An Analysis of the Effect of Prices and Income on Food Consumption in Indonesia

by

Paul Wing-Fai CHEUNG
B. Econ (Tas)
Grad. Dip. Ecomt (ANU)

Dissertation submitted in fulfilment of the requirement for the degree of Masters of Economics in the Australian National University

July 1987

Acknowledgements

In preparing this thesis I have been indebted to several people for their encouragement and suggestions. Specifically, I am grateful to my supervisor Dr. Ray Byron for his guidance, encouragement, helpful and constructive criticisms and suggestions. My gratitude to him is totally beyond words.

I would like to thank Dr. Michael McAleer and Dr. John Beggs for their comments. Especially, I wish to thank Dr. Tony Hall for his help when I most needed.

I would express my thank to Dr. Bruce Chapman for his helpful comments, and advice. His kind and patient assistance is most appreciated. Also, I wish to thank Dr. Paul Miller, Mr. Sukkha Pan and Mr. Gordon Wong.

This acknowledgement would be incomplete without mentioning, with affection, my gratitude to four persons, my parents and my brother Alex who support both me and my studies in Australia, and my girlfriend Karen who encourages my studies all these years.

I declare that this work has been compiled wholly by myself except otherwise indicated.

July 1987 Paul Cheung

CONTENTS

CHAPTER 1

INTRODUCTION

This study examines the effect of price and income on the pattern of food consumption in Indonesia. There is a general belief that a large proportion of the population in Indonesia is suffering from malnutrition. In a comparative study, Knudsen and Scandizzo(1982) argue that 40 percent of the population in Indonesia have calorie consumption levels well below the recommendations of the Food and Agriculture Organization (FAO) and the World Health Organization (WHO) of the United Nations. A similar study by Chernichovsky and Meesook(1983) strongly suggests that there are widespread deficiencies of all nutrients in Indonesia. On the other hand, as pointed out by Klumper(1985), the present food supply in Indonesia exceeds the minimum requirements by more than 20 percent. Therefore, it appears that the problem is more likely one of the maldistribution than of an overall shortfall in the availability of foods. Obviously, there is scope for government intervention in the form of food and nutrition policies. To evaluate the social welfare effects of public policies such as tax reforms or subsidy programs on food items, it is important to determine how a consumer will be affected by changes in relative prices and income. Since all welfare measures presume a knowledge of consumer demand functions, the first step is to correctly specify and estimate a system of commodity demands. Subsequently, attention should be given to the estimation of price and income elasticities. This is, in fact, the major underlying reason motivating this study.

Applied researchers have been attempting to explain consumer behaviour for a long time. A consumer in this context is defined to be either an individual or a group with a common purpose (e.g. a family). Commodities are assumed to be non-negative, finite and perfectly divisible. They include all possible goods and services in the consumer choice set. Barten(1968) and Yoshihara(1969) suggest that demand studies of this kind may be classified into two broad categories. One is concerned with specifying a demand equation explaining the quantity consumed for a single commodity. The other addresses the problem of reallocating total expenditure to an exhaustive set of different commodities. In this study we shall adopt the second approach, the empirical analysis of a complete demand system.

Demand analysis has a well established theoretical and empirical literature. Brown and Deaton(1972) presented an early survey of the topic. Elementary but comprehensive texts about the theory and its application have been written by Phlips(1974) and Powell(1974). A masterpiece in the literature has been provided by Theil(1975 &1976), and a concise review of complete demand theory can be found in Barten(1977). Deaton and Muellbauer(1980a) also presented a general text on consumer behaviour. Another examination of the consumer theory was given by Barten and Bohm(1982). Two modern contributions to the empirical aspects of demand analysis have been written by Deaton(1986a) and Bewley(1986). A new text on applied demand analysis written by Theil and Clements(1987) is due to be published soon.

Generally, there are at least two basic, distinct, but related questions, of interest. First, given various market parameters such as prices and income, what commodities and in what quantities will the consumer purchase? Secondly, how are these decisions affected by changes in the parameters? To answer these questions, it is essential to specify and estimate a demand model. To estimate price elasticities, time series data are usually used. But few developing countries such as Indonesia possess long enough runs of comparable data to permit sensible estimation. However, many developing countries possess good household survey data (cross-sectional) on demand pattern. In this study, a cross-sectional approach is used to examine food consumption behaviour in Indonesia, a developing country in the Asian Pacific region. The cross-sectional data referred to are a collection of household budgets giving all expenditures on consumer goods and services made by individual families over a specific short period of time. Using the cross-sectional approach has some advantages. For example, in empirical demand analysis, one should be conscious of the implicit assumption that the utility function remains unchanged over the observation period. While this restriction is unacceptable over time, it is not an unreasonable assumption for a cross-sectional study. Hence we assume we can ignore changing preferences.

When estimating a system of demand functions, it is not unusual to reject the restrictions of demand theory; examples include Barten(1969), Byron(1970), Deaton(1972), Christensen, Jorgensen and Lau(1975), Deaton and Muellbauer(1980b). In this study, we find that demand conditions are significantly rejected. Instead of assuming, over-optimistically, the acceptance of the underlying demand theory (Phlips[1974]), or on the contrary, strictly rejecting its validity (Christensen, Jorgensen & Lau[1975]), we would question the reasons why the conditions are rejected. In particular, is the model suffering from any statistical weaknesses which affect inference results? We adopt an unusual practice in demand analysis, that is, we apply rigorously a series of diagnostic checking and correcting procedures to the maintained demand model. Surprisingly, to the author's knowledge, this kind of exercise has never been applied to empirical demand analysis and in fact, is found to be extremely important and useful. As a result of using these procedures we find that the test statistics are greatly improved. However, all the demand restrictions are still rejected.

Another possible reason for rejection is small sample bias in the asymptotic test statistics. Therefore, size correcting methods are also considered. In addition, a new "distribution free" approach is examined to determine the validity of asymptotic tests in finite samples. As a result, the asymptotic test statistics are found to be justifiable in this particular finite sample. Further, a Monte-Carlo experiment is used to examine the impact of non-normality on hypothesis testing. It shows that its effects on inference are not as serious as one might expect.

Although all the results are negative, we are not able to conclude that the theory is invalid. As argued by Simons and Weiserbs(1979) and Mattei(1986), the demand models may not be adequately specified and therefore we need a more realistic model before we can make a firm conclusion. Possible areas for improvement in model specification are also outlined in this study.

Due to the limited time and space of this thesis, we are unable to deal with those suggestions empirically. But at least we would like to point out the direction for future studies and investigation in similiar area.

The structure of this study is as follows. In Chapter II, we review some aspects of basic demand theory. Phlips(1974) argues that since prices should be constant for a short time span in a cross-sectional survey, all restrictions (i.e. homogeneity, symmetry and negativity of the own substitution effect) of demand theory in terms of price derivatives disappear, the only restriction remaining being the adding-up condition. However, this need not be the case in practice because geographical and regional variations in prices still exist within the sample. As pointed out by Deaton(1986b), because of high transport costs, there is substantial spatial variation in prices especially in many developing countries. If different households in the budget study face different prices then estimating a complete demand system using budget data is not much different from using time series (Pollak and Wales[1978]). Hence, all demand restrictions will be considered. One should note that in a study confined to the demand for food there is an implicit assumption of a weakly separable utility function between different group of commodities. This condition will also be discussed in the chapter.

The data is described in Chapter III. They are found originally in the Indonesia household survey records, the SUSENAS tapes.

As pointed out by Brown and Deaton(1972), the choice of the demand model itself has important implications as strong *apriori* notions are built into the analysis by the choice of model and these will interact with the data to yield results which, to some extent, will be affected by the model chosen. Also, as illustrated by Park(1969), the parameters and estimated elasticities can be very sensitive to the model specified. Hence, in Chapter IV, we will discuss, estimate and evaluate the performance of several demand models to determine the maintained model. Based on the empirical results, we examine the usual elasticity estimates and test the validity of the underlying demand theory.

In Chapter V, an unusual exercise in diagnostic checking of the maintained model is presented. As will be argued, it is dangerous to accept a model without subjecting it to proper diagnostic testing. Surprisingly, this kind of exercise has never been applied to empirical demand analysis.

Apart from the standard parametric techniques, a new "distribution free" approach on testing demand restrictions will be discussed in Chapter VI. This type of non-parametric technique is becoming more popular in applied studies, especially when the parametric assumptions are difficult to justify. A conclusion is then presented in Chapter VII.

CHAPTER II

BASIC DEMAND THEORY

The origin of demand theory lies in the development of two basic concepts: utility and optimisation of consumption. Utility refers to the satisfaction associated with consumption. As pointed out by Katzner(1970), it was Edgeworth(1881) who expressed utility as a general function of all commodities. The concept of demand goes back at least to King(1696), who computed a demand schedule for wheat and derived the famous "King's Law" which is an inverse statistical relationship between price and quantity. Walras(1874) was the first to succeed in bridging the gap between utility and demand. From utility maximization, he derived demand as a function of all prices and initial endowments. Marshall(1890), with the assumption that all other commodity prices held constant, derived demand functions depended only on the prices of the commodity in question and income. He also demonstrated that such hypotheses imply demand curves are downward sloping. Then it was Slutsky(1915) who finalized the theory and transformed "Classical" demand theory into what it is today. The practical difficulty of this Classical demand approach is to specify explicitly a given individual's utility function. In this chapter, the theory of utility and the complete demand system approach will be reviewed briefly. We will examine some general and specific restrictions on demand functions. In addition, another important aspect in modelling demand functions, the flexible functional form approach, will also be discussed.

2.1 The Utility Function

Assume that for each consumer there exists a continuous utility function, μ, for a finite number of commodities, n.

$$\mu = f(q_1, q_2, ..., q_i, ..., q_n) \tag{2.1}$$

where q_i is the quantities consumed of commodity i. Thus, μ is a numerical representation of a preference ordering.

While the term commodities should be characterized by strictly non-negative numbers and should be perfectly divisible, the specification of preferences is assumed to be comparable, transitive, continuous, monotonic, strictly convex and differentiable. In other words, the consumer is able to judge whether a commodity (or commodity bundles) is (are) preferred. His preferences are assumed to be rational and consistent. He always prefers more to less (i.e. non-satiable). Also, if two commodity bundles are indifferent, then a linear combination of the two bundles is always preferred to either one of the single bundle.

From these, the utility function μ is assumed to be :

1) a continuous function of the quantities consumed. In a commodity set Q where q_0 is any commodity or bundle, the set of bundles not preferred to q_0 and the set of bundles to which q_0 is not preferred are both closed in Q, for any q_0.

2) a strictly increasing function of the quantities consumed

3) a strictly quasi-concave function. That is

$$\mu(a^*q_1+(1-a)^*q_0) \geq \mu(q_0)$$

where $\mu(q_0) = \mu(q_1)$ and $0 \leq a \leq 1$

4) a twice differentable function where the first order partial derivative, $\partial\mu/\partial q_i$, is called the marginal utility of good i. Because of the assumed monotonicity, marginal utility is always positive. Also, from Young's Theorem, the matrix of the second order derivatives, the Hessian, is symmetric. i.e.

$$\partial^2\mu/\partial q_i\partial q_j = \partial^2\mu/\partial q_j\partial q_i$$

In addition, the Hessian is non-singular and negative definite because the utility function is (strictly) quasi-concave.

5) a function defined up to a monotonic increasing transformation. This means that it is always defined to be ordinal.

As pointed out by Deaton(1986a), it is common in empirical work, to assume that the utility function is strictly quasi-concave so that for $0 < a < 1$, the equality in statement (3) above is strict. But it is restrictive in a sense that strictly quasi-concavity rules out the possibility of perfect substitutes.

2.2 The Complete Demand System Approach

The complete systems approach to demand analysis constitutes a joint analysis of the expenditure or consumption volume of those commodities which make up total private consumption. On the basis of a system of demand functions, mostly developed from Classical demand theory, demand is explained by income and price changes. The main advantage of the complete systems approach compared to an analysis of each single commodity is the increased efficiency of the estimation. The aggregation, homogeneity and symmetry constraints of Classical demand analysis add degrees of freedom to the analysis. When estimating a complete demand system, joint treatment of all commodities also makes it possible to take advantage of the correlation between commodities. This will be examined in more detail in Chapter 4.

Let

$$q_i = f(p_1, p_2, \ldots, p_i, \ldots, p_n, y) \qquad (2.2)$$

where q_i = quantities consumed for the i'th commodity

p_i = quantity price per unit of the i'th good

y = income or total expenditure

by definition

$$y = \Sigma_i p_i q_i \qquad (2.3)$$

Since consumers are price-takers, they are unable to influence prices. Therefore prices are assumed to be exogenous. Also, income y is fixed. The definition in (2.3) makes (2.2) a complete set of demand functions.

Applied consumption analysis is mainly concerned with the estimation of the parameters of one or all equations in (2.2). For a long time, researchers have been attempting to specify and estimate such demand systems. Examples of this activity are, the Linear Expenditure System within the utility framework proposed by Klein and Rubin(1947) and implemented by Stone(1954), the Translog

System by Christensen, Jorgensen and Lau(1975), and the Almost Ideal Demand System by Deaton and Muellbauer(1980b).

There are some basic assumptions used in complete demand analysis. First, income is measured by summing the expenditure for an exhaustive set of different commodities and we are discussing models concerning with the allocation of total expenditure to that particular set of commodities. This implies that the utility of the services yielded by saving and those yielded by current consumption are separable. Hence saving need not be considered.

The second assumption relates to the problem of identification. The supply functions are supposed to contain determinants which are absent from equation (2.2). It is implicitly assumed that at given fixed existing prices, consumers can buy what they can afford and the demand equations are then written with quantity dependent on prices and income. Otherwise, the function estimated may be a supply curve or a mixture of demand and supply.

Thirdly, it is generally assumed that the commodities are weakly separable. It is this condition that justifies the commodity, hence, makes demand studies empirically operational. This assumption will be discussed in further detail in section 2.4 of this chapter.

