this paragraph. The earlier collectors of family budget data like Eden (1797), Rowntree (1901) and Bowley and Burnett-Hurst (1915) were motivated in the same manner. Most of these surveys were on the poor members of the community. An important conclusion of Engel on Belgian working-class households was that a smaller proportion of expenditure was devoted to food at the end of the nineteenth century than at its beginning, in consequence of the general rise in the standard of living.[1] The Nigerian situation with regards to studies of family budgets has not been encouraging. Family budget data were first collected officially through surveys in 1953[2] for the main purpose of getting weights for the cost of living index as the consumer price index numbers were then called. Empirical studies using the Nigerian data are few or non-existing if interest is in rigorous approach.[3] There was no statistical information on family budgets that could be used to reasonable advantage after the civil war when rationing was necessary in the three Eastern states which were most affected. Even the present National Development Plan for 1970-74 devoted very little space to personal expenditure. Possibly, the main reasons may be that of lack of data and the general notion that there is excess demand in the Nigerian economy. But it should be realised that efforts directed

---

[1] Cited in Prais and Houthakker (1955), p.3.

[2] See F.O.S. (1956).

[3] This will be fully discussed later in Chapter 5.

to the study of family budgets, especially for low income earners, can never be wasted or unrewarding because later the tables will turn against excess demand, as we find in the developed countries. Then, the valuable knowledge derived in continuous empirical work on family budgets will be very helpful since what the industries produce ultimately find their ways in one way or the other to the consumers. In addition, it is hoped that this study will influence positively the approach to data collection, storage and publication in Nigeria.

This study is wholly carried out as an empirical econometric exercise, or an applied problem in statistics, in the sense that it attempts to apply measurement and statistical techniques in explaining economic behaviour of consumers. Quantification of the behaviour of consumers in response to different economic environments constitutes the oldest and perhaps the largest field of applied econometric studies. It is still growing despite the massive work done so far, and the general attitude that all that could be said has been said. It seems this is not so. If economics is to develop as a quantitative science, empirical work in it must be made not only in different economic environments in a narrow sense but in different so called non-economic environments in a rigorous manner covering all stages of economic development. To be able to achieve much, in the evolution of the developing countries, a lot of quantitative materials will be required. Even though developing countries need the statistical materials more than the developed, the acquisitionoof these materials is very much in the reverse order - a paradox of life.

Leser (1968) grouped applied econometric studies in consumers' behaviour into three broad divisions according to the level of aggregation adopted, as follows: "Firstly, there are consumption function - or, which comes to the same, savings function studies, in which the relationship between personal expenditure on one hand and personal income plus other variables on the other hand is intensively investigated, with emphasis on correct specification. Secondly, the distribution of consumption expenditure over a number of commodity groups is considered as a function of a limited number of variables, most of them common to all commodity groups. Thirdly, the demand for individual commodity is studied with the help of simple or elaborate models designed specifically for the commodity in question."[1] The scope of this thesis falls under division two above. Commodity groups are defined and, using budget data for selected urban areas of Nigeria which are finally pooled together, numerical contents are given to the concepts of consumer economic theory. This is after the relationships between commodity expenditures and total expenditure are established. Our main interest will be on income elasticities since studies using family budget data for one time period assumes that the data only show variation in income while prices are usually constant. Later, with the help of assumptions based upon observation of certain kinds of preference ordering of consumers, particularly when commodities are broadly classified, attempt is made to derive and estimate price elasticities, using the proportions of total expenditure on the commodities and the income elasticities (Engel elasticities) estimated from the budget data.

---

[1] Leser (1968), p. 544.

The other chapters of this thesis are made up of the following: The next chapter covers the economic basis of the study. The usual theory of consumer behaviour is restated to stress points that are relevant to this study. Emphasis is put on recent development concerning the relationship between theory and empirical work in which Leser (1941-42), Frisch (1959), Houthakker (1960b) and Barten (1964b) have played very important parts. Chapter 3 discusses models of errors in variables, regression analysis for grouped data, method of instrumental variables and comparison of regression lines. At the close of the chapter, the computational procedures adopted are discussed. In this chapter, an attempt is made to bridge the gap between the theory and the data available. Chapter 4 deals with the data, in particular with problems of data collection and urban consumer surveys in Nigeria, and the derivation of the data in the form used for this study. Some basic features of the results already published officially and of Nigeria are summarised as background information. The next two chapters (5 and 6) deal with the results and their interpretations. And, in the concluding chapter, attempts are made to give a brief summary of the results, to deal with some problems of analysing consumption patterns in developing countries, especially Nigeria and to outline some directions for further work.

CHAPTER 2

ECONOMIC BASIS[1]

2.0 INTRODUCTION

This chapter deals with the economic theory of the study. It is about the theoretical framework for the empirical investigation of consumers' behaviour. The framework is formulated to be useful for observations on consumer expenditure patterns. Its purpose is to translate propositions about preferences (which are not directly observable) into propositions about expenditures on specific commodities or commodity groups.

Even though applications of ultimate interest are to the aggregate behaviour of a community of people having certain characteristics in common, the theory, in the usual practice, starts with a single rational consumer who is either an individual or a group with common purpose, who purchases commodities in the market. Commodities include goods and services among which consumers may choose: yam, cassava, rent, taxi fare etc. A convenient mathematical device for summarising all relevant information about a consumer's behaviour is to construct demand functions, based upon constrained utility maximisation. The constructed demand functions show, for each possible set of market parameters that collection of commodities the consumer buys.

---

1 In writing this section, I have gained much from the recent contribution of Katzner (1970).

2.1 HISTORICAL NOTE

An important purpose of this section is to list most of the concepts used in the other sections which are not given any detailed development in the body of the study. The historical origins of demand theory lie in the discovery of two concepts: utility and demand. For a long time these two generally appeared only in separate contexts. Focus on one seemed to preclude consideration of the other. However during the second half of the 19th century, a link between them became apparent and at this point theoretical content was given to the study of demand. During the process of development, many concepts were introduced into the theory. In the early stage, starting with utility theory we have additivity, diminishing marginal utility and cardinality. It was Edgeworth (1881) who introduced the idea of a general utility function, as opposed to additive utility, and of an indifference map in order to analyse barter in the non-additive case. Other early important concepts include ordinal utility |Slutsky (1915) and Pareto (1896)| and marginal rate of substitution |Hicks and Allen (1934)|. More recently Samuelson (1947) proposed the concept of revealed preference in an attempt to avoid the problem of "integrability".

The concept of demand goes back to King (1696) who actually computed a demand schedule for wheat. The relationship between demand and utility goes both ways and so it is not surprising that while Walras (1874) was explaining the former in terms of the latter, Antonelli (1886) was examining the construction of indifference curves from demand. The utility theory approach to demand analysis involves tackling many problems. For example, the movement from preferences to utility functions |Debreu (1954)|; the impact of changes in an individual's preferences on his behaviour |Ichimura (1951)|; alternative forms of the utility explanation of consumer behaviour |Hotelling (1932) and Roy (1943), who dealt with the problems of indirect utility|; the equivalence between maximising direct utility with respect to commodity and minimising indirect utility with respect to prices and income, with budget constraint imposed in both instances |Court (1941)|; and direct and indirect additive utility |Konus (1939) and Houthakker (1960b)|. This last topic is further discussed in section 2.4.

In somewhat different vein, economists have also noted that the demand for certain commodities exhibits special traits when combined into groups such as food and clothing. Hicks (1939) shows that if prices of goods within any group always move in the same proportion, then an individual's demand for the group, as a whole, has properties identical to that for a single commodity. Another notion along this line is that of separability, introduced independently by Sono (1945) and Leontief (1947), to deal with the problem of tastes. Thus if an individual's preferences within a particular group do not depend on quantities of commodities outside it, his demand for that group often can be defined with similar properties without the assumption of price proportionality. It means that consumers' decision-making processes can be broken down into steps: he may first decide how much to spend on the group itself and then determine the way expenditure is to be allocated within it.

Analysis of the consumer's behaviour under various environments of uncertainty seems to have stemmed from Bernoulli (1738) where the St. Petersburg paradox was discussed and solved using assumptions based on diminishing marginal utility of wealth and expected utility rather than expected winnings; and the subsequent developments that followed von Neumann and Morgenstern (1944). Development along this line is far ahead in industrial decision-making than in consumer decision-making.[1] The details involved in all these problems and concepts, and others not mentioned here, will not be discussed but the ideas involved will be used in many parts of what follows.

2.2 THE THEORY OF CONSUMER'S BEHAVIOUR

It is quite reasonable to suppose that the characteristics of a consumer's likes and dislikes determine the purchases he makes in the market. And it is fairly generally agreed that the rationality of his behaviour may be described by postulating that the consumer has a definite preference over all conceivable commodity bundles and that he chooses those commodity bundles that are optimum with respect to his preference subject to budgetary constraints. Making a number of axioms on the preference ordering |Debreu (1954) and Uzawa (1959)| leads to the possibility of formulating a utility function which is at least twice differentiable.

$$U = U(q_1, q_2, q_3, \ldots, q_n) \qquad 2.2(1)$$

where U is utility derived from a certain bundle of quantities of n commodities and q_i is the quantity of commodity i in the bundle (i = 1,2,...n).

---

[1] A possible source of friction is highlighted by a collection of contributions by Wold, Shackle, Savage, Manne, Charnes, Samuelson and Malinvaud published in the Econometrica, Vol. 20, No. 4 (1952), pp. 661-679, concerning "strong independence assumption" if applied to the theory of consumer demand.

The axiom of rationality says that the consumer will maximise his utility subject to his limited resources given by

$$\sum_{i=1}^{n} p_i q_i = x \qquad 2.2(2)$$

where p_i is the price of commodity i and x is income. Constrained utility maximisation, using the method of Lagrange multiplier, leads to

$$\frac{\partial U}{\partial q_i} = \lambda p_i \qquad 2.2(3)$$

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

and 2.2(2) given above.

At the left-hand side of 2.2(3) we have the marginal utility of commodity i and λ, the Lagrange multiplier, usually interpreted in economic terms as marginal utility of income because

$$\frac{\partial U}{\partial x} = \lambda.$$

From 2.2(3) we have

$$\lambda = \frac{u_i}{p_i} \qquad 2.2(4)$$

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

$$\text{where } \frac{\partial U}{\partial q_i} = u_i \ .$$

Equations 2.2(2) and 2.2(3) are (n+1) equations in 2(n+1) variables: the n quantities of the commodities, the n prices, λ and income, x. These equations form the set of necessary conditions for constrained utility maximisation. The (n+1) equations can be solved for all q_i and λ and expressed as functions of prices and income, with the first n of such functions for the n quantities written as:

$$q_i = q_i(p_1, p_2, \ldots p_n, x) \qquad 2.2(5)$$

$$i = 1,2,\ldots n.$$

In the course of the development of the theory of demand, some restrictions on these equations are derived. To be able to do this we go back to equations 2.2(2) and 2.2(3) and differentiate them with respect to income and the prices now treated as parameters as in 2.2(5). Using matrix notation, for brevity, we have

$$\begin{vmatrix} U & -p \\ -p' & 0 \end{vmatrix} \; | \; Q_p \quad q_x \; | = \begin{vmatrix} \lambda I & 0 \\ q' & -1 \end{vmatrix} \qquad 2.2(6)$$

$$\text{where } U = \begin{vmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & & & \vdots \\ u_{n1} & \cdots\cdots & & u_{nn} \end{vmatrix}, \quad p = \begin{vmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{vmatrix},$$

$$Q_p = \begin{vmatrix} \partial q_1/\partial p_1 & \partial q_1/\partial p_2 & \cdots & \partial q_1/\partial p_n \\ \partial q_2/\partial p_1 & & & \vdots \\ \vdots & & & \vdots \\ \partial q_n/\partial p_1 & \cdots\cdots\cdots & & \partial q_n/\partial p_n \end{vmatrix}, \quad q = \begin{vmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{vmatrix}, \quad q_x = \begin{vmatrix} \partial q_1/\partial x \\ \partial q_2/\partial x \\ \vdots \\ \partial q_n/\partial x \end{vmatrix},$$

$$I = \begin{vmatrix} 1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & 1 & 0 & \\ & & 0 & & \ddots & \\ & & & & & 1 \end{vmatrix}, \quad 0 = \begin{vmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{vmatrix}$$

$$u_{ij} = \frac{\partial^2 U}{\partial q_i \partial q_j}, \quad i,j = 1,2,\ldots n$$

λ, 0, 1 are scalars, while $H = \begin{bmatrix} U & -p \\ -p' & 0 \end{bmatrix}$ is called bordered Hessian matrix.

The second order conditions for the maximisation, given above, require the bordered Hessian matrix to be negative definite, which means that

$$u_{ii} < 0, \quad \begin{vmatrix} u_{ii} & u_{ij} \\ u_{ji} & u_{jj} \end{vmatrix} > 0, \quad \begin{vmatrix} u_{ii} & u_{ij} & u_{ik} \\ u_{ji} & u_{jj} & u_{jk} \\ u_{ki} & u_{kj} & u_{kk} \end{vmatrix} < 0, \text{ etc.}$$

The main restrictions imposed on the demand functions can be derived from 2.2(2) and 2.2(6). Consider that there is a change in the j^{th} price only, then 2.2(6) reduces to

$$\begin{bmatrix} U & -p \\ -p' & 0 \end{bmatrix} \begin{vmatrix} \partial q_1/\partial p_j \\ \partial q_2/\partial p_j \\ \vdots \\ \partial q_n/\partial p_j \end{vmatrix} = \begin{vmatrix} 0 \\ \vdots \\ \lambda \\ 0 \\ 0 \\ \vdots \\ q_j \end{vmatrix} \longleftarrow \text{for the } j^{th} \text{ element} \qquad 2.2(7)$$

The effect of this change on commodity i is known by solving 2.2(7) and picking the corresponding solution for $\partial q_i/\partial p_j$. This can be done easily by using Cramer's rule which leads to

$$\partial q_i/\partial p_j = (-1)^{j+i} \lambda \frac{\Delta_{ji}}{\Delta} + q_j \frac{\Delta_{ni}}{\Delta} \qquad 2.2(8)$$

where Δ is the determinant of H
and $\Delta_{k\ell}$ is the minor for element (k,ℓ) of H.

In similar manner the change in commodity i due to change in income gives

$$\partial q_i/\partial x = -\frac{\Delta_{ni}}{\Delta} . \qquad 2.2(9)$$

Substitute 2.2(9) into 2.2(8), we have

$$(-1)^{j+i}\frac{\Delta_{ji}}{\Delta} = \partial q_i/\partial p_j + q_j\frac{\partial q_i}{\partial x} . \qquad 2.2(10)$$

Consider the effect on the demand for commodity i as a result of a change in p_j with utility and other prices constant but with a compensating change in income. Under this condition, we have

$$dq_i = \frac{q_i}{p_j}dp_j + \frac{q_i}{x}dx.$$

But from 2.2(2), $\frac{dx}{dp_j} = q_j$. Then the above expression, after substituting for dx, becomes

$$dq_i = \left(\frac{\partial q_i}{\partial p_j} + q_j\frac{\partial q_i}{\partial x}\right)dp_j$$

and

$$\therefore \frac{dq_i}{dp_j} = \frac{\partial q_i}{\partial p_j} + q_j\frac{\partial q_i}{\partial x} .$$

Define k_{ij} as substitution effect i.e.

$$k_{ij} = \frac{dq_i}{dp_j} = \left(\frac{\partial q_i}{\partial p_j}\right)_{\text{(U and the other p's constant)}} \qquad 2.2(11)$$

then

$$k_{ij} = (-1)^{j+i}\lambda\frac{\Delta_{ji}}{\Delta}$$

and 2.2(8) can then be written as

$$\partial q_i/\partial p_j = k_{ij} - q_j\,\partial q_i/\partial x . \qquad 2.2(12)$$

Equation 2.2(12) is the usual Slutsky equation. It is a partition of the effect of a change in the price of j on commodity i into substitution effect, k_{ij} and income effect. This partition is not a restriction in itself. Its importance lies in the restrictions imposed on the k_{ij}, the substitution term. In the first place, k_{ij} multiplied by appropriate prices, sum up to zero (a property of the determinant of H) i.e.

$$\sum_{j=1}^{n} k_{ij} p_j = 0 \qquad 2.2(13)$$

$$i = 1,2,3,\ldots n.$$

This is the homogeneity property which implies the absence of a money veil. There are n of these as indicated above. Also it can be shown that

$$k_{ij} = k_{ji} \qquad 2.2(14)$$

which provides the symmetry property. These are non-trivial for $i \neq j$ and there are $\frac{n(n-1)}{2}$ of them. There is the non-negativity property for the substitution part[1] of the own-price derivative viz

$$k_{ii} < 0 \qquad 2.2(15)$$

The consequence of this is that, in the case where the income derivative of 2.2(12) is positive or zero, i.e. when the commodity is not an inferior one, an increase in the price of this commodity leads to a decrease in the quantity bought. This provides the analytical justification of the intuitive notion of the negative

---

[1] Allen and Mishan (1965) have shown that if the concept of an increase in price is properly defined, this property remains invariant when one removes the linearity property of 2.2(2).

relationship between quantity and price. Then there is the adding-up property:

$$\sum_{i=1} p_i \frac{\partial q_i}{\partial x} = 1 \qquad 2.2(16)$$

which is derived by differentiating 2.2(2) with respect to income, x.

The corresponding elasticity approach to all the restrictions given above and some others is very valuable in empirical studies. Frisch (1959) and Leser (1941-42) carried out their analysis using the elasticity approach. Because of the usefulness of restrictions based on elasticity approach in what follows, some of them are now derived.

Given the demand function in the form 2.2(5), we define the following concepts:

$y_i = p_i q_i$ = expenditure on commodity i

$w_i = \frac{y_i}{x}$ = proportion of income expended on commodity i (budget proportion)

$E_{i0} = \frac{x}{q_i} \frac{\partial q_i}{\partial x}$ = income elasticity of demand for commodity i

and

$E_{ij} = \frac{p_j \partial q_i}{q_i \partial p_j}$ = j^{th} price elasticity of demand for commodity i

$i,j = 1,2,3,\ldots n.$

Since our analysis concerns static demand theory, elasticity of demand is a concept of comparative statics. It will not describe the short-run effect of price change, but the effect on the equilibrium which will be established after some time. Using 2.2(12) to substitute for k_{ij} in 2.2(13) leads to

$$\sum_{j=1}^{n} p_j \, {}^{\partial q_i}/_{\partial p_j} + x \, {}^{\partial q_i}/_{\partial x} = 0 \qquad 2.2(17)$$

which shows that 2.2(5) satisfies Euler's theorem |see Allen (1947)| for homogenous function. It follows that dividing each term of 2.2(17) by q_i and expressing it in elasticity form, gives

$$\sum_{j=1}^{n} E_{ij} + E_{i0} = 0 \qquad 2.2(18)$$

$$i = 1,2,\ldots n.$$

This is the corresponding homogeneity condition in elasticity form. Differentiating 2.2(2) with respect to price j we have

$$\sum_{i=1}^{n} p_i \, {}^{\partial q_i}/_{\partial p_j} + q_j = 0.$$

On multiplying each term of this expression by $\frac{p_j}{x}$, it becomes

$$\sum_{i=1}^{n} \frac{p_i p_j}{x} \frac{\partial q_i}{\partial p_j} + \frac{p_j q_j}{x} = 0$$

and, if the numerator and the denominator of the term under the summation sign are multiplied by q_i and re-arranged, using the fact that $w_i = \frac{p_i q_i}{x}$, we have

$$\sum_{i=1}^{n} w_i \frac{p_j}{q_i} \frac{\partial q_i}{\partial p_j} + w_j = 0. \qquad 2.2(19)$$

If 2.2(19) is expressed in elasticity form, it leads to

$$\sum_{i=1}^{n} w_i E_{ij} + w_j = 0 \qquad 2.2(20)$$

$$j = 1,2,\ldots n.$$

This is the Cournot aggregation condition - with respect to j.

From 2.2(16), it can be shown that

$$\sum_{i=1}^{n} w_i E_{i0} = 1 \qquad 2.2(21)$$

which is the Engel aggregation condition. If 2.2(12) is solved for k_{ij}, we have

$$k_{ij} = \frac{\partial q_i}{\partial p_j} + q_j \frac{\partial q_i}{\partial x}. \qquad 2.2(22)$$

Multiplying 2.2(22) by $^{p_j}/_{q_i}$, we get

$$k_{ij}\, ^{p_j}/_{q_i} = \frac{p_j}{q_i}\frac{\partial q_i}{\partial p_j} + \frac{p_j q_j}{q_i}\frac{\partial q_i}{\partial x}$$

which can be written as

$$\varepsilon_{ij} = \frac{p_j}{q_i}\frac{\partial q_i}{\partial p_j} + w_j \frac{x}{q_i}\frac{\partial q_i}{\partial x}$$

where $\varepsilon_{ij} = k_{ij}\frac{p_j}{q_i}$.

Following the definition of k_{ij} given by 2.2(11), ε_{ij} can be defined as substitution effect elasticity. Hence from the expression above, putting the right-hand terms in elasticity form, leads to

$$\varepsilon_{ij} = E_{ij} + w_j E_{i0} \qquad 2.2(23)$$

which is what Frisch (1959) called Slutsky Elasticity. Frisch (1959) remarked that ε_{ij} is related to Hicks-Allen Elasticity[1] of Substitution, σ_{ij} as

$$\sigma_{ij} = \varepsilon_{ij}/w_j.$$

---

[1] Hicks and Allen (1934), p. 199.

Since σ_{ij} is symmetric in i and j, we have

$$E_{io} + E_{ij}/\omega_j = E_{jo} + E_{ji}/\omega_i. \qquad 2.2(24)$$

All these restrictions, 2.2(13) through 2.2(24) are invariant under general transformation of the utility function.

What we have been discussing so far comes under the direct utility function. As mentioned under section 2.1, there are at least two concise ways of describing a given preference ordering in terms of utility function and the optimisation principles under the two are equivalent. The other way is indirect utility function (see sections 2.4, 2.5).

So far no assumption has been made about the preferences of the consumer whether for direct or indirect utility functions. Contributions to modern demand theory have considered various possibilities usually encountered in empirical work as indicated in section 2.1 above. There are results concerning the assumptions that utility functions can be sub-divided into separate groups and maximisation in these groups is independent of maximisation of utility in the other groups. The restriction on the utility functions which is of immediate importance to this study is that of additivity |Houthakker (1960b)| or want-structure independence between commodities |Frisch (1959)|. These topics form part of the main discussion of section 2.4 and no further discussion is given here, except to remark that the commodities on which observations exist for this study are in broad categories, suitable for this kind of assumptions on preference ordering. In addition, there is reason to believe that when commodities are in broad categories even though they all compete for the consumer's income, there is no likelihood that the kind of substitution existing between commodities like tea and sugar, or yam and cassava, exists between them.

This analysis has been developed for a consumer but what is of interest is a collection of consumers. The movement from analysis for a consumer to that of a group of consumers is the object of discussion in most economics textbooks, however elementary |see Prais and Houthakker (1955) and Pearce (1964), to mention two, for some serious discussion of this point|. The very weak assumption made here is that there is no reason for the behaviour of an individual in a group to be different from that of the group particularly if the group is reasonably homogeneous with regards to many factors that are likely to affect the preferences of the consumers (same utility function for every consumer in the group). The results given above will be regarded as holding for each group of consumers studied.

2.3 ENGEL FUNCTIONS

At this point it is necessary to go back to equation 2.2(5) and digress a little. Instead of considering commodity i as a function of all prices and income, we consider it as related only to income, with all prices given, as

$$q_i = Q_i(x)$$

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

which can be transformed to give

$$y_i = Y_i(x) \qquad 2.3(1)$$

$$i = 1,2,\ldots n.$$

This is to take note of the fact that we are relating expenditure on commodity i on total income. The functions 2.3(1) are called Engel functions after the name of the scholar who first made systematic study of the relationship between income and expenditure on food.[1] This kind of digression is necessary if we take seriously, as we should, an observation made by Houthakker that, "Theoretical considerations suggest, however, that the shape of the Engel curve to a large extent determines the form of the price-quantity relations, so that it is perhaps justified to concentrate on the former."[2] To lend support to this, is the opening remark by Leser (1941-42) that, "Market data lend themselves to a direct approach to the measurement of price elasticities of demand, but their use is associated with some statistical difficulties; data are usually only available for the demand of the wholesaler, not of the consumer, and tastes may change significantly during the period over which the observations have to be extended. Family budgets, on the other hand, do not present these difficulties but they give only Engel functions i.e. income - expenditure relations and thereby income elasticities of demand."[3] In addition the present study is based on budget data. Of recent years this kind of data available for different time periods is being used in place of the usual time series on prices and quantities |see Leser (1960) and Powell (1966) in particular|.

Equation 2.3(1) is the general form which Engel functions usually take. There are many aspects of Engel functions for consideration. Demand theory usually regards prices and income to be the most important factors affecting demand. The next candidate is 'tastes'.

---

[1] See Houthakker (1957), p. 532.

[2] Houthakker (1960a), p. 10.

[3] Leser (1941-42) Op. cit. p. 40.

'Tastes' is the most difficult to deal with among the three. By interpreting the concept of 'tastes' widely enough, all differences in consumption patterns could be attributed to differences in 'tastes'. 'Tastes' includes such things as location, family size and family composition, climate, social class etc. It is therefore a catch-all for those determinants of consumption which, except in the long run, are not themselves determined by economic factors. On the basis of this statement it is necessary to adjust 2.3(1) to take account of some of the determinants making up 'tastes'. This can be done in one of two ways. Either we specify 2.3(1) in some detailed form to include certain parameters to represent each of the determinants of tastes or we classify, into groups, all the consumers involved in the analysis to make each group of consumers homogeneous with regards to the determinants. There is the third possibility of combining the two alternative approaches given above. The first approach provides too many parameters to be estimated and in some cases it may be difficult to identify some of them. The second requires more data than may be forthcoming. It seems the third approach is more plausible. This means to classify the consumers with regards to some natural factors which are easily identified - location, occupation, social class etc. - and within each group, have 2.3(1) which will be specified to take note of family size and composition. A form of 2.3(1) that takes note of family size is

$$\frac{y_{ij}}{n_j} = Y_i\left(\frac{x_j}{n_j}\right) \qquad 2.3(2)$$

$$i = 1,2,...n; \quad j = 1,2,...N$$

where n_j = the size for family j in the group of consumers.

To take account of family composition possibly in terms of age and sex can lead to the form

$$\frac{y_{ij}}{B_{ijk}} = Y_i\left[\frac{y_j}{K_{jk}}\right] \qquad 2.3(3)$$

$i = 1,2,\ldots n$ (no. of commodities)

$j = 1,2,\ldots N$ (no. of families)

and $k = 1,2,\ldots T$ (no. of types of persons in each family)

The family effects are B_{ijk} and K_{jk}. B_{ijk} represents the effect on the expenditure on commodity i of the k^{th} type of person in family j, while K_{jk} represents the income effect of the k^{th} type of person in family j. Various forms of B_{ijk} and K_{jk} have been proposed and used. Prais and Houthakker (1955) used the linear form while Forsyth (1960) used the exponential form.

Apart from all the determinants put together as 'tastes' there is the possibility of allowing for economies of scale in consumption. The concept of economies of scale gives expression to the possibility that, with given levels of income per person, a large household may be able to attain a higher standard of living than a smaller household. This is easily seen with regards to a commodity like food. If 2.3(2) is adjusted to take note of this concept we have

$$y_{ij}/n_j^{\theta_i} = Y_i\left[x_j/n_j^{\theta_0}\right] \qquad 2.3(4)$$

where n_j is raised to power θ_i or θ_0 and $1 - \theta_i$ is the measure of the specific economies of scale for commodity i and $1 - \theta_0$ is the measure of the income economies of scale. We can take θ_0 as a weighted average of θ_i $(i = 1,2,\ldots n)$.

If there are no economies of scale $\theta_i = 1$ $(i = 0,1,\ldots n)$, similar adjustment can be made on 2.3(3) to incorporate the effect on the household (family) composition. If this is done we have

$$y_{ij}/B_{ijk}^{\theta_i} = Y_i\left[x/K_{jk}^{\theta_0}\right] \qquad 2.3(5)$$

with the $1 - \theta_i$ and $1 - \theta_0$ having the same interpretations as those given for 2.3(4).

Engel function in the form 2.3(5) is in a reasonably general form. For estimation to be carried out we need many assumptions which are mostly dictated by the kind of data available.

The next aspect of Engel functions is the specific form which such functions should take. Among the simple functions usually used - linear, semi-logarithmic, double-logarithmic, hyperbolic etc. - there is not a single one that satisfies all the important criteria usually put forward. The criteria by which the suitability of a kind of Engel function is to be judged are as follows in decreasing order of importance:[1] goodness of fit in a wide range of data, computational simplicity, theoretical plausibility with adding-up condition as a special case. Adding-up condition in this sense means that $\sum_i^n Y_i(x) \equiv x$. As far as goodness of fit in a wide range of data is concerned, linear and hyperbolic forms get rather low score, semi-logarithmic does well for food stuffs and the double-logarithmic is good for nearly all commodities. Aitchison and Brown (1957) have some reasonable diagrams for lognormal Engel functions. On computational simplicity, with the advent of electronic computers, this can now be achieved by all the forms particularly when the non-linear ones can be easily transformed

---

[1] See Houthakker (1960a), pp. 10-11.

to linearity. The lognormal Engel curve, which requires a rather involved iterative treatment, can be estimated more quickly on the electronic computer. Theoretical plausibility covers must of the restrictions derived in section 2.2. This implies in particular that $\Sigma Y_i(x) \equiv x$ (c.f. 2.2(2)). Moreover the Engel functions after using it to derive demand functions with prices allowed to vary, should be capable of being integrated into a system of indifference curves. In the case of linear form, theoretical plausibility criterion is satisfied, $\Sigma Y_i(x) \equiv x$ and the utility function generated is a quadratic one. All the other functions mentioned above failed to satisfy this criterion. As a result of the poor performance of these simple forms in the three criteria there is need to examine more complicated ones which will do well in all counts. The trend of this development is picked up in section 2.5. After discussing in section 2.4 the various means of deriving demand functions from budget data, section 2.5 starts with a review of different systems of demand functions recently employed in empirical demand studies.

2.4 DERIVING DEMAND FUNCTIONS FROM FAMILY BUDGET DATA

The utility approach to demand has many advantages. Firstly, it helps to throw light on certain problems which could not be solved by a purely objective, behavioural and operational approach. In this sense, by providing a rational foundation for the law of demand and with a generalization of the law, it helps to bring out the interdependence of the various economic factors. In so doing the classification of commodities into useful and significant categories is made, the negatively and positively sloping demand curves are explained and so providing useful background to statistical work. Secondly, from the various restrictions given in section 2.2, utility theory helps to determine whether certain forms of demand functions are consistent with given assumptions. These restrictions on the demand functions for individual and related commodities help to show which commodities are competing or complementing. The data best suited to utility approach for derivation of demand functions are the family budget data. It is not easy to derive a demand function from statistical data without making some assumptions regarding the nature of the preference ordering and how these affect the related utility functions. Another feature of the derivation of demand functions from family budget data which must come to mind is that the approach is implicitly or explicitly on the general static demand function as discussed in section 2.2 with no time element entering into the analysis in an essential way.

A demand function is completely determined, under the usual assumptions, by the income, own-price, and cross-price elasticities of demand. Hence this section deals with the derivation of these concepts, in particular, from budget data. The usefulness of these concepts, particularly for the developing countries, is given clearly by Frisch (1959) when he said in the opening paragraph of his paper that, "In A DECISION MODEL where one will attempt to solve the implementation problem as far as possible by arranging for zero demand pressure in the consumer markets, it becomes an important statistical and econometric problem to estimate the demand elasticities."[1] From

---

[1] Frisch (1959), p. 177.

budget data, proportions of income expended on the items of commodities and the Engel elasticities are estimated more easily than the own-price elasticities, while own-price elasticities are easier to estimate than the cross-price elasticities. The problem is 'How far is it possible to go in the direction of drawing conclusions about the price elasticities from the budget proportions and the Engel elasticities?' This section deals with various methods of answering the question put forward and how these methods can be adapted for the Nigerian situation.

The first attempt at deriving demand functions from family budget data was made by Pigou (1910). His method was meant to derive elasticities between income groups, and ratios of two elasticities of demand could only be derived because of an indeterminate parmmeter. Since later developments do not approach the problem in the same way as Pigou did, it is not necessary to give his detailed approach. Friedman (1935) made very critical appraisal of Pigou's method and attempted to improve on Pigou's method but ended in deriving own-price elasticity of a commodity in terms of another which also meant that an indeterminate parameter exists in his solution. Since a better re-arrangement of his solution is given by Houthakker (1960b) which will be mentioned later in this section, it will not be necessary to discuss Friedman's method. Another early contributor was Marschak (1931). The rationale of his procedure was that family budget data gives the relation between money income, quantities of the various commodities, and price paid; that for households of fixed size, the relationship between these three quantities, using our previous notations, is

$$q_i = q_i\left(^x/p_i\right); \qquad 2.4(1)$$
$$i = 1,2,\ldots n$$

that the quantity consumed may be varied, either by varying income or price alone, or by varying both, but not in the same proportion; that when the money income and all prices are varied by the same relative amount there is, in the long run, no change in the quantity consumed, since the ratio $^x/p_i$ is unchanged; and that since at the time of budgetary survey the prices are fixed then we can put $p_i = 1$ and the relation, 2.4(1) becomes

$$q_i = q_i(x) \, . \qquad 2.4(2)$$
$$i = 1,2,\ldots n.$$

This curve may be taken as a 'single' market demand curve if we assume that all prices rise and fall together (parallel-reaktion assumption) and that there are no important substitutes for the commodity in question. Then an infinitesimal increase in income by a small relative amount is equivalent to a decrease in the price by the same amount and vice versa. The final market demand is the summation of the demand curves of the individual households. This means that income elasticity is worked out before the price elasticity. From this brief summary we see that Marschak's method is based on the assumption that the difference between income and price elasticities of demand for a commodity is negligible, if the amount spent on it is small compared with total expenditure. But Frisch (1932) showed that the difference could not be safely neglected, and that the observed income-elasticity could, therefore, not be interpreted as price elasticities. This is an important contribution as will be seen below, that substitution between commodities is reduced as we have broad categories of commodities i.e. if the amount spend on each commodity is large, compared with total expenditure.

In an attempt to deal with the problem of estimating elasticities (income and prices) from family budget data, Leser (1941-42) decided to deal with all commodity groups together. With the aid of 2.2(18), 2.2(20) and the assumption that the cross-elasticities of demand depend only on the nature of the commodity whose price changes and not on the nature of the commodity for which the effect is studied, he derived his basic relation:

$$\frac{1}{1-\omega_i}E_{ii} = \frac{1}{1-\omega_j}E_{jj} + \frac{\omega_j}{1-\omega_j} - \frac{\omega_i}{1-\omega_i} + E_{j0} - E_{i0} \qquad 2.4(3)$$

His additional assumption stated above, in notation, is

$$E_{ij} = E_{kj} \qquad 2.4(4)$$

$$i \neq j,\ k \neq j.$$

The implication of this assumption is that the change in the price of a good causes the expenditure on it to change in the same direction as price, if the demand is price inelastic; or in the opposite direction, if the demand is price-elastic. The expenditure for the remaining commodities has therefore to be adjusted. The assumption then is that these adjustments are made in such a way, that demand for all those commodities increases or decreases by the same proportion. This means that the assumption excludes a high degree of substitutability or complementarity. The relation 2.4(3) can be expressed as

$$E_{ii} = -\omega_i - (1-\omega_i)(E_{i0} + G) \qquad 2.4(5)$$

$$\text{where } G = -E_{j0} - \frac{(\omega_j + E_{jj})}{1-\omega_j}.$$

In this form, the own-price elasticity of a commodity is expressed as a function of its budget proportion, Engel elasticity and a constant as far as the commodity is concerned, since the constant depends on the parameters of another commodity. The value of G is likely to be small, possibly always less than unity. A combination of 2.2(20) and 2.4(4) leads to

$$E_{ij} = \frac{-\omega_j}{1-\omega_j}(1 + E_{jj}) \qquad 2.4(6)$$

and substituting 2.4(5) into 2.4(6), we have

$$E_{ij} = -\omega_j(1 - E_{j0} - G) \qquad 2.4(7)$$

With the aid of 2.4(5) and 2.4(7), the own-price and the cross-price elasticities can be derived, except for the unknown G. These results are also subject to the stability conditions which are based on 2.2(13), 2.2(14) and 2.2(15). Even though Leser (1941-42) varied the assumption given by 2.4(4), such that this relation became the expected value of the adjusted assumption, his basic result did not change so that 2.4(5) still holds.

Houthakker (1960b), discussed similar problem, assuming an additive preference ordering giving rise to either a direct or an indirect additive utility function. He was mainly interested in the comparison of the performance of the two kinds of additive utility functions. His final results with regards to indirect additive utility functions are very close to 2.4(5) and so it is not necessary to go into their details.

Frisch has done more than any other scholar in tackling the problem posed at the beginning of this section, but we shall restrict ourselves to his last and most important contribution |Frisch (1959)| as far as the approach of this study is concerned. His approach is similar to the others in the making of assumptions about the consumers' preferences. He called his main assumption, on which the work rests, want-independence. The concept of want-independence is closely related to the idea of utility trees developed by Strotz (1957). Put simply, this means that a commodity i is want-independent of all the other goods if the marginal utility of commodity i depends only on the quantity of commodity i and not on any other quantity, or commodity i is want-independent of all the other commodities if the quantity of commodity i influences only the marginal utility of commodity i and not any of the other marginal utilities. With the aid of most of the restrictions given in section 2.2, in addition, defining the concepts of utility acceleration and want elasticities, he set out a group of formulae which can be used to get price-elasticities if budget proportions, Engel elasticities and utility elasticity as a result of change in income[1] are known. From 2.2(4), we know that the marginal utilities are functions of quantities, i.e.

$$u_i = u_i(q_1, q_2, \ldots q_n) \qquad 2.4(8)$$

$$i = 1,2,\ldots n.$$

The inverse functions can then be expressed as

$$q_i = q_i(u_1, u_2, u_3, \ldots u_n) \qquad 2.4(9)$$

$$i = 1,2,\ldots n.$$

Then utility acceleration is defined as

$$u'_{ik} = \frac{\partial u_i(q_1 q_2 \ldots q_n)}{\partial q_k} \times \frac{q_k}{u_i(q_1, q_2, \ldots q_n)}$$

$$i,k = 1,2,\ldots n$$

and want elasticities as

---

[1] This is usually the form in which Frisch defined the concept of money flexibility as shown below.

$$q'_{ik} = \frac{\partial q_i(u_1 u_2 \ldots u_n)}{\partial u_k} \times \frac{u_k}{q_i(u_1, u_2, \ldots u_n)}$$

$$i, k = 1, 2, \ldots n.$$

These two concepts are not invariant under a general transformation of u. The want elasticity q'_{ik} will indicate whether the want for commodity k is elastic or not with respect to the quantity of i. This means that if there were a given change in the marginal utility of commodity k, do we have to change the quantity q_i little or much. After showing several properties of these two concepts and their relation to the restrictions of section 2.2 he derived the following formulae:

$$E_{ii} = - E_{i0}\left(\omega_i - \frac{1 - \omega_i E_{i0}}{\check{\omega}}\right) \qquad 2.4(10)$$

i being want-independent of all other goods and $\check{\omega} = \frac{\partial \lambda}{\partial x} \cdot \frac{x}{\lambda}$ (money flexibility as defined by Frisch)

$$E_{ik} = E_{i0}\omega_k\left(1 + \frac{E_{k0}}{\check{\omega}}\right) \qquad 2.4(11)$$

if $q'_{ik} = 0$ for this particular (i,k) combination. No further assumption is needed.

$$\check{\omega} = \frac{E_{i0}(1 - \omega_i E_{i0})}{E_{ii} + \omega_i E_{i0}} \qquad 2.4(12)$$

i being want-independent of all the other goods

$$\check{\omega} = - \frac{\omega_k E_{i0} E_{k0}}{E_{ik} + E_{i0}\omega_k} \qquad 2.4(13)$$

if $q'_{ik} = 0$ for this particular (i,k) combination. No further assumption is needed.

and

$$E_{ik} = - E_{i0}\omega_k \frac{1 - E_{kk}}{1 - \omega_k E_{k0}}, \quad i \neq k \qquad 2.4(14)$$

k being want-independent of all other goods,
whether i is want-independent or not.