Also equation (2.3) provides a budget constraint on equation

(2.2). Further restrictions from the consumer demand theory are assumed so as to arrive at results which can be given a theoretical interpretation, and to reduce the dimensionality of the estimation problem. The more restrictive are the conditions, the greater the chances are that the model will be rejected by the data and the greater is the confidence we may attach to our estimates if they nevertheless turn out to be valid.

2.3 General Restrictions on Demand Functions

2.3.1 Homogeneity Condition

Every demand equation must be homogeneous of degree zero in income and prices. That is, if all prices and income are multiplied by a positive constant k, real income and relative prices would not change and the quantity demanded remains unchanged.

Suppose $q = f(y, p)$ and

$$q^{*} = f(yk,\ pk),$$

then q is homogeneous of degree r in y and p if

$$q^{*} = k^{r} f(y,\ p) = k^{r} q$$

In this case $r = 0$.

In applied work, only those mathematical functions which have this property can be considered as demand functions.

The homogeneity condition is based on the assumption that the individual consumer makes his decision irrespective of the monetary unit of account. It implicitly means that q in equation (2.2) does not contain pure monetary goods. The use of this condition has the consequence of eliminating the effect of inflation. That is, consumer demand is assumed insensitive to inflationary movements. The absence of money illusion is an attractive property for demand functions but it may nevertheless be untrue.

To make the condition operational, apply Euler's Theorem that, if a function, $z = f(q,y)$ is homogenous of degree r, then

$$q(\partial z/\partial q) + y(\partial z/\partial y) = rz$$

Applying this theorem to demand equation $q_i = f(p_1, p_2 \ldots p_n, y)$ gives, in general

$$\Sigma_j p_j(\partial q_i/\partial p_j) + y(\partial q_i/\partial y) = 0 \qquad (2.4)$$

Dividing all elements in (2.4) by q_i, we obtain,

$$\Sigma_j[(p_j/q_i)(\partial q_i/\partial p_j)] = -(y/q_i)(\partial q_i/\partial y)$$

That is, the sum of all direct and cross elasticities with respect to prices of any commodity i has to be equal to the minus of its income elasticity. The relationship can be denoted as

$$\Sigma_j e_{ij} = - E_i \qquad (2.5)$$

It is worth emphasizing that the demand equations will automatically satisfy the homogeneity restriction (as well as the other general restrictions) when the demand system is obtained by constrained maximization of an algebraically specified utility function.

2.3.2 Adding-up Condition

The budget constraint in (2.3) has to be satisfied over the observed or predicted range of variation of prices and income. Therefore, the demand equations have to be such that the sum of the estimated or predicted expenditure on the different commodities equals total expenditure in any period, i.e. equation (2.3) must hold. Differentiating the budget constraint with respect to y, we get

$$\Sigma_i p_i(\partial q_i/\partial y) = \Sigma_i(\partial(p_i q_i)/\partial y) = 1 \qquad (2.6)$$

where $\partial(p_i q_i)/\partial y$ is called the marginal propensity to consume good i, or its marginal budget share. According to equation (2.6), the marginal propensities to consume must sum to one. In other words, an increase in total expenditure must be entirely allocated to the different commodities. Note that the adding-up condition will be automatically satisfied if the demand system is derived by constrained maximization of a specific utility function.

2.3.3 Symmetry Condition

The basic idea is that the price derivatives of a demand equation can be decomposed, as developed by Slutsky(1915), into an income effect and substitution effect. In mathematical form, the total effect of price change is,

$$\partial q_i/\partial p_j = (dq_i/dp_j)_{y'} + [-(\partial q_i/\partial y)(dy/dp_j)]$$

where $(dq_i/dp_j)_{y'}$, known as the income compensated substitution effect, is the response of q_i to a compensated price change, evaluated at $y' = y + dy$. However, the compensation is such that $dy/dp_j = q_j$. Therefore, denoting the substitution effect as K_{ij}, the equation can be written as,

$$\partial q_i/\partial p_j = K_{ij} + [- q_j(\partial q_i/\partial y)] \qquad (2.7)$$

i.e. Total effect = Substitution effect + Income effect

Equation (2.7) is known as the "Slutsky Equation".

The symmetry condition is the restriction related to the substitution effect of price changes, K_{ij}. If the consumer is to behave consistently the income compensated substitution effect on the number of units bought of good i in response to a change in the price per unit of good j must be the same as the substitution effect on good j of the same change in the price per unit of good i, no matter how the units are defined. That is

$$K_{ij} = K_{ji}$$

It should be noted that the sign of the cross substitution effect K_{ij} is not determined. Following a standard definition by Hicks(1936), K_{ij} is positive if i and j are substitutes and is negative if they are complementary. If i and j are independent of each other K_{ij} is zero.

On the other hand, there is no general restriction on the sign of the income derivative $\partial q_i/\partial y$. It is negative if i is an inferior good and positive if it is not. Notice that the income derivative and the income effect always have the opposite signs.

2.3.4 <u>Negativity</u>

This condition implies that the substitution matrix K is negative semi-definite, i.e. the diagonal elements must be negative

$$K_{ii} < 0$$

The proof is based on the fact that the utility function is quasi-concave and continuous. Hence the matrix K is symmetric and negative semi-definite. The condition directly implies an increase in price with utility held constant must cause demand for that good to fall. This is the famous 'Law of Demand' which states that the own price compensated elasticities of demand are negative, or that compensated demand curves can never slope upwards. Even K_{ii} is always negative we still cannot determine the sign of the total effect on price change in (2.7) because the income effect may be positive, which happens only if the income elasticity is negative (i.e. the commodity i is an inferior good). Thus a positive price derivative (uncompensated) can only occur if the good is highly inferior and if it is purchased in large quantities. It means that its positive income effect is greater than the negative substitution effect in absolute term. Such a good is known as a "Giffen good" and indeed is extremely rare. Note that while all Giffen goods must be inferior, the converse is not true.

To summarize the four basic general properties of demand functions: they add up, they are homogenous of degree zero in prices and total expenditure, and their compensated price responses are symmetric and form a negative semi-definite matrix. While the adding-up and homogeneity conditions are consequences of the specification of a linear budget constraint, the symmetry and negativity conditions are derived from the existence of consistent preferences and the fact that utility is maximized. Violation of the symmetry and negativity conditions means that consumers are irrational as they made inconsistent choices and they failed to maximize their utility (or minimize their costs).

2.4 Some Specific Restrictions

2.4.1 Separability, Utility-tree Approach

The concept of separability arises from the independent work of Leontief(1947) and of Sono(1961). The usefulness of this condition is to partition the consumption set into subsets which would include commodities that are closer substitutes or complements to each other. Commodities may be grouped in such a way that goods which interact closely in the yielding of utility are grouped together while goods which are in different groups interact only in a general way. For example, different types of food may go into one group while other types (e.g. related to entertainment) go into another. It is expected that if there exists a relationship between one type of food and one type of entertainment, then that relationship will be much the same for all pairs of commodities chosen from the two groups.

That is,

$$\mu = f(q_1, q_2, q_3, q_4, q_5)$$

may be expressed as

$$\mu = F(A, B)$$

where $A = f_a(q_1, q_2)$

$B = f_b(q_3, q_4, q_5)$

This implies that a grouping of the variables should not modify μ, since if this is not the case, the utility function, μ, is not separable.

There are two definitions of the separable utility function, proposed by Strotz(1959). They are summarised below.

2.4.1.a <u>Weak Separability (Leontief,1947)</u>

Let the n commodities be partitioned into m mutually exclusive and exhaustive branches or groups and let there be n_r $(r=1,\ldots,m)$ commodities in each group such that $n = \Sigma_r n_r$.

The utility function, μ, in (2.1) is weakly separable if it can be expressed as

$$\mu = f(q_{11},\ldots,q_{1n1},q_{21},\ldots,q_{2n2},\ldots,q_{r1},\ldots,q_{rnr},\ldots,q_{m1},\ldots,q_{mnm}),$$

or $\quad \mu = F[\, f_1(q_1),f_2(q_2),\ldots,f_r(q_r),\ldots,f_m(q_m)\,] \qquad (2.8)$

where each f_r is a branch utility function or specific satisfaction function of group r, and each q_r is a function of $q_{r1},q_{r2},\ldots,q_{rnr}$.

Maximization of the overall μ implies maximization of the branch utility functions, $f_i(q_i)$, where i=1,..., m. Hence, it implies the existence of subgroup demands, i.e.

$$q_i = g_i(y_i, p_j) \qquad j = 1, 2, ..., nr$$

where y_i is the expenditure spent on group i.

The necessary and sufficient condition for a function to be weakly separable is that the marginal rate of substitution between any two variables belonging to the same group be independent of the value of any variable in any other group. That is,

$$\partial[(\partial\mu/\partial q_{ri})/(\partial\mu/\partial q_{rj})] \,/\, \partial q_{sk} = 0$$

where q_{ri}, q_{rj} refer to commodity i and j of the same group, R, and

q_{sk} refers to commodity k in group S where $S \neq R$

2.4.1.b <u>Strong Separability</u>

A utility function, μ, in (2.1) is strongly separable (block independent) in the branches if it can be written as (2.8) such that the marginal rate of substitution between any two goods belonging to two different groups is independent of the consumption of any good in any third group. That is,

$$[(\partial\mu/\partial q_{ri})/(\partial\mu/\partial q_{sk})] \,/\, \partial q_{tl} = 0$$

where q_{tl} refers to commodity l of group T, and q_{ri} and q_{sk} refer to commodity i and k of group R and S respectively and $R \neq S \neq T$. As a result, the marginal utility of a commodity in one group is independent of the consumption of any good in any other group.

If there are only two branches, weak and strong separability are identical. If this is not the case, the weak and strong conditions differ.

Assuming no homogeneity, a utility function with strong separability may be transformed into an additive function which is written as

$$\mu = f_1(q_1) + f_2(q_2) + ... + f_m(q_m)$$

and is referred as an additive separable utility function.

Although additivity in each individual commodity is known to have unacceptable empirical implications, it is less stringent to assume that a utility function is additive in branch utilities (additively separable) than in the utilities of each and every commodity (Strotz[1959]).

Separability is important to the hypothesis of the utility tree as the demand for a commodity in a branch can be expressed as a function of the prices in and the budget allotment to that branch. That is,

$$q_{ri} = f_{ri}(p, y) = F_{ri}(p_r, y_r)$$

where p is the vector of all prices,

y is total income or expenditure,

p_r is the vector of commodity prices in group r, and

y_r is the budget allotment to group r.

It is not the case that the quantities demanded in one branch are independent of the prices of commodities in other branches or of total expenditure. More accurately, total income and the prices of goods in the branch only transmit their effect on the budget allotment to that branch. Therefore, when the budget allotment to the branch is known, the prices of goods outside the branch can be ignored as their impacts felt through the income effects.

Another implication of (weak) separability is its consistency with the two-stage maximization procedure discussed by Strotz(1957). The (weakly) separable utility function appears, in the terminology of Strotz, as a utility tree with branches corresponding to $f_1, f_2, ..., f_m$. Households are assumed to proceed in two steps. The first is an optimal income allocation process among broad commodity groups. The second step is to decide the optimal spending of each budget allotments in the branch. Weak separability is both necessary and sufficient for the second stage of this two-stage budgeting process.

Under separability, the problem of aggregation over commodities finds a natural solution. It justifies the application of commodities grouping in empirical analysis.

Separability assumptions have been used frequently, but their usefulness depends on the ability to classify goods into groups for which the separability assumption may be considered empirically valid.

Gorman(1959) argues that, in order to make the budgeting process justifiable, it is necessary to go further and assume that a utility function falls into two parts, the first of which is additively separable, while the second is a function of homogeneous (of degree one) specific satisfaction functions. If a utility function is said to be weakly separable, the marginal rates of substitution between any two items in the same group must be independent of the consumption of goods in the other group. Additivity merely extends this postulate to items from a pair of different groups. If the specific satisfaction functions have to be homogeneous, it is inappropriate to group luxuries, near-luxuries and necessities together.

2.4.2 Aggregation Problem

There are two types of aggregation problem. One is aggregation of demand function over individual demands and the other is aggregation over commodities. Both are examined below.

2.4.2.a Aggregation Over Individuals

The theoretical basis of systems of demand functions is based on the theory of optimization behaviour for individual agents. But the statistical data used for empirical applications usually refer to markets in which several individuals operate (e.g. the household budget survey) or the demand for consumer goods in a whole economy (e.g. in time series study). A question of concern is to what extent micro theory can be considered relevant for the description of aggregate demand behaviour. Aggregation over individuals clearly creates some problems, as is demonstrated below.

Let the subscript h(1... H) refer to individual behaviour, the absence of the subscript indicating the corresponding average.

By definition

$$q = (1/H)^{*} \Sigma_h q_h \qquad (2.9)$$

$$y = (1/H)^{*} \Sigma_h y_h = (1/H)^{*} \Sigma_h (\Sigma_i p_i q_{ih}) = \Sigma_i p_i q_i$$

Micro theory is relevant for the individuai demand system

$$q_h = f_h(p_1, p_2, ..., p_n, y_h) \qquad (2.10)$$

While in empirical work, one is interested in the aggregated system of the economy as a whole, which is

$$q = f(p_1, p_2, ..., p_n, y) \qquad (2.11)$$

Substituting (2.10) into (2.9), we have

$$q = (1/H)^* \Sigma_h f_h(p_1, p_2, ..., p_n, y_h) \qquad (2.12)$$

Equations (2.12) and (2.11) are only equivalent under restrictive assumptions that the variations in y_h and $f_h(.)$ across individuals are the same. If this is the case, then the averaged consuming unit can be regarded as a representative household. Assume that the representative household allocated its expenditure so as to maximize welfare. As a consequence, the representative household behaves in the same way as an individual maximizing a utility function.

Another so called the convergence approach to the aggregation of linear equations is examined by Theil(1975). Suppose an individual h has a behavioural equation of the form

$$y_h = \alpha_h + \beta_h x_h + e_h \qquad h = 1, ..., H$$

To derive a linear equation in per capita variables,

i.e. $\overline{y} = (1/H)\ \Sigma_h y_h \qquad \overline{x} = (1/H)\ \Sigma_h x_h$

we have the following,

$$\overline{y} = \alpha + (1/H)\ \Sigma_h \beta_h x_h + (1/H)\ \Sigma_h e_h$$

$$\Rightarrow \qquad \overline{y} = \alpha + (\ \Sigma_h \beta_h x_h / \Sigma_h x_h\ )\ \overline{x} + (1/H)\ \Sigma e_h$$

where $\alpha = (1/H)\ \Sigma_h \alpha_h$

By treating the H individuals as independent random drawings from a population, the associated distribution of β's is then the theoretical counterpart of a discrete cumulated distribution function (by pre-arranging β's according to increasing magnitude). If the mean and the standard deviation of this theoretical distribution are respectively β and σ_β, and by assuming that the x's are non-stochastic and β_i and β_j are independent if $i \neq j$, the slope coefficient of $\overline{x}$ is then a random variable with expectation

$$\beta\, \Sigma_h x_h \,/\, \Sigma_h x_h = \beta$$

and variance

$$(\sigma_\beta^{\,2}/H) \ * \ \{\ 1 + [\ (1/H)\ \Sigma_h (x_h - \overline{x})^2\ ] \,/\, \overline{x}^2\ \}$$

The variance of the slope coefficient converges to zero as H tends to infinity. It follows from Chebyshev's inequality[1] that the coefficient of $\overline{x}$ converges in probability to its expectation β. All these conditions imply that for H sufficiently large, the linear equation in per capita variable can be simplified to

$$y \approx \alpha + \beta\overline{x} + (1/H)\ \Sigma_h e_h$$

Although the result is extremely attractive because of its simplicity, the approach suffers from some limitations which should be kept in mind. As pointed out by Theil(1975), the approach basically depends on the crucial assumption that there is independence of the factors determining the behaviour of the individuals (the values taken by the micro variables) and the way in which they react to these factors (the micro parameters).

2.4.2.b Aggregation Over Commodities

One of the justifications for this type of aggregation is due to the Composite Commodity Theorem which asserts that if a group of prices move in parallel, the corresponding group of commodities can be treated as a single good. Deaton and Muellbauer(1980a, Chapter 5.1) prove that if

$$p_i = kp_i^*, \qquad p_j = kp_j^* \qquad (2.13)$$

where p_i^*, p_j^* are base period prices for good i and j, and

k is a fixed ratio which varies with time but is common to both prices, then

$$p_i/p_j = p_i^*/p_j^*.$$

Provided that (2.13) holds, p_i and p_j can be treated as composite commodity prices and the new preferences will lead to the same choices as the original ones.

The usefulness of this theorem is that if relative prices are stable and largely independent of the pattern of demand, then commodity groups should be chosen so that close substitutes are grouped together. Barten and Turnovsky(1966) demonstrate that for composites of any number of elementary commodities, all macro parameters are sums of the corresponding micro parameters only if there is no specific interaction between elementary commodities of different composite exists.

The other justification for commodities aggregation is due to the hypothesis of separable preference (preference independence) analysed above. If that condition holds, the commodities can be partitioned into groups so that preferences within groups can be described independently of the quantities in other groups.

It may be important to note that the empirical significance of possible distortion caused by aggregation errors is largely unknown.

2.5 Flexible Functional Forms

The advantage of using a flexible form is its capability of providing a (local) second-order approximation to any unknown arbitrary utility function. The basic idea (Deaton[1986]) is that the choice of functional form should be such as to allow at least one free parameter for the measurement of each effect of interest. For example, the basic linear regression model with intercept is a flexible functional form. Even if the true data generation process is not linear, the linear model without parameter restrictions can offer a first-order Taylor approximation around at least one point.

Flexible functional forms can be constructed by approximating preferences or demands. By Shephard's Lemma, an order of approximation in prices or quantities, but not in utility, is lost by passing from preferences to demands, so that in order to guarantee a first-order linear approximation in the demands, second-order approximation must exist in preferences. Beyond that, one can choose to approximate the direct utility function, the indirect utility function or the cost function.

The definition of second-order approximation is given by Barnett(1983). In mathematics, V^* is a second-order local approximation to V at the point v_0 if

$$[V^*(v) - V(v)] \;/\; ||v-v_0||^2$$

tends to 0 as v tends to v_0 where $\|\ .\ \|$ designates the Euclidian norm.

Barnett(1983) demonstrate that the two definitions of second-order approximation in economics proposed by Diewert(1971), Christensen, Jorgensen and Lau(1973) and Christensen(1975) are equivalent to the usual mathematical definition. These are now considered.

2.5.1 <u>Definition proposed by Diewert(1971)</u>

Define V^* to be a second order approximation to V at v_0 if

$$(1)\quad V^*(v_0) = V(v_0)$$

$$(2)\quad \partial V^*/\partial v\ |v{=}v_0 = \partial V/\partial v\ |v{=}v_0$$

$$(3)\quad \partial^2 V^*/\partial v \partial v'\ |v{=}v_0 = \partial^2 V/\partial v \partial v'\ |v{=}v_0$$

If V^* possesses these capabilities at v_0 for any V, then Diewert(1971) calls V^* a (locally) flexible form. This is regarded as the standard definition.

2.5.2 <u>Definition proposed by Christensen, Jorgensen and Lau(1973) and Christensen(1975)</u>

This refers to the existence of some neighbourhood J of v_0 and some constant k such that

$$|V^{*}(v)-V(v)| \leq k \, ||v-v_0||^3/||v_0||^3 \quad \text{for all v belongs to J.}$$

The use of flexible functional forms might be rationalized as a convenient method for summarizing market behaviour. But they do not necessarily provide a good approximation over a range of observations, as measured by their ability to satisfy the regularity conditions required for utility maximization. Wales(1977) argues that the extent to which the flexible form will satisfy the desired regularity conditions depends on the parameters of the true utility function, the particular choice of functional form, and on variation in the determining variables. For example, the Translog approximations are generally better the smaller the variation in the independent variables. This suggests that the use of flexible forms with aggregated time series data may be more appropriate than is the case with cross-sectional data.

Apart from the common Taylor's series approximation, there is another method, which is known as the general class of Fourier series approximations suggested by Gallant(1981). He argues that Taylor's theorem only applies locally, and that the theorem fails as a means of understanding the statistical behaviour of parameter estimates and test statistics. Consequently, tests of hypotheses based on Taylor approximation properties may be seriously misleading. The reason (White[1980b]) is that when parameters are estimated by least squares methods the OLS estimates do not necessarily provide reliable information about the local properties (derivatives, elasticities) of unknown functions. Besides, the inexactness of the Taylor approximation interpretation is evidenced by a lack of agreement about the point of expansion. Usually, the approximation is considered to be taken at the mean of the explanatory variables. Gallant(1981) explains this failure as being a consequence of the fact that statistical regression methods expand the true function in a general Fourier series, (not in a Taylor's series). The former attempts to minimize the average prediction bias arbitrarily by increasing the number of terms in the expansion. Although this form seems to be essentially unbiased, as argued by Gallant(1981), its highly non-linear and complex structure imposes a heavy computational burden. It is worth commenting that White(1980) and Gallant(1981) did overstate their case. White's(1980) numerical example, which Gallant(1981) uses as his numerical justification, was incorrect. Byron and Bera[1983] show that, when arithmetic errors are corrected, White's approximation is actually quite good.

Despite the possible disadvantage of the Taylor expansion, it is the most commonly used method in demand theory. The best known of these approximations is the Translog model, which is a second-order approximation to a utility function. Another flexible form, the Almost Ideal Demand System (AIDS model), which is similiar to the Translog model, is the second-order approximation to a cost function. Although the Translog considerably predates the AIDS model, the latter is a good deal simplier to estimate. Both Translog and AIDS models yield demand functions which are first-order flexible. Another interesting property for the models is the capability of testing the validity of demand conditions described above. This will be discussed in more detail in Chapter 4.

Footnotes

1. Chebyshev's inequality:

    Let x be a (scalar) random variable with mean μ and variance σ^2 and let $k > 0$ be a real number. Then

    $$\Pr\{ | x - \mu | > k \} \leq \sigma^2/k^2$$

    The proof of this proposition can be found from Dhrymes(1978).

CHAPTER III

THE DATA SET

The original data file is from SUSENAS tapes which records a large-scale nation wide cross-sectional family budget survey in Indonesia for the year 1981. A brief description of the data tapes can be found in Byron(1983).

There are 18 types of information stored in the tapes, of which record type 10 and type 21 are the most relevant to our analysis. Record type 10 is "Household Identifier". It gives social and demographic information about the household, such as urban/rural area, province, Kabupaten (a large municipal unit), number of household members, and household income. Record type 21 is "Food Purchase and Consumption", and records each individual household's expenditure on every food item and the value of home produced food for own private consumption, during the week prior to the survey. For a detailed breakdown of the two information types, refer to Appendix 1.

It is useful to have a general understanding of Indonesia before beginning our investigation. An overview of relevant aspects is presented in section 3.1. A second section describes the data.

3.1 Indonesia, a General Description

Indonesia can be regarded as a developing country and is located in the Southern Asia region on the equator. It is a tropical island country and with a population of 151 million people (in 1981)[1] and an average annual population growth rate of about 2.2%.

The country consists of 6 main groups of islands, they are: Sumatra, Java, Nusa Tenggara, Kalimantan, Sulawesi, and Maluku and Irian Jaya (they are referred as islands for simplicity). Each island composes several provinces, of which in total there are 27 provinces. Appendix 2 shows the detail break down of provinces in each island and a general map of Indonesia is presented in Appendix 3. Each province is sub-divided into Kabupatens (equivalent to a large municipal unit), of which there are a total of 246 Kabupatens in the country. Each of these can be separated into urban and rural regions. In 1978, there were about 27,777,000 households living in Indonesia, of which 17% were in the urban regions and 83% were in rural regions. The SUSENAS sample also provide a representative geographical distribution of the population in Indonesia (Chernichovsky and Meesook[1983]).

Table 3.1 gives the percentage of area and population density of Indonesia, by island, in 1980. Kalimantan is the largest island in

Indonesia, but it has the second lowest percentage of population. On the other hand, while Java is the second smallest major island, it is the most populated area, being occupied by 62% of the total population. Consequently Java has the highest population density in Indonesia, while Maluku and Irian Jaya the lowest.

Table 3.2 gives the percentage of land area which was utilized for residential purpose in 1980. It shows that, on average, only about 3.5% of the total land area in Indonesia was in residential usage, and Java alone had the largest percentage of almost 13% compared to its area.

Religion is an important influence in Indonesia. Since 87% of the total population in Indonesia are Islamic, most Indonesians do not drink alcohol.

Also relevant to consumption patterns is racial make-up. The country is dominated by two groups, the Malay (Asian origin) and the Native (Non-Asian origin). The majority of the Asians live in the western part of Indonesia while the Non-Asian group are mainly in the eastern and southern islands. Generally speaking, the Asian people have higher income than the Non-Asians. Further, their diets are different. For example, rice is the major food crop for the Malay while tubers are the main crop for the Native. In fact, this may be due to their differences in customs, habits and income level.

3.2 The Data

The raw information on food expenditure is classified into 188 different food items in the SUSENAS survey (refer to Byron[1983]). In order to facilitate the discussion, it is necessary to aggregate those commodity items into a smaller and more manageable size, say in this case, 17 mutually exclusive and exhaustive food groups (with altogether 156 different sub-items). This categorization is based on the assumption that the utility function is additive separable (the rationale for which is considered below). Twelve "sub-totals" and twenty "other" items were excluded, the latter because there is no quantity information was provided. They account for only about 6 per cent of the total consumer expenditure on food. For a detail description of this reclassification, see Appendix 4.

Since the SUSENAS survey only records the quantity of and expenditure on each item (note that it is the usual practice in budget surveying), individual prices are calculated by dividing expenditure over quantity consumed, which means that derived prices differ between consumers. Prices are zero if no expenditures are reported for those items. Also, when aggregating 156 commodities into 17 groups, group prices are constructed using value weights, $p_i q_i$, as follows,

$$P_I = \frac{\sum_{i \in I} p_i (p_i q_i)}{\sum_{i \in I} p_i q_i} \qquad I = 1, 2, ..., 17$$

where P_I is the weighted group price for group I, and

p_i is the price of the i'th item in group I, and

q_i is the quantity consumed on commodity i in group I.

In calculating the index number, as a basic rule, quantity weights are used to weigh together average prices (of goods of the same kind) whereas value weights relating to the reference base period are used to weigh together price relatives or index number (from different kinds of goods)[2]. For the same reason, since the sub-items in a group are not necessarily of the same kind, the quantities are not directly comparable. Using quantity weights seems implausible and inappropriate, therefore the value weights have been preferred when measuring group prices.

Note that it is quite common for an agricultural household to have a combine production and consumption decision. As pointed out by Strauss(1984), households may produce commodities solely or partly for their own consumption and sell the surplus. Therefore the "quantity consumed" and "quantity purchased" differ. For our purposes, we record quantities expenditure, the amount and value of the household's own production contributing to its total consumption. Hence, the household's production for its own consumption is added to the amount of quantity and value purchased.

In order to reduce the complexity of the data, only part of the household information is selected. The variables are region (urban or rural), subround (quarter), province, Kabupaten, total number of household members, total number of adult household members, household income and expenses. Every household record contains the household information listed above and 17 commodity group prices and quantities consumed during the survey week. The SUSENAS tapes recorded about 59,000 different households[3]. Zero group prices are possible as households need not consume each and every food group during the survey week.

Obviously, the sample size at this stage is far from manageable in part because individual household expenditure information is highly disaggregated. To reduce the dimension of the data matrix and deal with the problem of zero expenditures on individual items in the recording period, we aggregate individual households according to Kabupaten (village) and region (urban/rural), henceforth, referred as "location". All the household information, such as number of members, number of adult members, income, expenses, commodity group prices and quantities consumed, are aggregated without regard to their subround. While households' commodity group prices are averaged by the number of non-zero prices so as to generate representable group prices for each location, all the other aggregated figures are normalized by the number of households in that location. After aggregation the sample contains 424 observations and every observation can be regarded as a representative household of the corresponding location. Although the aggregation will average out some of the variation between the households within the same location, it should still reflect the general characteristics of individual household's consumption behaviour.