Frisch (1959) suggested that if commodity i is want-independent of all the other goods, while k is not, we can use 2.2(24) together with 2.4(14).

For practical purposes there are several ways of applying 2.4(10) to 2.4(14). In the first place if there is a satisfactory estimate of ω then the own-price elasticities can be estimated using 2.4(10) and the cross-price elasticities using 2.4(11) or using 2.4(14) together with 2.2(24) depending on the assumption made. If the ω is unknown, it may be necessary to begin by making independent estimate for ω which we may apply to either 2.4(12) or 2.4(13) as the case may be. Apart from this it is a promising research work to determine ω for different groups of consumers since it serves as a measure of differences in expenditure patterns between the different groups.

Apart from dealing with this kind of want-independence, Frisch (1959) extended his analysis to cover the case where there are groups of commodities which are such that there might exist any sort of want dependence within each group, but no want dependency between a commodity in a given group and a commodity outside of the group, a kind of separability of utility function or the almost additive preference of Barten (1964b). The present study is not extended to cover this.

From the general review of all these past studies, a common thinking is revealed. There is the common assumption about the degree of substitutability or complementarity between commodities and, in all, the presence of an unknown parameter. An interesting investigation will be to find out how significantly different are the results, at least empirically. In section 2.5, a specific function will be considered.

For completeness, it is necessary to touch on the recent contribution of Barten (1964a) in which the main theme is the derivation of price-elasticities without apparently imposing any serious restrictions on the assumed behaviour of the consumer. He reformulated and analysed the kind of equilibrium analysis of consumer demand as shown in section 2.2 to include family effects, and tried to show the similarity between price and family effects for infinitesimally small change. He derived a pseudo Slutsky equation of the kind given by 2.2(12) for family effects. This means that family effects can be decomposed into two parts, the specific effect and the overall effect, in which the overall effect is the sum of weighted specific effects. The weights are the price elasticities. In addition, he indicated how to estimate simultaneously the family effects and price elasticities from cross-section data. This approach will not be given detailed attention in this study because of the type of data available for use at present. It is certainly a useful approach if there are suitable data. Such data should contain no zero observations on the variables. This is at present not possible for Nigerian data.

2.5 CHOICE OF A SYSTEM OF ENGEL FUNCTIONS

In section 2.3 an attempt was made to choose a set of Engel functions which will satisfy all the criteria set out in the section - goodness of fit for a wide range of data, computational simplicity and theoretical plausibility. Linear and double-logarithmic forms failed in at least one criterion each. The goodness of fit for the linear form is poor relatively while adding-up condition, a special case of theoretical plausibility, is not fulfilled by the double-logarithmic form. No choice was then made. It is necessary to search for the 'ideal' functions by considering different systems of demand functions which have some elements of linear or double-logarithmic forms. It seems that any function that satisfies these criteria must have some elements of at least one of these two functions. Hence a variety of complete systems of demand functions that are proposed in one way or the other are derived to remove the defects of linear or double-logarithmic functions. Many systems have been proposed and used by many scholars. In recent empirical work, Parks (1969), Yoshihara (1969) and most of the recent references they cited, examine the following systems of demand functions theoretically and empirically:

The system of Double-logarithmic Functions

$$q_i = A_i x^{b_{i0}} \prod_{j=1}^{n} p_j^{b_{ij}} \qquad 2.5(1)$$

$$i = 1,2,\ldots n.$$

Theil-Barten's Differential System

$$w_i d(\log q_i) = b_i \left| d(\log x) - \sum_{k=1}^{n} w_k d(\log p_k) \right|$$

$$+ \sum_{j=1}^{n} c_{ij} \left| d(\log p_j) - \sum_{k=1} b_k d(\log p_k) \right| \qquad 2.5(2)$$

$$i = 1,2,\ldots n.$$

Stone's Linear Expenditure System

$$y_i = c_i p_i + b_i (x - \sum_{j=1}^{n} c_j p_j) \qquad 2.5(3)$$

$i = 1,2,\ldots n.$

Houthakker's Indirect Addilog. System

$$q_i = \frac{a_i \left(\frac{x}{p_i}\right)^{\beta_i + 1}}{\Sigma a_j \left(\frac{x}{p_j}\right)^{\beta_j}} \qquad 2.5(4)$$

$i = 1,2,\ldots n.$

All the variables used in 2.5(1) to 2.5(4) have the meanings as used in the previous sections, while the A's, b's, a's, β's, c's are parameters to be estimated; d(log x), d(log q_i) and d(log p_k) are differentials with respect to the variables they contain. On theoretical grounds, each must satisfy the restrictions, 2.2(2), 2.2(18), 2.2(19), 2.2(21) and 2.2(24). In addition, for empirical purposes, Yoshihara (1969) introduced a further restriction that all income elasticities must not be equal to one. For, if they are equal to one, the proportions of total expenditure spend on different goods remain the same when prices do not change. Although expenditure shares change in reality, this does not provide a complete justification for introducing the new restriction, since the changes in shares might be due to changes in prices. But since the expenditure pattern of a consumer would probably change when he is presented with rising income and constant prices, it is reasonable to introduce the restriction that income elasticities are not all equal to one i.e.

$$E_{io} \neq 1 \qquad 2.5(5)$$

for some i.

The system of double-logarithmic functions, 2.5(1) fails to satisfy either 2.2(2) or 2.5(5) and because of this failure, it is not satisfactory theoretically. The Theil-Barten differential system, 2.5(2), fails to satisfy 2.5(5) on theoretical grounds, while 2.5(3) and 2.5(4) are satisfactory all round on theoretical grounds. Parks (1969) ignored condition 2.5(5) and examined 2.5(2), 2.5(3) and 2.5(4) while Yoshihara (1969) examined only 2.5(3) and 2.5(4). On empirical grounds there are no conclusive results on which one to choose. Parks (1969) found the Theil-Barten Differential System to fit the data used, best statistically. But Yoshihara (1969) favoured that of the Linear Expenditure System. They both used time series data from different countries, and the method of estimating the parameters of each system is different. As a result of all these, comparison is not easy.

For family budget studies, it seems that, apart from theoretical grounds, the next thing that must be considered is the form to which each of 2.5(1) to 2.5(4) reduces when considered as Engel functions. In this case, 2.5(1) reduces to the set of ordinary transformed double-logarithmic functions, 2.5(2) reduces to the set of ordinary transformed differential double-logarithmic functions, 2.5(3) to the set of ordinary linear functions and 2.5(4) becomes

$$q_i = \frac{a_i' x^{\beta_i + 1}}{\sum_{j=1}^{n} a_j' x^{\beta_j}} \qquad 2.5(6)$$

$$i = 1,2,\ldots n.$$

The system of functions, 2.5(6) satisfies all the conditions laid down above including 2.5(5) and so is the system of choice.

As a system of Engel functions, 2.5(6) satisfies the theoretical plausibility mentioned in section 2.3. With regards to computational simplicity, parameters can be estimated by using estimates derived from the ordinary double-logarithmic Engel functions. This means that not much extra work is involved. Houthakker maintained that it could be shown that this procedure of using the estimates of the ordinary double-logarithmic Engel functions gives maximum likelihood estimates.[1] Since the estimates of the parameters are derived from the ordinary double-logarithmic Engel functions, the goodness of fit for a wide range of data will be reasonably good. The Indirect Addilog. System as a system of Engel functions was developed and used originally by Leser (1941-42) and further developed and used by Wit and Somermeyer (1956), Houthakker (1960a,b) and Somermeyer et. al. (1962). The rest of this section is devoted to its development following the approach adopted by Houthakker (1960a). To begin with, an Engel function $f_i(x)$ which lacks the adding-up condition in the sense defined in section 2.3 can be transformed into one that satisfies it by writing

$$g_i(x) = \frac{x f_i(x)}{\sum_{j=1}^{n} f_j(x)} \qquad 2.5(7)$$

$$i = 1,2,\ldots n.$$

This means that $\sum_{i=1} g_i(x) \equiv x$. If this is applied to the double-logarithmic function in nontransformed form, we have

$$y_i = \frac{\alpha_i x^{\beta_i + 1}}{\sum_{j=1}^{n} \alpha_j x^{\beta_j}} \qquad 2.5(8)$$

$$i = 1,2,\ldots n.$$

Because of the unweildiness of the parameters, we consider the ratio for two commodities, say i and j and the resulting relation is

[1] Houthakker (1960b), p. 254.

$$y_i/y_j = \frac{\alpha_i}{\alpha_j} x^{\beta_i - \beta_j}. \qquad 2.5(9)$$

Taking the logarithm of 2.5(9) we have

$$\log y_i - \log y_j = \log \frac{\alpha_i}{\alpha_j} + (\beta_i - \beta_j)\log x \qquad 2.5(10)$$

This means that the ratios between any two expenditures have a constant elasticity with respect to total expenditure. But the income elasticity of demand[1] of 2.5(8) is:

$$E_{i0} = \beta_i + 1 - \sum_{j=1}^{n} \beta_j w_j. \qquad 2.5(11)$$

---

1 This result is derived by considering 2.5(8), in which case

$$\frac{\partial y_i}{\partial x} = \frac{\alpha_i(\beta_i+1)x^{\beta_i} \sum_{j=1}^{n} \alpha_j x^{\beta_j} - \alpha_i x^{\beta_i+1} \sum_{j=1}^{n} \alpha_j \beta_j x^{\beta_j - 1}}{\left(\sum_{j=1} \alpha_j x^{\beta_j} \right)^2}$$

$$= \frac{(\beta_i+1)\alpha_i x^{\beta_i}}{\sum_{j=1} \alpha_j x^{\beta_j}} - \frac{\alpha_i x^{\beta_i+1}}{\sum_{j=1} \alpha_j x^{\beta_j}} \cdot \frac{\Sigma \beta_j \alpha_j x^{\beta_j - 1}}{\sum_{k=1}^{n} \alpha_k x^{\beta_k}}$$

$$= (\beta_i+1)\frac{y_i}{x} - \frac{y_i \sum_{j=1}^{n} \beta_j y_j}{x^2}$$

$$\therefore E_{i0} = \frac{x}{y_i}\frac{\partial y_i}{\partial x} = \beta_i + 1 - \sum_{j=1}^{n} \beta_j \frac{y_j}{x}$$

$$= \beta_i + 1 - \sum_{j=1}^{n} \beta_j w_j.$$

This result gives an important difference between 2.5(8) and the ordinary double-logarithmic function which has constant elasticity. The difference depends on $1 - \sum_{j=1}^{n} \beta_j \omega_j$. The closer $\sum_{j=1} \beta_j \omega_j$ is to unity, the closer 2.5(11) is to the constant elasticity of the ordinary double-logarithmic function.

Since our interest is not only in income elasticity, but also in the price elasticities, we next derive the price elasticities for the corresponding demand system. The corresponding indirect utility function is

$$v(p_1, p_2, p_3, \ldots p_n, x) = \Sigma \alpha'_i \left(\frac{x}{p_i}\right)^{\beta'_i} \qquad 2.5(12)$$

and the constrained utility optimisation leads to the following system of demand functions:

$$q_i = \frac{\alpha'_i x^{\beta'_i+1} p_i^{-\beta'_i-1}}{\sum_{j=1}^{n} \alpha_j x^{\beta'_j} p_j^{-\beta'_j}} \qquad i = 1,2,\ldots n. \qquad 2.5(13)$$

This relation, 2.5(13), is the same as 2.5(4) as it should be. Following the method adopted in deriving 2.5(11), the following own-price elasticity for commodity i is

$$E_{ii} = -\beta'_i - 1 + \beta'_i \omega_i \qquad 2.5(14)$$

and the cross-elasticity with respect to commodity i as a change in the price of commodity j is

$$E_{ij} = \beta'_j \omega_j \qquad 2.5(15)$$

which is free of commodity i. This satisfies the relation, 2.4(4). But it is not useful getting these two results, 2.5(4) and 2.5(15), from family budget data unless the β' 's are estimated with the aid of data including price series. This means that we can not go far with these two relations (2.5(14) and 2.5(15)) using the present data. We have to resort to the general results given by 2.4(5) and 2.4(7). Substituting 2.5(11) into 2.4(5) to replace E_{i0}, we have

$$E_{ii} = -\beta_i - 1 + \beta_i\omega_i + (1 - \omega_i)\sum_{j=1}^{n}\beta_j\omega_j - (1 - \omega_i)G. \qquad 2.5(16)$$

Since G by its definition is likely to be small and since $1 - \omega_i$, also by definition is less than unity in absolute value, the term $(1 - \omega_i)G$ is likely to be very small and less than unity in absolute value. As a first approximation the last term may be dropped[1] and 2.5(16) becomes

$$E_{ii} = -\beta_i - 1 + \beta_i\omega_i + (1 - \omega_i)\sum_{j=1}^{n}\beta_j\omega_j. \qquad 2.5(17)$$

This approach seems reasonable, since 2.5(11) can be legitimately estimated from budget data. From the relation 2.5(16), the defect of using 2.5(14) easily showed itself. In the same vein, 2.4(7) can be replaced by

$$E_{ij} = \omega_j\beta_j - \omega_j\sum_{j=1}\beta_j\omega_j + \omega_j G. \qquad 2.5(18)$$

And because of the smallness of $\omega_j G$ it may be dropped too and this leads to

$$E_{ij} \quad \omega_j\beta_j - \omega_j\sum_{j=1}\beta_j\omega_j. \qquad 2.5(19)$$

With the knowledge of some estimates of E_{i0}, E_{ii} and ω_i $(i = 1,2,\ldots n)$; Frisch's money flexibility, ω can be estimated by using 2.4(12); which can then be used to estimate E_{ii} and E_{ij} using Frisch's formulae 2.4(10) and 2.4(11). The two methods of estimating E_{ii} and E_{ij} can then be compared.

---

[1] Different values of $(1-\omega_i)G$ between 0 and 1 can be used to test the sensitivity of the estimates.

CHAPTER 3

STATISTICAL BASIS AND COMPUTATIONS

3.0 INTRODUCTION

A comparison of real life facts of consumers' behaviour with theory is usually made on the basis of statistical models and inferences. It seems it is the only approach. This chapter forms the bridge between the theory of chapter 2 and the data used, to be discussed in chapter 4. This chapter examines the concepts measured; briefly discusses factors, other than income and prices, affecting consumption patterns; and in great detail discusses the statistical model adapted and the various statistical inference approaches used in the study.

Various problems are raised in the last chapter which have to be put into some systematic form. The basic relationship given in chapter 2 is, that each specific expenditure depends on income, i.e. relation 2.3(1). This has to be adorned with the usual error term to take the place of errors of specification and other errors, after putting the relation in an explicit form, of the kind given by 2.5(8). Then the variables are expressed in terms of per consumer unit to remove family effects as far as possible, as discussed below. Two kinds of consumer units are examined. The first kind is 'per person' calculation, in which the expenditures and income for each family are each divided by the number of persons in the family. This is based on the so-called homogeneity hypothesis. It is represented by 2.3(2). Then there is the concept of 'adult equivalences' calculation, in which the expenditures and income of each family are each divided by a measure which is based on regarding each member of the family as a fraction of an adult. This is represented by 2.3(3). These concepts with regard to consumption units are discussed fully in section 3.4. There is the problem of choice of Engel functions which was discussed in chapter 2 and a choice of 2.5(8) was made. It was stated in section 2.5 that the parameters of this system of Engel functions could be estimated using the ordinary double-logarithmic functions, and that such estimates possess maximum likelihood properties as stated by Houthakker (1960b). The main difference is in the elasticities derived from them. As a result of this, we shall estimate the parameters of the relationship of the form

$$y_{ij} = \beta_{0i} + \beta_{ii}x_j + \epsilon_{ij} \qquad 3.0(1)$$

$$i = 1,2,\ldots n; \quad j = 1,2,\ldots N$$

where y_{ij} is the logarithm of expenditure on commodity i and x_j is the logarithm of income all per consumer unit in family j; β_{0i} and β_{ii} are parameters relating to commodity i while ϵ_{ij} is the error term.

We assume that ϵ_{ij} satisfies all the usual assumptions that allow the use of ordinary least squares method to be applied to 3.0(1), namely:

$$E(\epsilon_{ij}) = 0 \qquad 3.0(2)$$

$$E(\epsilon_{ij}\epsilon_{ik}) = \sigma^2 \quad \text{for } j = k \qquad 3.0(3)$$

$$= 0 \quad \text{for } j \neq k$$

$$E(x_j\epsilon_{ij}) = 0 \, . \qquad 3.0(4)$$

And for the purpose of statistical tests of hypothesis that ϵ_{ij} is a random variable normally distributed with mean zero and variance σ^2. The relation given by 3.0(1) gives the impression that the kind of data available are observed individually and that they are not grouped. This is not so as shown below.

Apart from considering how family effects are incorporated into the analysis by using different concepts of consumer units, it is worth mentioning here, how factors like location and occupational differences are brought into the analysis. Relations of the kind given by 3.0(1) are fitted for each group of families. A group may be made homogeneous with regards to location, and/or occupation. By so doing we have sets of regression lines for each commodity expenditure - income relation, which are meant to indicate differences between location and/or occupation if there are any. As a result of this, a method of comparing sets of regression lines is developed and used to analyse the factors.

In the concluding section of this chapter, the computational procedures adopted are stated. It is not possible to discuss all the statistical techniques used in this chapter. What is discussed here is only basic to the study. If there is need to use any technique not discussed here, possibly dictated by some results from the data, some effort is taken to explain such technique before application. Needless to say that various simple statistical tests are used too.

3.1 REGRESSION ANALYSIS OF GROUPED DATA

Most budget data available for econometric studies are in grouped form. Even when they are not, it may be essential and necessary to group them. An important objective of statisticians is data reduction so as to make information supplied by a mass of data digestible and meaningful to the users. This is achieved by grouping and ultimately estimating a few parameters that will characterise the set of observations. Apart from this, public policy, with regards to data collection, requires that the data referring to individuals may not be published and must be grouped. Grouping may even reduce computation, so that there is less laborious numerical computation of individual observations. By grouping, as adopted in this study, following the work of Prais and Aitchison (1954), we mean a number of observations are taken together and the average values of all the variables are calculated for the group. Each average value may then be treated for all purposes as a repeated observation, the number of repetitions being equal to the number of observations in the appropriate group. Prais and Aitchison (1954) maintained that this kind of grouping leads to unbiased estimators even though the variance is not minimum since it will always be greater than that to be attained from the ungrouped form of data. The loss of efficiency is very small. Prais and Aitchison (1954) showed that in the case of one determining variable, the optimum method of grouping is obviously to arrange the observations in order of magnitude of that variable; and if they are equally spaced then, for a given number of groups, the method is to choose the number of observations that are to fall into each group so as to maximize the between groups variance which is the same as minimizing the within groups variance (this is akin to stratification principle in sampling theory). This general principle suggests that it is desirable to group together observations which are as homogeneous as possible with respect to the determining variables. The loss of efficiency depends on the determining variables only. An advantage of this approach is that the usual Shephard's correction is not required. As will be shown below, in this section, the unbiased estimator of σ^2 is given by the residual weighted sum of squares divided by the number of degrees of freedom; here, equal to the number of groups less the number of fitted parameters.

There are two major methods of grouping. First, the method of grouping usually associated with Shephard's correction, i.e. for the regression of y on x, the range of x is divided into intervals and each value of y falling within an interval is associated with the mid-point of that interval. And the second, the method in which the observations are grouped according to some scheme and the value of the variables are replaced by their mean values. A special case of this method is grouping which, while considerably reducing the amount of computation, leads to the same arithmetic results as would an ungrouped method. This is always achieved when the determining variable is discrete and there is only one determining observation per interval. The work of this section is of the second method. Most grouped family budget data are usually of this type and the data available for this study are not exceptions. They are all grouped according to the second method.

Grouping, as studied in this section, is a kind of matrix transformation of the observations and so can be studied by an extension of the matrix treatment of the general regression theory. Following Prais and Aitchison (1954), we consider the grouping of five observations, y_1, y_2, y_3, y_4, y_5 into two groups with three of them in the first group and two in the other.

Let $\bar{y}_1 = \frac{1}{3}(y_1 + y_2 + y_3)$ and $\bar{y}_2 = \frac{1}{2}(y_4 + y_5)$

What is carried out above can be put simply in matrix form as

$$\begin{vmatrix} \bar{y}_1 \\ \bar{y}_2 \end{vmatrix} = \begin{vmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{vmatrix} \begin{vmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{vmatrix}$$

or in some better notational form as

$$\bar{y} = Gy$$

where

$$\bar{y} = \begin{vmatrix} \bar{y}_1 \\ \bar{y}_2 \end{vmatrix}, \quad G = \begin{bmatrix} G_3 & 0_2 \\ 0_3 & G_2 \end{bmatrix} \text{ and } y = \begin{vmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{vmatrix}$$

$$\text{with } G_3 = \begin{vmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{vmatrix}, \quad G_2 = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \end{vmatrix}, \quad 0_3 = |000| \text{ and } 0_2 = |00|$$

In general, assume we have N observations to be grouped into m groups with $f_1, f_2, \ldots f_m$ in groups $1,2,3,\ldots,m$ respectively, in such a way that $\sum_{j=1}^{m} f_j = N$

then we have

$$\bar{y} = Gy \qquad 3.1(1)$$

where

$$\bar{y} = \begin{vmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \vdots \\ \bar{y}_m \end{vmatrix}, \quad G = \begin{vmatrix} G_{f_1} & & & 0 \\ & G_{f_2} & & \\ & & \ddots & \\ 0 & & & G_{f_m} \end{vmatrix} \text{ and } y = \begin{vmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{vmatrix}$$

with $G_{f_j} = \left[\frac{1}{f_j} \ \frac{1}{f_j} \ \frac{1}{f_j} \ \ldots \ \frac{1}{f_j} \right]$, a row vector of dimension f_j.

It is easy to show that

$$GG' = \begin{vmatrix} {}^{1}/f_1 & & & 0 \\ & {}^{1}/f_2 & & \\ & & \ddots & \\ 0 & & & {}^{1}/f_m \end{vmatrix} \qquad 3.1(2)$$

In the case of ungrouped data where we are faced with estimating the parameters of 3.0(1), given assumptions 3.0(2), 3.0(3) and 3.0(4), it is well known that the ordinary least squares method gives best linear unbiased estimators (BLUE) of the parameters |see Johnston (1963) and Malinvaud (1970) for some rigorous exposition of this method|. If we put 3.0(1) in matrix form and allowing for k independent variables together with their observed values, we have

$$y = X\beta + \varepsilon \qquad 3.1(3)$$

where

$$y = \begin{vmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{vmatrix}, \quad X = \begin{vmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & & & \\ x_{N1} & & \cdots & x_{Nk} \end{vmatrix},$$

$$\beta = \begin{vmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_k \end{vmatrix}, \text{ and } \varepsilon = \begin{vmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_N \end{vmatrix}$$

and the related assumptions become

$$E(\varepsilon) = 0 \qquad 3.1(4)$$

$$E(\varepsilon\varepsilon') = \sigma^2 I_N \qquad 3.1(5)$$

$$E(X'\varepsilon) = 0 \qquad 3.1(6)$$

The ordinary least squares estimators (representing these with "^" on the parameters) are:

$$\hat{\beta} = (X'X)^{-1}X'y$$

$$\text{Cov}(\hat{\beta}) = \sigma^2(X'X)^{-1} \qquad 3.1(7)$$

and

$$\hat{\sigma}^2 = \frac{1}{N-k}\left|y'y - \hat{\beta}'X'y\right|$$

But if assumption 3.1(5) is not satisfied and instead we have

$$E(\varepsilon\varepsilon') = \sigma^2\Omega, \qquad 3.1(8)$$

where σ^2 and Ω are known values but Ω is a non-unit square matrix;

it is known that by using Aitken's Generalised Least Squares Method |see Aitken (1934)|, the following results corresponding to 3.1(7) can be derived

$$\hat{\beta} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y$$

$$\text{Cov}(\hat{\beta}) = \sigma^2(X'\Omega^{-1}X)^{-1} \qquad 3.1(9)$$

and

$$\hat{\sigma}^2 = \frac{1}{N-k}\left|y'y - \hat{\beta}'X'\Omega^{-1}y\right|$$

In the case of grouped data, since grouping is a kind of matrix transformation as shown above, if the observations are grouped means as defined by 3.1(1), then 3.1(3) will be transformed to

$$\bar{y} = \bar{X}\beta + \bar{\varepsilon} \qquad 3.1(10)$$

where

$$\bar{y} = Gy, \quad \bar{X} = GX \text{ and } \bar{\varepsilon} = G\varepsilon.$$

The corresponding assumptions about $\bar{\varepsilon}$ will be:

$$E(\bar{\varepsilon}) = 0 \qquad 3.1(11)$$

$$E(\bar{\varepsilon}\bar{\varepsilon}') = \sigma^2 GG' \qquad 3.1(12)$$

and

$$E(\bar{X}'\bar{\varepsilon}) = 0 \qquad 3.1(13)$$

Because of the form of 3.1(12), the ordinary least squares method can no longer apply but we can apply the generalised least squares method as given above. From 3.1(2) and 3.1(8), it means that

$$\Omega = GG' = \begin{vmatrix} {}^1/f_1 & & & & \\ & {}^1/f_2 & & & \\ & & {}^1/f_3 & & \\ & & & \ddots & \\ & & & & {}^1/f_n \end{vmatrix}$$

and it can be shown quite easily that

$$\Omega^{-1} = (GG')^{-1} = \begin{vmatrix} f_1 & & & \\ & f_2 & & \\ & & \ddots & \\ & & & f_m \end{vmatrix} .$$

This means that

$$\hat{\beta} = |\bar{X}'(GG')^{-1}\bar{X}|^{-1}\bar{X}'(GG')^{-1}\bar{y}$$

$$\text{Cov}(\hat{\beta}) = \sigma^2|\bar{X}'(GG')^{-1}\bar{X}|^{-1} \qquad 3.1(14)$$

and

$$\hat{\sigma}^2 = \frac{1}{m-k}\,|\bar{y}'\bar{y} - \hat{\beta}'\bar{X}'(GG')^{-1}\bar{y}|$$

In the special case of a dependent variable regressed on an independent variable using data grouped into m groups as above, we have

$$\bar{y} = \beta_0 + \beta_1 \bar{x} + \bar{\varepsilon} \qquad 3.1(15)$$

where

$$\bar{y} = Gy = \begin{vmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \vdots \\ \bar{y}_m \end{vmatrix}, \quad \bar{x} = Gx = \begin{vmatrix} 1 & \bar{x}_1 \\ 1 & \bar{x}_2 \\ 1 & \bar{x}_3 \\ \vdots & \vdots \\ 1 & \bar{x}_m \end{vmatrix} \text{ and } \bar{\varepsilon} = G\varepsilon = \begin{vmatrix} \bar{\varepsilon}_1 \\ \bar{\varepsilon}_2 \\ \vdots \\ \bar{\varepsilon}_m \end{vmatrix}$$

then

$$\begin{vmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{vmatrix} = \begin{vmatrix} \sum_{j=1}^{m} f_j & \sum_{j=1}^{m} f_j \bar{x}_j \\ \sum_{j=1}^{m} f_k \bar{x}_j & \sum_{j=1}^{m} f_j \bar{x}_j^2 \end{vmatrix}^{-1} \begin{vmatrix} \sum_{j=1}^{m} f_j \bar{y}_j \\ \sum_{j=1}^{m} f_j \bar{x}_j \bar{y}_j \end{vmatrix}$$

leading to the following results

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\hat{\beta}_1 = \frac{\sum_{j=1}^{m} f_j \bar{x}_j \bar{y}_j - \sum_{j=1}^{m} f_j \bar{x}_j \sum_{j=1}^{m} f_j \bar{y}_j / \sum_{j=1}^{m} f_j}{\sum_{j=1}^{m} f_j \bar{x}_j^2 - (\sum_{j=1}^{m} f_j \bar{x}_j)^2 / \sum_{j=1}^{m} f_j}$$

$$\text{Var}(\hat{\beta}_1) = \sigma^2 / (\sum_{j=1}^{m} f_j \bar{x}_j^2 - (\sum_{j=1}^{m} f_j \bar{x}_j)^2 / \sum_{j=1}^{m} f_j) \qquad 3.1(16)$$

and

$$\hat{\sigma}^2 = \frac{1}{m-2} \left| \sum_{j=1}^{m} f_j \bar{y}_j^2 - (\sum_{j=1}^{m} f_j \bar{y}_j)^2 / \sum_{j=1}^{m} f_j - \hat{\beta}_1^2 (\sum_{j=1}^{m} f_j \bar{x}_j^2 - (\sum_{j=1}^{m} f_j \bar{x}_j) / \sum_{j=1}^{m} f_j) \right|$$

All the results for the special case can be meaningfully displayed in the usual analysis of variance table as shown in table 1 below:

Table 1. ANALYSIS OF VARIANCE

Source of Variation	d.f	Sums of squares
Total after grouping	m	$\sum_{j=1}^{m} f_j \bar{y}_j^2$
Residual after common mean	m-1	$\sum_{j=1}^{m} f_j \bar{y}_j^2 - (\sum_{j=1}^{m} f_j \bar{y}_j)^2 / \sum_{j=1}^{m} f_j = A$
Due to Regression	1	$\dfrac{\left[\sum_{j=1}^{m} f_j \bar{x}_j \bar{y}_j - \sum_{j=1}^{m} f_j \bar{x}_j \sum_{j=1}^{m} f_j \bar{y}_j / \sum_{j=1}^{m} f_j\right]^2}{\sum_{j=1}^{m} f_j \bar{x}_j^2 - (\sum_{j=1}^{m} f_j \bar{x}_j)^2 / \sum_{j=1}^{m} f_j} = \dfrac{C^2}{B}$
Residual after grouped regression	m-2	$A - \dfrac{C^2}{B}$

From this table, the $\hat{\sigma}^2$ and R^2 can be estimated as

$$\hat{\sigma}^2 = \frac{1}{m-2}\left[A - \frac{C^2}{B}\right]$$

and

$$R^2 = \frac{C^2}{B \times A}$$

where R^2 is defined as the proportion of total variation explained by fitting a regression line on the data.

There is another problem similar to the above in the sense that it results from the breakdown of assumption 3.1(5), although from a different source. This happens when, instead of 3.1(5), we have

$$E(\varepsilon\varepsilon') = \sigma^2\Omega = \begin{vmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_N^2 \end{vmatrix} \qquad 3.1(17)$$

This is the problem of heteroscedasticity. It is serious only when the range of variation in the variables is large. In such a situation, there is need to investigate the distribution of the residual variance. In family budget data, it is usual to find that the variance within each group of data is approximately proportional to the square of the mean expenditure within the group [see Theil (1951) and Brady (1938)]. This means that 3.1(7) is replaced by

$$\sigma^2\Omega = \sigma^2 \begin{vmatrix} E(y_1)^2 & & & \\ & E(y_2)^2 & & \\ & & \ddots & \\ & & & E(y_N)^2 \end{vmatrix}$$

In the case of grouped data, this expression becomes

$$\sigma^2 G\Omega G' = \begin{vmatrix} \bar{y}_1^2/f_1 & & & \\ & \bar{y}_2^2/f_2 & & \\ & & \ddots & \\ & & & \bar{y}_m^2/f_m \end{vmatrix}$$

which will give the same results as 3.1(14) except that $\sigma^2 G\Omega G'$ now replaces $\sigma^2 GG'$. When we consider the simple regression given by 3.1(15), f_j is replaced by $f_j/\bar{y}_j^2$. This assumption can be generalised as

$$E(\epsilon_j^2) = \{E(y_j)\}^r \qquad j = 1,2,\ldots N.$$

In some situations, there is the possibility that grouping may even help to reduce the problem of heteroscedasticity. This is the hope in this study since this problem is not fully examined further.

3.2 COMPARISON OF REGRESSION LINES[1]

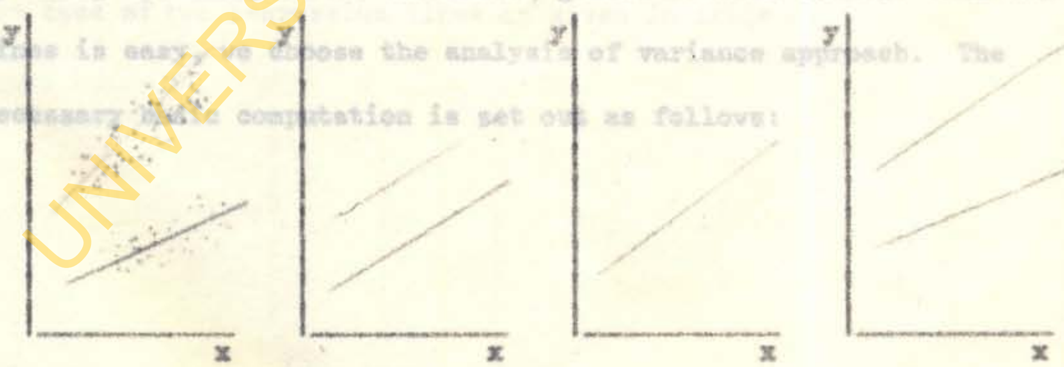
As indicated in the introductory consideration of this chapter, given above; several regression lines of the form 3.1(15), where the variables are defined as in 3.0(1), are fitted. It is necessary to test whether their slopes, at least, are significantly different. It may be necessary, for example, to test whether Engel elasticity for food among clerks is the same as that for farmers; or even further whether this elasticity is the same for clerks, farmers, traders and labourers in the same community. Such analysis helps the pooling of original data together particularly if the elasticities are not significantly different. This problem is essentially that of comparison of regression lines, or a more general case of the analysis of covariance,[2] since other tests are involved. It is not enough to test just the elasticities which are just the slopes of the regression lines

---

[1] This section is based on the notes taken from Dr. Quenouille's lectures, (as an undergraduate) at the London School of Economics (University of London) in 1962/63 session. An elementary and practical treatment of this technique is in Quenouille (1966), chapter 7.

[2] See Malinvaud (1970), pp. 271-275.

if 3.1(15) is used as defined above. It is possible to have two or more parallel lines i.e. lines with the same slope, but which are separate and distinct. For this reason, it is then necessary to test whether the distances between the lines are significantly different from zero. It is not possible to test the distances between the lines unless the lines are parallel. This means that test between regression lines is consequently carried out in steps. First, the residual mean squares in the regressions are compared using the variance-ratio test to see whether the variations about the lines are comparable (fig. 1a). If they are, they may then be pooled. Second, the regression coefficients are compared by using variance-ratio test (or t test in the case of two regression lines). If they are not significantly different (fig. 1b), the joint regression coefficient may be used. Then at the third step, the distances between the lines are tested to see if they are significantly different from zero (fig. 1c). The problems involved are shown diagramatically in fig. 1 in the case of two regression lines. Fig. 1d. is the case where



(a) Same variation about regression lines (b) Same slope (c) Same regression line (d) Different regression lines

Fig. 1 Comparison of Regression Lines

all the hypotheses are rejected. In what follows, the situation depicted by fig. 1a will not be considered. We shall assume that the variations in the regressions are the same. The other situations as we shall see later corresponding to testing between slopes, between means and between regressions.

As a starting point, we consider the case of two regression lines of the form 3.1(15). The two regressions are

$$\text{Set 1:} \quad \bar{y}_1 = \beta_{10} + \beta_{11}(\bar{x}_1 - \bar{\bar{x}}_1) + \varepsilon_1$$

$$\text{Set 2:} \quad \bar{y}_2 = \beta_{20} + \beta_{21}(\bar{x}_2 - \bar{\bar{x}}_2) + \varepsilon_2$$

$$\text{where} \quad \bar{\bar{x}}_i = \sum_{j=1}^{m} f_j \bar{x}_{ij} / \sum_{j=1}^{m} f_j.$$

$$i = 1,2$$

We are interested in examining whether there is any significant difference between the two lines assuming the variations in the regressions are the same. There are two ways of approach. There is the use of t and F tests combined according to the steps discussed above. But analysis of variance approach can be adopted. Since by using the analysis of variances, generalisation to more than two lines is easy, we choose the analysis of variance approach. The necessary basic computation is set out as follows:

Sums of Squares and cross products	Set 1	Set 2	Combined Data
$\sum_{j=1}^{m} f_j\bar{y}_j^2 - (\sum_{j=1}^{m} f_j\bar{y}_j)^2 / \sum_{j=1}^{m} f_j$	A_1	A_2	$A \geq A_1 + A_2$
$\sum_{j=1}^{m} f_j\bar{x}_j^2 - (\sum_{j=1}^{m} f_j\bar{x}_j)^2 / \sum_{j=1}^{m} f_j$	B_1	B_2	$B \geq B_1 + B_2$
$\sum_{j=1}^{m} f_j\bar{y}_j\bar{x}_j - \sum_{j=1}^{m} f_j\bar{y}_j \sum_{j=1} f_j\bar{x}_j / \sum_{j=1}^{m} f_j$	C_1	C_2	C
Sample size	m_1	m_2	$m_1 + m_2$
Regression Coefficient	C_1/B_1	C_2/B_2	C/B

This is followed by the analysis of variance of each set and the combined data separately. On the basis of such separate analyses, a combined analysis for comparing the regression lines is carried out. The case of two regression lines is given in table 2.

Table 2. FORMAT FOR COMBINED ANALYSIS OF VARIANCE FOR COMPARING TWO REGRESSION LINES

Source	d.f.	s.s.	m.s.*
After fitting different means and different slopes (Residual)	$m_1 + m_2 - 4$	$A_1 + A_2 - \frac{C_1^2}{B_1} - \frac{C_2^2}{B_2} = D$	s^2
After fitting different means and common slope	$m_1 + m_2 - 3$	$A_1 + A_2 - \frac{(C_1+C_2)^2}{B_1 + B_2} = E$	
After fitting a common mean and a common slope	$m_1 + m_2 - 2$	$A - \frac{C^2}{B} = F$	
Due to fitting a common slope	1	$\frac{C^2}{B}$	Same as s.s.
Due to difference between slopes	1	$E - D$	Same as s.s.
Due to difference between means	1	$F - E$	Same as s.s.
Due to regression	3	$F - D + \frac{C^2}{B}$	$\frac{1}{3}(F - D + \frac{C^2}{B})$

* Value under s.s. is divided by corresponding value under d.f.

Each component of table 2, particularly the last three can then be compared with the first component as is usual in analysis of variance.

In the general case where there are r regression lines, the necessary sums of squares and cross-products are computed; and the residual sum of squares for each set and the combined data separately are derived. Then the corresponding table for the combined analysis of variance for comparison of r regression lines is given below as table 3.

Table 3. FORMAT FOR COMBINED ANALYSIS OF VARIANCE FOR COMPARING r REGRESSION LINES

Source	d.f.	s.s.	m.s.*
After fitting different means and different slopes (Residual)	$\sum_{h=1}^{r} m_h - 2r$	$\sum_{h=1}^{r} A_h - \sum_{h=1}^{r} \frac{C_h^2}{B_h} = D$	S^2
After fitting different means and a common slope	$\sum_{h=1}^{r} m_h - r - 1$	$\sum_{h=1}^{r} A_h - \frac{\left[\sum_{h=1}^{r} C_h\right]^2}{\sum_{h=1}^{r} B_h} = E$	
After fitting a common mean and a common slope	$\sum_{h=1}^{r} m_h - 2$	$A - \frac{C^2}{B} = F$	
Due to common slope	1	$\frac{C^2}{B}$	
Due to difference between slopes	r - 1	E - D	S_1^2
Due to difference between means	r - 1	F - E	S_2^2
Due to regressions	2r - 1	$F - D + \frac{C^2}{B}$	S_3^2

* Value under s.s. is divided by the corresponding value under d.f.

Again only the last three components need to be compared with the first component for the usual variance-ratio test. The three components correspond with the hypotheses illustrated for the case of two regression lines by figure 1 (b. to d.).