The notion of a representative consumer has to be stressed again. As emphasized by Hicks(1956):

> "The statistical information on consumers' behaviour, which is available to us, always relates to the behaviour of groups of individuals - such, for instance, as the consumers of a particular commodity in a particular region. It is always material of this chapter which we have to test: and indeed it is material of this kind which

we want to test, for the preference hypothesis only acquires a *prime facie* plausibility when it is applied to a statistical average. To assume that the representative consumer acts like an ideal consumer is a hypothesis worth testing; to assume that an actual person, the Mr. Brown or Mr. Jones who lives round the corner, does in fact act in such a way does not deserve a moments' consideration."

This argument is theoretically hard to justify (Deaton and Muellbauer[1980a]). However, when considering the practical difficulty in empirical analysis, we may have to live with this weakness even it is theoretically vulnerable.

Another question is the treatment of zero prices. Even though prices are averaged, some zero group prices still exist. The data reveal that some of the commodity groups, for example chicken and alcohol are not consumed in some locations during the survey week. To eliminate those zero prices, they are replaced by the closest non-zero substitute. Following the geographical hierarchy, the replacement value of each zero locational price is searched level by level, beginning from the regional (urban/rural) price of the corresponding Kabupaten. If that is zero then the average mean price of that Kabupaten is used. If that is zero again the mean price of the corresponding province is used. The justification for grouping the items into 17 commodity groups is now considered.

While rice and grain are the most important food crops in most Asian countries, tubers are the second most fundamental crop type. Tubers are normally inferior to rice and grain in the sense that they are a secondary kind of crop and consumers apparently prefer rice and grain. It is important to examine the actual relationship between these two types of food.

Clearly, it is an advantage to possess a rich natural resource in marine products. Fish naturally is one of the most common foodstuffs in an island country such as Indonesia. Further, having dried fish is a common practice in most developing countries simply because of the ease of storage.

In Indonesia, rearing poultry, especially chicken, for private consumption is not uncommon. It is logical and acceptable to separate chicken from the meat group even though it is almost a standard practice to group them together in food analysis. The same argument is also applied to the separation between eggs and milk, and vegetables and legumes. Besides, the consumption of milk by Asians is generally less than that of Westerners (an outcome that is probably due to differences in custom and culture). Hence combining milk with eggs would bias the estimates as they are dissimilar goods.

In a tropical country like Indonesia, there is a favourable climate for the production and consumption of fruit. It is expected that fruit is one of the basic foods in daily consumption. Condiments, cooking oil, additives and prepared food each occupy only a minor portion of the total expenditure, but they are regarded as basic ingredients for daily cooking and diet.

As mentioned earlier, most of the population in Indonesia is Islamic (about 87 percent in 1980). Given their religion, most of the people do not drink alcohol. Hence it is wise not to put alcohol and tobacco in one group even though this is a common practice in demand analysis.

The descriptive statistics of each commodity group's expenditure are presented in Table 3.3. Note that rice occupies the largest share in total expenditure on food, accounting for almost 28%. Surprisingly, tobacco is the second largest with a share of 16%. This is followed by additives, vegetables, fish, etc. As expected, alcohol has the smallest share among all the food groups. Also, for reference purpose, the descriptive statistics of commodities' prices and quantity consumed are given in Table 3.4 and Table 3.5 respectively.

Footnotes

1. All statistical figures in this section are sourced from " *The Statistical Pocketbook of Indonesia, 1983* ".

2. Sourced from the Australian Bureau of Statistics' publication, "*A Guide to CPI* ", 1981.

3. After deleting the mistaken records, there are about 55744 households remained. I have to thank Dr. R.P. Byron for providing the original data tapes.

Table 3.1
Percentage of Area and Population Density of Indonesia, 1980

Island	(km^2) Area	Percentage of Total Area	Percentage of Total Population	(per km^2) Population Density
Sumatra	473,606	24.67	19.00	59
Java	132,187	6.89	61.88	690
Nusa Tenggara	88,488	4.61	5.76	96
Kalimantan	539,460	28.11	4.56	12
Sulawesi	189,216	9.85	7.05	55
Maluku & Irian Jaya	496,486	25.87	1.75	5
Indonesia	1,919,443	100	100	100

source : "*The Statistical Pocketbook of Indonesia, 1983*"

Table 3.2
Land Utilization for Residential Usage, by Island, 1980
(measured in hectare, 1 hectare ≈ .01 km^2)

Island	Household Compound	Percentage of Total Land	Total Land Area
Sumatra	1,668,698	3.6	45,851,234
Java	1,553,665	12.8	12,173,465
Nusa Tenggara*	125,919	1.9	6,579,269
Kalimantan	627,146	1.3	46,696,556
Sulawesi	567,592	3.03	18,723,862
Makulu & Irian Jaya	n.a.	n.a.	n.a.
Indonesia	4,543,020	3.5	130,024,000

* Timor Timur is not included
source : "*The Statistical Pocketbook of Indonesia, 1983*"

Table 3.3

Descriptive Statistics of Commodity Expenditures

No.	Group	Mean Expenditure	Standard Deviation	Minimum Expenditure	Maximum Expenditure	Budget Share %
1	Rice	2801.4	779.73	921.06	7000.4	27.825
2	Tuber	177.26	201.43	0.0000	2003.9	1.7606
3	Fish	839.23	721.02	0.0000	3493.7	8.3356
4	Dri. Fish	542.11	379.15	11.405	1987.6	5.3845
5	Meat	285.64	324.00	0.0000	2709.1	2.8371
6	Chicken	145.65	141.30	0.0000	1234.0	1.4467
7	Eggs	179.11	127.79	0.0000	674.37	1.7790
8	Milk	93.421	112.93	0.0000	804.45	0.9279
9	Vegi.	849.55	437.62	102.92	3078.8	8.4381
10	Legumes	279.15	195.75	0.0000	824.52	2.7725
11	Fruit	368.65	209.84	16.119	1456.8	3.6616
12	Condi.	300.60	135.43	78.989	116.82	2.9857
13	Cook. Oil	341.61	146.42	0.0000	1102.2	3.3930
14	Additive	1004.0	416.14	234.11	2753.8	9.9722
15	Pre. Food	232.25	574.60	0.0000	5769.6	2.3068
16	Alcohol	8.4115	29.619	0.0000	334.96	0.0836
17	Tobacco	1620.4	891.85	31.983	5184.9	16.095
Total Exp.		10068.	3542.0	3319.5	21992.	100.00

Table 3.4

Descriptive Statistics of Commodities Prices

No.	Group	Mean Price	Standard Deviation	Minimum Price	Maximum Price
1	Rice	2.2987	.31857	1.3357	3.3941
2	Tuber	1.2850	.67259	.25750	4.4231
3	Fish	6.8520	1.9980	2.3506	13.429
4	Dried Fish	2.0428	1.1848	.63090	8.0772
5	Meat	19.505	5.8234	5.0000	40.320
6	Chicken	14.956	4.2184	3.8333	31.667
7	Eggs	.75584	.18120	.43570	1.4825
8	Milk	12.692	2.8909	2.5000	24.356
9	Vegetable	1.3277	.39329	.65450	2.6918
10	Legumes	4.1704	1.1450	1.7441	9.4000
11	Fruit	2.0034	.63003	.53940	4.2454
12	Condiment	.72362	.19480	.21710	1.5995
13	Cook oil	5.6648	.69154	3.6681	7.7348
14	Additive	1.1248	.30363	.61080	3.4030
15	Pre. Food	.91374	.35628	.14410	4.4650
16	Alcohol	7.8805	2.4928	1.8055	27.091
17	Tobacco	1.0612	.44533	.24120	2.3308

Table 3.5

Descriptive Statistics of Quantity Consumed for each Commodity Group

No.	Group	Mean Quantity	Standard Deviation	Minimum Quantity	Maximum Quantity
1	Rice	1219.6	301.56	415.92	2767.6
2	Tuber	182.19	236.20	.00000	1576.9
3	Fish	131.03	127.68	.00000	1044.1
4	Dried Fish	273.58	163.98	5.6818	879.67
5	Meat	16.424	24.271	.00000	265.53
6	Chicken	9.6125	8.6520	.00000	56.250
7	Eggs	236.94	164.73	.00000	914.87
8	Milk	7.1063	8.2471	.00000	68.571
9	Vegetable	635.53	249.68	88.362	1859.6
10	Legumes	75.038	59.922	.00000	247.30
11	Fruit	188.86	111.90	13.073	1199.0
12	Condiment	427.22	183.76	124.78	1310.1
13	Cook oil	60.947	26.679	.00000	195.91
14	Additive	927.15	380.31	71.142	2313.7
15	Pre. Food	305.73	838.52	.00000	9383.4
16	Alcohol	1.1110	3.9040	.00000	51.754
17	Tobacco	1834.3	1208.0	15.125	6120.0

CHAPTER IV

MODEL ESTIMATION AND EVALUATION

Demand analysis is an intensively studied area in the economics literature, and there are many demand models developed and proposed by applied researchers. However not all of them are useful for this study. For example, the Translog and Rotterdam models are not suitable for the analysis, and reasons will be discussed later in this chapter. Three basic demand models are investigated. They are the Linear Expenditure System, the Double-Log System, and the Almost Ideal Demand System.

Estimation methods such as the Maximum Likelihood or the Zellner(1962) variant of Generalized Least Squares (Seemingly Unrelated Regression) are the most widely used estimators in demand analysis. Such simultaneous estimation approaches require the existence of a non-singular covariance matrix estimator. However as explained by Barten(1977), the budget constraint in equation (2.3) implies a linear dependence of the joint distribution of the disturbances if y and p are exogenous. Consequently, in theory, the covariance matrix is singular. This problem is usually solved by deleting an equation from the system. One can delete any of the n relations from equation (2.2) without losing information on the demand behaviour for the good. As well, the estimates are invariant to the particular equation which is selected for omission. Barten(1964) has shown that the resultant estimates from the Full Information Maximum Likelihood procedure are invariant with respect to the equation deleted[1]. However, the requirement that the estimated covariance matrix of the reduced system is non-singular implies that the number of observations be at least as large as the number of equations in the system, thus setting a natural upper limit to the degree of commodity disaggregation.

Differences in model specification, estimation method, definition of commodities and data compilation may help to explain the sometimes conflicting results. A general finding is that the parameters and estimated elasticities are very sensitive to and highly dependent on the model specified. Park(1969) estimated several different demand systems using the same data and found that the elasticities varied substantially with the model estimated. Kiefer and Mackinnon(1976) suggest such a difference may have two explanations. One is the structuralist reason that the demand systems which have been estimated simply do not describe the data adequately, either because the functional forms are too restrictive, or because aggregation across individuals, simultaneity with the supply side of the economy, durable goods and other complications have been ignored. The other explanation is stochastic; that the functional forms are in fact good enough to characterize the true model but the modest number of observations and fairly small relative price variation in the data series cause the estimates to be imprecise. They experimented with the Linear Expenditure System and Translog model using a simulation study with the same data (40 observations) and found that if the true system is not known, estimates are likely to be biased and unreliable. Their findings reinforced the structuralist explanation and suggested that demand systems such as the Linear Expenditure System or Translog models cannot be expected to perform well when the data were not generated by those models.

Another general finding by Klevmarken(1981) is that in most studies models which do not imply an additive utility function show a closer fit to the data than those in which such a function is implied. Also, the constant elasticity of demand system or the double logarithmic model usually obtain a relatively good fit. Generally, the use of expenditure shares as the dependent variable is more stable than using expenditure, volume, or rate of change in volume.

In comparative studies by Brown and Deaton(1972), Deaton(1974) and Theil(1975), it is argued that variables on the left hand side of the competing models should be comparable. In the present study, this recommendation is not totally followed because of the structural difference in their functional forms. Instead each model is judged in its most commonly applied form, which implies that the stochastic structure may vary from one model to another.

In this chapter several estimated models are reported and each of them is tested against the demand theory. The model which explains the data best will be chosen as the maintained model for further analysis. In order to validate the model, it is necessary to rely on some test statistics which are applicable to the multivariate situation. These hypothesis testing techniques are discussed below.

4.1 Hypothesis Testing

Engle(1984) argues that:

> " If the confrontation of economic theories with observable phenomena is the objective of empirical research, then hypothesis testing is the primary tool of analysis. "

This view is the fundamental objective of most applied demand researchers. Undoubtedly it is also one of the main goals of this paper, which is to test the validity of demand theory with reference to the data set described in the previous chapter.

Since a demand equation system deriving from maximizing a specific utility function subject to the budget constraint will automatically satisfy the basic demand conditions such as adding-up, symmetry and homogeneity, it is therefore impossible to test these hypotheses with the well known Linear Expenditure System (LES model). But there are many other demand models which enable us to test the theory. Among them is the Almost Ideal Demand System (AIDS model). As well, the Double-Logarithmic Demand System (DLOG model) is another candidate. Each of these models will be discussed in more detail later in this chapter.

The statistical tests commonly used in demand analysis are based on the Wald (W), Likelihood Ratio (LR) or Lagrange Multiplier (LM) principles. Since all of the three test statistics rely on Maximum Likelihood methods, only asymptotic properties can be expected for these tests. Note that Nonlinear Least Squares estimates are equal to Maximum Likelihood estimates if the error disturbances are normally distributed.

4.1.1 <u>The Wald, Likelihood Ratio and Lagrange Multiplier Tests</u>

The Wald test is based upon Wald's(1943) elegant analysis of the general asymptotic testing problem. It is the asymptotic approximation to the familiar t and F tests.

The Likelihood Ratio test is based upon the difference between the maximum of the likelihood under the null and under the alternative hypotheses. Wilks(1938) was the first to derive its general limiting distribution.

The Lagrange Multiplier test is derived from the constrained maximization principle and is based upon the Lagrange Multipliers by Aitcheson and Silvey(1958) and Silvey(1959). It is identical to that based upon the score as originally proposed by Rao(1948).

The three principles are based upon different statistics which measure the distance between the null and the alternative hypotheses. The Lagrange Multiplier test starts at the null hypothesis and evaluates whether movement toward the alternative would be a significant improvement, while the Wald approach starts at the alternative and considers movement towards the null. The Likelihood Ratio method compares the two hypotheses directly on an equal basis. Engle(1984) provided a clear and detailed explanation and comparison of the three statistical tests in the context of univariate model.

In the case of multivariate linear regression model, the three test statistics can be calculated as follows (Berndt and Savin

[1977]):

$$LR = T * \ln [|\tilde{\Omega}| / |\hat{\Omega}|]$$

$$W = T * tr [\hat{\Omega}^{-1} (\tilde{\Omega} - \hat{\Omega})]$$

$$LM = T * tr [\tilde{\Omega}^{-1} (\tilde{\Omega} - \hat{\Omega})]$$

where $\hat{\Omega}$ and $\tilde{\Omega}$ are the unrestricted and restricted variance covariance matrix of the residuals. The random disturbance are assumed to be independently and, identically normally distributed with zero mean vector and unknown but positive definite covariance matrix Ω.

The three test statistics all have the same limiting χ^2 distribution, and hence they employ the same asymptotic critical region. That is, they are all asymptotically distributed as Chi-square with k degrees of freedom, where k is the number of restrictions imposed. Under general conditions, it is proved that the properties of the three test statistics are asymptotically equivalent and are different only for finite samples (Engle[1984]). Also, on statistical grounds no one test procedure is uniformly most powerful against all alternatives. In fact, they all share the property of being asymptotically locally most powerful and invariant.

There is a well known numerical inequality among the test statistics, which was originally established by Savin(1976) and Berndt and Savin(1977). The relationship is:

$$W \geq LR \geq LM$$

Generally, the inequality holds for single equations or systems of linear equations, but its validity for non-linear models has not yet been established and appears unlikely to hold.

As illustrated by Berndt and Savin(1977) the Wald criterion has the largest size, followed by the Likelihood Ratio and Lagrange Multiplier criteria. The three criteria will have the same value only when the null hypothesis is exactly true in the sample. It implies that in practice there will always exist a level of significance for which these tests will yield conflicting inferences. However, when the null hypothesis is true, the dispersion between the test statistics will tend to decrease as the sample size increases.

4.1.2 GJ, an analog of the LR Test

Apart from the three common test statistics mentioned above, Gallant and Jorgenson(1979) have developed a test statistic, denoted as GJ, which is an analog of the likelihood ratio test in the case of a system of simultaneous and non-linear equations based on three stage least square estimators. The test statistic is also valid in the situation of non-linear SUR estimation.

The test statistic GJ is

$$GJ = (T^{*}\tilde{S}) - (T^{*}\hat{S})$$

where T is the number of observations,

$\hat{S}$ and $\tilde{S}$ denote the unrestricted and restricted minimums of the objective function $S(\theta)$ respectively,

In a non-linear SUR model

$$Y = f(\theta^{0}) + e$$

the objective function $S(\theta)$ is defined as

$$S(\theta) = (1/T) * (Y - f(\theta))' (\hat{\Omega} \otimes I) (Y - f(\theta))$$

GJ is distributed asymptotically as Chi-square with (r-s) degrees of freedom under the null where r and s are the numbers of estimated parameters in the unrestricted and restricted models.

As emphasized by Gallant and Jorgenson, the estimated covariance matrix $\hat{\Omega}$ must be fixed throughout when computing the test statistic, GJ. It means that $\hat{\Omega}$ of the unrestricted and restricted models are forced to be the same. The test statistic can be easily calculated using the econometric package SAS/ETS, version 5. The estimated covariance matrix is stored from the first regression of the unrestricted model, then carried to the second regression of the restricted model and fixed during estimation. The test statistic GJ will be equal to the difference of the two $T^{*}S(\theta)$

values generated from the estimations.

The four test statistics, the Wald, the LR, the LM and the analog of the LR (GJ), are reported frequently in the chapter. Each estimated model is now examined.

4.2 <u>The Naive Models</u>

According to basic demand theory, the quantity demanded for commodity i depends on its own price, other commodities' prices and consumers' income (or total expenditure). That is:

$$q_i = f(p_1, p_2, ..., p_i, ..., p_n, y) \qquad (4.1)$$

Deaton(1986) points out that even a linear function with an intercept is a flexible form and the linear model without parameter restrictions can offer a first-order Taylor approximation around at least one point. Expanding the function in (4.1) using a first-order Taylor series, gives

$$q_i = \beta_{i0} + \Sigma_j \partial q_i/\partial p_j (p_j - p_{j0}) + \partial q_i/\partial y (y - y_0) + e_i$$

where i,j = 1, 2..., n

By setting p_{j0} and y_0 equal to zero, the expansion becomes a first-order MacLaurin series and we may estimate the simple linear model of the following format

$$q_i = \beta_{i0} + \Sigma_j \beta_{ij} p_j + \alpha_i y + e_i \qquad (4.2)$$

where $i = 1, 2, ..., n$ and $n = 17$ in this case.

The technique used to estimate the model is the single equation Ordinary Least Squares (OLS) estimator. Consider a model:

$$Y_i = X_i \beta_i + u_i \qquad i = 1, 2, ..., n$$

where Y_i is a vector of sample observations of the dependent variable with dimension (T*1),

X_i is a non-stochastic design matrix of all explanatory variables with dimension (T^*K_i), and

β_i is a vector of unknown parameters with dimension (K_i*1)

u_i is an unobservable random disturbance vector with dimension (T*1) and is approximated by the residuals component e, where

$$\hat{e}_i = Y_i - X_i \hat{\beta}_i$$

and assuming the residuals possess the properties that

$E(e) = 0$, and $E(ee') = \sigma^2 \psi$.

There are n equations in the model. Thus,

$$Y = X\beta + u$$

where the dimensions of Y, X, β, and u are, respectively, (nT*1), (nT*K), (K*1), (nT*1), with $K=\Sigma_i K_i$. In this case T=424, n=17 and K_i=19 (17 prices, income and a constant term) for all i.

Assume that there is correlation between the error terms e in different equations. In time series studies, the correlation at a given point in time is known as contemporaneous correlation. There is no reason why this cannot be applied in a cross-sectional study with t referring to individuals. The assumption of contemporaneous disturbance correlation, but no correlation over time, implies that $E(e_{it},e_{js}) = \sigma_{ij}$ if t=s, but = 0 if t≠s. Alternatively, $E(e_i,e_j) = \sigma_{ij}I_t = \phi = \Omega \otimes I_t$, where Ω is the variance-covariance matrix of residuals and is positive definite and non-singular.

Applying the single equation OLS estimator, $\hat{\beta}_{OLS}$ to each commodity,

$$\hat{\beta}_{OLS} = (X'X)^{-1}XY$$

yields the minimum variance, linear unbiased estimator for each separate equation. Efficiency is gained by jointly considering all the equations using the Generalized Least Squares (GLS) estimator, $\hat{\beta}_{GLS}$, which is a wider class of linear unbiased estimators. GLS can improve the estimation because it both allows for the correlation between e_i and error vectors of the other equations and uses information on explanatory variables that are included in the system but excluded from a single equation.

By definition the linear GLS estimator is

$$\hat{\beta}_{GLS} = (X'(\Omega^{-1} \otimes I)X)^{-1} X'(\Omega^{-1} \otimes I)Y$$

Since in most applications Ω is unknown we can only apply the Estimated Generalized Least Squares (EGLS) estimator, $\hat{\beta}_{EGLS}$, which replaces Ω by $\hat{\Omega}$ where the estimator $\hat{\Omega}$ is based on LS residual $\hat{e}_i$. It has elements given by

$$\hat{\sigma}_{ij} = T^{-1} \hat{e}_i' \hat{e}_j \qquad i,j = 1, 2, \ldots, n.$$

The estimator $\hat{\beta}_{EGLS}$ is also frequently referred to as Zellner(1962)'s Seemingly Unrelated Regression (SUR) estimator, $\hat{\beta}_{SUR}$.

SUR will be more efficient than OLS provided that the correlation between the disturbances in different equations is not too low and the regression matrixes for different equations are sufficiently different (Judge, Hill, Griffiths, Luthepohl and

Lee[1982]). But in this case, $\hat{\beta}_{i,SUR}$, the i'th vector component of $\hat{\beta}_{SUR}$ is identical to $\hat{\beta}_{i,OLS}$ and there is no gain in efficiency because $X_1=X_2=\dots=X_m=X_0$ so that:

$$X = (I_m \otimes X_0).$$

Hence, the SUR estimator becomes:

$$\hat{\beta}_{SUR} = [\,(I\otimes X_0)'(\hat{\Omega}^{-1}\otimes I)(I\otimes X_0)\,]^{-1} * (I\otimes X_0)'(\hat{\Omega}^{-1}\otimes I))Y$$

$$= [\,\hat{\Omega}\otimes(X_0'X_0)^{-1}\,] * (\hat{\Omega}^{-1}\otimes X_0)Y$$

$$= (I\otimes(X_0'X_0)^{-1}X_0)Y$$

$$= (X'X)^{-1}X'Y$$

$$= \hat{\beta}_{OLS}$$

Therefore, in this case the SUR estimator, $\hat{\beta}_{SUR}$, is identical to the single equation OLS estimator, $\hat{\beta}_{OLS}$ and there is no gain in efficiency.

The simple linear model was estimated with single equation OLS and the results of the estimated model are presented in Table 4.1. The F statistics show that the explanatory variables are jointly significant in each equation but the R^2 statistics vary from 0.14 (prepared food) to 0.75 (tobacco). The estimated coefficients in the table are the derivatives of quantity consumed with respect to prices and income. We may then calculate price and income elasticities.

Define e_{ij} as the price elasticity of the demand for commodity i with respect to the price of commodity j, so that:

$$e_{ij} = (\partial q_i/\partial p_j) * (p_j / q_i)$$

Also let E_i as the income (total expenditure) elasticity of commodity i where:

$$E_i = (\partial q_i/\partial y) * (y / q_i)$$

Sometimes $E_i > 1$ and $E_i < 1$ are used to define luxuries and necessities, respectively. The justification will be discussed later when we look at the univariate Double-Logarithmic model. Also if $E_i < 0$, it means the purchase of the commodity declines absolutely (not just proportionally) as y increases. The commodity is referred to as an "Inferior Good".

The cross and own price elasticities and income elasticities (calculated at means) for the simple linear model are presented in Table 4.2. We may calculate the Slutsky Substitution Coefficient K_{ij} between commodity i and j using the "Fundamental Equation" in (2.7), and we can also measure the income effect. Both results are recorded in Table 4.3.

From the tables, it is encouraging to find that all the own substitution coefficients are significantly negative; that is, the negativity condition is satisfied comfortably. On the contrary, most of the elasticity figures are not significantly different from zero. For example, only rice's own price elasticity (uncompensated) is significantly negative and inelastic meaning that it is a normal and necessity commodity. This hypothesis is further supported by its significantly positive income elasticity which is less than one in absolute term. All the other commodities' own price elasticities are insignificantly negative. Milk is found to be quite abnormal as its own price is (insignificantly) positive meaning it has an upward sloping demand curve. But it is not inferior (hence not a Giffen good neither) since its income elasticity is non-negative. Besides, its own substitution effect and income effect are respectively -0.05 and -.011. The total price effect is therefore negative and contradictory to its positive own price elasticity. This unusual finding may be due to the simplicity of the model which cannot describe the data adequately. Another unusual finding is that the income elasticity for vegetables is significantly greater than one implying that it is a luxury good. With the exception of legumes, all goods have positive income elasticities indicating that they are normal goods. Legumes are an inferior good and a necessity because the income elasticity is negative (but insignificant) with an absolute value less than one. This is not too surprising as legumes are a basic kind of agricultural food crop in Indonesia. Once a family can afford better, it may change to other, higher quality, substitutes such as rice. As the total price effect of legumes is still negative, they are not "Giffen good". Also, fish, meat, chicken, milk, vegetables, prepared food, alcohol and tobacco are luxuries as their income elasticities are all greater than one (and insignificantly different from zero except for vegetables).

From the 136 pairs of cross substitution coefficients, 81 have inconsistent signs, 38 are positive (substitutes) where 18 of them are statistically significant, and 17 are negative (complements) where 8 of them are significant. There are some expected results from the figures. For example, rice and tubers are significant substitutes, as are rice and legumes. But the relationship between tubers and legumes is inconclusive. On the other hand, fish and dried fish are also significant substitutes. Interestingly fish and eggs, and dried fish and eggs, are both substitutes (and statistically significant too) meaning that fish, dried fish and eggs are all substitutes for each other. This implies an interesting question : If good A and B are substitutes, and good B and C are also substitutes, does it follow that A and C are also substitutes? While it seems logical to accept this hypothesis, theoretically it is not true. We cannot be conclusive without referring to the substitution coefficients. Empirical findings positively rejected this relationship. For instance we found that all fish, dried fish and egg, and fish, dried fish and legumes are two groups of significant substitutes. But eggs and legumes are (insignificant) complementaries. Further fish and legumes, and fish and prepared food are significant substitutes, but legumes and prepared foods are complements (statistically significant too).

Apart from the simple linear model, we may modify model (4.2) by taking the natural logarithm on each variable and estimate the equations with an univariate double-logarithmic demand model given by:

$$\ln q_i = a_{i0} + \Sigma_j e_{ij} \ln p_j + E_i \ln y + e_i \qquad (4.3)$$

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

where e_{ij} is the cross-price elasticity of the j'th price on commodity i, E_i is the income elasticity of i, and e_i is the error term (the term "ln" is referred as the natural logarithm).