3.3 ERRORS IN VARIABLES AND SIMULTANEOUS EQUATION PROBLEMS

In section 3.1, we examine how a single equation can be estimated, particularly when the observations on the variables are group means. In actual fact the single equation is one of a set of equations. In such a situation, assumption 3.1(13) for equation 3.1(10) is not likely to be satisfied. For example, if equation 3.1(10) is a member of a system of equations, some of the variables $\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots \bar{x}_k$ which have the matrix of observation $\bar{X}$ are likely to be endogenous. In which case those which are endogenous will be correlated with the error term associated with 3.1(10), hence 3.1(13) will not be satisfied. Another problem, which at first looks different but can be solved in the same way, as done below, concerns the fact that the analysis, so far, is based on the assumption that the true values of the variables are measured with no error. This is seldom correct, particularly in the kind of data available for econometric studies. These two points create similar problems hence they are treated together.

The ordinary least squares approach to regression analysis will result in biased estimates, even asymptotically. This means that even large samples cannot rescue the estimates from being biased. Summers (1959) was the first to point out the simultaneous equation trouble with regards to family budget data analysis, using the work of Prais and Houthakker (1955) as example. Liviatan (1961) developed the problem further and tried to solve it by the method of instrumental variables. They both assessed the gravity of the problem of bias involved. The problem of errors in variables and the use of instrumental variables are well discussed by Durbin (1954) while the method of instrumental variables is further developed by Sargan (1958). This section is mainly application of these sources to the simple case, involving only one instrumental variable.

Summers (1959) based his criticism on the assumption, which is usually made, that a household's expenditures on commodities in its budget depend upon its income (and other variables) but that its income does not depend upon its expenditures. If this is the case, the least squares method will be unbiased; but he maintained that there are some households which decide upon the amount of effort they will expend in order to earn income on the basis of how much they wish to spend. In the absence of empirical evidence, either way, it is not unreasonable to assume for statistical purposes that income does not depend upon expenditures. But Summers (1959) went further to say, "that the assumption that there is no 'feed back' between expenditure on individual items and total expenditure is not acceptable."[1] Prais (1959), in his reply to Summers, put up an alternative assumption that a household first decides on its total expenditure and then allocates it to the various commodities, comprising its budgets while one of the items acts as buffer absorbing the net balance of positive and

---

[1] Summers (1959), p. 122.

negative deviations from 'expected' expenditures on other commodities. Prais maintained that under such a situation, the ordinary least squares will not lead to any bias. In addition, it is useful to consider the concept of income and its relation to consumers' purchasing decision processes.

There are various hypotheses that have been put up since the pioneering work of Keynes (1936) who maintained that current consumption expenditure is a highly dependable and stable function of current income. This is the absolute income hypothesis. It is a first approximation of real life situation and so it is no surprise that it has remained as the springboard of other hypotheses as shown below. Brady and Friedman (1947) and to a large extent Duesenberry (1949) suggested that a consumer unit's consumption depends not on his absolute income but on his position in the distribution of income among consumer units in his community - the relative income hypothesis. While Tobin (1951) put up the wealth-income hypothesis, maintaining tentatively that changes in wealth may explain the rough constancy over time in the fraction of income saved. And Friedman (1957) propounded the permanent income hypothesis. The gist of this hypothesis is that consumption is not continually being adapted to casual changes in income. Each household fixes its expenditure according to the level which it considers normal for its income. An exceptionally high level of income in any one period will lead to savings against future needs, and conversely a temporary decrease in income will be compensated by the use of previously accumulated savings. The level of income which is considered normal is therefore more significant than the amount actually received.

It is not possible to incorporate all these ideas into the model of this study as that would extend it beyond its scope. At least an important fact comes out clearly from the discussion above i.e. that income whatever the meaning attached seems to be an exogenous variable which consumers adjust to, whether immediately or in the future. A useful model that can be adopted must relate specific expenditures on the commodities to total expenditure, and total expenditure related to income. In so doing commodity expenditures and total expenditure are treated as endogenous and income as exogenous. It is a kind of Engel-Keynes model. The reduced form of the model as dictated by the theory of consumer's behaviour is:

$$y_i' = \theta_{0i} + \theta_{1i}x' + v_i' \qquad i = 1,2,\ldots n$$

$$z' = \theta_{00} + \theta_{10}x' + v_0' \qquad 3.3(1)$$

$$\sum_{i=1}^{n} y_i' = z'$$

where y_i', z' and x' are the 'true' values of expenditure on commodity i, total expenditure and income respectively; the θ's are parameters and the v''s are the random error terms.

The following assumptions are made

$$E(v_i') = 0 \quad \text{for } i = 0,1,2,\ldots n$$

$$E(v_i'v_j') = \sigma_i^2 \quad \text{for } i = j \quad (i,j = 0,1,2,\ldots n)$$

$$= 0 \quad \text{for } i \neq j$$

$$E(xv_i') = 0 \quad \text{for all } i \quad (i = 0,1,2,\ldots n).$$

The corresponding structural form of 3.3(1) is

$$\left.\begin{aligned} y_i' &= \beta_{0i} + \beta_{1i} z' + v_i \\ & \qquad i = 1,2,\ldots n \\ z' &= \beta_{00} + \beta_{10} x' + v_0 \\ \Sigma y_i' &= z' \end{aligned}\right\} \quad 3.3(2)$$

where $\beta_{0i} = \theta_{0i} - \frac{\theta_{1i}\theta_{00}}{\theta_{10}}$, $\beta_{1i} = \theta_{1i}/\theta_{10}$

$\beta_{00} = \theta_{00}$, $\beta_{10} = \theta_{10}$

$v_i = v_i' - \beta_{1i} v_0'$

$v_0 = v_0'$

and

$E(v_i) = 0$

$E(v_i v_j) = \sigma_i^2 + \beta_{1i}\sigma_0^2$ for $i = j$ $(i,j=1,2,\ldots n)$

$= 0$ for $i \neq j$

$E(z'v_i) = \beta_{1i}^2\sigma_0^2$ for $i = 1,2,\ldots n.$

From these results it means that the application of ordinary least squares method to any of the first n equations (which are equations of interest) of 3.3(2) will lead to biased estimates since

$$E(z'v_i) \neq 0 \qquad 3.3(3)$$

$$i = 1,2,\ldots n.$$

In this situation, the assumption of Prais (1959) discussed above will not be correct theoretically since z' can no longer be treated as exogenous as shown by 3.3(3).

In real life, we do not have the 'true' values of the variables. Error of measurement is to be expected, at least for the near future in most economic data. It may be reduced, it seems it can not be eliminated. Measurement error in this sense is not the same as error due to grouping usually adjusted for by using Shephard's correction. As pointed out by Durbin (1954), neither can the method of grouping discussed in section 3.1 above remove or reduce measurement error theoretically. On the basis of this discussion, 3.3(2) has to be re-cast in terms of what are actually observed. Following the approach adopted by Durbin (1954) we define:

$$y_i = y_i' + y_i''$$
$$z = z' + z'' \qquad 3.3(4)$$
$$x = x' + x''$$

where the observed values y_i, z and x are partitioned into 'true' values y_i', z', x' and the error parts y_i'', z'' and x'' respectively.

Then the following further assumptions are made:

$$E(zy_i'') = E(y_i''x') = E(y_i''x'') = E(x''x') = E(z''x) = 0. \qquad 3.3(5)$$

Substituting 3.3(4) into 3.3(2) leads to

$$y_i = \beta_{0i} + \beta_{1i}z + W_i$$
$$i = 1,2,\ldots n$$
$$z = \beta_{00} + \beta_{10}x + W_0 \qquad 3.3(6)$$
$$\Sigma y_i = z$$

where $W_i = v_i' - \beta_{1i}v_0' + y_i'' - \beta_{1i}z''$

and

$W_0 = v_0 + z'' - \beta_{10}x''$.

It is quite easy to verify that

$$E(W_i z) \neq 0$$

as in the case of 3.3(2). Which means, again that the ordinary least squares approach will lead to bias estimates. The purpose of statistical analysis is to measure the relationship between y_i' and z' i.e. β_{1i}: but what is feasible is to use 3.3(6) which is just identified econometrically. In actual fact it is a recursive model.[1] This makes the estimation of the parameters quite easy.

There are three methods which can be applied and because the model is just identified and recursive, the three methods give equivalent results. They all have the disadvantage that they are biased, but asymptotically unbiased. The methods are indirect least squares, 2-stage least squares and method of instrumental variables.[2] Indirect least squares method is feasible when the structural model is exactly identified. In such a situation there is a one-one correspondence between the reduced form parameters and the structural parameters. If the reduced form parameters are estimated, they can then be used to get the estimates of the structural parameters. In the case of the 2-stage least squares method, the basic idea is to replace z in the first n equations of 3.3(6) by its estimated value (derived by regressing z on x) from $(n+1)^{th}$ equation in the model and then β_{1i} $(i = 1,2,...n)$ is estimated by regressing y_i on the estimated value of z instead of z itself. In effect we are using a corrected version of z which is made uncorrelated with W_{i}. For the method of instrumental variable, a variable which is correlated with z' but not with z", v_i', v_0 and y_i'' is sought for. This is then used to estimate β_{1i} $(i = 1,2,...n)$. In the case of a model specified by the various income hypotheses discussed above and from the assumptions

---

[1] See Summers (1959), p. 123.

[2] See Malinvaud (1970), pp. 611-628, 706-717.

given by 3.3(5) and those given for 3.3(1). When x is used as an instrumental variable, it is easily seen that the three methods are the same.

Using x as the instrumental variable and putting "~" on a parameter to indicate that the method of instrumental variable is used to estimate it, we have

$$\tilde{\beta}_{1i} = \frac{\sum_{j=1}^{m} f_j \bar{y}_{ij}(\bar{x}_j - \bar{\bar{x}})}{\sum_{j=1}^{m} f_j \bar{z}_j(\bar{x}_j - \bar{\bar{x}})} \qquad 3.3(7)$$

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

where $\bar{\bar{x}} = \sum_{j=1}^{m} f_j \bar{x}_j / \sum_{j=1} f_j$ and we assume we are using grouped data.

It can be shown that 3.3(7) is asymptotically unbiased since $\tilde{\beta}_{1i}$ tends to β_{1i} in probability sense, i.e.

$$\text{p.lim}(\tilde{\beta}_{1i}) = \beta_{1i}.$$

The variance of $\tilde{\beta}_{1i}$ is

$$\text{Var}(\tilde{\beta}_{1i}) = \frac{\sigma^2 \sum_{j=1}^{m} f_j(\bar{x}_j - \bar{\bar{x}})^2}{\left[\sum_{j=1}^{m} f_j \bar{z}_j(\bar{x}_j - \bar{\bar{x}})\right]^2} \qquad 3.3(8)$$

Using the analysis of variance approach, the residual sum of squares after fitting a regression line with slope $\tilde{\beta}_{1i}$ will be

$$\sum_{j=1}^{m} f_j\left[\bar{y}_{ij} - \bar{\bar{y}}_i - \tilde{\beta}_{1i}(\bar{z}_j - \bar{\bar{z}})\right]^2$$

$$= \sum_{j=1}^{m} f_j(\bar{y}_{ij} - \bar{\bar{y}}_i)^2 + \tilde{\beta}_{1i}^2 \sum_{j=1}^{m} f_j(\bar{z}_j - \bar{\bar{z}})^2 - 2\tilde{\beta}_{1i} \sum_{j=1}^{m} f_j \bar{y}_{ij}(\bar{z}_j - \bar{\bar{z}}).$$

$$\text{Since } \hat{\beta}_{1i} = \frac{\sum_{j=1}^{m} f_j \bar{y}_j (\bar{z}_j - \bar{\bar{z}})}{\sum_{j=1}^{m} f_j (\bar{z}_j - \bar{\bar{z}})^2}$$

then

$$\sum_{j=1}^{m} f_j \left[(\bar{y}_{ij} - \bar{\bar{y}}_i) - \tilde{\beta}_{1i}(\bar{z}_j - \bar{\bar{z}}) \right]^2$$

$$= \sum_{j=1}^{m} f_j (\bar{y}_{ij} - \bar{\bar{y}}_i)^2 - (2\tilde{\beta}_{1i}\hat{\beta}_{1i} - \tilde{\beta}_{1i}^2) \sum_{j=1}^{m} f_j (\bar{z}_j - \bar{\bar{z}})^2.$$

This means that

$$\sum_{j=1}^{m} f_j (\bar{y}_{ij} - \bar{\bar{y}}_i)^2 = (2\tilde{\beta}\hat{\beta}_{1i} - \tilde{\beta}_{1i}^2) \sum_{j=1}^{m} f_j (\bar{z}_j - \bar{\bar{z}})^2 +$$

$$+ \sum_{j=1}^{m} f_j \left[\bar{y}_{ij} - \bar{\bar{y}}_i - \tilde{\beta}_{1i}(\bar{z}_j - \bar{\bar{z}}) \right]^2. \qquad 3.3(9)$$

On the right hand side of 3.3(9), the first term is the amount of variation explained by fitting a regression line with slope $\tilde{\beta}_{1i}$ and the next term is the residual sum of squares after fitting the regression line. So that 3.3(9) gives the usual partition of total variation in regression analysis. This implies that

$$R_i^2 = \frac{(2\tilde{\beta}_{1i}\hat{\beta}_{1i} - \tilde{\beta}_{1i}^2) \sum_{j=1}^{m} f_j (\bar{z}_j - \bar{\bar{z}})^2}{\sum_{j=1}^{m} f_j (\bar{y}_{ij} - \bar{\bar{y}}_i)^2} \qquad 3.3(10)$$

and an unbiased estimate of σ_i^2 is given by

$$s_i^2 = \frac{1}{m-2} \sum_{j=1}^{m} f_j \left[\bar{y}_{ij} - \bar{\bar{y}}_i - \tilde{\beta}_{1i}(\bar{z}_j - \bar{\bar{z}}) \right]^2 \qquad . \ 3.3(11)$$

Liviatan (1961) maintained that the fact that $\hat{\beta}_{1i}$ is biased does not necessarily imply that it should be rejected because it has a smaller variance than $\tilde{\beta}_{1i}$. This means that in computing $\tilde{\beta}_{1i}$ instead of $\hat{\beta}_{1i}$ there is some amount of loss of efficiency which was first shown by Durbin (1954). However the bias of $\hat{\beta}_{1i}$ seems to outweigh its smaller variance since its expected square error, $Var(\hat{\beta}_{1i}) + |E(\hat{\beta}_{1i}) - \beta_{1i}|^2$, will in most cases be much larger than that of $\tilde{\beta}_{1i}$ because $E(\tilde{\beta}_{1i}) - \beta_{1i}$ tends to zero asymptotically. Asymptotic unbiasedness of $\tilde{\beta}_{1i}$ can be assessed by setting up the confidence limits for β_{1i} in small samples as computed by Durbin (1954). He made use of the assumption that the instrumental variable is not correlated with the random error term. In the present case, this means that the error term in each relation of the first n relations of 3.3(6). Then using the result that the true regression of $y_j - \beta_{1i}z$ on x (see 3.3(6)) is zero; and with the assumptions of 3.3(5) and those for 3.3(1), a random variable can be defined as

$$r_i^2 = \frac{\left\{\sum_{j=1}^{m} f_j \left[\bar{y}_{ij} - \bar{\bar{y}}_i - \beta_{1i}(\bar{z}_j - \bar{\bar{z}})\right]\bar{x}_j\right\}^2}{\sum_{j=1}^{m} f_j(\bar{x}_j - \bar{\bar{x}})^2 \sum_{j=1}^{m} f_j\left[\bar{y}_{ij} - \bar{\bar{y}}_i - \beta_{1i}(\bar{z}_j - \bar{\bar{z}})\right]^2} \qquad 3.3(12)$$

which is distributed as the square of a correlation coefficient based on m - 2 degrees of freedom. If we denote the α significance level of r by r_0 then

$$\text{Prob}\ (r_i^2 \leq r_0^2) = 1 - \alpha$$

which gives a quadratic form in β_{1i}. We can then solve for β_{1i} to get the two roots which may be interpreted, as was done by Liviatan (1961) in the case of α = 0·05, as the limits of the α-confidence interval for β_{1i}. The values of $\tilde{\beta}_{1i}$ and $\hat{\beta}_{1i}$ can then be examined to determine how many fall within the limits.

The above has been developed with the assumption that all the relations in 3.3(6) are linear. In fact the same results hold for the indirect addilog functions which are of interest in this study, since the adding-up criterion is satisfied.

3.4 ESTIMATION OF FAMILY EFFECTS

Family effects, in broad terms, cover nearly all factors affecting consumption except income and prices. The theory of consumers' behaviour is developed in terms of a rational individual who, even when he pulls his resources together with others to form a consumer unit, does not have interest different from the unit. This is not the situation in real life. The behaviour of a family will depend upon what Friedman (1952) called "culture complex prevailing within the social-economic group of which it is a member."[1] Family effects reflect the influence of the types of social class, occupation, location, age, sex, ethnic grouping and many others. The family effects will also determine to what extent economies of scale exist in the consumption patterns of the families. Hence there is the need to deflate observable variables by some of the determinants of 'culture complex' which usually show up as family effects and in this way to give the proper behaviour of consumer units. Then better estimates of Engel elasticities and the other elasticities will be derived. The extent to which all these factors or elements can be pursued depends very much on the available data size as will be found in chapter 5 below.

---

[1] Friedman (1952), p. 11.

The fact that an important element in government policies involves family size and composition e.g. family allowance, gives added incentive for investigating family effects. One group of factors determining family effects is used in classifying the families into groups, so that each group is homogeneous in terms of the factors used - social class, occupation and location. Comparison of regression lines is used to rank the effects of these factors. The next group of factors comprises family size and composition. Family size refers to the number of persons in the family irrespective of age, sex or any other means of distinguishing them. As far as this study is concerned, family composition is defined with respect to age and sex.

The first step in the elimination of these effects is the consideration of the family size. This is so because family size is positively correlated with income and a very high proportion of variation in expenditure patterns has much to do with family size. In some cases it is not impossible that variation accounted for by family size is more than that accounted for by income.[1] A very simple hypothesis concerning family size is the so-called homogeneity assumption which takes consumption per person as dependent only on income per person. It is the assumption of constant returns to scale as given by 2.3(2). This assumption is not appropriate because a family is composed of people different in many ways. But it is still a worth while hypothesis because it gives some interesting properties of Engel functions. One of these is that Engel curves for households of different sizes are likely to cross if taken over their full range.

---

1 See Prais and Houthakker (1955), p. 88.

Another is that the elasticities derived can be used to classify commodities as luxury or necessary.[1] This is determined through the family size elasticity of demand (see section 5.5). A commodity is a luxury if this elasticity is negative, and is a necessity if it is positive. The parameters of 2.3(2), after this relation has been specified as indirect addilog function, are estimated as discussed in 3.1 and 3.3. All estimates made according to this form are referred to as 'per person estimates' instead of homogeneity hypothesis estimates.

The analysis of family composition is a refinement of the analysis of family size. This means that there must be some improvements in the estimates over those of the 'per person estimates' if further analysis done is to be justified. The question worth asking in studying family composition effects according to Prais and Houthakker (1955) is "'How in fact does household composition influence consumption patterns?', rather than "How should consumption be affected by composition?'"[2] This statement helps to determine in what ways family composition effects should be estimated. The usual approach to this problem is to define a scale of equivalences based on either nutritional requirements or actual family behaviour with regards to consumption. A scale of equivalences consists of regarding other members of a family as equivalent to a fraction of a standard person in the family, usually the adult male. In such a case, a child

---

[1] See Prais and Houthakker (1955), p. 90. Note the distinction between luxury and necessity expressed in terms of income effects [see Allen and Bowley (1935), p. 5] which is adopted in this study.

[2] Prais and Houthakker (1955), p. 125.

may be regarded as equivalent to, say, ¾ of an adult male in the consumption of food in the family. Sometimes the scale may not take an adult male as standard but may define a standard family - Allen (1942) used the 2-adult family as standard while Forsyth (1960) used a couple-family as standard. As can be observed from the questions given above, scales based on nutritional requirements cannot be appropriate for econometric studies. On the basis of this, the present study does not make use of scales based on nutritional requirements. There are varied opinions as to the need for investigating family composition effects at all. There is the opinion that the phenomena involved are too complicated to be treated by the simple device of a scale of equivalences and as such that the problem is more or less insoluble. There is the other opinion that it requires a much more intensive investigation in which the device of a scale of equivalences is unnecessary and possibly misleading. Others believe that some attempts can be made and should be made to use knowledge of family composition, at least to improve estimates of Engel elasticities. What is done as far as this work is concerned is to attempt to improve the estimates of Engel elasticities using the major contribution by Prais and Houthakker (1955). Many attempts to deal with this problem abound in the literature [for example, see Sydenstricker and King (1921), Allen (1942), Friedman (1952), Forsyth (1960), and Barten (1964a) to mention a few, in addition to Prais and Houthakker (1955)]. It does not seem that the problem has been solved. Much will depend on the availability of good data. On the basis of Prais and Houthakker (1955), 2.3(3) is re-cast as an indirect addilog function. Following their approach and the analysis given in 3.1 and 3.3, family composition, based on age and sex, are incorporated into the analysis to get estimates based on 'per adult equivalence'. For the estimates, relationship considered is of the form:

$$\bar{y}_{ij} / \sum_{t=1}^{T} \alpha_{it} \bar{s}_{tj} = A(\bar{x}_j / \sum_{t=1}^{T} \alpha_{Ot} \bar{s}_{tj})^{\beta_i} \qquad 3.4(1)$$

$i = 1,2,\ldots n; \quad j = 1,2,\ldots m.$

where s_{tj} is the number of persons of type t in family j and $\bar{s}_{tj}$ is the group mean, β's, A and α's are parameters.

To derive α_{it} $(t = 1,2,\ldots T)$, 3.4(1) is expressed as

$$\bar{y}_{ij}/(\bar{x}_j/\Sigma\alpha_{Ot}\bar{s}_{tj})^{\beta_i} = A \sum_{t=1}^{T} \alpha_{it}\bar{s}_{tj} \qquad 3.4(2)$$

the α_O's are regarded as weighted averages of the α_i's $(i = 1,2,\ldots n)$.

This means that

$$\bar{p}_{ij} = a_{i1}\bar{s}_{1j} + a_{i2}\bar{s}_{2j} + \ldots a_{iT}\bar{s}_{Tj} \qquad 3.4(3)$$

$i = 1,2,\ldots n; \quad j = 1,2,\ldots m$

where

$$\bar{p}_{ij} = \bar{y}_{ij}/(\bar{x}_j / \sum_{t=1}^{T} \alpha_{Ot}\bar{s}_{tj})^{\beta_i}$$

and $a_{it} = A\alpha_{it}$.

If we know $p_{i1}, p_{i2}, p_{i3}, \ldots p_{im}$ for commodity i then since $\bar{s}_{1j}, \bar{s}_{2j}, \ldots \bar{s}_{Tj}$ $(j = 1,2,\ldots m)$ are known, regression analysis can be performed for 3.4(3) to estimate the set of parameters $a_{i1}, a_{i2}, \ldots, a_{iT}$ for commodity i. In such a case, by the method developed in 3.1, we have

$$\hat{a}_i = \left|\bar{S}'(GG')^{-1}\bar{S}\right|^{-1}\bar{S}'(GG')^{-1}\bar{p}_i \qquad 3.4(4)$$

where $\hat{a}_i = \begin{bmatrix} \hat{a}_{i1} \\ \hat{a}_{i2} \\ \vdots \\ \hat{a}_{iT} \end{bmatrix}$, $\bar{S} = \begin{bmatrix} \bar{s}_{11} & \bar{s}_{12} & \cdots & \bar{s}_{1T} \\ \bar{s}_{21} & & & \\ \vdots & & & \\ \bar{s}_{m1} & & \cdots & \bar{s}_{mT} \end{bmatrix}$ and $\bar{p}_i = \begin{bmatrix} \bar{p}_{i1} \\ \bar{p}_{i2} \\ \vdots \\ \bar{p}_{im} \end{bmatrix}$

Since observations for $\bar{p}_i$ $(i = 1,2,\ldots n)$ are not available, an iterative method is adopted. The computational procedure is explained in section 3.5.

Another aspect of family effects which is worth examining is the concept of economies of scale. This concept gives expression to the possibility that, with given levels of income per person, a large family may be able to attain a higher standard of living than a small family, e.g. with regards to a commodity group like food. In considering non-food commodities, it is more difficult to give meaning to economies of scale in a simple manner as given above. This is because many commodities are only bought for families within given size ranges and composition e.g. it will be odd for a one-person family to purchase items like baby food and probably we can expect a large family, say, 6-person family to purchase a dining table with six chairs. There is another problem, that there may be some uncertainty as to whether to attribute a given phenomenon to the effect of family composition or to the effect of economies of scale. In principle, the issue can always be decided by investigating comparable phenomena in households of constant composition but of different sizes, possibly all adult families. Some attempts at this will be considered, but not into details, in chapter 5. To give detailed analysis of economies of scale we should require to estimate the parameters of 2.3(4) and 2.3(5). The form of the present data does not allow for this.

3.5 COMPUTATIONAL PROCEDURES

This study has benefitted greatly from the modern electronic computer, particularly with the extension of this facility to the University of Ibadan, Nigeria, where much of the spade work of this research was carried out. Without this facility not much could have been done within the period spent. The three programs used were written by me. They were first used on the University of Ibadan, IBM 360/25 and were later adjusted and used on the University of London, CDC 6600 to revise some of the calculations previously made.

All the three programs will be briefly described below, and they are included as Appendix A. Even though electronic computers are quite fast and can do mass routine jobs in extremely short periods, there is quite a lot that still has to be done on desk calculating machines, particularly when human judgment is required in the process of the computation. All the calculations regarding the analysis of variance for the comparison of regression lines are carried out in this way. For this, I have found the electronic desk calculators at the London School of Economics very useful. This is, of course, after all the sums of squares and cross products are provided as output using program I as discussed below.

Program I is mainly a regression analysis program, to calculate equations like 3.1(15) following the approach discussed in section 3.1 whether 'per person' or for 'adult equivalences'. It can be adapted, by changing one or two rows of the program, to deal with linear, semi-logarithmic or double-logarithmic forms. Information is input in sets of group means and all the data are stored before the computations begin. This makes it possible for each specific expenditure to be regressed on total expenditure and income separately, and total expenditure on income. By so doing estimates for ordinary least squares are made while making it possible for instrumental variable approach to be done easily on the desk calculating machine by simple divisions. For each regression line fitted, the output is made up of all sums of squares and cross products, regression coefficients and their standard errors, and R^2.

Program II is concerned with the computation of $\bar{S}'(GG')^{-1}\bar{S}$ and its inverse which are used for 3.4(4). Observations on the number of persons of different types as group means and the corresponding number of households (f) are stored. The covariance matrix $\bar{S}'(GG')^{-1}\bar{S}$ is printed while its inverse is punched as output. The inverse matrix is used as an input for Program III.

Program III is an extension of Program I. The inputs are of the same form as Program I except that in addition, the inverse matrix, $\left[\bar{S}'(GG')^{-1}\bar{S}\right]^{-1}$ is also an input. The main aim of this program is to find the best estimates of the parameters of 3.4(3) and those of

$$\log \frac{\bar{y}_{ij}}{\sum_{t=1}^{T} a_{it}\bar{s}_{tj}} = \beta_0 + \beta_i \log \frac{\bar{x}_j}{\sum^{T} a_{0t}\bar{s}_{tj}} + W_i \qquad 3.5(1)$$

$$i = 1,2,\ldots n; \quad j = 1,2,\ldots n.$$

The computational aspect of this program is better explained in stages:

Stage 1: The β_i $(i = 1,2,\ldots n)$ of 3.5(1) are computed under the assumption that the a_{it} $(i = 1,2,\ldots n;\ t = 1,2,\ldots T)$ and a_{0t} $(t = 1,2,\ldots T)$ are all unity (i.e. per person estimates). The R_i^2 $(i = 1,2,\ldots n)$ are also calculated and stored.

Stage 2: The β's estimated in Stage 1 are then used to compute $\bar{p}_{ij}$ ($i = 1,2,...n$; $j = 1,2,...m$) using the relation

$$\bar{p}_{ij} = \bar{y}_{ij} / \left(\sum_{t=1}^{T} a_{0t} \bar{s}_{tj} \right)^{\beta_i} . \qquad 3.5(2)$$

$i = 1,2,...n$; $j = 1,2,...m$

The computed $\bar{p}_{ij}$ ($i = 1,2,...n$; $j = 1,2,...m$) are used to derive the first estimates for the a_{it} ($i = 1,2,...n$; $t = 1,2,...,T$) from 3.4(4). When all of them are estimated, the first iteration estimate of a_{0t} ($t = 1,2,...T$) is given by

$$\hat{a}_{0t} = \sum_{i=1}^{n} \omega_i \hat{a}_{it}$$

$t = 1,2,...T.$

and for subsequent iterations, the estimate is given by

$$\hat{a}_{0t}^{(\ell+1)} = \sum_{i=1}^{n} \omega_i \left(\frac{\sum_{t=1}^{T} \hat{a}_{0t}^{(\ell)} \bar{s}_t}{\sum_{t=1}^{T} \hat{a}_{it}^{(\ell)} \bar{s}_t} \right) a_{it}^{(\ell+1)}$$

$t = 1,2,...T,$

and where ℓ refers to the ℓ^{th} iteration.

This is required to satisfy the condition given in page 130 of Prais and Houthakker (1955).

Stage 3: The $\hat{a}_{it}$ ($i = 1,2,...n$; $t = 1,2,...T$) and $\hat{a}_{0t}$ ($t = 1,2,...T$) are then used to carry out the first 'adult equivalence' estimates of β_i ($i = 1,2,...n$) by using 3.5(1). The corresponding estimates of R_i^2 ($i = 1,2,...n$) are also computed. The decision to repeat the process is based on a simple test which depends on the present values of R^2 and those found in Stage 1; or in subsequent iterations on those of the ℓ^{th} iteration and those of the $(\ell+1)^{th}$ iteration as follows:

$$R^2_{i,\ell} - R^2_{i,\ell+1} < 0 \text{ , there is improvement}$$
$$= 0 \text{ , no change}$$
$$> 0 \text{ , no improvement.}$$

If there is improvement for i^{th} commodity, the corresponding estimates of a_{it}, a_{0t} ($t = 1,2,...T$), R_i^2 and β_i for commodity i replace their old values.

This test is carried out for $i = 1,2,3,...n$ to complete an iteration. The process stops when there is 'no change' or 'no improvement' for all the commodities, or after 20 iterations. In all calculations, the maximum number of iterations was eight. At the end of an iteration, if there is at least one improvement, the process moves to stage 2 and another iteration is begun using only the stored $\hat{\beta}_i$ ($i = 1,2,...n$). The process then continues. At the end of each iteration, the estimated values of a_{it} ($i = 1,2,...n$; $t = 1,2,...T$), a_{0t} ($t = 1,2,...T$), β_i ($i = 1,2,...n$) and their standard errors, and R_i^2 ($i = 1,2,...n$) are printed as output. If we define h as the number of commodity groups with 'no improvement' in each iteration ($0 \leq h \leq n$), the iteration with the lowest value of h close to the end of the computation is chosen as the best iteration. The decision process adopted here is different from that used by Prais and Houthakker (1955). While they were interested in optimising the values of $\hat{a}_{it}$ ($i = 1,2,...n$; $t = 1,2,...T$), the interest here is in improving the good-ness of fit of the relations.

CHAPTER 4

THE DATA

4.0 INTRODUCTION

The results of any statistical analysis depend, to a great deal, on the available data. This chapter is devoted to a discussion of the data used, set against the Nigerian background. Much of the discussion deals with basic information about Nigeria and, in particular, the selected urban areas, including a summary of the results already published officially on the surveys (source of the data), to aid the understanding of this study. In addition, problems of data collection and urban consumer sample surveys in Nigeria, comments on concepts measured, non-economic factors used and restrictions imposed on the range of data to ensure stability of the parameters, and procedure adopted in cross-tabulating the original data are discussed. Under normal circumstances - where data available are in the appropriate and acceptable form - this detailed discussion will not be necessary. But the Nigerian situation, like situations in most countries with underdeveloped data collecting agencies, is not ideal, hence this digression.

4.1 THE NIGERIAN SETTING[1]

For the purpose of appreciating the results of this study, there is need to include sections dealing with background information about Nigeria. This section and the next are meant to do this. Nigeria with a population

---

[1] Much of what is covered in this section and the next one can be found in Helleiner (1966), Mabogunje (1968) and Buchanan and Pugh (1955).

of about 60 millions, covers an area of 356,669 square miles. She is the largest single political unit along the west coast of Africa. She occupies a position where the western parts of the African continent meet the Equator. She lies between longitudes 3° and 15°E of Meridian and between latitudes 4° and 14°N of the Equator, so that she is entirely within the tropical zones, extending northwards from the coast line for over 650 miles. From the western border to the East, there is a distance of 700 miles at the widest part. The Atlantic ocean washes the coastline for some 500 miles. Although Nigeria is wholly within the tropics, the climate varies from the typical tropical at the coast to sub-tropical further inland.

There is abundance and variety of land leading to a range of agricultural commodities. She grows sufficient food for the need of her population - a test of this was the contribution of local food production to the strength of the economy during and immediately after the civil war. She relies less on imported food, mostly processed ones. It is estimated that production of food before 1966 accounted for about 80% of the total agricultural production. Agriculture is really the basis of the strength of the economy, although this position is now being challenged by crude oil. The principal foodstuff are yam, cassava, plantains, bananas, rice, beans and fruits of different kinds. The fish industry is growing at a fast rate, so also is poultry. Cattle, sheep and goats form a very considerable part of the internal economy. Apart from consumer goods from agriculture, there are manufactured goods which, until recently, were mostly imported. There is a rapid increase in the value of the home-produced manufacturing products. A rate of growth of 8.2% was recorded in 1962-3 but manufacturing is still a very small proportion of the gross domestic product.

Her large population provides a very large market both for consumers and labour force. There are many ethnic groups with a variety of customs and traditions. Similarly there is a variety of religions of which the most important are Islam and Christianity of various denominations. With a country of this size and potentiality, different backgrounds, customs and traditions, it is not possible to avoid taking into consideration a lot of economic and non-economic factors. This, however, should not be stressed too much since they form the basis of consumption habits in all societies.

It is necessary to note that what the budget contains for a typical Nigerian family, like families in most developing countries, is not likely to be the same as what it contains for a typical family in the developed, industrialised countries. This is not to say that the quality is poor but that such a budget will be influenced by local conditions with very few industrial commodities. This is particularly the case with families analysed in this study. They are mostly from low income group. The group that is usually studied in most econometric research of family budgets.[1] Although, basically man's needs are the same everywhere.

This study is based on the urban areas of Nigeria. In that sense we have to have clearly in our mind what we mean by urban area or urbanisation. Historically, urbanisation, development and a rising standard of living have been closely associated. Various definitions of urbanisation are found in the literature.[2] The Federal Office of Statistics (F.O.S.) has a concept of urban centres based on the definition of a town framed as far back as 1931 and based on the Indian census. A town is defined as

---

1 See Houthakker (1957) for a list of statistical series of such studies.

2 See Mabogunje (1968) and the references cited.

"a continuous collection of houses";[1] although this definition has not always been followed rigorously because in Nigeria it is always difficult to distinguish an urban unit from a cluster of villages. In most cases towns with a population of over 20,000 inhabitants are usually regarded as urban since the only practical means of measuring urbanisation is to compare the aggregate population of towns of over 20,000 inhabitants. The working definition adopted here is the one given by Wirth (1938) that an urban centre is "a relatively large, dense and permanent settlement of socially heterogeneous individuals", except that it is a dynamic process. Mabogunje (1968) gives the major factors of urbanisation in Nigeria as due to: existence of a traditional system of towns and cities, and the dynamic impact of modernisation begun by the colonial regime and continued by the independent government of the country. Towns and cities are seen as crucial modes in the developing economy. Their size, number, distribution, characteristics and activities are assumed to be useful indications of the economy. The rate of urbanisation in Nigeria has been very impressive. In 1921 only 16 towns had a population over 20,000. By 1931 this figure rose to 24 with only two towns over 100,000 and 18 were over 50,000, with 47 over 20,000. Urban populations as a percentage of total population are as follows:[2]

Period	%
1921	4.8
1931	6.7
1952/53	9.6
1962/63	16.3

---

[1] Heads (1958), p.66.

[2] These estimates are computed from Heads (1958), Mabogunje (1968) and F.O.S. (1969).

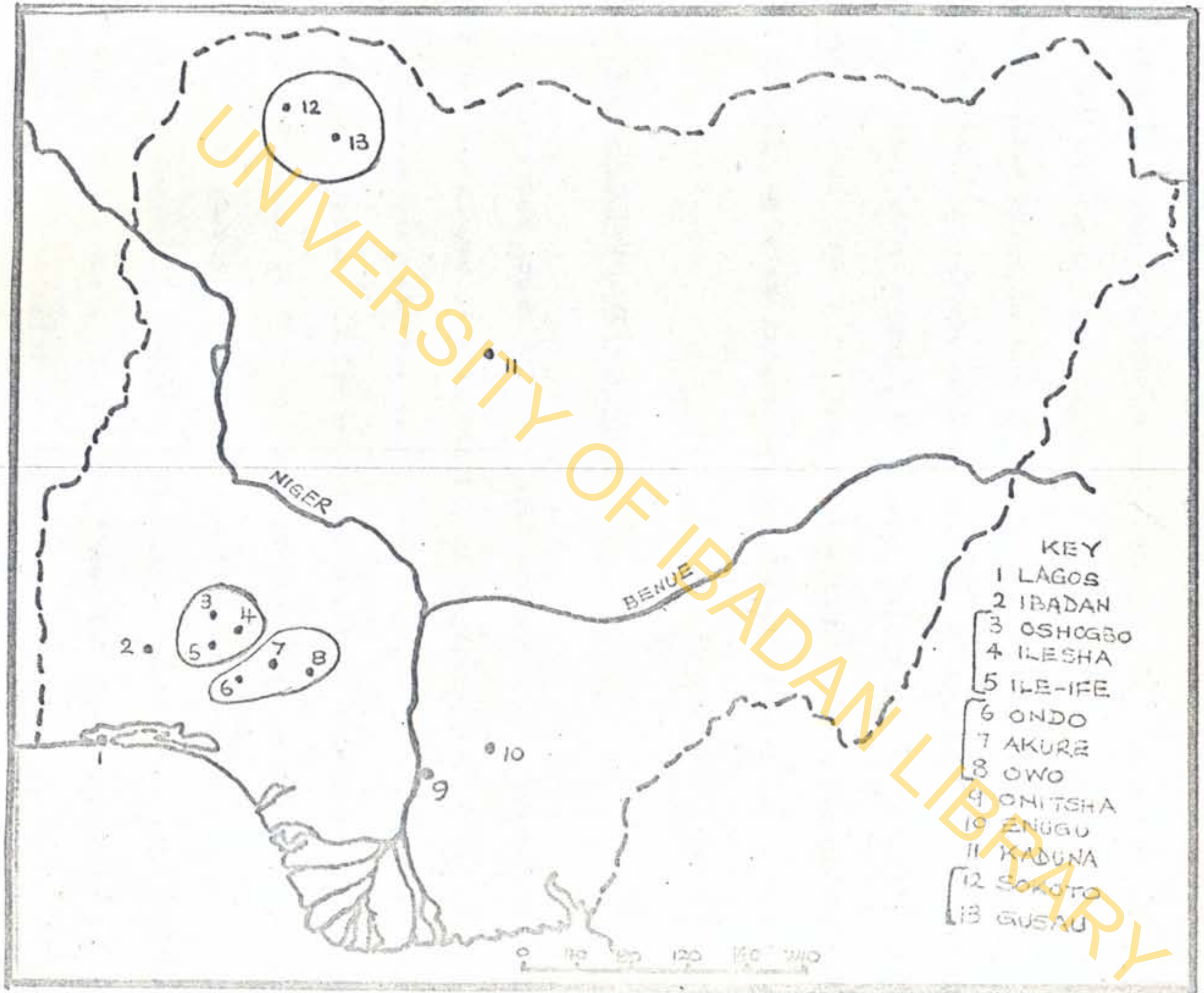


Fig. 2 Map of Nigeria showing the sampled Urban Areas

There is no doubt that this rate is tremendously increasing. With recent development programmes and improvement in education, health and sanitary facilities; the increase in the sense of collective responsibility and consciousness of self-help schemes, particularly in the rural areas; the creation of 12 states and local reforms creating more local 'capitals' and the drifts to these new centres; urbanisation in the country is bound to increase very rapidly within this decade.

It is proper, for completeness, to mention another definition of urbanisation which is more restrictive, and by no means, helpful and relevant to this study. Dickson (1947) defines an urban centre as 'a compact settlement engaged in non-agricultural occupations'. On the basis of this definition, some of the sampled areas discussed below will not qualify as urban centres as shown from table 4.

4.2 THE SELECTED URBAN AREAS

The urban areas used for this study are good representations of the various stages of urbanisation in Nigeria. The areas are Ibadan, Oshogbo, Ile-Ife, Ilesha, Akure, Ondo, Owo, Onitsha, Enugu, Kaduna, Sokoto and Gusau. As far as urbanisation is concerned, Mabogunje (1968) describes Ibadan as "the pinnacle of pre-European urbanism in Nigeria, the largest purely African city and the emporium for the commerce of an extensive region".[1] Enugu and Kaduna represent 'the metropolitan creation of the colonial era'. In between these two classes are the others which are local and provincial centres of education, administration, religion and commerce. All these towns, by African standard, are big. By 1963

[1] Mabogunje (1968), p.187

census [F.O.S.(1969)], they had between them a population of 1,968,524 distributed as follows:

Ibadan	627,379
Kaduna	149,910
Enugu	138,457
Onitsha	163,032
Oshogbo	208,966
Ile-Ife	130,050
Ilesha	165,822
Akure	71,106
Ondo	74,343
Owo	80,413
Sokoto	89,817
Gusau	69,231

Their relative locations are shown in fig.2.

Kaduna is an example of administrative centre (headquarters of the former Northern Region) with secondary functions as a railway junction. Its population includes a large number of government employees and a large proportion of the population of this town depends for its livelihood on this administrative group. Enugu owes its growth to the development of coal mines dating as far back as 1909. But later to its establishment as the headquarters of the former Eastern Region. Hence the increase in administrative population with a large expansion of the residential and administrative sectors. In contrast Ibadan has its commercial - administrative zone where foreign influence is dominant, quite distinct from the indigeneous urban pattern which is preserved with little modification. Even though Ibadan has a long history as an administrative centre (headquarters of the former Western Region) and as an educational centre, compared with Kaduna and Enugu, it is a rural city with about one fifth of its population engaged in farming. Ile-Ife, Oshogbo, Akure, Ondo and Owo are local administrative headquarters located on a stretch of land that forms the major cocoa producing sector of Nigeria. In this sense, it is not a surprise to find a high proportion of farmers and traders in the samples of these areas as shown in table 4. Compared with Kaduna and Enugu they have a low proportion of administrative population. Sokoto and Gusau are in the main cotton producing sector of the country. While Sokoto is the spiritual headquarters and first capital of the Fulani; Gusau, which lies east of the Sokoto River basin and on the railway line, is the second commercial centre of the northern part of Nigeria, after Kano. Onitsha lies on the highest east bank of the River Niger. With its huge market for goods, mostly imported, enhanced by the great concentration of roads upon it, Onitsha is also a major collector of palm oil, an important export of the East of the Niger.

The occupational distribution of the sampled households from these areas are shown in table 4. The occupational classes defined are: clerks, labourers, artisans, traders, farmers and others. The clerks are mainly 'white collar' workers including teachers and nurses. The labourers include porters, gardeners, cleaners, and shop attendants. The artisans are all skilled workers - carpenters, masons, fitters, mechanics, technicians, printers etc. Others are mainly professionals like arabic teachers and native doctors who usually supplement their earnings with other jobs, in most cases with farming, although they are not grouped with farmers. As the table shows, Ibadan, Kaduna, Enugu and Onitsha are sampled separately; Oshogbo, Ile-Ife and Ilesha form a unit; Akure, Ondo and Owo

TABLE 6. DISTRIBUTION OF THE SAMPLED HOUSEHOLDS BY OCCUPATIONS

	Clerks	Labourers	Artisans	Traders	Farmers	Others	All Low Income Group	Middle Income Group
Ibadan	51	65	146	53	15	13	343	119
Kaduna	121	55	119	42	-	7	344	117
Enugu	111	61	116	40	-	8	336	120
Onitsha	42	59	62	133	-	21	317	117
Os-Il-If	113	73	125	38	60	17	426	124
Ak-On-Ow	69	42	103	176	125	16	431	128
Sok-Gus	28	32	90	88	57	14	309	53
TOTAL	535	387	761	470	257	96	2406	778

form a unit; and Sokoto and Gusau another unit. Each unit which consists of more than one town is very contiguous - all towns forming the unit are within the radius of 30 miles with no major town separating them. If the samples are representative of the areas, it shows that Akure-Ondo-Owo unit has more farmers than any other, while Onitsha has the largest number of traders. There is a marked close resemblance between Kaduna and Enugu. For each occupation, figures for the two towns are quite similar. Ibadan has a high number of artisans, Sokoto-Gusau unit has a high number of artisans and traders, while Oshogbo-Ile Ife-Ilesha unit has high numbers of clerks and artisans.

4.3 PROBLEMS OF DATA COLLECTION IN NIGERIA

A very important deficiency in the developing countries is the absence of relatively accurate statistical data. The main contribution to this is lack of proper communication between the users and the producers of statistical data. Improvement in data collection and storage depends very much on this. The producer produces data which is not used and there is no incentive to improve. The users do not know that the statistical data they need are available. In some cases the users distrust the available statistical data and in the process of collecting their own, end up with worse ones with no sound theoretical basis. If relatively accurate data are to be available for use, the users and the producers must speak the same language, be able to communicate, and understand what is really needed. This study is partly an attempt to do this in the case of urban consumer surveys. It is hoped that by using these surveys and pointing out some of the defects when used for econometric studies, the producers will be able to appreciate the problems of a class of users of this important source of data in Nigeria.

Another aspect of data collection in Nigeria is that it is not easy to make national estimates for most variables that are of interest to social scientists, particularly economists. This is probably due to the size of the country and her rich varieties. Another reason is the high rate of illiteracy, particularly in the rural areas (about 75% for the country). It is physically impossible for the meagre statistical manpower available to cover the whole country in a national survey, going on concurrently in all parts. Up until now there has not been a successful national sample survey conducted in all parts of the country within the same period. As a result of this, surveys available on any subject matter are for different areas at different periods of time. This means, to pool them into one study, some assumptions had to be made.

Added to the two points covered above are the usual difficulties common with data collection by sample surveys irrespective of the level of development (except that, as a matter of degree, they are worse for the developing countries). There is the problem of non response usually due to suspicion expressed by respondents particularly if the survey has to do with probing private lives, information on income, and unjustified linking, by respondents, of surveys with taxation. Lack of frames from which samples can be selected is a serious problem. Language problem is likely to affect data collected by interviewing for a long time, since people in different parts of the country speak different languages, particularly at the local level. This problem is made worse by a high percentage of illiteracy mentioned above.

4.4 SOME CONCEPTS EXPLAINED

So far in this study, some terms have been used with no clear meanings attached. Since some of these terms and many others will crop up in the course of the next sections, it seems reasonable to go through some of the common ones. 'Household' and 'family' have been used often and in all cases interchangeably as they will continue to be. 'Income' has not been given any specific meaning. Sometimes it has been used to mean the same thing as total expenditure. There is a need to make a distinction between consumption, expenditure and intake.

In the Nigerian situation, the concept of family in the strict sense is not applicable as far as urban consumer surveys are concerned. By a household, we mean a group of people who live and pool their resources together, and eat from the same pot. It excludes family members, in the strict sense, who eat and sleep elsewhere, while it includes dependants not necessarily blood relations. This is the meaning given to family in this study. Two concepts of income are available in the results of the surveys. There is the concept of basic income which is the wage earner's salary plus overtime payment. This is summed for all the earning members of the household to get the household basic income. In the case of a self-employed household, basic income consists of profits from the members' gainful activities - trading, farming, etc. The second concept of income is 'total receipts' which, in addition to basic income, includes rent, cash gifts, loans, withdrawals, and other cash received. It is estimated that basic incomes forms about 80% of all receipts of most households. Because of the relative stability in basic income compared with 'total receipts', which is likely to fluctuate, sometimes violently, basic income is used as the exogeneous variable of the statistical model discussed in the last chapter. It seems to fulfil most of the requirements for an instrumental variable.

The comparison between expenditure, consumption and intake is necessary to assess the use of expenditure instead of any of the other two.

At any point in time these three words do not always mean the same thing. As a simple example, consider an item like yam. Consumption of yam in any given period of time by a household equals the stock of yam at the beginning of the period plus the quantity of yam bought during the period minus the stock at the end of the period. Expenditure, as we can see from this example, equals consumption minus stock variations. While intake equals quantity of yam consumed by each member of the household in the period whether obtained from their own purchases or from some other sources, allowing for wastage in the process of preparation. In a simple society, where no stock is kept, all the three will be equivalent. This, in actual fact, was the position in the early period of Crusoe's economy, because he was living from hand to mouth. The equivalence of the three also depends on the length of period. The longer the period, the more likely the equivalence. In modern society the Crusoe economy does not exist. The effect of stock variations will depend on many factors:- length of period considered as shown above; the stage of development of a society and the level of income of the households. As a society develops, there is an increasing use of processed commodities and preservation devices. People with low incomes are not likely to do much bulk purchasing for household use and so get involved with inventory control. Intuitively, it seems the effect of stock variations particularly for the low income groups will be very small. Another point of interest is the possibility of people consuming without spending, e.g. living in a house rent free. This is where the concept of intake will be most appropriate and imputed values can be used. In the case of housing no comparable figures are available to do this. The idea of consuming without spending is also examined later in Chapter 7, although first studied by me [Adamu (1970)] during the preliminary work on this thesis. This means that the concept of expenditure should be used with caution.

4.5 URBAN CONSUMER SURVEYS IN NIGERIA

Urban consumer surveys are among the most important surveys carried out by the Federal Office of Statistics (F.O.S.). They were started in 1953 when the first survey was conducted in Lagos. Since then surveys have been conducted in many parts of Nigeria: Ibadan, Enugu, Kaduna, Zaria, Sokoto, Akure, Onitsha, Aba and many other towns. Surveys have been conducted twice in some places. When weights for the consumer price index based on the first series started in 1953 became outdated, another series was also started with Lagos in 1959. Between 1961 and 1965 surveys were conducted in Enugu, Ibadan, Kaduna, Onitsha, Oshogbo-Ile Ife- Ilesha, Akure-Ondo-Owo and Sokoto-Gusau on a comparable basis with information sought on identical items. They were improvements on the 1959 Lagos survey. Hence the Lagos survey has been excluded here from the list of areas used. The earlier series were mainly concerned with low income employees, i.e. workers earning incomes not exceeding £350 in 1954, £400 in 1959; in recent series (mostly those after 1960) the limit has been put at £450 per annum.

In the surveys [F.O.S.(1963-66)] used for this study, the coverage includes low income groups (employees and self-employed) and middle income earners, i.e. those whose income exceeds £450 but not over £1,200 per annum. As mentioned earlier, the surveys were conducted, to be precise, within the period of March 1961 and October 1965 in a continuous form, and in most cases overlapping each other. The period of each survey is as follows:

Enugu - March 1961 to February 1962

Ibadan - August 1961 to July 1962

Kaduna - July 1962 to July 1963

Onitsha - January 1963 to January 1964

Oshogbo-Ile Ife-Ilesha - April 1963 to April 1964

Akure-Ondo-Owo - September/October 1964 to August/September 1965

Sokoto-Gusau - October/November 1964 to September/October 1965.

The surveys were spread over the periods and were conducted one after the other. Because it takes some time for expenditure patterns to change structurally, if there is any time effect on the different surveys at all, it is likely to be insignificant. Added to this is the fact that this period was a relatively stable one as far as social and economic activities in the country were concerned. It covered the early part of the first comprehensive development plan (1962-68 Development Plan) and it was the period before the series of crises that led to the civil war. 'Selected Areas', as used in the title of this thesis and as the title of section 4.2, is dictated by the available data and the constraint imposed that the period covered must be such that the range of data must allow for the stability of the parameters.

The main objective of the surveys was to provide information for the establishment of weights for urban consumer price indices. But since the surveys cover a wide variety of information - about various activities and basic demographic and social data of the households - there is no doubt about their usefulness in the various disciplines of social sciences. The data collected include information on age, sex, occupation, income, marital status of household members and housing conditions, apart from data on sources of earnings and their disbursement.

With regard to the sampling procedures, for the low income groups (earners and self-employed), a 2-stage systematic sampling method is adopted in most cases. A typical procedure is: "In the first stage, a comprehensive list of houses in the area is made and the required sample is selected systematically, i.e. suppose there are N houses listed and a sample of size n is required; one of the first $b = \frac{N}{n}$ houses is selected randomly by using the table of random numbers as a member of the sample; and the remaining n-1 members of the sample are determined at intervals of b. Then basic information is collected from the selected houses which makes it possible for frames for different low income groups to be prepared. At the second stage, a sample of the required number of households is then selected systematically from the listed households of each kind of low income group." In the case of the middle income earners, a 2-stage sampling design is also used. A list of establishments employing 10 or more persons served as the frame. The first stage is the random selection of establishments from the frame by simple random sampling using tables of random numbers. The second stage is the systematic selection of the required number of income earners from the pay-rolls of the establishments. In different cases, there were local variations.

The surveys of the three groups ran concurrently in the same area. Each household kept records of its daily receipts and payments for one month. There was a preliminary week, before the month actually used, during which households were studied with the same thoroughness as was required for the survey month. This was used to get the respondents accustomed to recording the required information. 30 households for low income groups and an average of 10 for middle income groups were studied during each of the 12 survey months in the same area. The only rejected schedules were due to incomplete information, except in the Sokoto-Gusau unit where there was a sizeable number of non-response due to people who moved away before and during the survey period. Detailed information about these movers was not available. The total number of households analysed by the F.O.S. was 2406 for the low income groups and 778 for the middle income earners. This means that about 114 low income households and 62 middle income households were rejected. As indicated below, not all the households are used in different stages of this study.

There are many ways that bias can be introduced into a survey - bias in selection and response, in recording, and in interpretation. Given the state of data collection in Nigeria, as discussed in section 4.3, these biases are not easy problems to tackle. The idea that a survey is accurate is a relative term. In terms of accuracy, there is not much that can be done at this stage. In further analysis of the data as used under cross-tabulated form, effort is made to point out possible inconsistencies and possibly reject some observed values. This, again, must be handled with care. Urban consumer surveys are still the best source of data for the kind of study carried out in this thesis and the Nigerian source is not an exception. It is hoped that by using these data much improvement will be made.

The surveys are about people in the urban areas. In any particular area, three categories of people are not covered, namely, employees earning over £450 per annum in establishments employing less than 10, employees earning over £1200 per annum in establishments employing 10 or more, and self-employed people earning over £450 per annum. If the results of this study are to be applied to a wider population than the ones actually covered this can be done in two possible ways: for the areas actually surveyed and for the whole country. In the first case, the simple assumption is that the patterns of expenditure of all categories discussed above are not likely to be significantly different. In the second case, if the trend of urbanisation, as described above, continues and there is not much change structurally with regard to expenditure patterns, such generalisation will not be off the mark. This is the area of extrapolation which is not examined in this study. Such extrapolation will be valuable, if empirically it is found that reaction to change in income by households of different kinds is not significantly different.

4.6 A SUMMARY OF PUBLISHED RESULTS OF THE SURVEYS

In section 4.5, a description and assessment of the urban consumer surveys were made. In this section a summary of the officially published version of the results of the surveys is given. In the published reports, all occupational classes discussed in section 4.2 are re-classified as employed and self-employed. Employed consists of mostly clerks and labourers while self-employed are mostly traders, farmers and other professionals. The artisans are divided between the two. The grouping of the goods and services into broad commodity groups follows the usual approach in classifying commodities for group index numbers in the consumer price index, hence approximately satisfying the Hicks (1939) condition for the concept of commodity.[1] This means that each commodity group is assumed to be a homogeneous collection.[2] In the official publication of the surveys, there are ten commodity groups, made up as follows:

Accommodation made up of
rent, water and house repairs.

Food made up of
staples, meat, fish, eggs, oil, fats, vegetables, fruits, nuts, snacks and babyfood.

Drinks consisting of
beer including stout, local and imported wine; hot drink (whisky, brandy etc.); local wine (palm wine and burukutu); and soft drinks.

---

[1] See Hicks (1939), p.33.

[2] Each commodity group satisfies some specific need of man - hunger, thirst, shelter, recreation etc.

Kola and Tobacco made up of

cigarette (local and imported), tobacco (local and imported) and kola nuts.

Fuel and light made up of

electricity, firewood and charcoal, kerosene, matches and candles.

Transport consisting of

buses, lorries, trains, auto- and bi-cycles including maintenance and hire, car, donkey, petrol and oil.

Clothing made up of

ready made (including shoes), materials, services - tailoring and shoe repairing.

Household goods consisting of

crockery (local and imported), glassware, cooking utensils, tables, bed, linen, furnishing materials, soap of different kinds and other durable materials.

Other purchases consisting of

Personal miscellaneous like stationery, toilet soap, books, newspaper, and medicine mainly from chemist shops.

Other services consisting of

entertainments, hospital fees, school fees, laundry, servants, etc.

The sum of expenditures on all the commodities is then called total expenditure.

All the information of each survey is summarised in tables, most of which are classified only by basic income. As a result of this, most of the calculations are averages (per household). The first or introductory part deals with the background, scope, and purpose of the enquiry. Concepts and definitions are also covered. Part II deals with the low income groups covering sample selection, organisation of the survey and analysis. The analysis contains distribution according to occupation and income, household size, accommodation, household receipts and payments, patterns of expenditure on goods and services, monetary transactions and weights for broad categories of commodities in consumer price indices. Part III is concerned with the middle income covering identical topics as in Part II. There remain the appendices which contain the detailed information required for econometric research, particularly appendices A, B and C. These form the source of data for the analysis based on data grouped only by income (this is called one-way classified data in the subsequent discussion). However the Survey Section of the F.O.S. gave further information on age and sex of members of the households for only the low income earners in Ibadan, Kaduna, Enugu and Oshogbo-Ile Ife-Ilesha units. In this case only four types of persons were defined; namely, adult male, adult female, male child, female child. This extra information was used to assess the improvement that might result if some information exists for one-way classified data. This extra information led to the first idea of cross-tabulating the original data as discussed in section 4.7.

As part of background information to this study it is necessary and useful to give a summary, in a comparative way, of the basic information in the publications. Generally, the middle income group spent less proportion of income on food (about 30%) than the low income groups, while wage earners spent less (about 40%) compared with self-employed (about 50%). In all cases, food ranks first in terms of proportion, followed by accommodation and transport. Clerks and artisans are usually close in their behaviour with regard to the above facts. The proportions given above vary also according to areas. Proportions spent on food in Ibadan and Sokoto-Gusau units are relatively higher than others with the relative position among the occupational groups maintained.

The average number of persons per household varies from area to area and especially between the occupational groups. Ibadan, Oshogbo-Ile Ife-Ilesha, and Sokoto-Gusau units have an average of 5 while Kaduna has 3 as the lowest. Generally traders and farmers have higher numbers (about 5) than clerks and artisans (about 3). The middle income group households generally have higher numbers (about 6).

With regard to accommodation, three aspects are worth considering. First, whether accommodation is owned, rent free or rented. Second, the kind of houses occupied - cement block or brick houses roofed with corrugated iron sheets or better roofing materials (CI), mud houses plastered with cement and roofed with corrugated iron sheets (CMI) or neither of the two, e.g. huts (C). Third, number of rooms occupied by the households. With regard to the low income group, a very high proportion live rent free in Akure-Ondo-Owo, Sokoto-Gusau, Oshogbo-IleIfe-Ilesha and Ibadan units compared with the other three. Most of the houses in Kaduna and Enugu are rent paying. Most of the households live in CI and CMI types of houses in all the areas except Sokoto-Gusau unit where type C is more common, particularly for self-employed. This is partly due to the climatic condition of the far north which is reasonably extreme (too hot). A high proportion in all areas, live in one-room houses. The corresponding analysis for the middle income group shows that high proportions are rent paying, live in CI type of houses and use more than two rooms. The proportion using electricity varies mostly according to areas. Kaduna, Enugu and Onitsha have a high proportion (about 75%) and there is not much difference between the low income groups. But the position is different for the other towns. There is a marked difference between the wage-earners (about 45%) and self-employed (about 15%) using electricity. In the case of the middle income group, nearly all households use electricity.

From this summary, it seems the main differences are caused by occupational factors.[1] This means that the difference between a town and another town with respect to any of the variables discussed, is due mainly to differences in occupational composition of the towns. Another advantage of this summary is that an assessment of the level of living of the households covered by the surveys is possible. This may help to explain what is usually regarded as erratic elements in expenditure patterns in the developing countries. This is so because there is likely to be a large number of zero values for some of the items of some of the households, in contrast with those of a developed country. There is a note of warning which concerns the level of aggregation of the published results. This may lead to suppression of vital information and distortion of basic facts. The one-way grouping approach adopted here is a source of such defects as will be shown later.

4.7 THE TWO-WAY CLASSIFIED DATA

In section 3.4, it was asserted that household size is highly correlated with income. This means that we cannot go too far in our analysis with one-way grouped data even with extra information on family composition. The discussion of the surveys in the last section was based on one-way grouped data and there is reason to believe that the results are not very satisfactory. Apart from the published results and the originally completed survey schedules, the F.O.S. has files containing summary results on each household. The object of this section is to discuss briefly how these files were used to obtain the two-way grouped data on which much of this work depends. Only the low income groups (earners and self-employed) were used due to the constraints

[1] In sociology, the first basic information required for stratification of a populaton is the occupation of the head of the family unit.

put by the available resources and, judging from past studies, it seems low income groups are more useful and important than the middle or the upper income group with regard to consumer behaviour.

In grouping the data, some of the comments made in the last section were taken into consideration. The most important one was that much source of differences was due to occupational differences. So we break from the traditional practice of the Federal Office of Statistics in grouping each unit separately under self employed and earners. Instead, all the households surveyed for all areas are pooled together and grouped according to five occupations - clerks, labourers, artisans, traders, farmers and others. In doing this, we take into consideration the comments made in section 4.5 concerning the survey periods and the idea that the range of data must allow for the stability of the parameters.

Not all households were used. All households with more than six members were excluded on the intuitive belief that most households of over six members are not likely to have useful patterns of spending for empirical studies and some empirical evidence that a group such as six and over members, is not homogeneous. These are probably true for the low income groups. Similarly, all households with incomplete or obviously inconsistent information are excluded, e.g. a household spending nothing on food in a month. When all these deductions are made, the total number of households finally used for the two-way grouped data is 1,918 made up of 379 clerks, 285 labourers, 628 artisans, 308 traders and 318 farmers and others. This means that about 80% of the originally analysed households by the F.O.S. are used.

Commodities are grouped in the same way as given in section 4.6 except in the case of 'kola and tobacco', now merged with drinks. This is done, partly because the proportion is relatively small and partly because it is nearer drinks than any other group since the item 'kola and tobacco' is made up of stimulants. The number of commodity groups then reduces to nine.

Reasonably detailed information exists for family composition with regard to age and sex of the members of the households. On the basis of this, seven types of persons are defined as:

S_1 = males aged 18 years and over

S_2 = females aged 18 years and over

S_3 = males aged 14-17 years

S_4 = females aged 14-17 years

S_5 = males aged 5-13 years

S_6 = females aged 5-13 years

S_7 = children under 5 years

This grouping of family members is reasonable in that most of the observations are not likely to be zero, neither will it allow too much information to be lost in grouping.

For each occupation, grouping is by basic income and household size, i.e. number of persons in the household. Seven income class intervals defined as:

under 150/-
150/- to 249/11d.
250/- to 349/11d.
350/- to 449/11d.
450/- to 549/11d.
550/- to 649/11d.
650/- and over

and six household sizes made up of 1-Person, 2-Person, 3-Person, 4-Person, 5-Person and 6-Person family types, which give a maximum of 42 grouped means for each variable. This was not achieved even for the all low income two-way classified data. Some of the cells were empty. Figures derived for basic income, each group expenditure, number of persons for each type of persons are grouped means and total expenditure is the sum of the group expenditures. When these results are appropriately weighted and summed over occupations, we have the corresponding data for 'all low income' two-way classified data. Since these data are not available elsewhere in their present form they are included in this study as appendix B and referred to, in the subsequent discussion, as two-way classified data.

CHAPTER 5

GENERAL ASSESSMENT OF THE ESTIMATES

5.0 INTRODUCTION

The main task of the estimating procedures fully discussed in chapter 3, is to estimate the β_{1i} $(i = 1,2, \ldots n)$ of 3.3(5) where the observed variables are in logarithmic form. Ordinary least squares method (O.L.S) and method of instrumental variable (M.I.V) estimating procedures, weighted by the number of households in each income class interval (or income-size cell in the case of two-way classified data) are used. The data are expressed in per person (in the case of one-way classified data, calculated by using average number of persons in the income class interval), 2.3(2) or per adult equivalence, 2.3(3) before applying the estimating procedures as indicated above. In order to reduce typing difficulties, it has been decided to drop the subscript i of β_1 in general whether we are referring to a single commodity group or many commodity groups. The main body of results connected with estimates of β_1 are set out in tables 5 to 20 which though located in different sections of this chapter, are generally placed reasonably close to the particular section that is most relevant. This principle is followed in all chapters of this thesis. The estimates are fully discussed in sections 5.2 to 5.4 and compared with other studies in section 5.5.

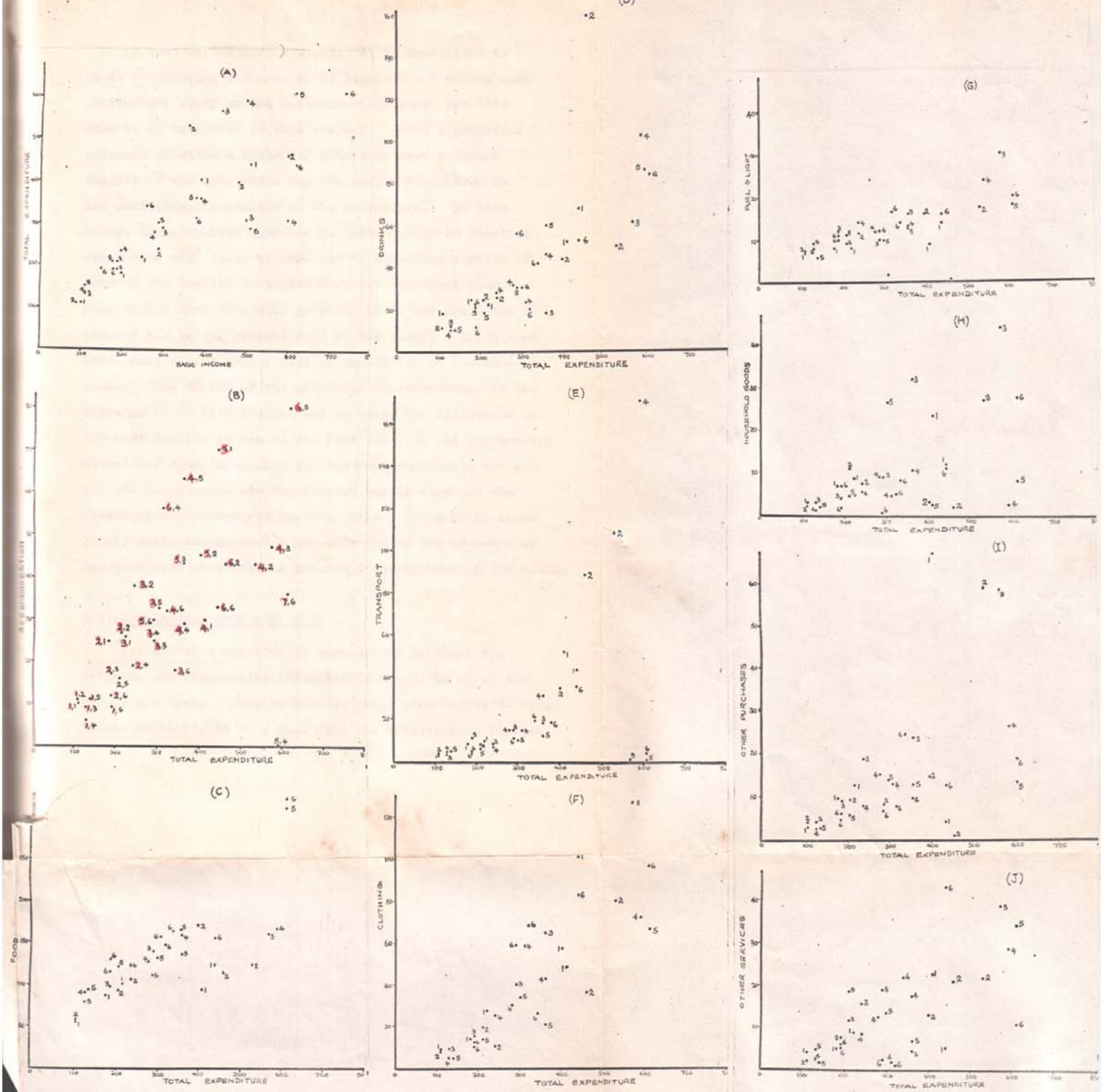


FIG 3 Graphical Representation of Expenditure Patterns forr Artisans
Based on Two-way Classified Data

In most statistical analysis, it is convenient to start by graphing the data to be analysed and making some preliminary study of the patterns displayed; and this work is no exception in that respect. Such a graphical approach provides a basis for detecting some valuable aspects of the data which may not reveal themselves in the statistical estimates of the parameters. In this study, the graphical approach is instrumental in specific ways which will later be examined in detecting sources of some of the erratic estimates that are sometimes made. Even though most data were graphed, it is not necessary to present all in the present form of the study. Apart from this, most of the graphs will be repetition of the same ideas. The choice of the graphical representation of the artisans as an illustration and as basis for discussion in the next section is due to the fact that, of all the two-way classified data (excluding all low income groups), the set of data on artisans has the largest sample size and the fewest zero observations for the cells. As will be shown later, estimates generally are affected by the presence of too many zero observations and small sample size in the cells.

5.1 A GRAPHICAL INTERPRETATION

All the data relating to expenditure patterns for artisans are graphically presented in fig.3, in which ten graphs are drawn. Fig.3a relates total expenditure to basic income while fig.3b to j deal with the relationship between commodity group expenditures and total expenditure. The axes are scaled differently for each commodity expenditure but same scale for basic income and total expenditure. For example, for accommodation, one inch represents 10 shillings while for food it is 50 shillings. The values are measured in shillings per month (20 shillings = £N1 = $2.80 or £N6 = £7). Each point on the graph shows household group means for commodity expenditure, total expenditure and size; except in the case of accommodation where a fourth dimension is added to indicate the basic income interval (the number indicating this is printed also in red). This shows that, for accommodation, a point (4.3), means that the household group is in basic income interval 4 (i.e. earning an income in the range of 350/- to 449/11d) with household size of 3. For the others, the number at each point only identifies the household size.

Starting with fig.3a, an assessment of the relationship between total expenditure and basic income can be made by drawing a line making an angle of 45° with the basic income axis, through the origin, and observing the number of households on both sides of the line. This means that all those households above the line have negative saving and those below, positive saving. If in addition we take half way through the whole range of basic income, say, 350 shillings and consider those earning income less or more than this amount, we find that of those less, 9 are above and 4 are below the line. And for those earning over

350 shillings, 4 are above and 11 are below the line. This means that at low income level, households have to subsidise their basic income to meet up with their needs for goods and services and, as income increases less is spent on goods and services and more is saved. Hence there is a tendency for a curve, concave to basic income axis, to develop rather than a straight line. This is in support of Keynes (1936) hypothesis mentioned above.

In general, all the graphs of fig.3b to j have some amount of curvatures. The curvatures for 'food' and 'fuel and light' are concave, while those of transport, drinks, clothing are convex to total expenditure axis. In terms of the values of β , for food, and fuel and light, β should be less than unity while those of the others should be greater than unity. In the case of accommodation, the curvature is very small not as pronounced. For household goods, other purchases and other services, the situations are not quite systematic. But for each type of household size, there is the general tendency for commodity expenditure to increase with increase in total expenditure with no general pattern for each commodity in terms of the household sizes.

Further examination of the graphs reveals certain features concerning commodity expenditures and household size for given total expenditure. This is carried out by examining points on the same vertical line. Two commodities, food and transport, are considered as examples. For food, in most cases, households with larger sizes spend more than those with smaller sizes. In the case of transport, a reverse pattern is common, the implication of this being that households with large size spend less on luxuries but more on necessities. In the case of 'economies of scale' such behaviour must be at a decreasing rate for necessities, as family size increases, at least for some observable range. In this study, a very simple example is found for food at a total expenditure of approximately 360 shillings for household size of 3, 4, 5; the distance between 3 and 4 is more than that between 4 and 5.

5.2 ESTIMATES BASED ON ONE-WAY CLASSIFIED DATA

As discussed in chapter 4, two variants of data from the same source are used - one is classified only by income (one-way classified data) and the other by income and household size (two-way classified data) with a cut-off point for the household size at 6. While the first kind of data suffers from the inevitable defects of high correlation between income and household size, the second is hindered by the lack of a sufficient number of households in the income-family size cells when a detailed analysis is made. The more severe problem is the case where, even though there are households in the cells, some commodity expenditures are at zero level. This forms the major source of erratic estimates and of those with negative signs.

Table 5. LEAST SQUARES (O.L.S) AND INSTRUMENTAL VARIABLE (M.I.V) ESTIMATES OF β_1 based on one-way classified data, per person

(a) Low Income Earners

	IBADAN		KADUNA		ENUGU		ONITSHA		OS-IF-IL		AK-ON-OW		SOK-GUS.	
	O.LS	M.I.V	O.LS	M.I.V	O.L.S	M.I.V	O.L.S	M.I.V*	O.L.S	MI.V	OL.S	M.I.V	O.L.S	M.I.V
1. ACCOM.	1.97	2.82	1.55	1.44	0.65	0.67	0.56		2.35	2.43	1.73	1.66	2.30	2.33
2. FOOD	0.43	0.35	0.26	0.13	0.37	0.36	0.79		0.27	0.18	0.54	0.54	0.63	0.56
3. DRINKS	2.00	2.00	1.88	2.20	1.51	1.55	0.43		1.74	1.83	0.62	0.66	0.21	2.83
4. KOLA-TOB.	-0.08	-0.62	-0.84	-0.71	0.58	-0.88	0.61		-2.27	-2.37	-0.91	-0.74	0.28	0.42
5. FUEL-LIGHT	0.29	0.34	-0.13	-0.34	0.91	0.91	0.87		1.12	1.18	0.95	1.00	0.77	0.77
6. TRANSPORT	1.33	1.39	2.84	3.07	2.25	2.13	1.64		2.59	2.81	2.37	2.44	4.78	6.47
7. CLOTHING	0.79	0.58	1.21	1.53	1.89	2.05	0.41		1.83	1.87	1.29	1.25	0.94	1.26
8. HOUSEHOLD G.	0.42	0.41	1.46	1.59	2.23	1.45	1.40		0.45	0.37	1.66	1.62	0.26	-0.59
9. O. PURCHASES	1.45	1.80	2.84	3.06	1.71	1.74	0.94		1.55	1.70	1.44	1.48	1.88	2.30
10.O. SERVICES	2.66	2.84	3.12	3.46	1.70	1.92	0.83		1.62	1.90	1.54	1.55	0.28	0.71

(b) **Low Income Self-Employed**

	IBADAN		KADUNA		ENUGU		ONITSHA		OS-IF-IL		AK-ON-OW		SOK-GU	
	O.L.S	M.I.V	O.L.S	M.I.V	O.L.S	M.I.V*	O.L.S	M.I.V*	O.L.S	M.I.V	O.L.S	M.I.V	O.L.S	M.I.V*
1. ACCOMM.	-1.13	-0.70	0.51	0.65	0.36		0.21		1.62	1.88	1.88	2.05	2.75	
2. FOOD	0.67	0.65	0.69	0.73	0.81		0.36		0.20	0.21	0.48	0.42	0.60	
3. DRINK	2.75	2.94	2.04	2.06	0.44		1.15		1.62	1.69	1.31	1.27	4.18	
4. KOLA-TOB.	0.04	0.17	0.35	0.24	-1.75		0.32		-1.21	-1.19	0.14	0.02	0.27	
5. FUEL-LIGHT	0.80	1.04	0.67	0.67	-0.25		0.96		0.46	0.50	0.24	0.21	0.56	
6. TRANSPORT	2.79	2.63	1.28	0.85	2.40		2.00		1.41	1.49	1.37	1.46	1.58	
7. CLOTHING	5.52	5.46	2.31	2.37	0.57		1.75		2.63	2.72	2.18	2.32	1.93	
8. HOUSEHOLD G.	2.29	1.36	2.52	2.51	1.99		1.10		0.55	0.66	0.55	0.53	1.10	
9. O.PURCHASES	0.86	0.71	0.74	0.65	5.00		2.88		2.00	1.93	1.41	1.45	1.97	
10. O.SERVICES	-0.92	-1.11	0.28	0.01	1.75		1.85		2.88	2.76	2.22	2.45	1.36	

(e) <u>Middle Income Earners</u>

	IBADAN		KADUNA		ENUGU		ONITSHA		OS-IF-IL		AK-ON-OW		SOK-GU	
	O.L.S	M.I.V	O.L.S	M.I.V	O.L.S	M.I.V	O.L.S	M.I.V	O.L.S	M.I.V	O.L.S	M.I.V	O.L.S	M.I.V*
1. ACCOM.	0.98	1.04	0.45	0.35	2.04	2.32	0.91	0.84	-1.14	-1.12	0.76	0.76	2.20	
2. FOOD	0.42	0.46	0.30	0.27	0.57	0.57	0.80	0.83	0.15	0.41	0.62	0.65	0.43	
3. DRINK	0.51	0.64	1.07	1.27	1.01	1.08	1.03	1.14	1.00	0.67	1.73	1.81	0.48	
4. KOLA-TOB.	0.37	0.54	-0.08	-1.71	0.34	-0.28	0.79	0.86	1.70	-4.39	0.88	0.94	0.19	
5. FUEL-LIGHT	0.74	0.96	0.51	0.45	1.02	1.08	1.09	1.10	0.71	1.41	0.35	0.56	0.98	
6. TRANSPORT	1.61	1.68	2.59	2.71	1.49	1.48	1.85	2.00	2.88	4.05	1.32	1.62	2.46	
7. CLOTHING	1.93	1.52	0.23	-1.14	0.61	0.29	0.41	0.30	1.11	0.98	1.64	1.35	0.87	
8. HOUSEHOLD G.	2.48	2.70	0.76	0.53	0.87	0.81	0.37	0.31	1.11	1.00	-0.27	-1.32	2.22	
9. O.PURCHASES	1.45	1.11	1.07	1.26	0.89	0.62	0.54	0.71	0.23	-0.46	0.35	0.18	-0.10	
10. O.SERVICES	3.24	3.22	2.06	2.52	1.18	1.16	1.03	0.72	1.52	0.37	1.08	0.92	0.15	

* Not computed because of poor relation between total expenditure and basic income involved.