The double-log demand model can be treated as a single equation approach to modelling commodity demand individually. This method has the great advantage of flexibility and is the best way of modelling the demand for an individual commodity in isolation. Martin and Porter(1985) applied the model to investigate the demand for meat in Australia. The model can be estimated by using the OLS procedure as explained above with the estimated results with reference to Indonesian data being presented in Table 4.4. Own price elasticities are again negative except for milk. Though it is statistically insignificant, it supports the earlier finding that milk has an upward sloping demand curve in the range of the data. Legumes are inferior as their income elasticity is (insignificantly) negative (income effect therefore is positive) which also confirms the results from the simple linear model.

The associated substitution matrix was calculated and is given in Table 4.5. Among 136 pairs of cross substitution coefficients, 76 are inconsistent in sign, 40 are substitutes and 20 are complements. Surprisingly only 2 pairs of those consistent substitutes are statistically significant. They are rice and additive, and additive and condiment. Negativity was not satisfied because milk has a positive own substitution effect of 0.1028 but is statistically insignificant. While its income effect is -0.015 (again insignificant), it is not inferior. This result is inconsistent with the previous result from the simple linear model. Also, the combined own price effect is 0.0878 which is positive, meaning its demand curve is upward sloping in the range of the data. This confirms the observed positive own price elasticity referred to earlier in the simple linear model.

If expressing (4.3) into budget share form, we have:

$$\ln w_i = \alpha_i + (E_i-1)\ln y + (e_{ii}+1)\ln p_i + \Sigma_{k\neq i} e_{ik} \ln p_k \tag{4.4}$$

We can see that the budget share of a good will increase (or decrease) with total expenditure y as E_i is greater than (or less than) unity. Therefore, if $E_i > 1$, it is defined to be luxury and is a necessity if $E_i < 1$. From Table 4.4, fish, meat, chicken, egg, milk, vegetables, fruit, prepared food and tobacco, are luxuries and significant which is also consistent with the finding from the simple linear model except that they are inconsistent before.

For a single equation demand function, we can test the homogeneity condition equation by equation, but not across equation restrictions such as the symmetry or the adding-up condition. Homogeneity requires the i'th equation to satisfy the condition in (2.5). To test homogeneity on a single equation model, it is necessary to impose the restriction:

$$\Sigma_j e_{ij} + E_i = 0$$

on the univariate double-log model. The condition can be tested using standard F test statistics. The results are shown in Table 4.6. Only 3 of the 17 equations significantly rejected the condition.

4.3 <u>The Linear Expenditure System</u>

Stone's (1954) Linear Expenditure System (LES model) is the system of demand equations derived from the well known Stone-Geary (or Klein-Rubin) utility function, which is:

$$\mu = \Sigma_i \beta_i \log(q_i - \gamma_i) \qquad (4.5)$$

with the normalizing assumption that $\Sigma_i \beta_i = 1$. The function is directly additive in nature.

Maximizing the utility function, μ, in (4.5) subject to the usual budget constraint, we can derive Stone's LES model in budget share form,

$$w_i = [p_i \gamma_i + \beta_i (y - \Sigma_j p_j \gamma_j)]/(\Sigma_i p_i q_i) \qquad (4.6)$$

where $w_i = p_i q_i/(\Sigma_i p_i q_i)$ and $i,j = 1, 2, ..., n$.

Stone's LES model is the first practical model to be based entirely upon the theory and it is the only type of the LES model globally compatible with the maximization of a classical utility function. It automatically satisfies the constraints of classical demand theory, i.e. homogeneity, adding-up, and symmetry.

The parameter γ_i's are often interpreted as minimum required quantities, or subsistence quantities, so that (4.6) has a very simple interpretation. It is that the committed expenditures $p_i\gamma_i$ are purchased first, leaving a residual "supernumerary expenditure", $(y - \Sigma_i p_i \gamma_i)$, which is allocated between the goods in a fixed proportion β_i. Also, the ordinary (uncompensated) price elasticities evaluated at the mean can be calculated as follows :

The own price elasticity e_{ii} is given by:

$$e_{ii} = [\ (1- \beta_i)\gamma_i \ / \ q_i \ ] \ -1$$

while the cross price elasticity e_{ij} is given by:

$$e_{ij} = - (\ \beta_i \ \gamma_j \ p_j) \ / \ (\ p_i \ q_i)$$

Further, the income elasticity for commodity i is:

$$E_i = (\ \beta_i \ / \ p_i\ ) \ ^* \ (\ y \ / \ q_i\ )$$

Obviously the ease of interpretation of the LES model makes it one of the most widely used demand models. There are, however, several problems with the model.

The first is that it has too few parameters to give it a reasonable chance of fitting the data. As pointed out by

Deaton(1986), the LES model does little more than fitting a bivariate regression between purchases and total expenditure.

The second problem is that the structure of the LES model does not allow complements or inferior goods. The non-satiety axiom of demand theory requires that:

$$\partial\mu/\partial q_i > 0$$

Applying this condition to the Stone-Geary utility function, we get:

$$\beta_i(q_i-\gamma_i)^{-1} > 0$$

The function is defined only if $q_i > \gamma_i$, therefore $(q_i-\gamma_i)^{-1}$ is positive. Hence, it requires that $\beta_i > 0$. The cost function of (4.6) is:

$$C(\mu,p) = \Sigma_i p_i\gamma_i + \mu\ \pi^k\ p_k{}^{\beta k} \qquad (4.7)$$

which is concave provided both that all β_i's are non-negative and q is not less than $\Sigma_i p_i\gamma_i$ so that $q_i \geq \gamma_i$ for all i, hence $(q_i-\gamma_i) \geq 0$. From the "Fundamental Equation of Value Theory" of Hicks(1946) in equation (2.7), the Slutsky substitution coefficient is:

$$K_{ij} = K_{ji} = \partial q_i/\partial p_j + q_j\ (\partial q_i/\partial y)$$

Applying this to the LES model,

$$K_{ij} = \beta_i\ (q_j-\gamma_j)\ /\ p_i$$

As $(q_i-\gamma_i)$, β_i and p_i are all positive, the substitution term K_{ij} is positive unambiguously. For concavity no two goods may be complements, so all goods must be substitutes. Also, inferiority can only occur for goods with β_i negative, but this violates the condition in the LES model and consequently rules out the existence of inferior goods.

The third problem of the LES model is that it is derived from a utility function belonging to the directly additive class. It implies that the marginal utility of one commodity is independent of every other commodity, that there are no specific substitution effects[2] and that the own price elasticity is approximately proportional to the income elasticity (Deaton[1974]). Houthakker(1960) has shown that under directly additive preferences, compensated cross-substitution effects are directly proportional to income derivatives.

All these disadvantages of Stone's LES model imply that the approach is very restrictive and, thus critical attention should be paid when applying the model.

Although Stone's LES model involves non-linear estimation, it is not difficult to implement. Broadly speaking, there are two standard estimating procedures we may use. Consider, for simplicity, a univariate non-linear model with additive error,

$$y = f(x,\beta) + e$$

where β is the parameter vector, and e is the random errors which are independently and identically distributed with mean zero and finite but unknown variance σ^2.

If the errors are normally distributed, we may apply the Maximum Likelihood estimation technique. The estimates are chosen such that the logarithm of a specific distribution function is maximized (or minimized). The likelihood function is based on the joint probability density of the sample and thus requires an exact knowledge of its distribution. Therefore, e is usually assumed to be normally distributed with mean zero and finite variance σ^2.

But if the exact form of the error distribution is unknown and we are not willing to assume normality, Non-Linear Least Squares (NLLS) estimation is an appropriate method. The estimates are chosen such that the sum of square error function, $S(\beta)$, is minimized.

We find that in a household budget data, the error distribution is highly non-normal. Therefore, normality is too strong *apriori* assumption to be made. Consequently, the least squares method is preferred.

The asymptotic properties of the NLLS estimator were rigorously examined by Jennrich(1969) and Malinvaud(1980). Jennrich(1969) proves that the NLLS estimator is asymptotically numerically stable and strongly consistent ($\hat{\beta}$ converging to the true values, β_0, almost surely). Malinvaud(1980), on the other hand, demonstrates its weak consistency ($\hat{\beta}$ converging to β_0 in probability). Weak consistency is more common in the econometric literature and is often called by the simpler name of consistency. The (weak) consistency of $\hat{\beta}$ is proven by showing that plim $T^{-1}S(\beta)$ is minimized at the true values β_0[3]. Strong consistency is proven by showing instead, that the same holds for the almost sure limit of $T^{-1}S(\beta)$. As was proved by Amemiya(1983), strong consistency implies weak consistency. Intuitively, it seems obvious that if $T^{-1}S(\beta)$ is close to plim $T^{-1}S(\beta)$ and if the latter is minimized at β_0, then $\hat{\beta}$, which minimized the former, should be close to β_0. Also, Jennrich(1969) rigorously proved the asymptotic normality of NLLS estimator.

Pindyck and Rubinfeld(1976) argue that, since q_i is a non-linear function of β_i, the linear least square theory is no longer valid. The statistical tests used to evaluate the linear regression equation are not directly applicable to a non-linear regression. The reason, as explained, is that we cannot obtain an unbiased estimate of σ^2. Even if the random disturbance term, u, is normally distributed with zero mean, the residual, e_t, given by:

$$e_t = y_t - f(x_t, \hat{\beta})$$

will not be normally distributed (nor will it have zero mean). Thus the sum of square residuals will not follow a Chi-square distribution. Hence, the estimated coefficients themselves will not be normally distributed, and standard t-tests and F-tests, therefore, cannot be applied. It can also be argued that since the true covariance matrix is unknown, instead of using t and F distributions for testing hypotheses, we should use, respectively, the standard normal and χ^2 distribution, because the t(T-k) distribution converges to the standard normal distribution and the F(J,T-k) distribution converges to a multiple (1/J) of the $\chi^2(J)$ distribution. In terms of asymptotic properties this is a valid argument, however, in terms of the finite sample performance of the tests the t and F distribution may be preferrable (Judge, Hill, Griffiths, Luthepohl & Lee[1982]). The arguments by Pindyck and Rubinfeld(1976) can be downplayed, given the findings of Amemiya(1983). He points out that in the process of proving asymptotic normality of NLLS estimator, it has shown that, asymptotically,

$$\hat{\beta} - \beta_0 \approx (G'G)^{-1}G'e$$

where $G = (\partial f/\partial\beta)_{\beta 0}$. Note that the condition exactly holds in the linear case,

i.e. $$\hat{\beta} - \beta_0 = (X'X)^{-1}X'e$$

where $X = (\partial f/\partial\beta)_{\beta 0}$. The practical consequence of the approximation is that all the results for the linear regression model are asymptotically valid for the non-linear regression model if we treat G as the regressor matrix. In particular, we can use the usual t and F statistics with an approximate precision.

Since there is in general no explicit formula for the NLLS estimation, the minimization of the objection function, $S(\beta)$, must usually be carried out by some iterative method. A brief description of the non-linear estimation techniques can be found from Goldfeld and Quandt(1972), Judge, Griffiths, Hill and Lee(1980), and Judge, Hill, Griffiths, Lutkepohl and Lee(1982). Two NLLS estimating procedures were tried for the LES model with both converging to the same solution[4]. These are considered below.

The first method was the standard non-linear estimation technique using the Gradient method and the optimization procedure used is the Newton-Raphson algorithm(or simply the Newton algorithm). It is an iterative procedure moving from one trivial set of coefficient values for β_i, say β_k, to a new set, say β_{k+1}, in such a way that the objective function, $S(\beta)$, is minimized. The objective function is the sum of square errors, which is:

$$S(\beta) = \text{Det}[\ y - f(x,\beta)\ ]'\ [\ y - f(x,\beta)\ ]\ /\ T \qquad (4.8)$$

where T is the number of observations and in this case T = 424.

The iterative procedure can be summarized as:

$$\beta_{k+1} = \beta_k + \zeta_k \qquad (4.9)$$

where $\zeta_k = -\ t_k\ d_k\ g_k$, and

ζ_k is known as step,

t_k is known as step length and is set to 1 for Newton algorithm,

d_k is direction and is equal to the inverse of the Hessian of $S(\beta)$ at β_k,

and g_k is the gradient, is equal to $\partial S/\partial\beta\ |_{\beta k}$.

The process iterates to convergence. Since the equations may have different local solutions, there is no guarantee that it will converge to a global minimum. Since different starting values may yield different solutions, using alternatives is one way to locate the global minimum. Two problems have been pointed out by

Maddala(1977). First, if the starting value is far from the minimum, the Newton-Raphson method may not converge and can keep moving in the wrong direction. The starting values are required to be sufficiently close to the true values (Jennrich[1969]). This can be resolved by estimating the model using a subset of the sample so that consistent and reliable starting values could be obtained. Second, the method requires the calculation of the second derivatives, the Hessian matrix, which causes substantial computing difficulty.

The estimation for Stone's LES model is constructed using a Fortran program of non-linear least squares procedure[5] on the Sperry Univac 1100 computer. The estimated model is presented in Table 4.7. Analytic derivatives were employed to calculate the gradient vector g_k and the direction matrix d_k. The equation for Tobacco (equation 17) was dropped to avoid singularity. Three different sets of starting values were arbitrarily chosen for β_i and γ_i. The first set of values for β_i and γ_i being respectively 1/17 and the averaged mean quantity consumed for all i. The second set of values are 1/17 and the minimum non-zero quantity consumed for i. The final set are fixed at the values 0.5 and 100, respectively, for all i. Importantly, all three set of values converged to the same solution. The cross and own price, and income elasticities are given in Table 4.8. All own price elasticities except legumes are negative. Legumes are insignificantly positive and elastic. Income elasticities are positive implying that all goods are normal (as mentioned earlier, no inferior goods can be included in the LES model). The luxury items included tubers, fish, dried fish, meat, chicken, egg, milk, vegetables, fruit, cooking oil, additive, prepared food, alcohol and tobacco, except that rice and condiment are necessities. Among them, only rice, vegetable, condiment and additive are statistically significant.