This section deals with estimates based on one-way classified data. The main results are given in tables 5 to 8. Table 5 contains the estimates of β_1 when the consumer unit is per person calculated by dividing each group mean by the average number of persons in the corresponding income interval. Table 6 gives estimates based on per adult equivalence for low income earners in Ibadan, Kaduna, Enugu and Oshogbo-IleIfe-Ilesha as explained in section 4.6; and table 7 sets out estimates for both kinds of consumer units in a comparative basis for Kaduna low income earners, as an illustration. In general, the estimates for 'kola and tobacco' have wrong signs and the R^2 is reasonably small as revealed by table 7. This probably means that it has no significant relationship with income, or the results are due to the usual under-estimation of this item in most sample surveys connected with family budgets. There is the general tendency for people to under-state their consumption of tobacco, in particular. At a glance, the estimates for Onitsha and to some extent Enugu and Sokoto-Gusau frequently fail to meet the required standard of consistency. As shown in table 7 for Kaduna low income earners, all estimates that are not positive have low good-ness of fit, judging from their low values of R^2. This also happens when poor estimates of β_1 are made, judging from what is expected of the values of β_1 theoretically. For broadly classified commodities, as will be revealed in the discussion later given in section 5.5, there should be reasonable agreement on estimates based on different sources.

For example, compare the estimates for food under 'per person' with those under 'adult equivalence' in table 7.

When some definition of family effects is introduced, as discussed in 4.6 resulting in estimates given in table 6, the estimates are much improved in terms of correct sign and increase in the values of R^2 for these estimates as shown in table 7. Similarly, there is not much improvement in the estimates for 'kola and tobacco' and the estimates are not statistically significant even at 5% level. The estimates of the family effect scales are given in table 18 where they are discussed in section 5.4 together with their counterpart estimates for two-way classified data.

Even though the estimates given in table 5 are not so very impressive, it seems useful to get some effects of the performance of 'other' factors, apart from income, family size and family composition, influencing consumption patterns; namely, area and occupation, in particular. In order to carry out this assessment, calculations based on the discussion of section 3.2 are shown in table 8. Variance-ratios are calculated to compare area regression lines for each commodity group. This is done for three broad occupational groups - low income earners, self-employed and middle income groups. These results should be interpreted with caution since the estimates are based on one-way classified data. The tabulated variance-ratios for 5% and 1% levels, together with their corresponding degrees of freedom, are included in the table to aid analysis.

Table 6. LEAST SQUARES (O.L.S) AND INSTRUMENTAL VARIABLE (M.I.V) ESTIMATES OF β_1 BASED ON one-way classified data, per adult equivalence, for low income earners only

	IBADAN		KADUNA		ENUGU		OS-IF-IL	
	O.L.S	M.I.V	O.L.S	M.I.V	O.L.S	M.I.V	O.L.S	M.I.V
1. Accom.	1.97	2.25	1.21	1.19	0.88	0.90	1.71	1.70
2. Food	0.60	0.62	0.66	0.62	0.56	0.56	0.62	0.62
3. Drink	1.67	1.57	1.56	1.64	1.47	1.47	1.38	1.39
4. Kola & Tob.	0.08	-0.05	0.15	0.15	0.07	0.08	-0.61	-0.58
5. Fuel & Light	0.83	0.84	0.36	0.29	1.09	1.11	1.09	1.10
6. Trans-port	1.20	1.21	1.90	2.01	1.50	1.33	1.94	1.94
7. Clothing	0.38	0.33	1.08	1.21	1.64	1.63	1.43	1.42
8. House-hold G.	0.77	0.88	1.32	1.34	1.54	1.30	0.67	0.66
9. O. Pur-chases	1.89	1.87	1.69	1.87	1.59	1.61	1.42	1.43
10. O. Ser-vices	2.38	2.22	1.84	2.06	1.61	1.65	1.44	1.49

In general, average commodity expenditures (in log. form) are significantly different from area to area except for clothing and household goods for low income earners; and drinks, fuel and light, transport, and other purchases for middle income earners at 5% level. The difference in average commodity expenditure is more pronounced for low income self-employed. This is probably because self-employed occupational groups are not evenly distributed among the areas. For

Table 7. LEAST SQUARES (O.L.S) AND INSTRUMENTAL VARIABLE (M.I.V) ESTIMATES OF β_1 based on one-way classified data for Kaduna low income earners

	per person						per adult equivalence					
	O.L.S			M.I.V			O.L.S			M.I.V		
	β_1	S.E(β_1)	R^2	β_1	S.E(β_1)	R^2	β_1	S.E(β_1)	R^2	β_1	S.E(β_1)	R^2
1. Accom.	1.55	0.24	0.89	1.44	*	*	1.21	0.16	0.92	1.19	0.18	0.90
2. Food	0.26	0.22	0.22	0.13	0.24	0.20	0.66	0.13	0.85	0.62	0.19	0.71
3. Drink	1.88	0.57	0.68	2.20	0.72	0.91	1.56	0.27	0.87	1.64	0.29	0.85
4. Kola & Tob.	-0.84	1.25	0.08	-0.71	1.27	0.11	0.15	0.73	0.01	0.15	0.74	0.01
5. Fuel-Light	-0.13	0.44	0.02	-0.34	0.46	*	0.36	0.23	0.33	0.29	0.23	0.32
6. Transport	2.84	0.68	0.78	3.07	*	*	1.90	0.42	0.81	2.01	0.43	0.80
7. Clothing	1.21	0.82	0.31	1.53	0.79	0.38	1.08	0.44	0.55	1.21	0.45	0.53
8. Household G.	1.46	0.75	0.43	1.59	0.62	0.58	1.32	0.40	0.69	1.34	0.41	0.67
9. O. Purchases	2.84	0.75	0.74	3.06	0.16	0.99	1.69	0.52	0.68	1.87	0.54	0.66
10. O. Services	3.12	0.98	0.67	3.46	0.59	0.89	1.84	0.64	0.63	2.06	0.66	0.61

* Negative due to truncated sums of squares, hence rejected.

example Onitsha has a high concentration of traders while Akure-Ondo-Owo unit has the largest number of farmers. This is a clear indication that the areas are not homogeneous in the specific respect just stated. Differences revealed in these tests are not likely to be due to area alone but to other factors too, of which occupation is an important one.

The tests for between slopes of low income earners' expenditure patterns show that the differences between areas are often not significant. The variance-ratios for accommodation, drinks and transport are significant at 1% level; while in addition only 'other services' is significant at 5% level. For low income self-employed, accommodation, drinks and clothing are highly significant at 1% and, in addition, only 'other purchases' is significant at 5% level. As far as middle income earners are concerned, only 'accommodation' is significant at 1% level with the addition of 'other services' at 5% level. Generally ignoring estimates for 'kola and tobacco' because of under-estimates as shown above, estimates for food, fuel and light, and household goods are not significantly different from area to area. The tests for overall regression lines are not important because of their dependence on the other two sets of tests given above. The separate performances of the tests for low income earners, low income self-employed and middle income earners, whether for between means or between slopes, are worth noting. There are less number of significant cases for middle income earners compared with any of the other two. This is probably because this group is homogeneous in terms of occupation - the households contain mainly 'white collar' workers. This is not so, for example, with the self-employed group as mentioned above.

From the results given in section 4.6, the significant differences between means achieved above are no surprise. But there is no conclusive evidence in table 8 that the slopes for each commodity group are significantly different. Apart from non-homogeneous nature of the households in the low income groups, a source of caution is the possible effect of the high correlation between income and household size on these estimates. More will be said about this in section 5.3. As a result of this analysis no corresponding comparison of regression lines is made for table 6 since the patterns of variations follow those of table 5a, tested under low income earners, with the exception of fuel and light where there are improvements in the estimates. For example, the negative sign for Kaduna has been removed.

A very important advantage of the estimates given in table 6 is that it shows the type of improvement that can result when some information is given about the family composition of the income intervals in one-way classified data. As will be shown later in section 5.3, this does not contradict the fact that the better approach is cross-classification of the data with respect to income and

Table 8. **CALCULATED VARIANCE-RATIOS FOR COMPARING AREA REGRESSION LINES based on one-way classified data, per person**

	LOW INCOME EARNERS			LOW INCOME SELF-EMPLOYED			MIDDLE INCOME EARNERS		
	Between means	Between slopes	Between Regressions	Between means	Between slopes	Between Regressions	Between means	Between slopes	Between Regressions
Degrees of freedom	6,31	6,31	12,31	6,28	6,28	12,18	6,35	6,35	12,35
Tabulated Variance Ratios at 5%	2.41	2.41	2.08	2.44	2.44	2.12	2.35	2.35	2.04
Tabulated Variance Ratios at 1%	3.45	3.45	2.82	3.53	3.53	2.90	3.36	3.36	2.74
1. Accomm.	8.27	3.66	5.97	37.18	6.23	21.703	7.20	3.58	5.39
2. Food	12.26	1.34	6.80	18.34	1.44	9.89	12.58	1.24	6.90
3. Drink	2.71	3.92	3.31	103.48	12.87	58.18	1.52	0.92	1.22
4. Kola & Tobacco	7.56	0.52	4.04	2.80	0.63	1.71	11.23	0.35	5.79
5. Fuel and Light	5.62	2.07	3.85	5.93	0.81	3.37	0.96	0.86	0.91
6. Transport	9.54	3.39	6.47	19.35	0.52	9.94	0.04	1.49	0.76
7. Clothing	1.14	1.30	1.22	3.51	23.21	33.36	3.75	1.69	2.72
8. Household Goods	1.52	1.68	1.60	4.00	0.40	2.20	7.93	2.34	5.13
9. O. Purchases	7.91	0.73	4.32	20.08	3.12	11.60	1.68	0.84	1.26
10. O. Services	5.55	2.60	4.08	11.96	1.82	6.89	9.07	3.29	6.18

household size. The advantages effected are shown clearly in table 7 in terms of the standard errors of the estimates and their corresponding R^2. Comparing the ordinary least squares estimates under 'per person' with those under 'per adult equivalence', the standard errors are generally greater in the former than in the latter for corresponding estimates, so also are the R^2 smaller in the former than in the latter. The same happens, to a reasonable extent, for the method of instrumental variable.

In concluding this section, it is necessary to make some remarks with regards to differences in the estimates of ordinary least squares and the method of instrumental variable. As stated in section 3.3, the method of instrumental variable is asymptotically unbiased while this property is not possessed by the ordinary least squares. But method of instrumental variable is less efficient. The relative inefficiency of this method is given clearly in table 7, if we compare the estimates of the standard errors for the two approaches under 'per adult equivalence'. But the estimates derived, using the two approaches are quite close. It may be that when estimates are based on the adult equivalence of the consumer unit, not much loss will result if the ordinary least squares method is used.

Table 9. LEAST SQUARES (L.S) AND INSTRUMENTAL VARIABLE (I.V) ESTIMATES OF β FOR DIFFERENT FAMILY GROUPS based on two-way classified data per person

(a) CLERKS

	1-PERSON		2-PERSON		3-PERSON		4-PERSON		5-PERSON		6-PERSON	
	L.S	I.V	L.S	I.V	L.S	I.V	L.S	I.V	L.S	I.V	L.S	I.V
1. ACCOM.	1.03	0.89	1.66	1.63	1.30	1.33	1.57	1.60	-0.86	-1.59	5.05	4.80
2. FOOD	0.83	0.86	0.40	0.39	0.44	0.42	0.33	0.28	0.54	0.16	0.72	0.76
3. DRINK	1.54	1.56	1.01	1.04	1.26	1.07	0.99	1.39	1.20	2.09	0.95	1.05
4. FUEL-LIGHT	0.54	0.53	0.50	0.48	0.99	1.16	0.46	0.44	1.06	1.48	1.25	1.35
5. TRANSPORT	0.72	1.23	1.81	1.62	1.66	1.71	2.22	2.96	2.00	2.03	1.02	0.87
6. CLOTHING	0.38	0.02	4.37	3.32	2.81	2.71	0.20	0.03	2.70	2.03	2.44	2.19
7. HOUSEHOLD G.	0.90	0.86	1.20	1.80	1.30	0.65	-1.21	-0.39	0.18	0.06	0.95	1.35
8. O.PURCHASES	0.87	0.62	0.62	0.61	1.22	1.41	0.95	0.41	1.08	1.24	1.20	1.54
9. O.SERVICES	1.18	0.96	1.29	0.91	1.74	2.00	1.36	0.84	-0.49	0.88	1.54	1.53

(b) LABOURERS

	1-PERSON		2-PERSON		3-PERSON		4-PERSON		5-PERSON		6-PERSON	
	L.S	I.V	L.S	I.V	L.S	I.V	L.S	I.V	L.S	I.V	L.S	I.V
1. ACCOM.	1.05	1.06	0.75	0.74	3.49	*	1.92	1.43	2.26	3.51	2.34	2.32
2. FOOD	0.65	0.69	0.61	0.53	0.68	*	0.43	0.47	0.51	0.51	0.55	0.53
3. DRINK	1.00	1.04	1.96	2.07	1.39	*	1.42	1.68	2.23	-0.20	0.80	0.84
4. FUEL-LIGHT	0.06	-0.37	1.09	1.09	0.34	*	0.36	0.88	0.40	0.59	1.18	1.15
5. TRANSPORT	0.72	-0.03	-0.20	-0.18	6.43	*	-0.58	-0.75	-10.76	-6.70	-5.27	-5.14
6. CLOTHING	0.58	-0.80	2.12	2.30	3.48	*	0.85	0.51	3.16	0.01	0.44	0.51
7. HOUSEHOLD G.	-0.08	-0.68	0.57	0.68	-2.07	*	1.36	1.87	4.13	0.94	1.56	1.56
8. O.PURCHASES	1.25	1.18	1.28	1.35	1.21	*	0.56	0.94	3.87	0.62	2.78	2.73
9. O.SERVICES	2.40	1.64	0.65	0.61	2.10	*	-4.03	-2.91	0.84	-1.83	1.93	1.95

(c) ARTISANS

	1-PERSON		2-PERSON		3-PERSON		4-PERSON		5-PERSON		6-PERSON	
	L.S	I.V	L.S	I.V	L.S	I.V	L.S	I.V	L.S	I.V	L.S	I.V
1. ACCOM.	1.35	1.43	0.97	1.06	1.33	1.37	1.07	1.28	1.38	1.44	1.21	1.52
2. FOOD	0.62	0.66	0.57	0.62	0.62	0.62	0.47	0.47	0.66	0.64	0.54	0.42
3. DRINK	0.64	0.60	1.03	1.07	1.08	1.06	1.52	1.45	1.29	1.33	1.39	1.53
4. FUEL & LIGHT	0.33	0.34	0.43	0.37	0.59	0.60	0.41	0.20	0.85	0.81	0.52	0.56
5. TRANSPORT	1.93	1.90	1.81	1.65	0.05	0.60	1.71	1.78	-1.62	-1.35	1.93	1.88
6. CLOTHING	1.51	1.50	1.36	1.24	1.84	1.83	2.38	2.36	1.11	1.29	2.09	2.46
7. HOUSEHOLD G.	1.49	1.50	1.25	1.09	1.60	1.60	2.06	2.04	0.83	0.99	1.15	0.95
8. O. PURCHASES	1.97	1.98	1.36	1.28	1.20	1.22	1.35	1.17	0.77	0.96	1.45	1.06
9. O. SERVICES	0.89	0.84	1.57	1.68	1.35	1.25	1.58	1.35	1.14	1.19	1.13	0.41

(d) TRADERS

1. ACCOM.	0.18	0.19	1.10	1.06	0.87	0.96	0.77	0.79	1.42	1.52	-3.02	1.78
2. FOOD	0.42	0.38	0.05	0.07	0.59	0.57	0.61	0.61	0.56	0.53	0.43	0.74
3. DRINK	-0.64	-0.30	0.89	0.93	1.36	1.34	1.59	1.62	1.46	1.41	c0.93	1.48
4. FUEL & LIGHT	0.12	0.03	0.86	0.89	0.30	0.26	0.95	0.83	0.52	0.51	0.43	-0.87
5. TRANSPORT	1.61	1.88	0.52	0.63	2.71	2.77	0.87	0.87	0.96	1.25	3.58	2.51
6. CLOTHING	0.30	3.13	1.78	1.96	0.69	-0.23	1.71	1.21	0.14	-0.32	0.39	0.94
7. HOUSEHOLD G.	-2.10	-0.63	1.83	2.00	-0.83	-0.53	2.16	2.12	1.51	1.15	-0.39	-1.94
8. O. PURCHASES	2.20	2.33	2.78	2.44	1.90	1.98	1.49	1.41	1.99	2.65	2.09	1.98
9. O. SERVICES	-4.65	-0.68	1.41	1.10	1.48	1.41	0.46	0.48	2.33	2.00	4.71	3.05

(e) FARMERS

	1-PERSON		2-PERSON		3-PERSON		4-PERSON		5-PERSON		6-PERSON	
	L.S	I.V	L.S	I.V	L.S	I.V	L.S	I.V	L.S	I.V	L.S	I.V
1. ACCOM.	1.19	0.93	0.81	0.86	-1.31	0.57	0.08	-0.06	0.73	0.74	0.74	-0.08
2. FOOD	0.38	0.39	0.57	0.55	0.52	0.50	0.89	0.82	0.37	0.36	0.66	0.60
3. DRINK	0.53	0.69	0.46	0.46	0.81	0.77	1.08	1.16	0.99	0.86	0.62	0.64
4. FUEL & LIGHT	0.26	0.21	-0.44	-0.40	0.46	0.45	0.41	0.42	-1.01	-1.04	0.29	-0.01
5. TRANSPORT	-0.18	-0.16	1.17	1.36	2.00	2.01	0.57	0.32	1.89	1.94	1.26	1.27
6. CLOTHING	1.49	1.63	1.96	1.89	0.77	0.39	1.58	1.74	1.76	2.07	1.63	1.74
7. HOUSEHOLD G.	2.26	2.64	2.68	2.88	0.72	0.63	1.45	1.48	-0.12	-1.71	1.12	1.03
8. O. PURCHASES	1.25	1.70	2.41	2.54	1.70	1.56	1.68	1.72	0.96	0.96	1.61	1.85
9. O. SERVICES	1.90	2.04	2.06	2.03	2.03	2.16	0.69	0.39	1.92	2.13	2.36	2.25

* Not computed because of poor relation between total expenditure and basic income for 3-PERSON FAMILY FOR LABOURERS.

5.3 ESTIMATES BASED ON TWO-WAY CLASSIFIED DATA

In section 4.7, the main reasons for computing two-way classified data were discussed. The results of table 8, even though not quite conclusive gave justification for the tabulations in the way they were carried out. In addition most empirical studies are based on this approach. An outstanding and impressive list of series in this form can be found in Houthakker (1957).

In this section the main results for estimates based on two-way classified data are analysed. The relevant results are set out in tables 9 to 17. There are two parts to the discussion of this section. In the first part, discussion is concentrated on estimates for different family types (1-PERSON, 2-PERSON, etc. families) whether for occupations or for all low income. The results for this part are shown in tables 9 to 14. The second part is concerned with the estimates for the occupational groups and all low income ignoring family types except that each group mean is divided by the corresponding family size to derive estimates based on 'per person'. The corresponding results are shown in tables 15 to 17. The discussion of family composition effects are given in section 5.4 and comparison with other studies given in section 5.5.

With regards to analysis of estimates for different family types, as an illustration detailed analysis is given only for artisans (as illustrating the properties of the others) and the all low income groups. The main reasons for carrying out estimates for different family types are to examine differences in the estimates of β_1 for different family sizes, and to examine the effects of zero observations and small sample size on the estimates. As far as this last reason is concerned, this is easier with the family type estimates. The estimates of β_1 are shown in tables 9 to 12. If the observations in appendix B are compared with the results in these tables, we find that most of the estimates with negative signs are those with zero observations in the data used, particularly for upper income class intervals. For example, the Artisans' 5-person group has negative estimates for transport because its last two income intervals contain zero observations for transport expenditure. As previously stated these negative estimates are usually not significant statistically. Apart from this problem, the estimates are reasonably good considering the values of their R^2 although they can be improved with increase in the sample size. From the results in table 10, it seems not much is gained in deriving estimates based on adult equivalence, in terms of improvement in the good-ness of fit, when two-way classified data are used. The reason for the ineffectiveness of adult equivalence is probably due to the two-way classification of the data which must have obscured much of the variation of family composition in the classification. Family size is likely to be closely related with the size of each type of persons and income of a household. This problem is examined further below in section 5.4. On the basis of this, only estimates based on 'per person' are given for all low income group family types.

Table 10. LEAST SQUARES ESTIMATES based on two-way classified data for Artisans only

(a) Per person

	1-PERSON		2-PERSON		3-PERSON		4-PERSON		5-PERSON		6-PERSON	
	β_1	R^2	β_1	R^2	β_1	R^2	β_1	R^2	β_1	R^2	β_1	R^2
1. Accom.	1.35	0.86	0.97	0.89	1.33	0.96	1.07	0.14	1.38	0.82	1.21	0.74
2. Food	0.62	0.87	0.57	0.89	0.62	0.96	0.47	0.83	0.66	0.82	0.54	0.71
3. Drink	0.64	0.74	1.03	0.94	1.08	0.89	1.52	0.97	1.29	0.89	1.39	0.68
4. Fuel & Light	0.33	0.85	0.43	0.16	0.59	0.63	0.41	0.14	0.85	0.71	0.52	0.85
5. Transport	1.93	0.98	1.81	0.86	0.05	0.23	1.71	0.81	-1.62	0.16	1.93	0.63
6. Clothing	1.51	0.93	1.36	0.87	1.84	0.98	2.38	0.95	1.11	0.11	2.09	0.82
7. Household G.	1.49	0.97	1.25	0.60	1.60	0.88	2.06	0.85	0.83	0.06	1.15	0.25
8. O.Purchases	1.97	0.81	1.36	0.71	1.20	0.72	1.35	0.24	0.77	0.07	1.45	0.64
9. O.Services	0.89	0.67	1.57	0.83	1.35	0.74	1.58	0.75	1.14	0.09	1.13	0.16

(b) Per adult equivalence

	2-PERSON		3-PERSON		4-PERSON		5-PERSON		6-PERSON	
1. Accommodation	1.00	0.94	1.78	0.91	1.17	0.17	1.28	0.78	1.10	0.83
2. Food	0.63	0.79	0.58	0.86	0.51	0.63	0.70	0.88	0.56	0.82
3. Drink	1.08	0.95	1.59	0.90	1.52	0.92	1.20	0.84	1.32	0.73
4. Fuel & light	0.45	0.18	0.78	0.36	0.32	0.10	0.81	0.68	0.89	0.30
5. Transport	1.57	0.81	0.43	0.01	1.67	0.84	-2.04	0.19	1.42	0.47
6. Clothing	1.35	0.86	1.89	0.97	2.51	0.97	0.96	0.07	1.66	0.71
7. Household Goods	1.38	0.76	1.64	0.79	2.03	0.83	0.86	0.06	1.34	0.47
8. Other Purchases	1.33	0.68	1.68	0.68	1.52	0.28	0.59	0.04	1.12	0.56
9. Other Services	1.65	0.94	2.40	0.88	1.53	0.74	1.01	0.06	0.68	0.09

Table 11. LEAST SQUARES ESTIMATES OF β_1 FOR ALL LOW INCOME FAMILY GROUPS based on TWO-WAY CLASSIFIED DATA PER PERSON

	1-PERSON		2-PERSON		3-PERSON		4-PERSON		5-PERSON		6-PERSON	
	β_1	R^2	β_1	R^2	β_1	R^2	β_1	R^2	β_1	R^2	β_1	R^2
1. Accom.	1.03	0.86	1.11	0.97	1.32	0.93	1.76	0.95	1.39	0.35	1.43	0.86
2. Food	0.53	0.96	0.44	0.88	0.50	0.93	0.54	0.93	0.54	0.97	0.57	0.98
3. Drink	0.86	0.92	1.09	0.96	1.25	0.95	1.36	0.94	1.22	0.78	1.18	0.94
4. Fuel & Light	0.35	0.75	0.54	0.96	0.50	0.70	0.60	0.68	0.57	0.74	0.75	0.77
5. Transport	1.61	0.94	1.84	0.85	1.86	0.89	1.60	0.89	1.62	0.86	1.88	0.90
6. Clothing	1.55	0.94	1.75	0.91	1.89	0.95	1.69	0.90	2.15	0.92	1.80	0.96
7. Household G.	1.78	0.74	1.88	0.85	1.18	0.89	2.03	0.90	0.85	0.50	0.93	0.38
8. O.Purchases	2.01	0.89	1.90	0.94	1.72	0.88	1.68	0.81	1.81	0.92	1.91	0.86
9. O.Services	1.72	0.97	1.39	0.66	1.65	0.71	1.06	0.78	1.80	0.77	1.81	0.86

Table 12. METHOD OF INSTRUMENTAL VARIABLE ESTIMATES OF β_1 FOR ALL LOW INCOME FAMILY GROUPS BASED ON TWO-WAY CLASSIFIED DATA PER PERSON

	1-PERSON	2-PERSON	3-PERSON	4-PERSON	5-PERSON	6-PERSON
1. Accom.	1.043	1.117	1.349	1.784	1.540	1.480
2. Food	0.539	0.444	0.512	0.541	0.531	0.569
3. Drink	0.844	1.097	1.210	1.389	1.310	1.188
4. Fuel & Light	0.376	0.553	0.538	0.604	0.545	0.731
5. Transport	1.616	1.757	1.887	1.597	1.738	1.853
6. Clothing	1.542	1.702	1.838	1.692	2.084	1.802
7. Household G.	1.832	1.936	1.193	2.058	1.705	1.028
8. O.Purchases	1.978	1.937	1.749	1.688	1.814	1.899
9. O.Services	1.741	1.392	1.590	1.020	1.768	1.754

Table 13. BUDGET PROPORTIONS FOR ALL LOW INCOME FAMILY TYPES based on TWO-WAY CLASSIFIED DATA

	1-PERSON	2-PERSON	3-PERSON	4-PERSON	5-PERSON	6-PERSON
1. Accom.	0.101	0.108	0.089	0.079	0.070	0.059
2. Food	0.404	0.405	0.490	0.479	0.514	0.523
3. Drink	0.111	0.106	0.100	0.108	0.092	0.105
4. Fuel & Light	0.037	0.042	0.048	0.063	0.055	0.050
5. Transport	0.049	0.066	0.049	0.047	0.043	0.047
6. Clothing	0.117	0.106	0.094	0.100	0.093	0.086
7. Household G.	0.048	0.038	0.034	0.036	0.023	0.018
8. O. Purchases	0.058	0.057	0.043	0.041	0.039	0.049
9. O. Services	0.053	0.071	0.052	0.048	0.071	0.062

Table 14. CONFIDENCE LIMITS FOR THE ESTIMATES OF ALL LOW INCOME 1-PERSON GROUPS

Expenditure group (i)	O.L.S estimate	M.L.V estimate	95% limits for β_i Lower	Upper
1. Accommodation	1.028	1.043	0.413	1.561
2. Food	0.531	0.539	0.260	0.650
3. Drink, Tobacco & Kola	0.863	0.844	0.344	1.062
4. Fuel and Light	0.354	0.376	0.097	0.631
5. Transport	1.612	1.616	0.818	2.150
6. Clothing	1.554	1.542	0.682	2.086
7. Household Goods	1.778	1.832	0.353	3.257
8. Other Purchases	2.006	1.978	0.880	2.786
9. Other Services	1.722	1.741	1.099	2.207

Table 15. ESTIMATES FOR SEPARATE OCCUPATIONS AND ALL LOW INCOME based on POOLED 2-WAY CLASSIFIED DATA PER PERSON

	CLERKS			LABOURERS			ARTISANS		
	O.L.S Est. β_1	O.L.S Est. R^2	M.I.V Est. of β_1	O.L.S Est. β_1	O.L.S Est. R^2	M.I.V Est. of β_1	O.L.S Est. β_1	O.L.S Est. R^2	M.I.V Est. of β_1
1. Accommodation	1.450	0.48	1.422	1.780	0.90	1.809	1.293	0.75	1.331
2. Food	0.744	0.92	0.730	0.777	0.94	0.773	0.744	0.93	0.743
3. Drink	1.105	0.79	1.149	1.266	0.64	1.217	1.215	0.93	1.217
4. Fuel & Light	0.784	0.77	0.800	0.694	0.32	0.673	0.781	0.70	0.756
5. Transport	1.147	0.58	1.236	0.802	0.40	1.007	1.207	0.41	1.223
6. Clothing	1.535	0.41	1.387	0.963	0.22	0.809	1.337	0.72	1.364
7. Household G.	1.306	0.43	1.321	1.061	0.35	0.809	1.413	0.71	1.399
8. O. Purchases	1.263	0.79	1.234	1.370	0.64	1.247	1.456	0.65	1.333
9. O. Services	1.159	0.51	1.154	1.058	0.19	0.792	1.138	0.56	1.094

	TRADERS			FARMERS			ALL LOW INCOME		
1. Accommodation	1.232	0.71	1.180	1.091	0.17	1.004	1.400	0.89	1.409
2. Food	0.758	0.83	0.726	0.728	0.87	0.698	0.729	0.94	0.713
3. Drink	1.108	0.77	1.121	0.886	0.84	0.896	1.138	0.96	1.144
4. Fuel & Light	0.625	0.34	0.537	0.468	0.23	0.387	0.676	0.88	0.668
5. Transport	1.113	0.67	1.182	0.815	0.24	0.871	1.313	0.89	1.348
6. Clothing	1.262	0.42	1.230	1.399	0.62	1.460	1.397	0.92	1.420
7. Household G.	0.943	0.27	0.981	1.232	0.65	1.228	1.457	0.85	1.466
8. O. Purchases	1.163	0.55	1.267	1.244	0.72	1.357	1.418	0.89	1.454
9. O. Services	0.789	0.20	0.712	1.329	0.56	1.347	1.151	0.79	1.171

Table 16. BUDGET PROPORTION FOR OCCUPATIONS AND ALL LOW INCOME based on TWO-WAY CLASSIFIED DATA

	CLERKS	LABOURERS	ARTISANS	TRADERS	FARMERS	ALL LOW INCOME
1. Accom.	0.104	0.104	0.102	0.081	0.066	0.095
2. Food	0.359	0.500	0.460	0.498	0.471	0.441
3. Drink	0.108	0.098	0.118	0.100	0.088	0.108
4. Fuel & Light	0.036	0.051	0.047	0.054	0.041	0.045
5. Transport	0.069	0.045	0.053	0.043	0.035	0.052
6. Clothing	0.126	0.098	0.096	0.087	0.124	0.108
7. Household G.	0.050	0.029	0.034	0.028	0.064	0.040
8. O.Purchases	0.073	0.030	0.050	0.045	0.043	0.053
9. O.Services	0.075	0.045	0.040	0.064	0.068	0.058

Table 17. CALCULATED VARIANCE-RATIOS FOR COMPARISON OF OCCUPATION REGRESSION LINES USING TWO-WAY CLASSIFIED DATA PER PERSON

	Between means	Between slopes	Between Regressions
Degrees of Freedom	4, 148	4, 148	8, 148
Tabulated Variance Ratio } at 5%	2.43	2.43	2.00
at 1%	3.44	3.44	2.62
1. Accommodation	1.876	0.814	1.345
2. Food	6.940	0.172	3.559
3. Drink	2.383	2.179	2.280
4. Fuel and Light	1.102	1.093	1.097
5. Transport	1.672	0.307	0.990
6. Clothing	0.529	0.489	0.509
7. Household Goods	1.583	0.543	1.063
8. Other Purchases	5.474	0.197	2.836
9. Other Services	2.484	0.437	1.461

Estimates for all income family types are given in tables 11 and 12. Compared with tables 9 and 10, there is no doubt that much improvement can be achieved with increase in the sample sizes. For the first time, we have tables of results with no negative values and all estimates are reasonably consistent for nearly all commodities. The R^2 are almost all very high. 'Food' and 'fuel and light' are in all cases necessities with the estimates of β_1 generally increasing with family size. Since larger families are likely to be relatively poorer than the smaller families, no doubt such patterns of the estimates should be expected. The fact that these estimates are not regularly increasing from the small family size to larger size shows that there are some amount of economies of scale. This is further discussed later in this chapter. The budget proportions are shown in table 13. Generally for necessities the proportion increases with increase in household size while for luxuries, in most cases, the proportion decreases. All these are empirical results which will require further work possibibly using larger sample size.

The most important advantage of the first part of this section is the detection and analysis of the source of erratic estimates of β_1, which is generally regarded as a feature of underdeveloped economy. But most studies usually deal with broad group of households ignoring estimates for separate family types while using family size to get the cell means per person. Partly as a means of comparing this work with other studies and partly to derive more useful results e.g. comparison of occupational regression lines, estimates of β_1 and other results for different occupations and all low income were made. These are shown in tables 15 to 17. Food and fuel and light are consistently necessities while others - except for labourers, traders and farmers in some cases - are luxuries. The results for all low income seem the best set in terms of the higher values of R^2. The necessary variance-ratios for comparing the occupational regression lines were computed and set out in table 17 as in the case of table 8. There is a high significant difference between mean expenditure on food, drinks, other purchases and other services. As far as the slopes of the regression lines are concerned, none shows any significant difference. These results seem more plausible than those given in table 8. They seem to give some insight into the place of interdependence of preference orderings among consumers in the same community. Even though they may not be at the same level of consumption, which is probably justified by the significant differences between means, there is no reason to believe that their reactions to a change in income should be significantly different. Table 16 which shows budget proportions for occupations and all low income also supports the results given by the variance-ratios calculated for differences between means. From these results derived from table 17 compared with those given for table 8 we also have some support for the view generally held that cross-classified data are more suitable for econometric study than the one-way form. It shows also that it is better to pool observations together and cross-classify them rather than to group them only by income but for different areas and occupations. If the sample size is large enough to allow for this other kind of grouping, this should be done by all means. Nevertheless, first priority should be given to cross-classification. Also it is better to classify the data for each occupation rather than for each area because it seems occupational groups are more homogeneous than areas.

In the process of the discussion of the estimates some comments have been passed on some of the statistical problems discussed in chapter 3. In particular, concerning bias and efficiency of the estimates, and heteroscedasticity. As was indicated in section 3.3, the problem of heteroscedasticity always arises when estimates are based on family budget data, unless some ways of removing this problem directly or indirectly are found. It is my belief that the effect of heteroscedasticity has been reduced very much, if not removed, in this study. This has been done indirectly in many ways. First, the observations used are transformed since they are all in logarithmic form. In this form, much of the variation in the original data will be removed. Second, the observations are grouped means, with the β_1's estimated by weighted ordinary least squares method. This, in itself, is a kind of generalised least squares method as shown in section 3.1. And third, the way the observations are reduced to per consumer unit either per person or per adult equivalence, helps to reduce the effect of heteroscedasticity. This is much more so when adult equivalence or the two-way classified data is used (cf. the standard errors in per person with those based on adult equivalence in table 7). Some comments have been made in section 5.2 with regards to bias and efficiency of the estimates by comparing the standard errors shown in table 7.

In section 3.3, an approach suggested by Durbin (1954) for setting confidence limits for β_1 was discussed. On the basis of this, the confidence limits for the β_1's using the 'all low income 1-PERSON results as an illustration were calculated and shown in table 14. This means that the estimates based on both ordinary least squares and instrumental variable methods are within the limits. This is not a surprise since the estimates are reasonably close. It does not necessarily mean that this must always happen.

Much of what will be analysed in chapter 6 will be based on the 'all low income two-way classified data' particularly the results based on this set of data shown in tables 12, 13, 15 and 16. This is because the data in this form are better than any other form as shown in these tables. Considering table 15 (cf. results for the occupations and all low income) and the fact that there is no significant difference between estimates of β_1 for each set relating to the occupations, it seems other factors like area and occupation, influencing consumption patterns, do not have much effects on β_1 but on the levels as shown in tables 8 and 17. It then means that not much will be lost if we use common estimates of β_1 as

Table 18. ESTIMATES OF FAMILY EFFECT SCALES based on one-way classified data*

	IBADAN					KADUNA				
		TYPES OF PERSONS					TYPES OF PERSONS			
	ω	S_1	S_2	S_3	S_4	ω	S_1	S_2	S_3	S_4
1. Accomm.	0.07	1.00	0.01	0.0	0.01	0.11	1.00	-0.01	0.0	0.0
2. Food	0.50	1.00	0.09	-0.05	-0.02	0.44	1.00	0.10	-0.05	-0.10
3. Drink	0.08	1.00	-0.01	0.01	0.0	0.08	1.00	-0.03	0.0	0.02
4. Kola & Tob.	0.02	1.00	0.0	0.0	0.0	0.02	1.00	0.0	-0.01	0.01
5. Fuel & Lt.	0.04	1.00	0.01	0.0	0.0	0.05	1.00	0.02	-0.01	-0.01
6. Transport	0.07	1.00	0.0	0.0	0.0	0.06	1.00	0.0	-0.01	0.02
7. Clothing	0.08	1.00	0.01	-0.01	-0.01	0.11	1.00	-0.02	0.05	-0.03
8. Household G.	0.19	1.00	0.01	0.0	0.0	0.04	1.00	0.0	0.01	0.0
9. O.Purchases	0.04	1.00	0.0	0.0	0.0	0.05	1.00	0.0	0.0	0.01
10. O.Services	0.07	1.00	-0.02	0.02	0.01	0.05	1.00	-0.04	-0.01	0.03
INCOME		1.00	0.04	-0.02	0.01		1.00	0.02	-0.01	-0.03

	ENUGU					OS-IF-IL				
1. Accomm.	0.12	1.00	-0.16	0.43	-0.08	0.08	1.00	0.0	0.0	0.0
2. Food	0.39	1.00	0.86	0.63	-0.56	0.46	1.00	-0.01	0.0	-0.01
3. Drink	0.09	1.00	-0.50	0.47	-0.04	0.09	1.00	0.0	0.0	0.0
4. Kola & Tob.	0.02	1.00	-0.02	0.02	0.01	0.02	1.00	0.0	0.0	0.0
5. Fuel & Lt.	0.04	1.00	-0.18	0.09	0.14	0.04	1.00	0.0	0.0	0.0
6. Transport	0.06	1.00	0.0	0.63	0.94	0.04	1.00	0.0	0.0	0.0
7. Clothing	0.12	1.00	-0.07	0.23	0.22	0.12	1.00	0.0	0.0	0.0
8. Household G.	0.04	1.00	-0.58	0.02	0.27	0.04	1.00	0.0	0.0	0.0
9. O.Purchases	0.05	1.00	-0.44	0.34	-0.06	0.06	1.00	0.0	0.0	0.0
10. O.Services	0.08	1.00	-0.36	0.38	0.07	0.05	1.00	0.0	0.0	0.0
INCOME		1.00	-0.42	0.45	0.18		1.00	0.0	0.0	0.0

* Note that: S_1 = male adult, S_2 = female adult, S_3 = male child, S_4 = female child

ω = budget proportion.

Table 19. ESTIMATES OF FAMILY EFFECT SCALES based on TWO-WAY CLASSIFIED DATA FOR ARTISANS ONLY*

	2-PERSON							3-PERSON						
	TYPES OF PERSONS							TYPES OF PERSONS						
	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_1	S_2	S_3	S_4	S_5	S_6	S_7
1.Accomm.	1.00	-0.01	-0.07	0.09	0.12	0.26	0.31	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2.Food	1.00	0.70	-1.33	0.13	-0.29	0.0	0.0	1.00	0.0	0.0	0.0	0.0	0.0	0.0
3.Drink	1.00	0.50	-1.19	0.59	-0.02	0.55	0.0	1.00	0.0	-0.47	-0.77	1.24	0.0	0.0
4.Fuel-Lt.	1.00	0.07	-0.05	-0.01	-0.03	0.29	0.03	1.00	0.0	-0.23	-0.88	1.96	0.0	0.0
5.Transport	1.00	0.0	-0.80	0.91	0.17	-1.04	-1.25	1.00	1.0	1.0	1.0	1.0	1.0	1.0
6.Clothing	1.00	1.34	-0.83	0.34	0.01	-0.37	0.41	1.00	0.0	0.62	1.70	0.0	0.0	0.0
7.Househd.G.	1.00	-0.09	0.47	-0.15	0.01	-0.52	-0.85	1.00	0.0	0.41	0.57	-0.91	0.0	0.0
8.O.Purchases	1.00	0.62	0.43	0.24	0.20	-0.71	0.0	1.00	0.0	0.20	0.09	-0.19	0.0	0.99
9.O.Services	1.00	-0.08	0.22	0.01	0.10	-0.26	-0.36	1.00	1.62	-0.26	0.07	-0.32	-0.78	0.55
INCOME	1.00	0.49	-0.81	0.26	-0.07	-0.11	-0.11	1.00	-0.09	0.04	0.09	0.14	-0.03	-0.04

continued...

Table 19 continued

	4-PERSON							6-PERSON						
	TYPES OF PERSONS							TYPES OF PERSONS						
	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_1	S_2	S_3	S_4	S_5	S_6	S_7
1.Accomm.	1.00	0.43	0.0	0.0	0.51	0.0	1.61	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2.Food	1.00	0.0	0.0	0.0	0.26	0.0	0.0	1.00	0.0	0.0	0.0	0.0	0.0	-0.13
3.Drink	1.00	1.15	0.0	0.0	0.43	-1.32	1.99	1.00	0.80	0.0	-1.62	0.0	-0.81	1.13
4.Fuel-Lt.	1.00	0.74	1.66	1.09	0.26	0.27	0.38	1.00	-0.31	0.0	-0.41	0.0	0.46	-0.46
5.Transport	1.00	1.09	0.50	0.0	0.84	-1.57	1.50	1.00	-0.12	0.0	1.18	0.0	0.45	0.09
6.Clothing	1.00	1.11	0.0	0.0	0.57	-1.51	0.0	1.00	1.0	1.00	1.0	1.00	1.00	1.00
7.Househd.G.	1.00	0.69	0.0	0.0	0.08	0.14	0.83	1.00	0.19	1.68	0.0	0.0	0.0	0.27
8.O.Purchases	1.00	0.08	0.39	0.0	-0.25	0.19	0.39	1.00	1.07	0.0	0.0	0.0	0.02	1.35
9.O.Services	1.00	0.16	0.14	0.0	-0.04	-0.27	0.48	1.00	1.06	0.0	0.0	0.0	0.82	1.06
INCOME	1.00	0.39	0.14	0.06	0.32	-0.35	0.49	1.00	0.03	0.01	-0.05	0.0	-0.04	-0.13

* Estimates for 5-PERSON FAMILY Are rejected because of the presence of negative estimates for corresponding columns in table 11.

And note that S_i (i = 1,2, ... 7) are as defined in section 4.7.

given by all low income two-way classified data when our main interest is in elasticities of demand.

3.4 FAMILY EFFECT SCALES

In the estimation of β_1 based on adult equivalences, types of persons defined by sex and age are used. In the case of one-way classified data, four types of persons were defined, and seven types were defined in the case of two-way classified ones. The methods of estimating the scales of the family effects, based on the types of persons and of estimating β_1 through adult equivalences, are discussed in sections 3.4 and 3.5. The scales estimated and used in this study are shown in tables 18 and 19. The main purpose of introducing the concept of adult equivalence is to eliminate family composition effects from the family budget data so as to get better estimates of the parameters. It is the same objective as the consideration of other factors like area and occupation as discussed above. As far as eliminating family composition effects is concerned, it is reasonably achieved for one-way classified data as shown in table 7. This is in the sense that standard errors of the estimates are reduced compared with estimates based on 'per person'. This is far less the case with two-way classified data; probably because of the close relationship between household composition, size and income, and the cross-tabulation as mentioned above in section 3.3. There is a tendency for large families to have high number of persons for each type of persons and to earn high income. This means that, for cross-classified data, much of the variation due to family composition must have been obscured in income and size used for grouping.

Another aspect is the estimation of adult equivalence scales. It is reasonable to expect adult equivalence scales to serve as a meaningful indication of the behaviour of the different types of persons in the family, particularly their consumption behaviour patterns. There is not much meaning that could be attached to the estimates of adult equivalence scales set out in tables 18 and 19 (for the details about methods used, see sections 3.4 and 3.5). This may be because the commodities are broadly classified. In such a situation, the impact of individual members or types of members in the family can not be easily felt. For adult equivalence scales to be meaningful with the right magnitudes and sign, the level of aggregation of the commodities to form single commodities must be such that it should be possible to identify what is required collectively and individually by the members of a household e.g. baby food is required by children and can only be bought by families with children, whereas tobacco is only required by some adults in the family. At such a level of aggregation, the scale parameters will be constrained by certain a priori information, in particular, some of the α's of 3.4(1) can be assumed to be zero. In the present study, the only constraint put on the scale parameters is that the total of adult equivalences for any household should be greater than zero and that no scale should be greater than 2 in absolute value.

The problem of getting meaningful adult equivalence scales is universal. Even Forsyth (1960) was not happy with his estimates of equivalence scales.[1] He regarded his equivalence scale hypothesis as being too simple. But as in this study, his commodities were broadly classified and this must have affected his results as shown above.

5.5 COMPARISON WITH OTHER STUDIES[2]

Comparison, in the sense in which one would expect it to be employed here, is not easy since it is difficult to get an identical study with which to compare the present results. In an ideal situation, there is need to compare them with previous work in Nigeria and with studies of other countries of identical and different stages of development. In the Nigerian situation, examples of previous studies are very scanty if not non-existing. In the case of other studies, two relatively recent studies will be mentioned - Houthakker (1957) and Forsyth (1960).

The Federal Office of Statistics (1966) made some estimates using the same one-way classified data for low income earners in Ibadan, Kaduna and Enugu and fitting a

---

1. See Forsyth (1960), p.387.
2. In this section the estimates of β_1 can be interpreted as income elasticities based on double logarithmic form to fit in with other studies used here.

linear Engel function for each of similar set of commodity groups studied in this work, except that 'kola and tobacco' was excluded. The results were poor, because no account was taken of the fact that the observations were grouped means and in addition, some of the considerations given in this study were neglected. As an illustration the estimates for Kaduna are given as:

	linear Regression co-efficient	R^2
Food	0.19	0.39
Drink	0.11	0.40
Fuel and Light	0.03	0.18
Transport	0.15	0.25
Clothing	0.14	0.28
Accommodation	0.10	0.39

A simple assessment of these results can be made by comparing their R^2 with those given in table 7 under 'per person'. The main criticism of the producers of these results is that no effort was made to examine the results critically.

The main discussion concerning other studies relevant to this work will be based on Houthakker (1957) and Forsyth (1960). Houthakker (1957) made an international comparison of Engel elasticities. Four broad commodities were examined - food, clothing, housing and miscellaneous goods. Only food and clothing fit well with the definitions used in this study. Housing, in addition to rent, covers fuel and light while the fourth commodity covers all the other groups. The relevance of Houthakker's study is seen in the aspect concerning development planning in the developing countries where he asserted, in the introductory part of his work, that: "Many countries are now engaged in the construction of development programs and to do this adequately it is clearly necessary to have some idea about the changes in consumption that are likely to occur with rising income levels. In many underdeveloped countries, however, the data for estimating changes in consumption are unfortunately lacking. There is a real question concerning the extent to which data from one country are applicable to conditions in another country; it is one of the principal aims of the present investigation to clarify this point".[1] In the closing remarks, relating to the statement quoted above, Houthakker maintained that: "if no data on the expenditure patterns of a country are available at all, one would not be very far astray by putting the partial elasticities with respect to total expenditure at 0.6 for food, 1.2 for clothing, 0.8 for housing and 1.6 for all other items combined and the partial elasticity with respect to family size at 0.3 for food, zero for housing and clothing, and -0.4 for miscellaneous expenditures. But it would be prudent not to use these guesses for wide extrapolations, and more prudent still to organise a survey and cross-classify the

1. See Houthakker (1957), p.533.

results".[1] Some approximate estimates for partial elasticities with respect to family size corresponding to estimates of β_1 given in tables 11 and 15 are set out in table 20[2]. Considering Houthakker's observations in the light of tables 15 and 20a, particularly for the results under 'all low income', even though his guesses are within reasonable range, it seems prudent for a country like Nigeria 'to organise a survey and cross-tabulate the results', for the sole purpose of estimating the required parameters.

Table 20b is useful for the light it sheds on the concept of economies of scale which can be examined by considering the changes in the estimates as the family size increases particularly in the case of 'food' and 'fuel and light'.

---

1. Ibid. p.550.
2. These estimates are derived by using the relation

$$1 - E_{io} = \eta_i$$

existing between income elasticity, E_{io} (if we interpret β as income elasticity, at least in the case of double log. function), and family size elasticity, η_i, for commodity i [see Prais and Houthakker (1955), p.90]. Similar results are achieved by considering the fitted regression

$$\log \left(y_i/_n\right) = \hat{\beta}_{oi} + \hat{\beta}_{1i} \log \left(x/_n\right) \quad i=1,2, \ldots. n$$

which can be expressed as

$$\log y_i = \hat{\beta}_{oi} + \hat{\beta}_{1i} \log x + \left(1 - \hat{\beta}_{1i}\right) \log n$$

Hence $\eta_i = 1 - \hat{\beta}_{1i}$ $\quad (i=1,2, \ldots. n)$

Table 20. **ESTIMATES OF FAMILY-SIZE CO-EFFICIENTS based on TWO-WAY CLASSIFIED DATA**

(a) DIFFERENT OCCUPATIONS AND ALL LOW INCOME

	CLERKS	LABOURERS	ARTISANS	TRADERS	FARMERS	ALL LOW INCOME
1. Accom.	-0.450	-0.780	-0.293	-0.232	-0.091	-0.400
2. Food	0.256	0.223	0.256	0.242	0.272	0.271
3. Drink	-0.105	-0.266	-0.215	-0.108	0.114	-0.138
4. Fuel-Lt.	0.216	0.306	0.219	0.375	0.532	0.324
5. Transport	-0.147	0.198	-0.207	-0.113	0.185	-0.313
6. Clothing	-0.535	0.037	-0.337	-0.262	-0.399	-0.397
7. Househd. G.	-0.306	-0.081	-0.413	0.057	-0.232	-0.457
8. O. Purchases	-0.263	-0.370	-0.356	-0.163	-0.244	-0.418
9. O. Services	-0.159	-0.058	-0.138	0.211	-0.329	-0.151

(b) ALL LOW INCOME FAMILY GROUPS

	1-PERSON	2-PERSON	3-PERSON	4-PERSON	5-PERSON	6-PERSON
1. Accom.	-0.03	-0.11	-0.32	-0.76	-0.39	-0.43
2. Food	0.47	0.56	0.50	0.46	0.46	0.43
3. Drink	0.14	-0.09	-0.25	-0.36	-0.22	-0.18
4. Fuel-Lt.	0.65	0.46	0.50	0.40	0.43	0.25
5. Transport	-0.61	-0.84	-0.86	-0.60	-0.62	-0.88
6. Clothing	-0.55	-0.75	-0.89	-0.69	-1.15	-0.80
7. Househd. G.	-0.78	-0.88	-0.18	-1.03	0.15	0.07
8. O. Purchases	-1.01	-0.90	-0.72	-0.68	-0.81	-0.91
9. O. Services	-0.72	-0.39	-0.65	-0.06	-0.80	-0.81

The work of Forsyth (1960) was based on United Kingdom data. Most of the commodity groups are the same as those defined in this study. His 'standard family', which is a couple, is similar to the 2-person group defined in this study. Hence for commodity groups of identical composition, Forsyth's estimates[1] of β_1 from the double logarithmic form are given below side by side with the relevant column of table 11:

	Forsyth's Estimates for 'standard family'	All low income 2-Person
Housing (Accommodation)	0.76	1.11
Food	0.53	0.44
Drink	1.25	1.09
Fuel and Light	0.35	0.54
Transport	1.91	1.84
Clothing	1.30	1.75
Household Goods	1.61	1.88
Other Goods (Other Purchases)	0.99	1.90
Other Services	2.12	1.39

The results are reasonably close apart from food which is under estimated for 2-person all low income in the present study (cf. this estimate with others in table 11). The difference between luxury and necessity is evident solely with respect to housing and other goods (other purchases). In the case of housing, this can be explained by climatic differences. Other differences will be due to different stages of development in view of the difference between U.K. and Nigeria.

1. See Forsyth (1960), table 1, col.2, model 3 for the "Forsyth's estimates" given.

CHAPTER 6

THE ELASTICITIES

6.0 INTRODUCTION

This chapter is related to chapter 2 as chapter 5 is to chapter 3 as far as empirical results of theoretical and methodological aspects of this work are concerned. In chapter 5, the estimates based upon the procedures outlined in chapter 3 were analysed and the performances of the alternative estimating procedures were assessed. Results based on different forms of the available data and the effects of various other factors (apart from income) are similarly analysed. Estimates based on cross-tabulated data, in particular those for all low income, were found to be statistically better. These estimates, shown in tables 12, 13, 15 and 16, are now used to derive elasticities as discussed in chapter 2, especially sections 2.4 and 2.5. The estimated elasticities form the basis of the discussion in this chapter. They are set out in tables 21 to 27. The main empirical contribution of this study to economic theory is found in this chapter. It is an attempt at using the meagre statistical information available to derive results characterising demand functions. Such results are important in aiding economic policies. In section 6.1, the indirect addilog elasticities, as given by 2.5(11) and 2.5(18) are discussed. Section 6.2 deals with elasticities based on Frisch's formulae, 2.4(10) and 2.4(11). Elasticities based on Frisch's formulae were made possible by using indirect addilog estimates of E_{io}, E_{ii} and ω_i (i=1,2, ... n) to derive Frisch's money flexibility, $\check{\omega}$, through formula 2.4(12). The justification for linking these two seemingly unrelated concepts is that Leser's assumption, 2.4(4), which is the basis of indirect addilog approach is a special case of Frisch's 'want-independence' assumption. For Leser's assumption, the cross-elasticity is the same for all other commodities when there is a change in the price of commodity i. For Frisch's want-independence, this is not necessarily so since it varies with the income elasticity of each of the other commodities. Although the substitutability or complementarity is likely to be low in both cases, it does not necessarily have to be the same for all other commodities in the case of Frisch's want-independence. Frisch's suggestion[1] that if $\check{\omega}$ is unknown it could be derived independently for estimating 2.4(10) and 2.4(11) buttresses the present approach. Tentatively, we would suggest that Leser-Frisch approach is a valuable contribution for an economy where only survey data with just one period level exist. It may even be used, where there are abundance of data, as check on some estimates from other sources. And in the concluding section, 6.3, the present results from sections 6.1 and 6.2 are compared with those estimated by Leser (1941-42), Frisch (1959) and others based on these sources.

---

1. See Frisch (1959), p.188.

Table 21. **INDIRECT ADDILOG INCOME ELASTICITIES FOR ALL LOW INCOME FAMILY GROUPS**

	1-PERSON	2-PERSON	3-PERSON	4-PERSON	5-PERSON	6-PERSON
1. Accom.	1.058	1.098	1.367	1.760	1.526	1.470
2. Food	0.554	0.425	0.530	0.517	0.517	0.559
3. Drink	0.859	1.078	1.228	1.365	1.296	1.178
4. Fuel & Light	0.391	0.534	0.556	0.580	0.531	0.721
5. Transport	1.631	1.738	1.905	1.573	1.724	1.843
6. Clothing	1.557	1.683	1.856	1.668	2.070	1.792
7. Household G.	1.847	1.917	1.211	2.034	0.691	1.018
8. O. Purchases	1.993	1.918	1.767	1.664	1.800	1.889
9. O. Services	1.756	1.373	1.608	0.996	1.754	1.744

6.1 INDIRECT ADDILOG ELASTICITIES

The indirect addilog income elasticities shown in table 21 are computed from tables 12 and 13, and those in table 27 based on relevant parts of tables 15 and 16 using 2.5(11). As was pointed out in section 2.5, the main difference between the ordinary double logarithmic income elasticity and the indirect addilog income elasticity is $1 - \sum_{j=1}^{n} \beta_j \omega_j$. It seems pertinent to ask whether estimates derived from one method are likely to be consistently larger, equal or smaller than those of the other i.e. whether

$$1 - \sum_{j=1}^{n} \beta_j \omega_j > 0 \qquad 6.1(1a)$$

$$= 0 \qquad 6.1(1b)$$

$$< 0 \qquad 6.1(1c)$$

If 6.1(b) is satisfied then condition 2.2(21) is fulfilled. Under such a situation the ordinary double-logarithmic approach will be satisfactory. Theoretically, condition 2.2(21) is not satisfied by the double-logarithmic approach while the indirect addilog does, since

$$\sum_{i=1}^{n} \omega_i \left[\beta_i + 1 - \sum_{j=1}^{n} \beta_j \omega_j \right] = 1$$

In all the estimates, no set actually show that 6.1(1b) is satisfied as evident by comparing tables 12 and 21 (if it does, corresponding columns will be the same). This shows a serious breakdown of stability conditions in the use of ordinary double-logarithmic approach. Two of the columns of table 21 (1-Person and 3-Person) satisfy 6.1(1a) and the

Table 22. **INDIRECT ADDILOG PRICE ELASTICITIES FOR ALL INCOME FAMILY GROUPS**

(a) Own-Price

	1-PERSON		2-PERSON		3-PERSON		4-PERSON		5-PERSON		6-PERSON	
1. Accom.	-1.052	-0.90G	-1.087	-0.89G	-1.334	-0.91G	-1.700	-0.92G	-1.489	-0.93G	-1.443	-0.94G
2. Food	-0.734	-0.60G	-0.658	-0.60G	-0.760	-0.51G	-0.748	-0.52G	-0.765	-0.49G	-0.789	-0.48G
3. Drink	-0.874	-0.89G	-1.070	-0.89G	-1.205	-0.90G	-1.326	-0.89G	-1.268	-0.91G	-1.159	-0.90G
4. Fuel-Lt.	-0.413	-0.96G	-0.553	-0.96G	-0.577	-0.95G	-0.607	-0.94G	-0.557	-0.95G	-0.735	-0.95G
5. Transport	-1.600	-0.95G	-1.689	-0.93G	-1.861	-0.95G	-1.546	-0.95G	-1.693	-0.96G	-1.803	-0.95G
6. Clothing	-1.492	-0.88G	-1.611	-0.89G	-1.775	-0.91G	-1.601	-0.90G	-1.970	-0.91G	-1.724	-0.91G
7. Househd. G.	-1.806	-0.95G	-1.882	-0.96G	-1.203	-0.97G	-1.997	-0.96G	-0.693	-0.98G	-1.017	-0.98G
8. O. Purchases	-1.935	-0.94G	-1.666	-0.94G	-1.734	-0.96G	-1.637	-0.96G	-1.769	-0.96G	-1.845	-0.95G
9. O. Services	-1.716	-0.95G	-1.346	-0.93G	-1.576	-0.95G	-0.996	-0.95G	-1.700	-0.93G	-1.698	-0.94G

(b) Cross-Price

	1-PERSON		2-PERSON		3-PERSON		4-PERSON		5-PERSON		6-PERSON	
1. Accom.	0.006	+0.10G	0.011	+0.11G	0.033	+ 0.09G	0.060	+ 0.08G	0.037	+0.07G	0.027	+0.06G
2. Food	-0.180	+0.40G	-0.233	+0.41G	-0.230	+ 0.49G	-0.231	+ 0.48G	-0.248	+0.51G	-0.230	+0.52G
3. Drink	-0.015	+0.11G	0.003	+0.11G	0.023	+ 0.10G	0.039	+ 0.11G	0.028	+0.09G	0.019	+0.11G
4. Fuel-Light	-0.022	+0.04G	-0.020	+0.04G	-0.021	+ 0.05G	-0.027	+ 0.06G	-0.026	+0.06G	-0.014	+0.05G
5. Transport	0.031	+0.05G	0.049	+0.07G	0.044	+0.04G	0.027	+ 0.05G	0.031	+0.04G	0.040	+0.05G
6. Clothing	0.065	+0.12G	0.072	+0.11G	0.081	+0.09G	0.067	+ 0.10G	0.100	+0.09G	0.068	+0.09G
7. Household G.	0.041	+0.05G	0.035	+0.04G	0.008	+ 0.03G	0.037	+ 0.04G	-0.007	+0.02G	0.001	+0.02G
8. O.Purchases	0.059	+0.06G	0.052	+0.06G	0.033	+ 0.04G	0.027	+0.04G	0.031	+0.04G	0.044	+0.05G
9. O.Services	0.040	+0.05G	0.027	+0.07G	0.032	+0.05G	0.00	+0.05G	0.054	+0.07G	0.046	+0.06G

rest 6.1(1c). This means that $1 - \sum_{j=1}^{n} \xi_j w_j$ has no systematic direction but its absolute value was never more than 0.02 for results in table 21.

As was presented in section 2.4 and 2.5, it is not easy to estimate price elasticities from family budget data even after making some assumptions about consumer preference orderings. There is still an indeterminate value G not known. Any results derived are conditioned by the value of G. Table 22 gives the general results with the indeterminate value, G for individual family types of all low income. It is clear from the table that the effect of G on own-price elasticities is greater than its effect on cross-price elasticities although the two price elasticities are not comparable in magnitude. In table 23, an attempt is made to illustrate how a feasible and workable value of G which will be reasonable can be located empirically. On the basis of this search, it is hoped that setting G = 0 (or close to it) will give a reasonable estimate. From table 23 it is observed that as the value of G increases from zero, all own price elasticities increase in absolute value and as G decreases from zero, all own-price elasticities decrease in absolute value. For cross-price elasticities, they vary directly with G although at G = 0, all necessaries are negative and all luxuries are positive. A useful and satisfactory value of G must, in broad terms, intuitively fulfil the condition that a commodity with elastic income elasticity should have elastic price elasticity, and one with inelastic income elasticity should have inelastic price elasticity irrespective of their magnitude. Another aspect of this problem to consider is the position of luxuries and necessities; it seems reasonable to assume that a commodity that is a luxury would have positive cross-price elasticity with respect to other commodities and a necessity, a negative cross-price elasticity with respect to other commodities. Suppose there is an increase in the price of food, such a change should have negative effect on other commodities since food is a necessity. On the basis of this discussion, the choice of G = o seems a reasonable first approximation.

With G = o in table 22, 'food' and 'fuel and Light' are inelastic while the others are elastic. The estimates of cross-price elasticities for 'food' and 'fuel and light' are also negative while for others they are positive. Apart from food, the cross-price elasticities are reasonably low. Similar estimates putting G = o are made for all low income based on tables 15 and 16 and shown in a section of table 27. These results can be interpreted in the same way as those made above but there is a reduction in the values for cross-price elasticity with respect to food compared with those shown in table 22.

Table 23. INDIRECT ADDILOG ELASTICITIES BASED ON SOME GIVEN VALUE OF G FOR ALL LOW INCOME 4-PERSON FAMILY GROUP

	G = -1		G = -½		G = 0		G = ½	
	own price	cross	own price	cross	own price	cross	own price	cross
1. Accom.	-0.780	-0.020	-1.240	0.020	-1.700	0.060	-2.160	0.100
2. Food	-0.238	-0.711	-0.488	-0.471	-0.748	-0.231	-1.008	0.009
3. Drink	-0.436	-0.071	-0.881	-0.016	-1.326	0.039	-1.771	0.094
4. Fuel & Lt.	-0.333	-0.087	-0.137	-0.057	-0.607	-0.027	-1.077	0.003
5. Transport	-0.596	-0.023	-1.071	0.002	-1.546	0.027	-2.021	0.052
6. Clothing	-0.701	-0.033	-1.151	0.017	-1.601	0.067	-2.051	0.117
7. Househd. G.	-1.037	-0.003	-1.517	0.017	-1.997	0.037	-2.477	0.057
8. O.Purchases	-0.677	-0.013	-1.157	0.007	-1.637	0.027	-2.117	0.047
9. O. Services	0.046	-0.050	-0.521	-0.025	-0.996	0.0	-1.471	0.025

6.2 ELASTICITIES BASED ON FRISCH'S FORMULAE

On the basis of the remark made in the introductory section of this chapter (section 6.0), concerning the estimation of Frisch's money flexibility from indirect addilog estimates, the results shown in table 24 were computed for different family types of all low income using equation 2.4(12) after putting G = 0. Apart from food, the estimates for different commodities are quite close. The corresponding estimates for all low income are shown in table 27 with estimates for all commodities including food quite close. It is assumed as before that the effect of setting G = 0 is very small. This is revealed by

$$\check{\omega} = \frac{\left(\beta_i + 1 - \sum_{j=1}^{n} \beta_j \omega_j\right)\left(\sum_{j=1}^{n} \beta_j \omega_j - \beta_i\right)}{-1 - \left(-\omega_i\right)\left(\beta_i - \sum_{j=1}^{n} \beta_j \omega_j + G\right)}$$

if we substitute for E_{io} and E_{ii} in 2.4(12), using 2.5(11) and 2.5(16) respectively. An assessment of the estimates of $\check{\omega}$ is given in section 6.3 where the present results are compared with values given by Frisch (1959).

For the family types of all low income, the own-price elasticities based on 2.4(10) and cross-price elasticities based on 2.4(11) after estimating $\check{\omega}$ as given above are shown in tables 25 and 26. As an illustration, cross-elasticities are only given for 4-Person family. Like the indirect addilog estimates, the own-price elasticities maintain the same direction with 'food' and 'fuel and light', as inelastic commodities while the others are elastic. Since Frisch's want-independent assumption is a more general one than that of Leser as pointed out in section 6.0, the interpretation of the cross elasticities based on it should be different. That a commodity is want-independent in relation to all other commodities does not necessarily mean that the degree of want-independence should be the same; or in terms of cross-elasticities with respect to other commodities should be the same. In actual fact the cross-price elasticities vary with income elasticities. A commodity that is want-independent with other commodities will then have cross-price elasticities which have the same direction but not necessarily the same magnitude. What table 26 has in common with table 22b is that, as in table 22, all cross-price elasticities with respect to a change in the price of a particular commodity have the same sign i.e. with respect to the price of a necessity, the sign is negative and for luxury, it is positive. Again, in magnitude, apart from food, all cross-price elasticities are reasonably low.

The results for all low income are given in table 27 and they seem to be more stable, particularly with reference to 'food', than those discussed in the last paragraph. The Frisch's money flexibility are quite close for all commodities. With regards to cross-price elasticities, it seems reasonable to say that a commodity is more want-dependent with respect to another commodity the nearer the income elasticities of the two commodities. Two commodities that

Table 24. ESTIMATES OF FRISCH'S MONEY FLEXIBILITY FOR ALL LOW INCOME FAMILY GROUPS

	1-PERSON	2-PERSON	3-PERSON	4-PERSON	5-PERSON	6-PERSON
1.Accom.	-1.00	-0.999	-0.990	-0.970	-0.986	-0.990
2.Food	-0.843	-0.724	-0.784	-0.778	-0.760	-0.796
3.Drink	-0.997	-0.999	-0.995	-0.988	-0.994	-0.997
4.Fuel-Light	-0.966	-0.984	-0.984	-0.980	-0.977	-0.994
5.Transp.	-0.987	-0.977	-0.977	-0.990	-0.986	-0.981
6.Clothing	-0.972	-0.965	-0.958	-0.969	-0.940	-0.966
7.House-hold G.	-0.980	-0.982	-0.999	-0.980	-0.997	-1.001
8.O.Purch.	-0.969	-0.972	-0.985	-0.988	-0.985	-0.978
9.O.Serv.	-0.981	-0.926	-0.987	-1.00	-0.974	-0.978
Average	-0.966	-0.948	-0.962	-0.960	-0.955	-0.965

are both necessaries are more want-dependent than any one of them will be with another commodity that is a luxury. Similar results will be obtained by interchanging necessity and luxury in the last sentence. Observation of the columns under cross-price elasticities in table 27 shows these clearly. One need hardly labour the point that cross-price elasticities under the assumption of want-independent are not symmetrical. Table 27 gives the summary of the main contribution of this chapter. From these results it seems, Leser-Frisch approach, whereby the money flexibility is estimated from indirect addilog and subsequently used to estimate own- and cross-price elasticities, is worth recommending for use with the kind of data used in this study.

Table 25. OWN-PRICE ELASTICITIES FOR ALL LOW INCOME FAMILY GROUPS USING FRISCH'S FORMULAE

	1-PERSON	2-PERSON	3-PERSON	4-PERSON	5-PERSON	6-PERSON
1.Accomm.	-1.085	-1.139	-1.370	-1.716	-1.534	-1.478
2.Food	-0.669	-0.543	-0.668	-0.653	-0.663	-0.702
3.Drink	-0.899	-1.121	-1.243	-1.160	-1.315	-1.193
4.Fuel & Light	-0.413	-0.573	-0.589	-0.619	-0.569	-0.756
5.Transp.	-1.633	-1.737	-1.889	-1.591	-1.746	-1.831
6.Clothing	-1.500	-1.637	-1.768	-1.614	-1.942	-1.725
7.House-hold G.	-1.831	-1.948	-1.248	-2.037	-0.728	-1.054
8.O.Purch.	-1.940	-1.912	-1.773	-1.683	-1.823	-1.868
9.O.Servs.	-1.742	-1.272	-1.615	-1.036	-1.732	-1.720

Table 26. CROSS-PRICE ELASTICITIES FOR ALL LOW INCOME 4-PERSON GROUP USING FRISCH'S FORMULAE

i \ j	1 Acc	2 Food	3 Drink	4 F & L	5 Transp	6 Clothg	7 Hshd G	8 O.Pur.	9 O.Servs.
1.Accom.	-	-0.389	0.080	-0.044	0.053	0.130	0.071	0.053	0.003
2.Food	0.041	-	0.024	0.013	0.016	0.038	0.021	0.016	0.001
3.Drink	0.108	-0.302	-	-0.034	0.041	0.101	0.055	0.041	0.002
4.Fl-Lt.	0.046	-0.128	0.026	-	0.017	0.043	0.023	0.017	0.001
5.Trans.	0.124	-0.348	0.072	-0.039	-	0.116	0.063	0.047	0.003
6.Clothg.	0.132	-0.369	0.076	-0.042	0.047	-	0.067	0.050	0.003
7.H.Goods	0.160	-0.450	0.093	-0.051	0.050	0.150	-	0.061	0.004
8.O.Purch.	0.131	-0.368	0.076	-0.041	0.061	0.123	0.067	-	0.003
9.O.Servs.	0.079	-0.220	0.045	-0.025	0.050	0.073	0.040	0.03.	-

Table 27. **ESTIMATES OF ELASTICITIES FOR ALL LOW INCOME GROUP based on two-way classified data per person**

	Indirect Addilog Est.*				Est. Based on Frisch's Formulae*									
						E_{ij}								
	E_{io}	E_{ii}	E_{ij}	$\check{\omega}$	E_{ii}	j=1 Acc	j=2 Food	j=3 Drink	j=4 F&L	j=5 Trans.	j=6 Cloth.	j=7 Hshd.	j=8 O.Pur.	j=9 O.Ser.
1. Accomm.	1.380	-1.34	0.036	-0.992	-1.349	-	-0.186	0.020	-0.222	0.024	0.061	0.025	0.033	0.013
2. Food	0.684	-0.824	-0.140	-0.914	-0.787	0.026	-	0.010	-0.011	0.012	0.030	0.013	0.016	0.006
3. Drink	1.115	-1.102	0.013	-0.999	-1.116	0.043	-0.150	-	-0.018	0.020	0.050	0.020	0.026	0.010
4. Fuel & Lt.	0.639	-0.655	-0.016	-0.991	-0.659	0.024	-0.086	0.009	-	0.011	0.028	0.012	0.013	0.006
5. Transport	1.319	-1.303	0.016	-0.995	-1.316	0.050	-0.178	0.018	-0.021	-	0.059	0.024	0.031	0.012
6. Clothing	1.391	-1.349	0.042	-0.986	-1.350	0.053	-0.187	0.020	-0.022	0.025	-	0.026	0.033	0.013
7. Househd. G.	1.437	-1.419	0.018	-0.995	-1.433	0.055	-0.194	0.021	-0.023	0.025	0.064	-	0.034	0.013
8. O.Purchases	1.425	1.403	0.022	-0.992	-1.413	0.054	-0.192	0.020	-0.023	0.025	0.063	0.026	-	0.013
9. O.Services	1.142	-1.134	0.008	-0.999	-1.149	0.044	-0.154	0.016	-0.018	0.020	0.051	0.021	0.027	-

average $\check{w}$ = -0.985

* G = 0

6.3 COMPARISON WITH OTHER STUDIES

In his pioneering work, Leser (1941-42) made estimates of the kind shown in table 22. Even though his method of estimation was different from that used above, his estimates presented as part of table 28a are quite similar except that his indeterminate value, A, is a bit different from G as used above. Table 28 contains results based on tables 2 and 3 of his work and they are comparably similar to those given in tables 22 above. From table 28 it can be easily observed that his results show that a value of A between 0.135 and 1.0 (as he, himself indicated) will be more appropriate. For example if A = ½ is used, all commodities which have income-inelastic demand have negative cross-price elasticities while all those with income-elastic demand have positive cross-price elasticities. Similarly all commodities that are income-inelastic have own-price inelastic demand and those that are income-elastic have also own-price-elastic demand. In general these results are quite in agreement with G = e as given in table 27 under indirect addilog elasticities. As a means of buttressing the idea of linking his work with that of Frisch as carried out above, table 28b was computed on the basis of table 28a. The computation puting A = ½ gives Frisch's money flexibility (average) of -1.043 for the U.S. 1918-19 households used. The estimates for price elasticities based on Frisch's formulae are quite consistent with the results reported above.

Table 28. INCOME AND PRICE ELASTICITIES OF DEMAND, U.S. 1918-19 based on Leser (1941-42), tables 2 and 3 *

(a) INDIRECT ADDILOG

(i,j) Commodity	w_j	E_{jo}	E_{jj}	E_{ji}	E_{jj} A=0.135	E_{jj} A=½	E_{jj} A=1	E_{ji} A=0.135	E_{ji} A=½	E_{ji} A=1
1. Food	0.389	0.70	-0.48 -0.61A	-0.33 +0.39A	-0.56	-0.79	-1.09	-0.28	-0.14	0.06
2. Clothing	0.162	1.39	-0.86 -0.84A	-0.03 +0.16A	-0.97	-1.28	-1.69	-0.018	-0.05	0.13
3. Rent	0.136	0.79	-0.33 -0.86A	-0.10 +0.14A	-0.45	-0.76	-1.20	-0.08	-0.03	0.04
4. Fuel & Light	0.054	0.56	-0.05 -0.95A	-0.05 +0.05A	-0.18	-0.53	-1.00	-0.04	-0.03	0.0
5. Furniture & Furnishing	0.050	1.42	-0.87 -0.95A	-0.01 +0.05A	-1.00	-1.35	-1.82	-0.003	0.02	0.04
6. Miscellaneous	0.209	1.38	-0.86 -0.79A	-0.04 +0.21A	-0.97	-1.25	-1.65	-0.012	0.07	0.17

* A > o, indeterminate

continued...

Table 28 continued

(b) FRISCH'S FORMULAE FOR $A=\frac{1}{2}$

i	ω	E_{ii}	E_{ij} j=1 Food	j=2 Cloth	j=3 Rent	j=4 Fuel	j=5 Furn.	j=6 Misc.
1. Food	-1.003	-0.761	-	0.038	-0.023	-0.018	0.013	0.047
2. Cloth.	-1.014	-1.258	-0.178	-	-0.046	-0.035	0.025	0.094
3. Rent	-1.081	-0.783	-0.101	0.043	-	-0.020	0.014	0.053
4. Fuel & Light	-1.087	-0.551	-0.072	0.030	-0.018	-	0.010	0.038
5. Furnit.	-1.031	-1.336	-0.182	0.077	-0.047	-0.036	-	0.096
6. Misc.	-1.042	-1.249	-0.177	0.074	-0.046	-0.035	0.025	-

Average ω = -1.043

In considering the estimates of Frisch's money flexibility it is necessary to bear in mind his assumption about this parameter that[1]:

"in most cases the money flexibility has values of the order of magnitude given below:

$\breve{\omega}$ = -10 for an extremely poor and apathetic part of the population.

$\breve{\omega}$ = -4 for the slightly better off but still poor part of the population with a fairly pronounced desire to become better off.

$\breve{\omega}$ = -2 for the middle income bracket, 'the median part' of the population.

$\breve{\omega}$ = -0.7 for the better-off part of the population.

$\breve{\omega}$ = -0.1 for the rich part of the population with ambitions towards 'conspicuous consumption'."

If we use the average of the values of $\breve{\omega}$ shown in table 27, which is -0.985 as the estimates of $\breve{\omega}$ for all low income group of the selected urban households in Nigeria, it seems by Frisch's assumption, that this group has $\breve{\omega}$ which falls between -0.7 and -2.0 with the interpretation given above. From table 28b, it is no surprise to see that the U.S. households of 1918-19 studied by Leser (1941-42) has an average of -1.043 for $\breve{\omega}$. These results definitely must have some policy implication which is outside the scope of this work.

CHAPTER 7

CONCLUSION

In order to give a brief indication of the problems endemic in a study of this kind and of feasible directions for further investigations along the lines we have been following, it would appear worthwhile to set out what would be regarded as the major contributions of this study. For these reasons, our main findings are:

The system of indirect addilog Engel functions gives reasonable fit for the kind of data used (see Table 15).

Effects due to different levels of factors like area and occupation are reasonably significant for average commodity expenditures (tests between means) but not for the slopes of commodity expenditures with respect to total expenditure (the β_i's).

Area effects are reasonably explained by occupational composition of areas (sections 4.6, 5.2 and 5.3).

The kind of variation of estimates of β_i as family size increases, confirms standard thinking about economies of scale particularly for food (sections 5.1, 5.3 and 5.5).

Family composition cannot be meaningfully incorporated into the present analysis because of broad grouping of commodities which makes individual and collective needs of the households inseparable (section 5.4).

Erratic estimates of budget study parameters in underdeveloped countries are due mainly to zero expenditures on certain commodities even though these commodities are consumed not at zero level, e.g. rent-free accommodation (section 5.3).

Superiority of two-way classified data over one-way classification is confirmed (Chapter 5).

And, last but by no means least, with the aid of the Leser-Frisch approach to the assumptions of substitutability or complementarity between broadly grouped commodities, some useful estimates of income and price elasticities of demand are made (Chapter 6).

The conclusion that different levels of area or occupation do not significantly affect β_1 is very important for economic policies. As a result of this, common estimates of β_1 based on data for all low income earners are very valuable. Such estimates are more stable because of the increase in sample size and they form the main numerical results of this thesis (see tables 15 and 27). In terms of what policy-makers are interested in, resulting estimates discussed in Chapter 6, based on the joint utilisation of Leser (1941-42) and Frisch (1959) are important. Estimates of income and price elasticities form basic and valuable economic information, which most underdeveloped countries need very much for planning the growth of their economies. This means that the underdeveloped countries do not have to wait for years, to allow data series of their economies to be long enough before making such estimates. Even though this approach cannot be a substitute for those based on more adequate data like combined use of cross-section and time series data, it will give reasonable guideline pending the availability of richer statistical materials. The emphasis of this study is in deriving numerical results but this can only be carried out by using appropriate statistical methods. A part of Chapter 3 is a synthesis of statistical methods from different sources. The numerical results show the reasonableness of the statistical methods. An important aspect of this study is the comparison with other studies. The studies compared seem reasonably consistent. The main differences are due to broad climatic conditions and stages of development of the various countries.

The form of data used is extremely important in family budget studies, particularly if the data are grouped as is generally necessary. But the benefit of grouping cannot be achieved if it is not appropriately carried out. Income and family size, which are usually highly correlated, seem to be the most important factors. Because of the high correlation between them, all data, when grouped, must be cross-tabulated by the two factors. The consideration of further classification, e.g. according to area or occupation must be based on the size of available data. Grouping of data by income according to area using broad occupational classifications as is the practice of the Federal Office of Statistics should be discouraged.