Table 4.9 shows the substitution and income effect of the LES model. All own substitution coefficients are negative therefore satisfying the negativity condition as expected. But only one of them is statistically significant. Cross substitution coefficients are all positive (they are not reported in the table) indicating the goods are all substitutes for each other, which supports the theory that no complementary items can be included in the model.

Apart from the Gradient method, the technique using Step-wise Least Squares procedure was used[6] (it is similar to the earlier approach used by Stone(1954) to estimate the LES model). This iterative procedure is, basically, divided into two steps, examined below.

The first step is to assume γ_i are known parameters and to estimate β_i for the n-1 equations. The model is:

$$(p_i q_i - p_i \gamma_i) = \beta_i \,(y - \Sigma_i p_i \gamma_i) + u_i$$

As the regressors on the right hand side are the same for each equation, single equation OLS will suffice.

If $\hat{\beta}_i$ and $\hat{\Omega}$ are estimated for β_i and the residual variance-covariance matrix, the second step is to estimate γ_i given these results. The model is:

$$(p_i x_i - \beta_i y) = [\, p_i(1-\beta_i) \,]\, \gamma_i + \Sigma_{j \neq i}\, \beta_i p_j \gamma_j + u_i$$

or $\quad A_i = W_i \gamma + u_i \qquad$ where i = 1,..., n-1

Suppose there are only two equations to be estimated, then

$$A_1 = W_1 \gamma + u_1 \text{ ,and}$$

$$A_2 = W_2 \gamma + u_2$$

The Generalized Least Squares estimator for γ is:

$$\gamma = [\, (W_1, W_2)\, \Omega^{-1}\, (W_1, W_2)' \,]^{-1} \, [\, (W_1, W_2)\, \Omega^{-1}\, (A_1, A_2)' \,]$$

The two steps iterative process continued until the solutions converged. The results derived from this estimator are not presented as the method converged to the same solution obtained from the Gradient method. Both estimators generate identical results. But the Step-wise Least Squares procedure is computationally much faster. In this particular example, it took about 25 minutes (CPU time) to estimate the LES model using the Gradient method. But with the step-wise process, it took only about 4 minutes (CPU time).

The LES model was also estimated by the Seeming Unrelated Non-Linear Regression estimator (non-linear SUR) using a common statistical package, the SAS/ETS version 5, on the FACOM M360 main frame computer. The FACOM is a virtual memory machine and is powerful enough to handle problems with a maximum of about 9000K whereas the SAS/ETS package is a fairly common statistical software which is able to solve systems of both linear and non-linear regression models.

The Non-linear SUR estimator, with a standard Gauss-Newton algorithm chosen as the optimization procedure, is an extension to Zellner's(1962) SUR method with the exception that non-linear response functions and non-linear parametric constraints are incorporated. The estimator is strongly consistent, asymptotically normally distributed (Ω unknown), and asymptotically more efficient than the single equation least squares estimator (Gallant[1975]).

Consider the estimation of the parameters of a set of n non-linear regression equations with additive error disturbances, with the responses contemporaneously but not serially correlated. That is:

$$y_{ti} = f_i(x_{ti}, \beta_i) + e_{ti} \qquad i=1,\dots, n \text{ and } t=1,\dots, T$$

The n-variate errors are assumed to be independent, each having a mean of zero, the same distribution function (not necessary normal) and positive definite variance-covariance matrix (Ω). In other words, the errors are independent but identically distributed. The estimation procedure of non-linear SUR is as follows:

(1) The first step is to obtain the least squares estimator $\hat{\beta}_i$ by minimizing

$$S_i(\beta_i) = 1/T * [y_i-f_i(\beta_i)]'[y_i-f_i(\beta_i)]$$

equation by equation.

(2) The second step is to form the residual vectors

$$\hat{e}_i = y_i - f_i(\hat{\beta}_i)$$

and to obtain the estimate of the residual covariance matrix $\hat{\Omega}$ by estimating the elements σ_{ij} where

$$\hat{\sigma}_{ij} = [1/(T-k)] * \hat{e}_i'\hat{e}_j \qquad i,j = 1,\dots, m$$

and k is the number of parameters

(3) The third step is to estimate the system

$$y = f(\beta) + e$$

in a single regression using the Aitken type estimator $\tilde{\beta}$ by minimizing

$$S(\beta) = 1/T * \{[y-f(\beta)]'[\hat{\Omega}^{-1} \otimes I][(y-f(\beta)]\} \qquad (4.10)$$

where $y = (y_1', y_2', ..., y_m')'$ which is (Tm*1),

$f(\beta) = [\ f_1'(\beta_1), f_2'(\beta_2), ..., f_m'(\beta_m)]'$ which is (Tm*1),

and $e = (e_1', e_2', ..., e_m')$ which is (Tm*1).

(4) The final step is to obtain an estimate of the covariance matrix of parameter estimates Ψ,

$$\Psi = (1/T)\ [\ F'(\tilde{\beta})\ (\hat{\Omega}^{-1} \otimes I)\ F(\tilde{\beta})\ ]$$

where $F(\tilde{\beta}) = \partial f_i(x_i, \beta_i)/\partial \beta_i$

The optimization procedure adopted is the Gauss-Newton Method (or simply the Gauss Method). In this algorithm the Hessian matrix is approximated by:

$$Z(\beta)'\Omega Z(\beta)$$

where $Z(\beta) = \partial f/\partial \beta$. Its iterative principle is similar to the Gradient Method with Newton Algorithm in (4.9) except that:

(1) the step length, t_k, is set to 1/2

(2) the direction matrix is

$$d_k = [\, Z(\beta_k)'\Omega Z(\beta_k) \,]^{-1}$$

$$\text{where } Z(\beta_k) = \partial f/\partial\beta|_{\beta k} \qquad \text{and } \Omega = I_T$$

(3) the gradient

$$g_k = \partial S/\partial\beta|_{\beta k} = -Z(\beta_k)'[\, y\text{-}f(x,\beta_k) \,]$$

The Gauss-Newton method also involves extra computational effort in computing and inversing the Hessian matrix $Z(\beta)'\Omega Z(\beta)$.

To estimate Stone's LES model, previous starting values were tried. Also a trial run based on the first 100 observations was tried in order to obtain a set of reliable starting values. Interestingly, both results converged to the same solution point which was different from the previous convergence point. The most likely explanation of this result is the difference in the specification of the objective function, $S(\beta)$, since equation (4.10) is different from (4.8). This can change the optimal profile when minimizing the objective function. Hence, even though the two approaches yield global minima their optimal solutions are different. This possibility was confirmed by substituting the optimal solution from one objective function into the other. As expected, the resulting solution is exactly the same as the one generated previously by the function. The estimates of the model using SAS/ETS are presented in Table 4.10. Comparing the results with Table 4.7, it was found that the beta coefficients are fairly similar but the gamma values are different. Apart from the large difference in magnitude, opposite signs are observed for vegetables and tobacco.

The associated elasticities, substitution and income effects are given in Table 4.11 and Table 4.12, respectively. The significance of own price elasticities is greatly improved as more than half of the equations are significantly different from zero. The demand curve for legumes is once again found to be upward sloping (but statistically insignificant). Income elasticities are all positive (hence income effects are all negative) but only few of them are significant. Negativity is also satisfied. As expected, no goods are inferior or complementary. Almost all commodities are luxury, except for rice and condiment which are necessities. Only rice, vegetables, condiments and additives are statistically significant. Comparing with Table 4.8 and Table 4.9, results are largely consistent except the own price elasticity for legumes is now positive. The difference in the estimated coefficients is due to the difference in estimation techniques. One is the standard Gradient method using Newton-Raphson algorithm and the other is non-linear SUR estimator using Gauss-Newton method. Most important is the difference in the specified objective functions. Therefore, the conflicting results occur because of the differences in objective function used. Since the SAS/ETS package will be used extensively in this study, for the sake of consistency, the non-linear SUR results are preferred.

4.4 The Double-Logarithmic Demand System

In the early section of this chapter, the results of the single equation double-log demand functions are reported. Although it is possible to test the homogeneity restriction equation by equation, it is structurally impossible to test across equation restrictions for the symmetry, adding up and the system-wide homogeneity conditions. Hence it is necessary to reconstruct the demand function in (4.3) and to set up a system of double-log demand models which allows a test of general demand theory[7].

The model in (4.3) is:

$$\ln q_i = \alpha_{i0} + \Sigma_j\, e_{ij} \ln p_j + E_i \ln y$$

Following Byron's(1968) approach, we multiply (4.3) by budget share w_i on both sides and get:

$$w_i \ln q_i = \alpha_{i0} w_i + \Sigma_j\, e_{ij} w_i \ln p_j + w_i E_i \ln y$$

then add and subtract the term, $\Sigma_j\, E_j w_i w_j \ln p_j$, on the right hand side, we have:

$$w_i \ln q_i = \alpha_{i0} w_i + \Sigma_j (e_{ij} w_i + w_i w_j E_i) \ln p_j + w_i E_i \ln y -$$
$$\Sigma_j w_i w_j E_i \ln p_j$$

or $$w_i \ln q_i = \alpha_{i0} w_i + \Sigma_j a_{ij} \ln p_j + b_i (\ln y - \Sigma_j w_j \ln p_j) \tag{4.11}$$

where $a_{ij} = e_{ij} w_i + w_i w_j E_i$ and $b_i = w_i E_i$.

The model in (4.11) is the unrestricted multivariate double-log demand system (DLOG model). Although the system is linear in nature, it was estimated using the non-linear SUR estimator from the SAS/ETS package. The reason is that across equation restrictions can then be imposed. Also the GJ test statistic (analog of the LR) is only available in the non-linear option. The estimated results are presented in Table 4.13.

We may now test demand theory based on the DLOG system. The adding-up condition requires the sum of the marginal propensity to consume (or the marginal budget share) to be equal to one. Applying the DLOG model means that:

$$\Sigma_i w_i E_i = 1 \quad (\text{ or } \Sigma_i b_i = 1)$$

which can be incorporated into model (4.11) straight-forwardly. Next, the homogeneity condition requires that:

$$\Sigma_j e_{ij} + E_i = 0$$

Since $a_{ij} = e_{ij}w_i + w_i w_j E_i$, by taking the summation on both sides, we have:

$$\Sigma_j a_{ij} = \Sigma_j e_{ij} w_i + \Sigma_j w_i w_j E_i$$

$$= w_i \Sigma_j e_{ij} + E_i w_i \Sigma_j w_j$$

$$= w_i \Sigma_j e_{ij} + E_i w_i \qquad (\text{ as } \Sigma_j w_j = 1 \text{ })$$

$$= w_i(\Sigma_j e_{ij} + E_i)$$

$$= 0 \qquad (\text{ homogeneity requires } \Sigma_j e_{ij} + E_i = 0 \text{ })$$

The last restriction is the symmetry condition with the substitution coefficients K_{ij}. By definition:

$$K_{ij} = (\partial q_i/\partial p_j) + q_j(\partial q_i/\partial y)$$

As K_{ij} is invariant to monotonic increasing transformation[8], multiplying the equation by the term $(p_j/q_i)(y/p_j q_j)$, it becomes

$$K_{ij} = (\partial q_i/\partial p_j)(p_j/q_i)(y/p_j q_j) + q_j(\partial q_i/\partial y)(p_j/q_i)(y/p_j q_j)$$

$$\Rightarrow \qquad K_{ij} = e_{ij}/w_j + E_i$$

The symmetry condition

$$K_{ij} = K_{ji}$$

means that:

$$e_{ij}/w_j + E_i = e_{ji}/w_i + E_j$$

$$\Rightarrow \quad w_i e_{ij} + w_i w_j E_i = w_j e_{ji} + w_i w_j E_j$$

$$\text{or} \quad a_{ij} = a_{ji}$$

Each demand condition is tested using the techniques mentioned in section A of this chapter. The results in Table 4.14 show that each demand condition was rejected significantly. The joint validity of the adding-up, symmetry and homogeneity conditions cannot be empirically tested in the DLOG model due to insufficient memory. A maximum memory size of 9000K was tried using the FACOM M360 machine but it still failed. Since each demand condition was rejected separately, the joint test should also be rejected. The elasticities associated with the unrestricted model are reported in Table 4.15. While all own price elasticities are negative, income elasticities are positive except for tubers and alcohol which implies that the last two are inferior (but are insignificant). As well, no significant luxury goods can be found from the model.

The substitution and income effects are presented in Table 4.16. Sixty-five pairs have inconsistent cross-substitution coefficients, 38 are substitutes (of which 19 are statistically significant) and 32 are complements (only 6 are significant).

Negativity appeared to be satisfied as all own substitution coefficients are negative. Although tubers and alcohol have negative income elasticities (inferior), they are not "Giffen goods" because the negative own substitution effect is greater than the positive income effect. Therefore the demand curves are still downward sloping.

4.5 The Almost Ideal Demand System

The Almost Ideal Demand System (AIDS model) proposed by Deaton and Muellbauer(1980b) is a flexible model which gives an arbitrary first-order approximation to any demand system. The demand system is derived from a general cost function or expenditure function which is a flexible form and general enough to act as a second-order approximation to any arbitrary cost function. The cost function itself is derived from a specific class of preferences known as the PIGLOG class (Muellbauer[1975] & [1976]) which permits exact aggregation over consumers. These preferences are represented via the cost function $c(\mu,p)$ which defines the minimum expenditure necessary to attain a specific utility level at given prices. Define the PIGLOG class by:

$$\ln c(\mu,p) = (1-\mu) \ln a(p) + \mu \ln b(p)$$

Deaton and Muellbauer(1980b) take the specific functional forms for a(p) and b(p) as:

$$\ln a(p) = a_0 + \Sigma_k \alpha_k \ln p_k + 1/2^* \Sigma_k \Sigma_j \gamma_{kj} \ln p_k \ln p_j$$

$$\ln b(p) = \ln a(p) + \beta_0 \Pi_k p_k{}^{\beta k}$$

From the fundamental property of the cost function its price derivatives are the quantity demanded

$$\partial c(\mu,p)/\partial p_i = q_i$$

By multiplying both sides by $p_i/c(\mu,p)$, we find that:

$$\partial \ln c(\mu,p)/\partial \ln p_i = p_i q_i/c(\mu,p) = w_i$$

where w_i is the budget share of good i. Hence the demand function in budget share form derived from the cost function is:

$$w_i = \alpha_i + \Sigma_j \gamma_{ij} \ln p_j + \beta_i \mu \beta_0 \Pi p_k{}^{\beta k} \qquad (4.12)$$

Since μ can be expressed as a function of p and y (the indirect utility function), by substituting that into (4.12) we can derive the AIDS demand functions in budget share form,

$$w_i = \alpha_i + \Sigma_j \gamma_{ij} \ln p_j + \beta_i \ln(y/P) \qquad (4.13)$$

where P is a price index defined by:

$$\ln P = \alpha_0 + \Sigma_k \alpha_k \ln p_k + 1/2^* \Sigma_k \Sigma_j \gamma_{kj} \ln p_k \ln p_j$$

Given that the α_i parameters act as intercepts, the AIDS model can thus provide a local first-order approximation to any true demand system, whether derived from utility maximizing behaviour or not.