The major problem encountered in this study concerns the data and this is about the presence of many zero observations of variables. For grouped data, in an ideal situation, this should not happen or at the worst, should be reduced to a minimum. It could happen at the lower income intervals but not at the upper ones. It can occur for disaggregated commodities like babyfood, sweets etc., but not for commodities broadly classified like food and accommodation. When data are grouped, it is assumed that zero observations occurring for a few households will be counteracted by a large number of other households with non-zero observations. In averaging over both kinds of households, zero observations are eliminated. Unfortunately this was not the case in some cells of the data used. A possible source is the sample size of the surveys which is quite small. This means that very few households may fall in some income-size cells while in some, none at all. Another reason is the effect of the stage of the economy. The more developed a country, the less the influence of culture complex. In Nigerian situation, a good example of its impact is shown in the case of accommodation, where there is a high proportion of households living rent free as discussed in section 4.6. As revealed in a study [Adamu (1970)] of the distribution of zero expenditures among a sample of households, concerning a selection of commodities including accommodation, there is a high connection between the degree of urbanisation, the proportion of the population who are earners and the proportion of households paying rent. The more urbanised, the higher the proportion employed compared with self-employed and the higher the proportion paying rents. This shows that the elimination of the problem of zero expenditure is a matter of time; the problem is likely to be reduced, or to be totally eliminated, as the economy becomes more monetized and there is rapid increase in the service (tertiary) sector of the economy.

The pioneering nature of this study, as far as Nigerian conditions are concerned, has led to some digressions to deal with such related topics as data collection. As a result of this, many openings for further research become evident. Only three broad areas are considered. These seem to encompass most of what is covered in Chapters 2, 3 and 4. There is need to increase the sample size of households covered in Nigerian sample surveys, to improve the quality and quantity of information so extracted and to use appropriate methods of grouping. There is need for National Surveys, not surveys in pocket areas with no view to national estimates. The national surveys should make possible the feasibility of providing both local (rural and urban) and national estimates. These data, if grouped for publication or storage, should be cross-tabulated by income and family size. There is no reason why such national surveys should not be on a continuous basis. These forms of data are important for the country if basic economic and social research is to be meaningful. It is hoped that studies of the kind covered in this thesis will influence the main data collecting agencies in the country to conduct national surveys with the possibility of deriving cross-tabulated results for rural and urban areas. Family budget data collection is a very expensive business which can be handled better by an agency like the Federal Office of Statistics with full co-operation of other agencies at the state levels. In addition, there is also the need for co-operation between the producers and the users of statistical information to allow for 'feed back' of ideas and for optimum use of the country's limited resources in data collection. Good budget data are needed before econometric studies of demand can be started.

The second direction for further investigation concerns the economic theory required for the kind of economic analysis used in this study. The study deviates from the traditional practice of Engel function analysis in many ways, as revealed by what is covered in sections 2.3, 2.5 and 3.3. This approach aims at linking theoretical concepts with empirical work in a systematic way. These sections and other related ones need to be re-examined. For example, the analysis is based on an assumption which allows for broad classification of the commodities. There is need to relax this assumption to allow for some amount of disaggregation. For example, food can be regrouped to consist of staples, fat and oil, vegetables and fruits, meat and fish, and other food. This will lead to the more general concept of utility tree [Strotz (1957)]. As evident from Frisch's work [Frisch (1959), p. 178], the present study contains empirical results of a special kind of utility tree, although with only primary branches and no secondary ones. Extension of the work in this form even though not leading to general utility functions (opinions differ whether this is even necessary) will make it possible for more meaningful results to be derived and allowing for better rigorous study of family effects.

The statistical methodology of this study is a very important direction of research. This is the direction of much concern to me. But unfortunately the development of this section depends very much on the other two discussed above. The statistical methods adopted in this study, as discussed in Chapter 3, are relatively simple but are dictated by the available data and the form of model 3.3(6). In a way, a kind of traditional approach to data analysis has been adopted - putting data from the same source in different forms to analyse different specific factors - a kind of experimentation or simulation. Single equation estimating procedure has been used throughout. It is doubtful whether a simultaneous estimating procedure could have done significantly better because of the kind of model specified by 3.3(6) - an exactly identified recursive model. But with the kind of extension which is likely to result from an extended utility tree, as mentioned above, there is need to use some kind of simultaneous estimating procedure. This is not likely to present any problem because of the great advancement in electronic computers with large core storage and high speed.

REFERENCES

ADAMU, S. O. (1970), "Some Aspects of Expenditure Patterns in Nigeria", The Nigerian Journal of Economic and Social Studies, Vol. 12, No. 3, pp. 265-288.

AITKEN, A. C. (1934), "On Least Squares and Linear Combinations of Observations", Proceedings of the Royal Society, Edinburgh, Vol. 55, No. 5, pp. 42-48.

ALLEN, R. G. D. (1942), "Expenditure Patterns in Families of Different Sizes", in Studies in Mathematical Economics and Econometrics, edited by O. Lange, et. al., Chicago, University Press.

ALLEN, R. G. D. (1947), Mathematical Analysis for Economists, London, Macmillan.

ALLEN, R. G. D. and MISHAN, E. J. (1965), "Is the Substitution Term Ambiguous?", Economica, Vol. 32, pp. 215-222.

ALLEN, R. G. D. and BOWLEY, A. L. (1935), Family Expenditure, London, King and Son.

ANTONELLI, G. B. (1886), Teoria mathematica della economica politica, Pisa; English Translation, "On the Mathematical Theory of Political Economy", in J. S. Chipman, et. al. (eds), Studies in the Mathematical Foundations of Utility and Demand, New York, Harcourt, Brace and World.

BARTEN, A. P. (1964,a), "Family Composition, Prices and Expenditure Patterns", in P. E. Hart, et. al. (eds), Econometric Analysis for National Economic Planning, London, Butterworths.

BARTEN, A. P., (1964,b), "Consumer Demand Functions under Conditions of Almost Additive Preferences", Econometrica, Vol. 32, pp. 1-38.

BERNOULLI, D. (1738), "Specimen theoriae novae de mensura sortis", in Commentarii academiae Scientiarum imperialis Petropolitanae; English Translation, "Exposition of a New Theory on the Measurement of Risk", Econometrica, Vol. 22, (1954), pp. 23-26.

BOWLEY, A. L. and BURNETT-HURST, A. R. (1915), Livelihood and Poverty, London, Bell.

BRADY, D. S. (1938), "Variations in Family Living Expenditures", Journal of the American Statistical Association, Vol. 33, pp. 385-389.

BRADY, D. S. and FRIEDMAN, R. D. (1947), "Savings and Income Distribution", in Studies in Income and Wealth, Vol. 10, pp. 247-265.

BUCHANAN, K. M. and PUGH, J. C. (1955), Land and People of Nigeria, London, University of London Press.

COURT, L. M. (1941), "Entrepreneurial and Consumer Demand Theories for Commodity Spectra", Econometrica, Vol. 19, pp. 135-162.

DEBREU, G. (1954), "Representation of a Preference Ordering by a Numerical Function", in R. M. Thrall, et. al. (eds), Decision Processes, New York, Wiley.

DICKINSON, R. E. (1947), City, Region and Regionalism, London, Routledge and Kegan Paul.

DUESENBERRY, J. S. (1949), Income, Savings and the Theory of Consumer Behavior, Cambridge, Mass., Harvard University Press.

DURBIN, J. (1954), "Errors in Variables", Review of the International Statistical Institute, Vol. 22, pp. 23-32.

EDEN, F. M. (1797), The State of the Poor, London, J. Davis.

ENGEL, E. (1857), "Die Productions - und Consumptions - verhältnisse des Königreichs Sachsen", reprinted in International Statistical Institute Bulletin, Vol. 9, Appendix.

EDGEWORTH, F. Y. (1881), Mathematical Psychics, London, Kegan Paul.

FEDERAL OFFICE OF STATISTICS, (1956), Urban Consumer Surveys, Lagos, Lagos, F.O.S.

FEDERAL OFFICE OF STATISTICS, (1963 - 1966), Urban Consumer Surveys, (published separately for Ibadan, Enugu, Kaduna, Onitsha, Oshogbo-Ilelfe-Ilesha, Akure-Ondo-Owo, and Sokoto-Gusau), Lagos, F.O.S.

FEDERAL OFFICE OF STATISTICS, (1966), "Expenditure Elasticities of Demand for Household Goods", Statistical Studies (mimeo), Lagos, F.O.S.

FEDERAL OFFICE OF STATISTICS, (1969), Population Census of Nigeria 1963, Vol. III, Lagos, F.O.S.

FORSYTH, F. G. (1960), "The Relationship between Family Size and Family Expenditure", Journal of the Royal Statistical Society, Vol. 123, Part 4 (Series A), pp. 367-397.

FRIEDMAN, M. (1935), "Professor Pigou's Method for Measuring Elasticities of Demand from Budgetary Data", Quarterly Journal of Economics, Vol. 50, pp. 151-163.

FRIEDMAN, M. (1952), "A Method of Comparing Incomes of Families Differing in Composition", in Studies in Income and Wealth, Vol. 15, pp. 10-24.

FRIEDMAN, M. (1957), A Theory of the Consumption Function, Princeton, University Press.

FRISCH, R. (1932), "Méthodes nouvelles pour mesurer l'utilité marginale", Revue d'économie politique, année 46, pp. 14-28.

FRISCH, R. (1959), "A Complete Scheme for Computing all direct and cross Demand Elasticities in a Model with many Sectors", Econometrica, Vol. 27, pp. 177-196.

HEADS, J. (1958), "Urbanisation and Economic Progress", N.I.S.E.R., 1958 Conference Proceedings, Ibadan, N.I.S.E.R.

HELLEINER G. K. (1966), Peasant Agriculture, Government and Economic Growth in Nigeria, Homewood, Irwin.

HICKS, J. R. (1939), Value and Capital, London, Oxford University Press.

HICKS, J. R. and ALLEN, R. G. D. (1934), "A Reconsideration of the Theory of Value", Economica, Vol. 1, (New Series), pp. 52-76, 196-219.

HOTELLING, H. (1932), "Edgeworth's Taxation Paradox and the Nature of Demand and Supply Functions", Journal of Political Economy, Vol. 40, pp. 577-616.

HOUTHAKKER, H. S. (1957), "An International Comparison of Household Expenditure Patterns", Econometrica, Vol. 25, pp. 532-551.

HOUTHAKKER, H. S. (1960,a), "The Influence of Prices and Incomes on Household Expenditures", Bulletin of the International Statistical Institute, Vol. 37, pp. 9-22.

HOUTHAKKER, H. S. (1960,b), "Additive Preferences", Econometrica, Vol. 28, pp. 244-257.

ICHIMURA, S. (1951), "A Critical Note on the Definition of Related Goods", Review of Economic Studies, Vol. 18, pp. 179-183.

JOHNSTON, J. (1963), Econometric Methods, New York, McGraw-Hill.

KATZNER, D. W. (1970), Static Demand Theory, London, Macmillan.

KEYNES, J. M. (1936), General Theory of Employment, Interest and Money, London, Macmillan.

KING, G. (1696), Natural and Political Observations and Conclusions upon the State and Condition of England in 1696, London, J. Stockdale.

KONUS, A. A. (1939), "On the Theory of Means", Acta Universitatis Asiae Mediae, Series Va, No. 24.

LEONTIEF, W. W. (1947), "Introduction to a Theory of the Internal Structure of Functional Relationships", Econometrica, Vol. 15, pp. 361-373.

LESER, C. E. V.(1941-42),"Family Budget Data and Price Elasticities of Demand", Review of Economic Studies, Vol. 9, pp. 40-57.

LESER, C. E. V. (1960), "Demand Functions for Nine Commodity Groups in Australia", Australian Journal of Statistics, Vol. 2, pp. 102-113.

LESER, C. E. V. (1968), "A Survey of Econometrics", Journal of the Royal Statistical Society, Vol. 131 (Series A), pp. 530-566.

LIVIATAN, N. (1961), "Errors in Variables and Engel Curve Analysis", Econometrica, Vol. 29, No. 3, pp. 336-362.

MABOGUNJE, A. L. (1968), Urbanisation in Nigeria, London, University of London Press.

MALINVAUD, E. (1970), Statistical Methods of Econometrics, Amsterdam, North-Holland.

MARSCHAK, J. (1931), Elastizitat der Nachfrage, Tubingen.

NEUMANN, J. von and MORGENSTERN, O. (1944), The Theory of Games and Economic Behaviour, New York, Wiley.

PARETO, V. (1896), Cours d'economie politique, Lausanne, Rouge.

PARKS, R. W. (1969), "Systems of Demand Equations: An Empirical Comparison of Alternative Functional Forms", Econometrica, Vol. 37, No. 4, pp. 629-650.

PEARCE, I. F. (1964), A Contribution to Demand Analysis, Oxford, Clarendon.

PIGOU, A. C. (1910), "A Method of Determining Numerical Values of Elasticities of Demand", Economic Journal, Vol. 20, pp. 636-640.

POWELL, A. (1966), "A Complete System of Consumer Demand Equations for the Australian Economy fitted by a Model of Additive Preferences", Econometrica, Vol. 34, No. 3, pp. 661-675.

PRAIS, S. J. (1959), "A Comment", Econometrica, Vol. 27, pp. 127-129.

PRAIS, S. J. and AITCHISON, J. (1954), "The Grouping of Observations in Regression Analysis", Review of the International Statistical Institute, Vol. 22, pp. 1-22.

PRAIS, J. S. and HOUTHAKKER, H. S. (1955), The Analysis of Family Budgets, Cambridge, University Press.

QUENOUILLE, M. H. (1966), Introductory Statistics, London, Pergamon.

ROBINSON, J. (1969), The Economics of Imperfect Competition, (2nd ed.), London, Macmillan.

ROWNTREE, B. S. (1901), Poverty: a study of town life, London, Macmillan.

ROY, R. (1943), De l'utilité, contribution à la théorie des choix, Paris, Hermann.

SAMUELSON, P. A. (1947), Foundations of Economic Analysis, Harvard, University Press.

SARGAN, J. D. (1958), "Estimation of Economic Relationships Using Instrumental Variables", Econometrica, Vol. 26, pp. 393-415.

SCHULTZ, H. (1938), The Theory and Measurement of Demand, Chicago, University Press.

SLUTSKY, E. E. (1915), "Sulla teoria del bilancio del consumatore", Giornale degli economisti, pp. 1-26; English Translation, "On the Theory of the Budget of the Consumer", in G. J. Stigler et. al. (eds.), Readings in Price Theory, Homewood, Irwin.

SOMERMEYER, W. H. et. al. (1962), "A Method of Estimating Price and Income Elasticities from Time Series and its Application to Consumers' Expenditure in Netherlands, 1949 - 1959", Statistical Studies, No. 13, pp. 30-52.

SONO, M. (1945), "The Effect of Price Changes on Demand and Supply of Separable Goods", Kokumin Keizar Zasski, pp. 1-51; English Translation, International Economic Review, (1961), Vol. 2, pp. 239-271.

STONE, R., et. al. (1954), The Measurement of Consumers' Expenditure and Behaviour in the U.K. 1920 - 1938, Vol. 1, Cambridge, University Press.

STROTZ, R. H. (1957), "The Empirical Implication of a Utility Tree", Econometrica, Vol. 25, pp. 269-280.

SUMMERS, R. (1959), "A Note on Least Squares Bias in Household Expenditure Analysis", Econometrica, Vol. 27, pp. 121-126.

SYDENSTRICKER, E. and KING, W. L. (1921), "The Measurement of the Economic Status of Families", Quarterly Publication of the American Statistical Association, Vol. 17, pp. 842-857.

THEIL, H. (1951), "Estimates and their Sampling Variance of Parameters of Certain Heteroscedastic Distributions", Review of the International Statistical Institute, Vol. 19, No. 2, pp. 141-147.

TOBIN, J. (1951), "Relative Income, Absolute Income and Savings", in Money, Trade, and Economic Growth, in honour of J. H. Williams, New York, Macmillan.

UZAWA, H. (1960), "Preference and Rational Choice in the Theory of Consumption", in Mathematical Methods in the Social Sciences, edited by Arrow, et. al., Stanford, University Press.

WALRAS, L. (1874), Eléments d'économie politique pure, Lausanne, Carbay; English Translation, Elements of Pure Economics, Homewood, Irwin.

WIRTH, L. (1938), "Urbanism as a way of Life", American Journal of Sociology, Vol. 44, No. 2, pp. 1-24.

WIT, J. W. W. A. and SOMMERMEYER, W. H. (1956), "An Allocation Model", mimeographed report M.14 of the Central Bureau of Statistics, (in Dutch).

WOLD, H. O. A. and JUREEN, L. (1953), Demand Analysis, New York, Wiley.

YOSHIHARA, K. (1969), "Demand Functions: An Application to the Japanese Expenditure Patterns", Econometrica, Vol. 37, No. 2, pp. 257 - 274.

APPENDIX A: HOUSEHOLD BUDGET STUDY PROGRAMS

```
C     HOUSEHOLD BUDGET STUDY PROGRAM 1
C     Y=A+BX
      DIMENSION VAR(19),DATA(4),X1(10,50),XN(10,50),DN(10),CN(10,10),Y(5
     10),X(50),Y1(50),X2(50),Y3(50),S1(50),W1(50),Y2(50)
      WRITE(3,222)
  222 FORMAT(1H1,56HADAMU, ENGEL ELASTICITY BY WEIGHTED LEAST SQUARES ME
     1THOD)
      WRITE(3,22)
   22 FORMAT(1HO,24HFOR DOUBLE LOG. RELATION)
  777 READ(1,4)NPAG
      IF(NPAG)776,790,776
C     LAST CARD OF DATA DECK IS BLANK
  776 READ(1,4)NG,NO,NW,NWW
C     NG IS THE NO.OF GROUPS OF COMMODITIES
C     NO IS THE NO. OF TYPES OF PERSONS
C     NW=0 IF THE DATA IS GROUPED AND 1 IF OTHERWISE
C     NWW=0 IF DATA IS IN FOR PER CAPITA ONLY
      IF(NWW)811,810,811
  811 READ(1,400)(DN(J),J=1,NO)
      DO 33 K=1,NG
   33 READ(1,400)(CN(J,K),J=1,NO)
   12 WRITE(3,321)NPAG
  321 FORMAT(1H1,96X,'PAGE',1X,I4)
  810 READ(1,401)(VAR(I),I=1,19)
  401 FORMAT(19A4)
      WRITE(3,402)(VAR(J),J=1,19)
  800 FORMAT(10F8.3)
  880 FORMAT(2F10.5,F5.2,F5.1,8F5.2)
  402 FORMAT(1HO,10X,19A4)
      WRITE(3,941)
  941 FORMAT(1HO,'PER CAPITA EST.')
      WRITE(3,666)
      READ(1,4)LX
    4 FORMAT(4I4)
  400 FORMAT(10F10.4)
      DO 91 I=1,LX
   91 READ(1,880)Y1(I),Y3(I),S1(I),W1(I),(XN(J,I),J=1,NO)
      READ(1,4)LX
      DO 324 I=1,LX
  324 READ(1,800)(X1(J,I),J=1,NG)
C     COMPUTING INCOME PER CONSUMER UNIT
      KKK=0
  937 LAA=0
      WRITE(3,940)
  940 FORMAT(1HO,'TOTAL EXP. AS INDEP. VARIABLE')
  936 NEXT=0
      NGG=NG+1
      DO 34 K=1,NGG
      IF(K-NGG)325,901,325
  325 IF(NEXT)20,72,20
   72 IF(KKK)13,14,13
```

```
   13 DO 46 IX=1,LX
   46 Y(IX)=0
      DO 49 IX=1,LX
      DO 45 J=1,NO
   45 Y(IX)=Y(IX)+XN(J,IX)*DN(J)
      IF(Y(IX))908,908,907
  908 Y(IX)=S1(IX)
  907 IF(LAA)903,904,903
  904 Y2(IX)=Y1(IX)
      GO TO 49
  903 Y2(IX)=Y3(IX)
   49 Y2(IX)=Y2(IX)/Y(IX)
      GO TO 20
   14 DO 970 IX=1,LX
      IF(S1(IX))35,36,35
   36 S1(IX)=0.0005
   35 IF(LAA)905,906,905
  906 Y2(IX)=Y1(IX)
      GO TO 970
  905 Y2(IX)=Y3(IX)
  970 Y2(IX)=Y2(IX)/S1(IX)
      GO TO 20
  901 DO 902 IX=1,LX
      IF(KKK)902,933,902
  933 Y(IX)=S1(IX)
  902 Y2(IX)=Y3(IX)/Y(IX)
   20 NEXT=NEXT+1
C     COMPUTING SPECIFIC EXPENDITURE PER CONSUMER UNIT
      IF(K-NGG)909,910,909
  909 IF(KKK)2,37,2
    2 DO 47 I=1,LX
   47 X(I)=0
      DO 50 IX=1,LX
      DO 48 J=1,NO
   48 X(IX)=X(IX)+CN(J,K)*XN(J,IX)
      IF(X(IX))50,39,50
   39 X(IX)=S1(IX)
   50 X2(IX)=X1(K,IX)/X(IX)
      GO TO 21
   37 DO 16 IX=1,LX
      IF(S1(IX))16,17,16
   17 S1(IX)=0.0005
   16 X2(IX)=X1(K,IX)/S1(IX)
      GO TO 21
  910 DO 911 IX=1,LX
      IF(KKK)911,913,911
  913 Y(IX)=S1(IX)
  911 X2(IX)=Y1(IX)/Y(IX)
C     COMPUTING THE ELASTICITIES AND R-SQUARED
   21 Z=1
      SUMC=0
      SUMY=0
      SUMX=0
      SUQY=0
      SUQX=0
      SUYX=0
```

```
      DO 25 IX=1,LX
      IF(X2(IX))61,26,61
   26 X2(IX)=0.0005
   61 IF(Y2(IX))62,27,62
   27 Y2(IX)=0.0005
   62 DATA(1)=ALOG(X2(IX))
      DATA(2)=ALOG(Y2(IX))
      IF(NW)571,572,571
  572 W=W1(IX)
      GO TO 747
  571 W=1
  747 SUMC=SUMC+W*Z
      SUMY=SUMY+DATA(1)*W
      SUMX=SUMX+DATA(2)*W
      SUQY=SUQY+DATA(1)*DATA(1)*W
      SUQX=SUQX+DATA(2)*DATA(2)*W
      SUYX=SUYX+DATA(1)*DATA(2)*W
   25 CONTINUE
      VARY=SUQY-(SUMY*SUMY)/SUMC
      VARX=SUQX-(SUMX*SUMX)/SUMC
      CVYX=SUYX-(SUMY*SUMX)/SUMC
      REGF=CVYX/VARX
      ADVA=VARY-REGF*CVYX
      XLX=LX
      SUMF=XLX-2.0
      DVAR=ADVA/SUMF
      SAPV=SQRT(DVAR/VARX)
      RVAR=1.0-ADVA/VARY
      IF(K-NGG)914,915,914
  915 WRITE(3,916)
  916 FORMAT(1H0,'REGR. OF TOTAL EXP. ON BASIC INCOME')
  914 WRITE(3,122)SUMC,SUMY,SUMX,SUQY,SUQX,SUYX,REGF,SAPV,RVAR
  666 FORMAT(1H0,10HNO OF OBS.,2X,8HSUM OF Y,4X,8HSUM OF X,4X,10HS. OF S
     1Q.Y,2X,10HS. OF SQ.X,2X,10HS. OF PROD,2X,8HEST.OF B,3X,9HS.E. OF B
     2,3X,9HR-SQUARED)
  122 FORMAT(1H0,10(E11.4,1X))
   34 CONTINUE
      IF(LAA)931,932,931
  932 LAA=LAA+1
      NPAG=NPAG+1
      WRITE(3,321)NPAG
      WRITE(3,666)
      WRITE(3,939)
  939 FORMAT(1H0,'BASIC INCOME AS INDEP VARIABLE')
      GO TO 936
  931 IF(KKK)934,935,934
  935 IF(NWW)812,934,812
  812 KKK=KKK+1
      NPAG=NPAG+1
      WRITE(3,321)NPAG
      WRITE(3,950)
  950 FORMAT(1H0,'ADULT EQUIVALENT')
      GO TO 937
  934 GO TO 777
  790 WRITE(3,99)
   99 FORMAT(1H1,'END OF COMPUTATION')
      CALL EXIT
      END
```

```
C     HOUSEHOLD BUDGET STUDY PROGRAM 11
      DIMENSION DN(12),U(24,12),VAR(19),XN(10,50),W1(50)
      COMMON FN,WIN,IX
      WRITE(3,100)
  100 FORMAT(1H1,'COVARIANCE MATRIX AND ITS INVERSE FOR TYPES OF PERSONS
     1')
      READ(1,144)NO,NW
C     NO=NO. OF TYPES OF PERSONS, NW=1 DATA NOT GROUPED
  144 FORMAT(2I4)
      READ(1,4)NPAG
  703 READ(1,400)(VAR(I),I=1,19)
  400 FORMAT(19A4)
    4 FORMAT(I4)
  401 FORMAT(1H1,15X,19A4,11X,'PAGE',I4)
  888 FORMAT(25X,F5.1,8F5.2)
      WRITE(3,401)(VAR(I),I=1,19),NPAG
      NPAG=NPAG+1
      READ(1,4) LX
      IF(LX-1000)13,790,13
   13 WRITE(3,111)LX
  111 FORMAT(1H0,I4,2X,'NO OF GROUPS OF OBSERVATIONS')
      DO 880 I=1,LX
      RHAD(1,888)W1(I),(XN(J,I),J=1,7)
  880 CONTINUS
C     COVARIANCE MATRIX CALCULATION
      SUM = 0
      DO 450 I=1,NO
      DN(I)=0
      DO 450 J=1,NO
  450 U(J,I)=0
      Z=1
      DO 451 IX=1,LX
      IF(NW-1)433,444,433
  433 WIN=W1(IX)
      GO TO 455
  444 WIN=1.0
  455 SUM=SUM+Z*WIN
      DO 451 I=1,NO
      DN(I)=DN(I)+XN(I,IX)*WIN
      DO 451 J=1,NO
  451 U(J,I)=U(J,I)+XN(I,IX)*XN(J,IX)*WIN
      FN=SUM
      DO 452 I=1,NO
      DO 452 J=1,NO
  452 U(J,I)=U(J,I)-DN(I)*DN(J)/FN
      NEXTT=1
      WRITE(3,17)FN
   17 FORMAT(1H0,5X,F10.0,'NO OF OBSERVATION')
      WRITE(3,275)
  275 FORMAT(1H0,30X,9HXX MATRIX)
  454 DO 462 I=1,NO
  462 WRITE(3,463)I,(U(J,I),J=1,NO)
  463 FORMAT(1H0,I1,1X,7E11.4/(1H0,2X,7E11.4))
  344 GO TO(464,703),NEXTT
  464 DO 465 I=1,NO
```

```
      DO 465 J=1,NO
      JJ=J+NO
      IF(I-J)466,499,466
  499 U(JJ,I)=1
      GO TO 465
  466 U(JJ,I)=0
  465 CONTINUE
      KK=NO+NO
      DO 470 I=1,NO
      M=I
      T=ABS(U(I,I))
      KZ=I+1
      IF(KZ-NO)471,471,472
  471 DO 474 J=KZ,NO
      IF(T- ABS(U(I,J)))473,474,474
  473 M=J
      T=ABS(U(I,M))
  474 CONTINUE
C     REPLACE LOWER WITH GREATER ROW
      IF(M-I)472,472,475
  475 DO 476 L=I,KK
      T=U(L,I)
      U(L,I)=U(L,M)
  476 U(L,M)=T
C     COMPUTING ROW WITH ONE
  472 IF(U(I,I))477,478,477
  477 T=1.0/U(I,I)
      M=I-1
      U(I,I)=1
      DO 479 L=KZ,KK
  479 U(L,I)=U(L,I)*T
C     REDUCING UPPER OF ONE TO ZERO
  481 IF(M)482,482,483
  483 T=U(I,M)
      U(I,M)=0
      DO 484 L=KZ,KK
  484 U(L,M)=U(L,M)-U(L,I)*T
      M=M-1
      GO TO 481
C     REDUCING LOWER OF ONE TO ZERO
  482 IF(KZ-NO)486,486,470
  486 DO 488 L=KZ,NO
      T=U(I,L)
      U(I,L)=0
      DO 488 M=KZ,KK
  488 U(M,L)=U(M,L)-U(M,I)*T
  470 CONTINUE
      DO 489 I=1,NO
      DO 489 J=1,NO
      K1=NO+J
  489 U(J,I)=U(K1,I)
  492 NEXTT=2
      WRITE(2,493)(VAR(I),I=1,19)
  493 FORMAT(1H0,19A4)
      DO 341 I=1,NO
  341 WRITE(2,463)I,(U(J,I),J=1,NO)
```

```
      GO TO 344
  478 WRITE(3,494)
  494 FORMAT(/16HNO INVERSE ERROR)
      GO TO 703
  790 CALL EXIT
      END
```

```
C     HOUSEHOLD BUDGET STUDY PROGRAM 111
C     COMPUTING SPECIFIC AND INCOME SCALES, AND ELASTICITIES BASED ON
C     EQUIVALENT ADULTS BY ITERATIVE METHOD USING LEAST SQUARES
      DIMENSION RE(10),REN(10),RV(10),RVN(10),U(10,10),X1(10,50),XN(8,50
     1),DATA(5),C(10,10),D(10),BN(10),W(10),Y3(50),SAP(10),Y(50),X(50),C
     2N(10,10),DN(10),SA(10),VAR(19),Y1(50),S1(50),W1(50),X2(50),Y2(50),
     3CE(10,10)
      DOUBLE PRECISION X,Y,X2,Y2,C,D
  441 READ(1,4)NPAG
   14 FORMAT(1H1,96X,'PAGE',I4)
      IF(NPAG)102,123,102
C     DECK OF DATA CARDS MUST END WITH A BLANK CARD
  102 READ(1,4)NG,NO,NW
C     NG FOR GROUPS OF COMMODITIES AND NO FOR NO OF TYPES OF PERSONS
C     NW=0 IF THE DATA IS GROUPED AND NW=1 IF OTHERWISE
      DO 7 I=1,NO
    7 READ(1,886)(U(J,I),J=1,NO)
      READ(1,400)(VAR(I),I=1,19)
  400 FORMAT(19A4)
      WRITE(3,402)(VAR(J),J=1,19),NPAG
  402 FORMAT(1H1,10X,19A4,10X,'PAGE',I4)
      WRITE(3,243)
  243 FORMAT(1H0,'DOUBLE-LOG. ENGEL FUNCTION')
      READ(1,4)LX
      DO 9 I=1,LX
    9 READ(1,880)Y1(I),Y3(I),S1(I),W1(I),(XN(J,I),J=1,NO)
      READ(1,4)LX
      DO 343 I=1,LX
  343 READ(1,888)(X1(J,I),J=1,NG)
  886 FORMAT(9X,8F9.4)
  888 FORMAT(10F8.3)
      LLL=0
      LAA=1
      NEXIT=0
      GO TO 870
  408 LL=0
  333 KK=0
      DO 3 K=1,NG
      DO 3 J=1,NO
      C(K,J)=0
    3 CONTINUE
      NEXT=0
      DO 999 K=1,NG
      IF(NEXT)588,6,588
    6 IF(LL)853,854,853
  854 FN=0
      SY=0
      DO 8 IX=1,LX
      IF(NW)616,8,616
  616 W1(IX)=1
    8 SY=SY+Y1(IX)*W1(IX)
  853 NEXT=NEXT+1
  880 FORMAT(2F10.5,F5.2,F5.1,7F5.2)
      DO 30 IX=1,LX
      IF(LL)66,67,66
   66 SINN=Y(IX)
```

```
       GO TO 920
    67 SINN=S1(IX)
   920 Y2(IX)=Y1(IX)/SINN
       IF(RE(K))851,852,850
   851 Y2(IX)=Y2(IX)**(-RE(K))
       GO TO 30
   852 Y2(IX)=1.0
       GO TO 30
   850 Y2(IX)=Y2(IX)**(RE(K))
       Y2(IX)=1.0/Y2(IX)
    30 CONTINUE
     4 FORMAT(4I4)
   588 DO 11 I=1,NO
    11 BN(I)=0
       W(K)=0
       DO 99 IX=1,LX
       W(K)=W(K)+X1(K,IX)*W1(IX)
       X2(IX)=X1(K,IX)*Y2(IX)
    99 X2(IX)=X2(IX)*W1(IX)
C      COMPUTING THE SPECIFIC SCALES
       DO 31 J=1,NO
       DO 31 IX=1,LX
    31 BN(J)=BN(J)+XN(J,IX)*X2(IX)
       DO 581 I=1,NO
       DO 581 J=1,NO
   581 C(K,I)=C(K,I)+U(I,J)*BN(J)
       IF(C(K,1))468,403,468
   403 C(K,1)=0.0005
   468 DO 535 I=1,NO
   535 C(K,I)=C(K,I)/C(K,1)
       JKK=0
       DO 908 I=1,NO
       CK=ABS(C(K,I))
       IF(CK-2.0)908,910,910
   910 JKK=JKK+1
   908 CONTINUE
       IF(JKK)909,999,909
   909 DO 912 I=1,NO
   912 C(K,I)=1.0
   999 CONTINUE
C      COMPUTING INCOME SCALES
       DO 33 I=1,NG
    33 W(I)=W(I)/SY
   663 DO 36 I=1,NO
    36 D(I)=0
       DO 37 I=1,NO
       DO 37 K=1,NG
    37 D(I)=D(I)+C(K,I)*W(K)
       IF(D(1))469,404,469
   404 D(1)=0.0005
   469 DO 545 I=1,NO
   545 D(I)=D(I)/D(1)
C      COMPUTING INCOME PER CONSUMER UNIT
   870 NEXT=0
       DO 991 K=1,NG
       IF(LLL)466,405,466
   466 IF(NEXT)71,72,71
```

```
   72 DO 46 IX=1,LX
   46 Y(IX)=0
      DO 49 IX=1,LX
      DO 45 J=1,NO
   45 Y(IX)=Y(IX)+XN(J,IX)*D(J)
      IF(Y(IX))855,855,323
  855 Y(IX)=S1(IX)
  323 Y2(IX)=Y1(IX)/Y(IX)
      IF(Y2(IX))62,27,62
   27 Y2(IX)=0.0005
   62 Y2(IX)=ABS(Y2(IX))
      Y2(IX)=ALOG(Y2(IX))
   49 CONTINUE
   71 NEXT=NEXT+1
C     COMPUTING SPECIFIC EXPENTUREPER CONSUME UNIT
      DO 47 I=1,LX
   47 X(I)=0
      DO 50 IX=1,LX
      DO 48 J=1,NO
   48 X(IX)=X(IX)+C(K,J)*XN(J,IX)
      IF(X(IX))856,856,431
  856 X(IX)=S1(IX)
      DO 906 I=1,NO
  906 C(K,I)=1.0
  431 X2(IX)=X1(K,IX)/X(IX)
      IF(X2(IX))61,26,61
   26 X2(IX)=0.0005
   61 X2(IX)=ABS(X2(IX))
      X2(IX)=ALOG(X2(IX))
   50 CONTINUE
C     COMPUTING ELASTICITIES AND R-SQUARED
      GO TO 407
C     COMPUTING ESTIMATES BASED ON PER CAPITA
  405 IF(NEXT)411,412,411
  412 WRITE(3,241)
  241 FORMAT(1HO,'ESTIMATES BASED ON PER CAPITA')
      WRITE(3,666)
      WRITE(3,303)
  303 FORMAT(1HO,'SPECIFIC EXP. ON TOTAL EXP.')
      DO 413 IX=1,LX
      Y2(IX)=Y1(IX)/S1(IX)
      IF(Y2(IX))205,206,205
  206 Y2(IX)=0.0005
  205 Y2(IX)=ALOG(Y2(IX))
  413 CONTINUE
      NEXT=NEXT+1
  411 DO 414 IX=1,LX
      X2(IX)=X1(K,IX)/S1(IX)
      IF(X2(IX))211,212,211
  212 X2(IX)=0.0005
  211 X2(IX)=ALOG(X2(IX))
  414 CONTINUE
  407 Z=1
      SUMC=0
      SUMY=0
      SUMX=0
```

```
      SUQY=0
      SUQX=0
      SUYX=0
      DO 25 IX=1,LX
      DATA(1)=X2(IX)
      DATA(2)=Y2(IX)
      IF(NW)757,767,757
  767 WIN=W1(IX)
      GO TO 747
  757 WIN=1
  747 SUMC=SUMC+G*WIN
      SUMY=SUMY+DATA(1)*WIN
      SUMX=SUMX+DATA(2)*WIN
      SUQY=SUQY+DATA(1)*DATA(1)*WIN
      SUQX=SUQX+DATA(2)*DATA(2)*WIN
      SUYX=SUYX+DATA(1)*DATA(2)*WIN
   25 CONTINUE
      VARY=SUQY-(SUMY*SUMY)/SUMC
      VARX=SUQX-(SUMX*SUMX)/SUMC
      CVYX=SUYX-(SUMY*SUMX)/SUMC
      REN(K)=CVYX/VARX
      ADVA=VARY-(CVYX*CVYX)/VARX
      XLX=LX
      SUMF=XLX-2.0
      DVAR=ADVA/SUMF
      SAP(K)=DSQRT(DVAR/VARX)
      RVN(K)=1.0-ADVA/VARY
      IF(NEXTT)872,464,872
  464 IF(LLL)991,872,991
  872 WRITE(3,122)SUMC,SUMY,SUMX,SUQY,SUQX,SUYX,REN(K),SAP(K),RVN(K)
  666 FORMAT(1H0,10HNO OF OBS.,2X,8HSUM OF Y,4X,8hSUM OF X,4X,10HS. OF S
     1Q.Y,2X,10HS. OF SQ.X,2X,10HS. OF PROD,2X,8HEST.OF B,3X,9HS.E. OF B
     2,3X,9HR-SQUARED)
  122 FORMAT(1H0,9(E11.4,1X))
  991 CONTINUE
      IF(LLL)409,416,409
  416 DO 415 K=1,NG
      RE(K)=REN(K)
  415 RV(K)=RVN(K)
      LLL=LLL+1
      GO TO 408
  409 IF(NEXIT)355,381,355
  355 NPAG=NPAG+1
      WRITE(3,14)NPAG
      GO TO 366
  381 IF(LL)874,89,874
   89 DO 476 J=1,NO
      DO 476 K=1,NG
  476 CN(K,J)=1.0
  874 DO 992 K=1,NG
      IF(RV(K)-RVN(K))878,595,595
  878 RV(K)=RVN(K)
      RE(K)=REN(K)
      DO 876 J=1,NO
  876 CN(K,J)=C(K,J)
      GO TO 992
  595 KK=KK+1
  992 CONTINUE
```

```
      NPAG=NPAG+1
      WRITE(3,14)NPAG
      WRITE(3,32)KK,LL
   32 FORMAT(1HO,'NO. OF NO CHANGE IS',I4,2X,'NO. OFROUNDS IS',I4)
      WRITE(3,661)
  661 FORMAT(1HO,'THE ELASTICITIES')
      WRITE(3,662)(REN(K),K=1,NG)
      WRITE(3,667)
  667 FORMAT(1HO,'THE STANDARD ERROR FOR CORRESPONDING ELASTICITIES')
      WRITE(3,662)(SAP(K),K=1,NG)
      WRITE(3,664)
  664 FORMAT(1HO,'THE R-SQUARED')
      WRITE(3,662)(RVN(K),K=1,NG)
  662 FORMAT(1HO,12(2X,F8.4))
  366 WRITE(3,665)
  665 FORMAT(1HO,'INCOME SCALES')
      WRITE(3,662)(D(J),J=1,NO)
      DO 669 K=1,NG
      WADA=W(K)*(-REN(K))
      WRITE(3,668)K,W(K),WADA
  668 FORMAT(1HO,'SPECIFIC SCALES FOR GROUP',I4,2X,'ITS PROPORTIONATE EX
     1PENDITURE=',F5.3,'AND ITS PRICE ELASTICITY=',F5.3)
  669 WRITE(3,662)(C(K,J),J=1,NO)
      IF(NEXIT)441,356,441
  356 IF(KK-NG)77,79,79
   77 IF(LL-20)455,79,79
  455 LL=LL+1
      GO TO 333
   79 NEXIT=NEXIT+1
      NPAG=NPAG+1
      WRITE(3,14)NPAG
      WRITE(3,242)
  242 FORMAT(1HO,'EST. BASED ON A.E. USING TOTAL EXP. AS INDEP')
      WRITE(3,666)
      DO 456 J=1,NO
      DO 456 K=1,NG
  456 C(K,J)=CN(K,J)
      GO TO 663
  123 WRITE(3,341)
  341 FORMAT(1H1,'END OF COMPUTATIONS')
      CALL EXIT
      END
```

APPENDIX B: 2-WAY CLASSIFIED DATA

KEYS TO THE TABLES

f = NUMBER OF HOUSEHOLDS

x = BASIC INCOME

z = TOTAL EXPENDITURE

S_1 = MALES AGED 18 YEARS AND OVER

S_2 = FEMALES " " " "

S_3 = MALES AGED 14 - 17 YEARS

S_4 = FEMALES " " " "

S_5 = MALES AGED 5 - 13 YEARS

S_6 = FEMALES " " " "

S_7 = CHILDREN UNDER 5 YEARS

y_1 = ACCOMMODATION

y_2 = FOOD

y_3 = DRINKS, TOBACCO AND KOLA

y_4 = FUEL AND LIGHT

y_5 = TRANSPORT

y_6 = CLOTHING AND FOOTWEAR

y_7 = HOUSEHOLD GOODS

y_8 = OTHER PURCHASES

y_9 = OTHER SERVICES

TABLE 1: CLERKS

INCOME →	Under 150/-																	
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	6	95.1	81.8	0.83	0.17	0.00	0.00	0.00	0.00	0.00	4.0	33.2	6.3	2.4	5.5	16.3	1.3	7.8	5.0
2	3	130.8	105.7	1.00	0.00	0.333	0.00	0.333	0.333	0.00	3.3	56.7	23.8	4.0	2.5	0.0	0.3	7.5	7.6
3	7	110.0	137.8	1.14	0.86	0.00	0.14	0.29	0.00	0.57	21.0	84.9	13.4	5.1	2.4	2.1	1.2	5.2	2.5
4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6	2	138.0	129.1	1.50	1.50	1.00	0.00	0.50	0.50	1.00	0.0	59.9	25.3	3.6	19.7	1.5	3.1	7.2	8.8

INCOME →	150/- to 249/11d																	
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	51	181.3	197.9	0.98	0.02	0.00	0.00	0.00	0.00	0.00	23.4	68.7	15.8	8.0	8.8	31.1	12.0	18.6	11.5
2	33	219.6	232.8	1.12	0.21	0.18	0.03	0.36	0.06	0.00	21.5	92.3	26.9	10.0	13.6	36.0	9.3	12.7	10.5
3	17	200.0	205.9	1.00	0.76	0.06	0.12	0.35	0.18	0.53	13.6	95.0	19.4	7.4	13.6	20.9	21.2	9.5	5.3
4	15	197.6	250.7	1.13	0.93	0.00	0.07	0.60	0.27	1.00	31.6	123.7	15.7	13.6	8.0	20.0	7.3	15.3	15.5
5	9	206.2	238.8	1.22	1.11	0.22	0.22	0.78	0.22	1.22	27.1	150.4	12.7	10.4	13.3	48.5	7.6	14.5	14.3
6	3	221.7	230.9	1.00	2.00	0.33	0.00	0.00	0.67	2.00	43.3	111.9	17.1	6.1	7.5	0.3	3.7	37.7	3.3

INCOME →	250/- to 349/11d																	
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	28	294.5	259.6	0.96	0.04	0.00	0.00	0.00	0.00	0.00	23.1	85.2	29.5	7.5	21.3	28.6	19.6	23.4	21.4
2	32	294.8	257.5	1.03	0.16	0.19	0.13	0.41	0.06	0.00	33.5	101.9	20.8	11.9	18.6	19.8	12.4	15.3	23.3
3	17	280.1	289.6	1.06	0.76	0.12	0.06	0.35	0.00	0.53	25.9	124.7	38.9	10.6	15.6	29.9	15.8	15.9	12.3
4	24	287.2	293.9	1.08	0.88	0.17	0.21	0.38	0.29	0.83	23.8	123.5	25.7	15.2	8.7	54.2	13.7	15.3	13.8
5	12	309.2	241.6	1.25	1.25	0.08	0.25	0.33	0.50	1.33	7.5	111.6	21.4	10.3	10.5	12.5	7.0	11.7	49.1
6	11	323.1	327.6	1.09	1.27	0.09	0.18	0.90	0.45	2.00	6.3	170.4	39.0	13.5	15.1	37.2	7.8	10.7	27.6

INCOME →	350/- to 449/11																	
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	8	370.5	441.6	1.00	0.00	0.00	0.00	0.00	0.00	0.00	30.1	136.0	73.8	8.1	4.0	84.1	11.4	63.0	31.1
2	13	418.2	470.7	1.08	0.00	0.23	0.08	0.46	0.15	0.00	65.0	118.5	60.5	12.3	55.2	74.0	6.5	20.4	58.3
3	11	395.9	286.8	1.27	0.64	0.36	0.00	0.45	0.18	0.09	35.0	97.8	44.3	11.6	30.4	27.5	2.1	16.8	21.3
4	10	418.1	356.0	1.00	0.80	0.60	0.20	0.30	0.30	0.80	53.3	129.7	33.1	16.1	59.9	20.5	7.5	12.8	23.1
5	6	389.3	406.5	0.83	1.00	0.17	0.33	0.51	0.83	1.33	38.3	170.0	52.6	12.5	24.0	64.5	7.7	23.1	13.8
6	2	378.0	490.1	1.00	1.00	0.50	0.00	0.00	1.50	2.00	51.0	120.1	39.4	8.7	163.5	39.5	3.0	4.5	60.4

INCOME →	450/- to 549/11																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	6	514.3	392.2	1.00	0.00	0.00	0.00	0.00	0.00	0.00	40.5	103.1	53.8	5.7	76.2	12.2	3.0	16.2	81.5
2	6	489.2	308.4	1.00	0.33	0.00	0.17	0.33	0.17	0.00	42.3	99.4	37.4	8.3	38.1	9.2	14.4	18.0	41.3
3	5	502.3	406.3	0.80	0.80	0.20	0.00	0.40	0.40	0.40	91.3	119.7	51.3	10.5	15.6	81.7	18.9	10.9	6.4
4	1	516.0	894.8	1.00	1.00	0.00	0.00	1.00	0.00	1.00	118.5	209.2	37.6	22.6	32.5	49.8	13.5	161.1	250.0
5	5	492.5	381.2	1.00	1.20	0.20	0.20	0.40	0.80	1.20	35.0	116.8	51.4	16.1	11.0	100.8	5.8	28.0	16.3
6	4	510.0	505.3	1.00	2.00	0.00	0.25	0.25	2.00	0.50	39.3	191.9	58.3	24.4	29.5	48.3	4.1	75.8	33.7

INCOME →	550/- to 649/11d																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	4	582.1	344.0	1.00	0.50	0.00	0.00	0.25	0.25	0.00	38.0	106.1	32.8	16.3	24.1	86.8	12.7	9.4	17.8
3	3	620.0	332.5	1.67	0.33	0.33	0.33	0.33	0.00	0.00	30.0	165.0	11.0	25.6	6.8	26.8	9.6	39.7	18.0
4	3	580.0	452.5	1.00	1.00	0.00	0.00	0.67	0.33	1.00	86.7	152.4	41.1	15.1	21.1	10.1	98.3	18.4	9.3
5	4	572.1	480.9	1.00	2.00	0.50	0.00	0.00	0.50	1.00	30.8	159.2	15.6	29.4	68.5	100.6	10.1	21.9	44.8
6	8	606.9	485.0	1.38	1.13	0.13	0.13	0.63	1.75	0.88	45.5	194.1	49.2	23.6	26.4	33.8	24.6	37.7	50.1

INCOME →	650/- and over																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	2	815.0	375.5	0.50	0.50	0.00	0.00	0.00	0.00	0.00	22.5	186.7	31.3	28.1	18.4	15.0	59.5	11.5	0.5
2	3	710.4	392.1	1.33	0.33	0.33	0.00	0.00	0.00	0.00	26.7	105.1	89.0	10.5	7.7	36.8	93.1	22.2	1.0
3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	1	671.0	699.5	1.00	0.00	1.00	0.00	0.00	1.00	1.00	160.0	155.7	35.8	31.5	65.0	153.0	0.0	29.0	69.5
5	1	685.0	778.2	1.00	1.00	0.00	0.00	1.00	0.00	2.00	0.0	247.9	159.1	28.8	109.8	135.0	7.3	30.3	60.0
6	3	690.0	552.5	1.33	0.67	0.33	0.33	0.67	1.33	1.33	56.7	221.8	123.2	16.2	36.0	4.2	10.0	61.6	22.8

TABLE 2: LABOURERS

INCOME →	Under 150/-																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	42	106.3	122.4	0.90	0.05	0.05	0.0	0.0	0.0	0.0	16.4	60.6	14.3	6.4	4.6	9.5	3.6	3.6	3.4
2	26	105.5	128.3	0.92	0.54	0.12	0.08	0.19	0.08	0.04	12.6	81.0	7.5	5.6	2.6	9.6	2.5	2.0	4.9
3	22	101.5	164.4	1.05	0.82	0.19	0.0	0.32	0.14	0.50	10.1	99.6	12.9	9.9	6.0	8.1	9.3	4.6	3.9
4	22	111.3	159.1	1.05	1.14	0.05	0.09	0.68	0.36	0.68	6.3	95.3	8.6	8.9	5.1	19.0	3.7	4.1	8.1
5	11	114.9	144.6	1.27	1.00	0.09	0.18	0.73	0.64	1.09	3.9	90.2	10.1	12.4	11.2	7.4	1.8	1.6	6.1
6	14	119.1	135.6	1.78	1.14	0.14	0.0	1.36	0.79	0.79	2.4	102.9	13.7	7.3	1.7	1.7	1.7	1.1	4.8

INCOME →	150/- to 249/11d.																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	24	185.9	203.8	1.00	0.0	0.0	0.0	0.0	0.0	0.0	23.0	84.5	20.8	9.0	10.3	29.0	6.0	6.7	14.5
2	13	169.1	168.1	1.00	0.62	0.15	0.0	0.15	0.08	0.0	14.7	83.9	13.3	8.0	8.2	20.5	4.1	4.9	10.5
3	22	193.6	198.5	1.27	0.73	0.05	0.14	0.09	0.41	0.32	20.8	112.3	16.5	10.6	7.0	16.0	6.5	5.6	3.2
4	23	189.8	211.5	1.13	1.09	0.09	0.13	0.39	0.22	1.00	15.6	115.7	19.8	11.8	13.3	14.5	5.5	6.8	8.5
5	15	200.4	273.9	1.07	1.00	0.13	0.20	0.60	0.73	1.27	15.3	122.6	20.1	15.1	6.6	41.0	25.7	20.9	6.6
6	17	198.3	229.2	1.24	1.47	0.24	0.24	1.12	0.53	1.06	10.1	137.4	25.6	14.8	5.1	16.8	3.1	5.3	11.0

INCOME →	250/- to 349/11d																	
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				s_1	s_2	s_3	s_4	s_5	s_6	s_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	1	275.0	161.3	1.00	0.0	0.0	0.0	0.0	0.0	0.0	15.0	104.3	26.2	11.1	0.0	0.0	0.8	3.9	0.0
2	6	274.4	235.5	1.00	0.67	0.17	0.0	0.17	0.0	0.0	20.0	108.0	30.4	10.3	10.7	45.1	3.5	3.3	4.2
3	3	298.6	170.6	1.00	1.00	0.0	0.0	0.0	0.33	0.67	27.4	90.9	10.9	10.7	1.3	12.9	13.3	3.2	0.0
4	6	282.9	204.5	1.17	1.00	0.17	0.0	0.50	0.33	0.83	4.7	113.4	12.5	15.6	16.2	22.2	9.2	5.8	4.9
5	8	288.5	272.5	1.13	1.00	0.25	0.13	0.50	0.38	1.63	29.9	127.4	36.1	18.4	22.5	21.4	6.5	4.4	5.9
6	3	290.7	374.8	1.00	1.67	0.33	0.0	2.00	0.0	1.00	25.1	200.6	21.2	26.8	12.7	22.0	9.8	21.9	34.7

INCOME →	350/- to 449/11d																	
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				s_1	s_2	s_3	s_4	s_5	s_6	s_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	1	360.0	250.7	1.00	0.0	0.0	0.0	0.0	0.0	0.0	37.0	112.4	24.5	0.0	34.0	0.0	0.0	2.5	40.3
2	1	355.5	471.9	1.00	1.00	0.0	0.0	0.0	0.0	0.0	35.0	247.2	58.0	24.5	0.0	70.0	3.8	11.3	22.1
3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5	1	375.0	161.9	1.00	2.00	0.0	0.0	2.00	0.0	0.0	60.0	86.1	0.0	15.8	0.0	0.0	0.0	0.0	0.0
6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

INCOME →	450/- to 549/11d																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				s_1	s_2	s_3	s_4	s_5	s_6	s_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	2	522.7	450.5	1.00	0.0	0.0	0.0	0.0	0.0	0.0	119.2	136.7	82.4	22.4	-	20.0	4.4	25.8	39.6
2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	1	450.0	714.0	1.00	1.00	0.0	0.0	1.00	0.0	1.00	75.0	126.8	34.0	9.7	14.5	171.0	23.5	4.8	0.0
5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

INCOME →	550/- to 649/11d																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				s_1	s_2	s_3	s_4	s_5	s_6	s_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6	1	600.0	543.4	1.00	1.00	0.0	0.0	1.00	1.00	2.00	40.0	179.7	52.4	25.3	12.8	82.0	18.4	28.8	104.0

TABLE 3: ARTISANS

INCOME →	Under 150/-																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	65	102.1	114.5	0.98	0.0	0.02	0.0	0.0	0.0	0.0	9.9	57.5	18.5	6.7	3.2	8.3	3.1	3.2	4.2
2	34	92.2	108.0	1.09	0.35	0.18	0.06	0.15	0.15	0.0	11.1	57.5	11.1	6.9	4.2	7.9	2.4	4.3	2.2
3	46	109.2	132.7	0.98	0.85	0.13	0.07	0.28	0.09	0.57	8.8	79.1	11.9	7.6	3.8	9.3	3.0	4.3	4.9
4	41	113.8	130.2	1.07	1.07	0.07	0.05	0.51	0.24	0.98	6.3	89.2	10.1	9.3	3.8	4.7	1.4	2.5	2.9
5	25	112.3	139.5	1.08	1.32	0.12	0.20	0.60	0.68	1.00	11.2	93.0	12.8	6.0	5.2	4.6	2.1	2.8	1.8
6	16	146.0	187.8	1.31	1.63	0.19	0.19	0.94	0.94	0.81	8.8	129.1	11.4	10.0	3.9	12.0	1.9	4.7	6.0

INCOME →	150/- to 249/11d																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	51	201.8	183.8	0.98	0.0	0.02	0.0	0.0	0.0	0.0	25.1	82.6	24.8	8.0	7.9	13.8	6.8	9.8	5.0
2	36	193.0	212.1	1.08	0.19	0.28	0.0	0.28	0.08	0.06	27.0	92.6	24.4	9.0	8.0	19.1	10.9	9.3	11.8
3	39	185.8	192.7	1.10	0.80	0.18	0.10	0.18	0.21	0.44	17.0	96.7	21.0	11.7	10.8	15.6	4.2	9.1	6.6
4	40	197.7	235.6	1.10	0.95	0.20	0.10	0.50	0.35	0.78	19.1	121.9	27.5	13.8	6.6	25.0	5.6	7.7	8.4
5	42	179.3	211.6	1.07	1.29	0.19	0.05	0.81	0.67	0.86	15.7	119.2	17.3	11.0	6.4	13.2	4.6	5.6	18.6
6	20	179.4	187.3	1.00	1.40	0.60	0.10	0.90	0.80	1.20	12.4	109.8	21.8	11.1	3.8	10.2	6.9	6.0	5.3

INCOME →	250/- to 349/11d																	
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	23	253.5	223.5	0.96	0.0	0.04	0.0	0.0	0.0	0.0	26.0	100.6	22.0	8.4	9.6	26.9	8.5	12.6	8.9
2	12	291.7	240.9	1.08	0.33	0.17	0.0	0.25	0.17	0.0	37.9	106.3	25.6	11.6	8.2	10.8	6.1	18.9	15.5
3	22	296.0	294.6	1.14	0.64	0.23	0.23	0.36	0.18	0.32	24.8	137.8	30.5	10.2	14.83	40.8	8.8	8.1	18.8
4	20	280.1	277.3	1.05	1.05	0.10	0.05	0.20	0.30	1.15	26.4	125.4	31.5	12.2	14.6	30.7	8.9	15.2	12.4
5	16	287.4	297.8	1.06	0.94	0.06	0.13	0.56	0.63	1.63	32.6	126.3	30.5	10.2	9.7	34.3	26.7	13.9	13.6
6	13	256.2	337.0	1.15	1.31	0.23	0.08	1.00	0.62	1.62	17.9	163.9	41.5	13.7	19.5	26.2	8.1	24.5	21.7

INCOME →	350/- to 449/11d																	
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	5	391.1	406.7	0.83	0.17	0.0	0.0	0.0	0.0	0.0	30.0	92.6	52.4	8.6	52.0	48.4	23.6	76.5	22.6
2	11	355.8	529.6	1.09	0.54	0.09	0.09	0.18	0.0	0.0	43.1	123.9	50.6	17.5	107.7	79.5	27.2	58.4	21.7
3	2	440.9	566.2	1.50	0.50	0.0	0.00	0.50	0.50	0.0	47.0	160.5	61.5	30.2	0.0	126.1	44.3	57.9	38.7
4	9	393.9	357.3	1.11	0.78	0.0	0.0	0.78	0.0	1.33	28.9	157.3	46.3	13.3	31.2	43.0	10.4	9.5	17.4
5	6	382.6	357.1	1.33	1.50	0.0	0.0	0.67	1.00	0.50	63.4	164.7	60.8	13.0	12.3	20.9	2.5	12.6	6.9
6	6	380.1	316.6	1.17	1.00	0.33	0.0	1.17	1.00	1.33	32.1	146.8	20.7	16.6	18.5	68.3	5.0	7.6	1.0

INCOME →	450/- to 549/11d																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				s_1	s_2	s_3	s_4	s_5	s_6	s_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	2	510.0	439.7	1.00	0.0	0.0	0.0	0.0	0.0	0.0	70.0	123.4	67.7	13.9	43.0	100.8	11.9	4.0	5.0
2	5	477.2	399.9	1.20	0.40	0.0	0.0	0.40	0.0	0.0	45.2	171.1	44.4	16.1	34.8	57.6	3.2	14.4	13.1
3	3	498.9	355.8	1.33	0.33	1.00	0.0	0.0	0.33	0.0	43.3	134.1	18.5	16.3	18.5	64.8	32.2	23.5	4.6
4	1	500.0	591.8	1.00	2.00	0.0	0.0	0.0	0.0	1.00	0.0	166.5	102.4	23.7	171.0	71.5	2.5	25.9	28.3
5	1	500.0	594.4	1.00	3.00	0.0	0.0	0.0	1.00	0.0	20.0	109.3	48.5	65.5	0.0	0.0	0.0	0.0	0.0
6	3	506.3	287.3	1.00	1.00	0.0	0.33	0.67	1.33	1.67	30.0	109.3	56.4	12.3	11.0	58.6	0.8	6.6	2.3

214

INCOME →	550/- to 649/11d																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				s_1	s_2	s_3	s_4	s_5	s_6	s_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	1	600.0	464.5	1.00	0.0	0.0	1.00	0.0	0.0	0.0	42.8	114.2	159.2	0.1	88.5	36.0	2.0	0.7	21.0
3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	3	596.9	306.0	1.00	1.00	0.33	0.0	0.33	0.33	1.00	55.7	156.2	23.4	1.9	14.5	58.5	4.7	12.8	1.7
5	2	624.1	614.4	1.00	1.00	0.50	0.50	1.00	0.50	0.50	80.0	309.0	85.7	18.2	0.0	66.5	8.0	13.0	34.0
6	4	620.2	441.5	1.25	1.25	0.0	0.0	0.75	0.50	2.25	33.0	155.2	53.0	16.1	35.0	82.8	11.0	12.4	43.0

INCOME →	650/- and over																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	1	720.5	460.4	3.00	0.0	1.00	0.0	0.0	0.0	0.0	39.0	77.2	79.9	116.5	18.0	48.0	88.5	0.0	8.5
5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6	-	736.2	612.7	1.00	1.00	0.0	0.0	1.00	1.50	1.50	36.0	316.4	83.7	19.6	3.7	96.3	27.6	16.5	10.9

TABLE 4: TRADERS

INCOME →	Under 150/-																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	37	92.4	119.1	0.78	0.22	0.0	0.0	0.0	0.0	0.0	10.2	70.2	11.7	5.2	4.1	8.4	2.4	2.0	4.9
2	27	93.3	152.1	0.92	0.30	0.19	0.04	0.37	0.07	0.11	13.9	82.9	15.8	6.9	6.9	5.2	1.5	4.1	14.9
3	18	97.7	131.7	1.00	0.61	0.17	0.32	0.39	0.22	0.28	12.1	78.9	10.1	8.9	3.4	5.7	3.3	3.4	5.9
4	26	100.3	160.2	1.15	0.89	0.27	0.20	0.42	0.15	0.92	14.3	87.5	15.8	9.9	5.8	10.0	2.5	3.6	10.8
5	23	120.2	189.5	1.00	1.17	0.09	0.0	0.87	0.78	1.09	8.1	119.2	15.0	13.7	7.3	7.6	3.2	3.5	11.9
6	11	124.2	211.5	1.18	1.09	0.09	0.18	1.27	1.00	1.18	21.8	118.4	17.4	11.2	7.3	8.2	11.0	4.9	9.0

INCOME →	150/- to 249/11d																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	13	182.1	163.1	0.85	0.15	0.0	0.0	0.0	0.0	0.0	17.0	81.8	15.1	5.3	9.3	22.2	4.3	4.8	3.3
2	10	201.8	216.9	1.00	0.30	0.20	0.20	0.20	0.10	0.0	21.7	109.9	32.0	15.5	6.0	7.1	6.4	6.1	12.2
3	18	179.9	188.4	0.94	0.83	0.17	0.06	0.39	0.33	0.28	12.8	105.8	18.8	11.5	13.0	11.5	2.1	7.0	5.9
4	21	178.7	241.4	1.05	1.05	0.24	0.05	0.48	0.33	0.81	14.3	109.4	24.5	29.7	7.0	19.4	9.6	15.3	12.2
5	28	169.5	209.3	1.07	1.21	0.21	0.21	0.86	0.64	0.89	11.8	127.1	12.7	16.3	5.5	18.6	2.9	7.9	6.6
6	17	180.8	227.3	1.06	1.29	0.35	0.18	1.00	0.82	1.29	11.8	136.7	20.5	13.5	7.3	19.4	2.9	7.4	7.8

INCOME →	250/- to 349/11d																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	2	271.3	213.3	1.00	0.0	0.0	0.0	0.0	0.0	0.0	1.5	62.3	9.0	3.6	13.7	88.4	26.2	7.4	1.2
2	9	320.4	349.4	1.33	0.22	0.44	0.0	0.0	0.0	0.0	37.9	93.9	36.2	12.0	12.5	23.1	3.0	66.3	64.5
3	4	280.0	248.3	1.50	1.00	0.0	0.0	0.0	0.50	0.0	23.0	110.6	39.0	5.6	14.7	24.5	5.0	15.1	10.8
4	4	292.3	311.4	1.00	1.00	0.50	0.0	0.50	0.25	0.75	18.1	126.7	31.2	25.0	9.5	59.8	9.0	17.5	14.6
5	4	324.3	263.0	1.25	1.00	0.25	0.0	0.50	0.75	1.25	13.6	129.9	28.0	14.0	31.9	0.8	2.5	18.5	23.8
6	6	279.4	305.7	1.11	1.56	0.0	0.44	1.11	1.00	0.78	4.7	140.2	23.5	24.9	28.4	13.6	6.1	12.4	51.9

INCOME →	350/- to 449/11d																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	1	382.0	397.9	1.00	0.00	0.0	0.0	0.0	0.0	0.0	25.0	152.2	1.7	9.0	11.0	181.0	0.0	18.0	0.0
2	3	402.9	280.1	1.67	0.0	0.0	0.0	0.33	0.0	0.0	20.8	73.9	25.0	7.2	39.0	97.7	3.5	4.4	8.6
3	3	377.8	335.1	1.33	0.67	0.33	0.0	0.33	0.0	0.33	23.7	140.6	26.8	17.1	36.3	34.1	0.5	14.3	41.7
4	3	393.8	308.8	1.33	1.00	0.0	0.0	0.33	0.33	1.00	20.0	128.6	71.7	9.9	8.8	34.5	8.3	12.4	14.6
5	4	398.6	405.8	1.00	0.50	0.50	0.0	0.25	0.25	1.25	25.6	172.8	46.0	22.0	17.2	16.6	8.7	45.9	51.0
6	4	398.8	269.5	1.00	1.00	0.25	0.0	1.00	1.00	1.75	31.3	165.5	33.7	0.8	7.6	12.0	3.6	7.1	7.9

INCOME →	450/- to 549/11d																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5	1	545.0	571.4	2.00	1.00	1.00	1.00	0.0	0.0	0.0	40.0	280.2	46.6	25.3	4.0	25.0	34.5	1.8	150.0
6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

INCOME →	550/- to 649/11d																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	2	588.0	475.7	0.50	0.50	0.50	0.0	0.0	0.50	0.0	42.5	70.2	31.5	32.3	4.0	15.2	115.2	76.0	88.8
3	1	600.0	284.8	1.00	0.0	1.00	1.00	0.0	0.0	0.0	62.0	98.8	16.0	10.7	39.0	0.0	15.0	23.0	20.3
4	3	603.0	526.4	1.67	0.67	0.67	0.0	0.67	0.0	0.33	43.3	178.2	197.0	29.8	16.8	20.0	11.8	17.3	12.2
5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

INCOME →	650/- and over																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5	-	800.0	569.7	2.50	0.50	1.00	0.0	0.0	0.0	1.00	52.3	205.1	61.2	20.5	22.2	36.5	127.0	8.2	36.7
6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

TABLE 5: FARMERS AND OTHERS

INCOME →	Under 150/-																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	21	81.3	87.5	1.00	0.0	0.0	0.0	0.0	0.0	0.0	4.0	58.4	8.2	3.8	1.5	8.0	0.7	1.5	1.4
2	18	89.5	120.9	1.33	0.39	0.22	0.06	0.0	0.0	0.0	14.8	62.6	15.5	9.8	5.1	6.1	1.3	2.7	3.0
3	28	76.5	115.5	1.00	0.93	0.0	0.11	0.29	0.21	0.57	8.2	68.3	11.3	6.5	2.3	8.4	2.5	2.2	5.7
4	50	88.1	108.6	1.02	1.08	0.16	0.04	0.40	0.42	0.88	6.3	70.5	9.4	9.2	2.6	3.7	1.7	1.9	3.3
5	33	93.6	125.9	1.03	1.21	0.12	0.06	0.94	0.67	0.91	2.3	81.8	12.5	10.5	3.3	5.5	3.6	3.9	2.5
6	29	91.5	113.7	1.17	1.35	0.24	0.24	1.31	1.00	0.69	1.6	75.8	11.9	6.1	3.4	6.3	3.0	2.7	2.9

INCOME →	150/- to 249/11d																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	6	275.1	121.4	1.00	0.0	0.0	0.0	0.0	0.0	0.0	2.5	60.3	15.5	3.6	3.5	18.1	3.2	10.7	4.0
2	7	215.1	204.9	1.00	0.0	0.29	0.14	0.43	0.14	0.0	17.4	67.1	15.2	7.3	7.5	50.6	23.0	10.4	6.4
3	8	178.7	210.5	1.13	0.63	0.25	0.13	0.50	0.13	0.25	16.1	114.6	15.6	8.8	16.4	8.8	2.3	4.4	23.5
4	18	195.1	310.6	1.06	1.17	0.0	0.11	0.56	0.56	0.56	7.8	216.3	23.5	13.6	7.3	11.9	7.8	9.5	12.9
5	14	185.6	176.2	1.14	1.43	0.43	0.07	0.71	0.71	0.57	4.7	100.0	12.4	8.3	9.0	22.8	3.1	6.5	9.4
6	8	161.4	146.3	1.13	1.25	0.38	0.13	1.63	0.63	0.88	0.0	71.7	22.0	5.8	3.4	24.6	1.5	14.7	2.6

INCOME →	250/- to 349/11d																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	4	365.5	498.6	1.00	0.0	0.0	0.0	0.0	0.0	0.0	32.5	96.2	20.3	7.1	15.8	69.3	200.0	11.7	45.7
2	12	264.2	242.7	0.83	0.33	0.17	0.17	0.25	0.25	0.0	24.7	97.9	24.1	6.4	12.8	16.4	7.7	12.1	40.6
3	9	324.0	289.8	1.33	0.56	0.11	0.0	0.22	0.56	0.22	17.8	98.9	18.5	9.3	11.2	47.3	3.8	5.4	78.4
4	7	271.9	310.5	1.43	1.00	0.0	0.0	0.43	0.14	1.00	15.0	139.8	34.8	13.0	7.3	60.8	7.8	20.9	11.1
5	2	286.5	291.5	1.50	1.50	0.0	0.50	1.00	0.0	0.50	37.5	108.3	10.2	10.7	30.3	43.1	12.1	17.8	21.5
6	12	296.0	269.3	1.17	1.17	0.33	0.08	1.50	0.67	1.08	13.4	125.9	20.8	17.7	9.4	34.3	7.2	18.7	21.9

INCOME →	350/- to 449/11d																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	2	380.0	350.0	1.00	0.0	0.0	0.0	0.0	0.0	0.0	30.0	77.9	6.4	2.1	0.0	200.3	7.6	20.9	4.8
2	8	390.2	245.4	1.00	0.0	0.25	0.0	0.63	0.13	0.0	30.9	85.6	22.3	8.4	24.3	17.6	18.3	23.9	14.1
3	2	403.2	520.5	1.00	0.50	0.0	1.00	0.50	0.0	0.0	0.0	157.2	68.6	14.4	36.3	56.5	21.4	143.3	22.8
4	3	378.9	393.3	1.00	1.00	0.0	0.33	0.0	0.33	1.33	25.0	209.6	85.6	22.1	8.8	21.1	4.4	11.1	5.6
5	2	380.0	371.9	1.00	0.50	0.50	0.0	1.00	0.50	1.50	15.0	187.1	90.5	5.7	43.5	3.8	2.5	15.5	8.3
6	6	391.8	397.3	1.00	1.33	0.17	0.33	1.17	1.00	1.00	16.7	198.8	30.1	13.7	14.1	42.6	10.8	10.8	59.7

INCOME →	450/- to 549/11d																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5	1	461.7	352.1	1.00	1.00	0.0	2.00	0.0	0.50	0.50	75.0	66.6	27.8	2.5	86.4	33.5	8.5	2.5	49.3
6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

INCOME →	550/- to 649/11d																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	2	586.7	412.0	1.00	0.0	0.0	0.0	0.0	1.00	0.0	40.0	146.5	22.6	5.9	6.9	141.3	8.0	37.2	3.6
3	1	600.0	315.3	1.00	0.0	1.00	0.0	1.00	0.0	0.0	75.0	121.3	37.3	10.6	30.0	0.0	4.0	19.1	18.0
4	1	560.0	458.1	1.00	1.00	0.0	0.0	1.00	1.00	0.0	0.0	105.3	34.6	14.0	0.0	137.0	148.0	19.2	0.0
5	1	570.0	424.7	1.00	2.00	0.0	0.0	1.00	0.0	1.00	0.0	87.1	71.3	0.5	1.0	250.0	0.5	6.5	7.8
6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

INCOME →	650/- and over																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				s_1	s_2	s_3	s_4	s_5	s_6	s_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	2	703.8	471.5	1.00	0.0	0.0	0.0	0.0	0.0	0.0	17.5	181.2	47.6	8.0	2.0	110.0	17.3	11.9	76.0
2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6	1	840.0	405.8	1.00	1.00	1.00	0.0	1.00	2.00	0.0	73.8	85.5	15.3	0.0	45.5	47.5	14.9	103.8	19.5

TABLE 6: ALL LOW INCOME

INCOME →	Under 150/-																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	171	98.2	113.0	0.92	0.06	0.02	0.0	0.0	0.0	0.0	10.6	60.3	14.3	5.8	3.6	8.9	2.7	3.0	3.8
2	108	96.3	126.1	1.07	0.37	0.18	0.06	0.19	0.09	0.05	12.6	70.5	12.5	7.0	4.6	7.1	2.0	3.5	6.3
3	121	98.6	134.6	1.03	0.85	0.08	0.09	0.31	0.14	0.50	10.1	80.6	11.8	7.8	3.7	7.9	4.0	3.8	4.9
4	139	101.6	130.6	1.07	1.05	0.10	0.09	0.48	0.31	0.89	5.8	83.1	10.7	9.3	3.9	7.6	2.0	2.8	5.4
5	92	110.6	146.5	1.07	1.26	0.11	0.10	0.77	0.70	0.99	6.4	95.2	12.9	10.3	4.5	6.0	2.9	3.2	5.1
6	72	115.3	148.4	1.17	1.43	0.21	0.17	1.21	0.82	0.99	6.4	99.0	13.3	7.9	4.2	6.8	2.4	3.3	5.1

INCOME →	150/1 to 249/11d																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				S_1	S_2	S_3	S_4	S_5	S_6	S_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	144	194.5	188.9	0.97	0.02	0.01	0.0	0.0	0.0	0.0	22.6	77.6	19.8	7.8	8.6	23.5	8.2	12.1	8.7
2	99	201.2	218.7	1.06	0.29	0.22	0.04	0.29	0.08	0.02	22.0	99.4	23.9	9.7	7.5	25.9	9.8	9.6	10.9
3	104	188.2	196.8	1.15	0.77	0.14	0.11	0.25	0.21	0.37	16.5	102.7	19.0	10.5	11.3	15.3	7.0	7.7	6.8
4	117	192.3	206.7	1.13	1.03	0.15	0.09	0.40	0.34	0.86	17.4	94.5	23.3	16.2	8.3	19.3	6.8	10.2	10.7
5	108	182.7	230.6	1.09	1.22	0.22	0.13	0.78	0.64	0.92	14.2	121.8	15.5	12.5	7.1	22.7	7.2	9.2	20.4
6	65	184.4	206.7	1.09	1.40	0.25	0.15	1.03	0.71	1.37	11.5	119.4	22.3	11.8	5.2	15.7	4.1	8.7	8.0

INCOME →	250/- to 349/11d																	
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)							
				s_1	s_2	s_3	s_4	s_5	s_6	s_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	58	287.0	263.0	0.97	0.02	0.02	0.0	0.0	0.0	0.0	24.4	93.2	25.6	7.9	15.9	32.9	28.0	17.7	17.4
2	71	290.6	261.9	1.04	0.27	0.21	0.09	0.30	0.10	0.0	32.2	101.5	24.9	10.8	14.4	20.2	8.6	20.8	28.5
3	55	294.6	281.9	1.16	0.64	0.22	0.11	0.36	0.22	0.29	24.0	122.7	30.7	9.8	13.7	35.7	10.1	10.3	24.9
4	61	283.0	282.7	1.05	0.97	0.15	0.10	0.34	0.44	0.95	21.4	125.2	27.7	14.7	11.2	44.5	10.7	15.1	12.2
5	42	297.2	264.1	1.17	1.00	0.12	0.17	0.62	0.52	1.40	23.4	121.8	27.7	12.2	16.4	22.8	4.1	12.1	23.6
6	48	288.0	309.4	1.13	1.33	0.19	0.17	1.43	0.67	1.08	12.1	153.8	31.1	17.6	17.2	28.1	2.4	17.5	29.6

INCOME →	350/- to 449/11d																	
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)							
				s_1	s_2	s_3	s_4	s_5	s_6	s_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	17	377.7	406.7	0.94	0.06	0.0	0.0	0.0	0.0	0.0	30.2	116.0	52.4	7.1	19.8	88.0	13.2	55.8	24.2
2	36	389.9	422.6	1.11	0.19	0.17	0.06	0.39	0.08	0.0	46.2	112.7	45.9	12.9	61.5	65.0	15.1	31.2	32.1
3	18	398.7	351.8	1.16	0.56	0.28	0.28	0.44	0.17	0.11	30.6	118.5	46.0	14.9	28.6	42.8	8.6	35.0	26.8
4	25	401.8	355.3	1.28	0.84	0.0	0.12	0.28	0.24	1.24	37.1	149.1	48.8	15.1	37.3	30.3	8.3	11.4	17.9
5	19	387.4	374.2	1.26	1.05	0.16	0.11	0.74	0.68	1.00	42.3	166.3	55.0	14.1	19.6	30.8	5.3	22.6	18.2
6	18	388.1	352.2	1.18	1.11	0.22	0.11	1.00	0.94	1.44	28.9	165.3	28.8	11.2	30.7	44.0	6.4	8.2	28.7

INCOME →	450/- to 549/11d																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				s_1	s_2	s_3	s_4	s_5	s_6	s_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	10	515.1	413.1	1.00	0.0	0.0	0.0	0.0	0.0	0.0	62.0	113.9	62.3	10.7	54.3	31.4	5.0	15.7	57.8
2	11	483.7	349.9	1.09	0.37	0.0	0.09	0.36	0.09	0.0	43.6	132.0	40.6	11.8	36.6	31.2	9.3	16.3	28.5
3	8	501.1	387.4	1.00	0.50	0.50	0.0	0.25	0.50	0.25	73.3	125.1	39.0	12.7	16.7	75.4	23.9	15.6	5.7
4	3	488.7	656.5	1.00	1.33	0.0	0.0	0.67	0.0	1.00	64.5	167.5	58.0	18.7	72.7	97.4	21.0	63.9	92.8
5	8	496.1	432.7	1.13	1.38	0.25	0.50	0.25	0.75	0.75	38.8	174.1	47.5	21.7	18.2	70.3	9.0	18.0	35.1
6	7	508.4	411.9	1.29	1.57	0.0	0.29	0.43	1.71	0.71	35.3	156.5	57.5	19.2	21.6	52.7	2.7	46.1	20.3

INCOME →	550/- to 649/11d																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				s_1	s_2	s_3	s_4	s_5	s_6	s_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	9	586.4	389.5	1.11	0.11	0.11	0.11	0.11	0.44	0.0	40.0	103.5	44.3	15.7	23.0	77.3	25.5	29.4	30.8
3	5	612.0	319.7	1.40	0.20	0.60	0.40	0.40	0.0	0.0	45.4	143.0	17.3	19.6	17.9	16.1	9.6	32.3	18.5
4	10	589.9	438.3	1.20	0.90	0.30	0.0	0.60	0.30	0.70	55.7	156.6	81.9	15.4	15.7	40.3	49.2	16.5	7.0
5	7	586.6	525.3	1.00	1.43	0.43	0.14	0.86	0.43	0.71	40.5	191.7	45.0	34.9	39.3	112.2	8.1	17.1	36.5
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

INCOME →	650/- and over																		
SIZE OF HOUSE-HOLD ↓	f	x	z	DEMOGRAPHIC DATA (Number Per Household)							SPECIFIC EXPENDITURE (IN SHILLINGS PER THOUSAND)								
				s_1	s_2	s_3	s_4	s_5	s_6	s_7	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
1	4	759.4	422.4	0.75	0.25	0.0	0.0	0.0	0.0	0.0	20.0	183.9	39.4	18.0	10.2	62.5	38.4	11.7	38.3
2	3	710.4	392.2	1.33	0.33	0.33	0.0	0.0	0.0	0.0	26.7	105.1	89.0	10.6	7.6	36.8	93.1	22.2	1.1
3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	4	747.9	654.8	2.25	0.25	1.00	0.0	0.0	0.25	0.25	76.1	240.9	59.5	47.3	29.8	68.5	83.6	11.3	37.8
5	1	685.0	778.2	1.00	1.00	0.0	0.0	1.00	0.0	2.00	0.0	247.9	159.1	28.8	109.8	135.0	7.3	30.3	60.0
6	6	730.4	548.1	1.17	1.00	0.17	0.17	0.83	1.50	1.17	52.6	230.6	92.1	14.6	26.8	42.1	16.7	54.3	18.3