The adding-up condition requires that:

$$\Sigma_i\alpha_i = 1, \quad \Sigma_i\gamma_{ij} = 0, \quad \Sigma_i\beta_i = 0$$

Since, by construction, the data add up automatically, this condition is not testable. Since the sum of the budget shares equals one, it follows that the contemporaneous covariance matrix is singular, a problem which is solved by deleting one redundant equation. In this case equation 17, the demand for tobacco, is dropped. Its coefficients can be recovered from the rest of the estimates.

The homogeneity condition requires that:

$$\Sigma_j\gamma_{ij} = 0$$

The last condition, symmetry, requires that:

$$\gamma_{ij} = \gamma_{ji}$$

In situations where prices are highly collinear, Deaton and Muellbauer(1980b) suggest that it may well be adequate to approximate P in (4.13) as proportional to some known index P^*, i.e. $P \approx \phi P^*$. One of the candidates is Stone's(1953) price index:

$$\ln P^* = \Sigma_k w_k \ln p_k$$

Then the model becomes:

$$w_i = (\alpha_i - \beta_i \ln \phi) + \Sigma_j\gamma_{ij}\ln p_j + \beta_i \ln (y/P^*) \qquad (4.14)$$

Let $\alpha_i^* = \alpha_i - \beta_i \ln \phi$. Note that as P were approximated by the known index P^*, the model is linear in parameters α, β and γ. It is shown by Deaton and Muellbauer(1980b) that the model in (4.14) provides an excellent approximation to the true model in (4.13). The approximated linear AIDS is commonly used in applied work. For example Ray(1980).

The Hicksian price and income elasticities for the AIDS model are[9]:

The own price elasticity

$$e_{ii} = -1 + (\gamma_{ii}/w_i) - \beta_i$$

, and the cross price elasticity

$$e_{ij} = (\gamma_{ij}/w_i) - \beta_i \ (w_i/w_j)$$

, and the income elasticity

$$E_i = 1 + (\beta_i/w_i)$$

Note that the changes in real expenditure operate through β_i coefficients. Hence, $\beta_i < 0$ for necessity (as $E_i < 1$) and $\beta_i > 0$ for luxury ($E_i > 1$).

Since we are dealing with 17 commodities, by following the true model in (4.13), we will have 17*17 = 289 cross price products in the true price index P. In order to reduce the complexity and the size of the model, we adopted the linear approximated model in (4.14).

The estimated coefficients for the unrestricted AIDS model are presented in Table 4.17. The model was first estimated using a standard linear SUR estimator[10]. Then the model was re-estimated using linear and non-linear SUR estimators from the SAS/ETS package. The estimates all converged to the same solution.

Since the adding-up condition is automatically satisfied when constructing the AIDS model, we can only test symmetry and homogeneity conditions. From Table 4.18, the test statistics are much larger than the critical values, and consequently the individual and joint restrictions are significantly rejected by the data. Especially, the test statistics are extremely large for across-equation restrictions which are about 5 times larger than the critical values. The elasticities for the unrestricted model are shown in Table 4.19. All own price elasticities except milk are negative. Milk has a positive but insignificant own price elasticity. It is not an inferior but a luxury because its positive income elasticity is greater than one. But it is insignificantly different from zero. However, its total price effect (substitution effect plus income effect) is insignificantly negative. This finding is the same as the result from the simple linear model mentioned earlier.

Generally, the income elasticities, except for meat and legumes, are all positive, indicating the goods are normal. Meat and legumes are inferior but insignificant. Luxury items include fish, milk, vegetables, prepared food, alcohol and tobacco. This results is supported by both values of β_i and E_i. Substitution coefficients and income effects of the unrestricted model are given in Table 4.20. The negativity condition is again satisfied though only 4 of them are significant. Among the 136 pairs substitution coefficients, 75 are inconsistent in signs, 45 are substitutes (only one of them is significant) and 16 are complements (all insignificant).

4.6 The Transcendental Logarithmic Model

This model was introduced by Christensen, Jorgenson and Lau(1975). They used a Taylor second order expansion series to approximate the negative of the logarithm of the direct utility function by a function quadratic in logarithm of the quantity consumed. That is:

$$- \ln U = \alpha_0 + \Sigma_i \ln q_i + (1/2) * \Sigma_i \Sigma_j \beta_{ij} \ln q_i \ln q_j \qquad (4.15)$$

Maximizing the utility function in (4.15) subject to the normal budget constraint:

$$\Sigma_i p_i q_i = y$$

yields the demand equations in budget share form of

$$w_i = (\alpha_i + \Sigma_j \beta_{ij} \ln q_j) / (\alpha_M + \Sigma_j \beta_{Mj} \ln q_j) \qquad (4.16)$$

where $j = 1, 2, ..., 17$

$$\alpha_M = \Sigma_k \alpha_k$$

$$\beta_{Mi} = \Sigma_k \beta_{ki}$$

The budget share equations can also be derived using the duality approach. In this case, they approximated the logarithm of the indirect utility function by a function quadratic in the logarithm of the ratio of prices to the value of total expenditure. If we replace the term " $\ln q_i$ " in (4.15) and (4.16) by " $\ln (p_i/y)$ ", we can then obtain the budget share equations corresponding to the indirect translog utility function. We can also test the validity of demand theory by imposing symmetry, additivity and homogeneity conditions.

However, several empirical problems should be noted. The number of parameters increase rapidly with the number of goods with the actual number of parameters to be estimated in an unrestricted Translog model of n commodities being (n*n)+(n-1). For a demand system of 17 commodities (only 16 are estimated because one of them is redundant), there will be 305 unrestricted parameters.

Based on a simulation exercise, Guilkey(1978) shows that the accuracy of the Translog approximation was promising as its performance deteriorates by only a small amount when the underlying true functional form becomes more complex. On the contrary, Kiefer and MacKinnon(1976) found that the Translog function could not approximate the priori specified underlying the LES function suggesting that the flexibility of the Translog function is doubtful. As well, the quality of the Translog approximation deteriorates rapidly with increased dispersion in the regressors (Wales[1977]). On the other hand, an efficient parameter estimator is generally more difficult to obtain if the variation in the independent variables is small. There is an obvious conflict between features desirable for approximation and efficient estimation. In this regard, the use of the Translog approximation may not be appropriate with cross-sectional data[11].

The next problem is due to the non-linear nature of the demand system. Unlike the AIDS model, the Translog system cannot be approximated by a linear model. When dealing with a large demand system of 17 commodities, the non-linear structure causes a heavy computational burden. It is practically impossible to estimate such a system using standard econometric packages. Test runs on the full model using respectively the SAS/ETS and SHAZAM on the FACOM

M360 and UNIVAC 1100 mainframe computers were unsuccessful due to insufficient memory. One possible solution is to use self-written computer programmes. However a difficulty is that the analytic derivatives of the Translog model are not easily defined, as a result, the Translog model was not used in this study.

4.7 The Rotterdam Model

The model was first proposed by Theil(1965) and Barten(1966)[12]. It is derived by total differentiating a general demand function for good i (i is assumed to be the first commodity without loss of generality), which is:

$$q_i = f(\ p_i,\ p_2,\ \ldots,\ p_m,\ y\ )$$

and let $k = 2, 3, \ldots, m$.

If we totally differentiate the model, we get:

$$d\ q_i = (\partial q_i/\partial p_i)\ d\ p_i + \Sigma_{k=2}^{m}\ (\partial q_i/\partial p_k)\ d\ p_k + (\partial q_i/\partial y)\ d\ y \qquad (4.17)$$

Since $d\ x_i = x_i\ d \ln x_i$, (4.17) then becomes:

$$q_i\ d \ln q_i = (\partial q_i/\partial p_i)(p_i\ d \ln p_i) + \Sigma_k\ [(\partial q_i/\partial p_k)(p_k\ d \ln p_k)] + (\partial q_i/\partial y)(y\ d \ln y)$$

By multiplying both sides by the term (p_i/y), we have:

$$w_i \; d \ln q_i = (\partial q_i/\partial p_i)(p_i^2/y) \; d \ln p_i + \Sigma_k \, [(\partial q_i/\partial p_k)(p_k p_i/y) \; d \ln p_k \,]$$

$$+ (\partial q_i/\partial y) \; p_i \; d \ln y \qquad (4.18)$$

Expanding k, then add and subtract the terms "$q_j(\partial q_j/\partial y)(p_i p_j/y) \; d \ln p_j$" with $j = 1,2,..,m$. Equation (4.18) becomes:

$$w_i \; d \ln q_i = (\partial q_i/y)(p_i \; d \ln y) + (\partial q_i/\partial p_i)(p_i p_i/y)d \ln p_i$$

$$+ q_i(\partial q_i/\partial y)(p_i p_i/y)d \ln p_i - q_i(\partial q_i/\partial y)(p_i p_i/y)d \ln p_i$$

$$+ (\partial q_i/\partial p_2)(p_i p_2/y)d \ln p_2 + q_2(\partial q_i/\partial y)(p_i p_2/y)d \ln p_2$$

$$- q_2(\partial q_i/\partial y)(p_i p_2/y)d \ln p_2 + \; \ldots$$

$$+ (\partial q_i/\partial p_m)(p_i p_m/y)d \ln p_m + q_m(\partial q_i/\partial y)(p_i p_m/y)d \ln p_m$$

$$- q_m(\partial q_m/\partial y)(p_i p_m/y)d \ln p_m$$

Note that the term

$$(\partial q_i/\partial p_j)(p_i p_j/y)d \ln p_j + q_j(\partial q_i/\partial y)(p_i p_j/y)d \ln p_j$$

$$= [(\partial q_i/\partial p_j) + q_j(\partial q_i/\partial y)] \; (p_i p_j/y)d \ln p_j$$

$$= K_{ij} \; (p_i p_j/y)d \ln p_j$$

where $j = 1, 2, ..., m$ and K_{ij} is the substitution coefficient.

Let $B_i = p_i(\partial q_i/\partial y)$. The equation becomes:

$$w_i \; d \ln q_i = B_i \; d \ln y + K_{ii}(p_i p_i/y)d \ln p_i + K_{i2}(p_i p_2/y)d \ln p_2$$

$$+ \ldots + K_{im}(p_i p_m / y) d \ln p_m - B_i(p_i q_i / y) d \ln p_i$$

$$- B_i(p_2 q_2 / y) d \ln p_2 - \ldots - B_i(p_m q_m / y) d \ln p_m$$

Let $S_{ij} = K_{ij}(p_i p_j / y)$. The equation finally becomes:

$$w_i \, d \ln q_i = B_i[\, d \ln y - \Sigma_j(w_j \, d \ln p_j) \,] + \Sigma_j(S_{ij} \, d \ln p_j) \qquad (4.19)$$

which is the Rotterdam model.

The Rotterdam model is a first-order approximation to any unknown demand system. For example, it can be derived from the Stone-Geary utility function.

By treating B_i and S_{ij} as constants, the model becomes linear and can be estimated by OLS equation by equation (Phlips[1973]). But as pointed out by Byron(1984), this linearization assumes B_i and S_{ij} to be constants but they can vary and differ at each observation. This results in the exclusion of terms from the equations which leads to biased and inconsistent estimation, although the bias in the parameter estimates is minimal if the coefficients are relatively constant.

Also, the Rotterdam model is derived from a differential equation and is a discrete approximation to a continuous time model. Therefore, we consider the model is theoretically implausible in a cross-sectional survey. However, Theil & Clements(1987) have applied it to international cross-section data. As pointed out by

Deaton(1974), the model may be integrated into any utility function over a sufficiently small time period and also when infinitesimal changes can be observed. This may probably be another practical weakness of the model.

4.8 Model Evaluation

To compare different demand systems is not straightforward. As efficiency will not be achieved in unrestricted models, we may use single equation R^2s when comparing models, a common practice by demand researchers. Or we may compute the system R^2 statistic. By definition, the system R^2 is:

$$SR^2 = 1 - (|e'e| / |y'y|)$$

Both single equation R^2 and system R^2 are not directly comparable as they do not account for the differences in degrees of freedom. Also they can not be used to evaluate models with different dependent variables. It means that we need an indicator to measure the goodness of fit for a demand system as a whole. Theil(1971) proposed a measure known as "Average Information Inaccuracy", $\hat{I}$, which is calculated as:

$$\hat{I} = \Sigma_t I_t/424$$

where $I_t = \Sigma_i w_{it} \ln (w_{it}/\hat{w}_{it})$

w_{it} and $\hat{w}_{it}$ are, respectively, the true and estimated budget shares of commodity i for the t'th observation. But since $\hat{I}$'s are not directly comparable, correction has to be made for the loss of degrees of freedom. As well, the statistic is biased as it does not distinguish the effects of over-prediction and under-prediction. Over-predicted budget shares will cause the logarithmic terms to become negative and vice versa for under-prediction. The two opposite effects will partially average out the total effect. Theil(1971) also suggested the "Corrected Average Information Inaccuracy", $I\hat{C}$ which is:

$$I\hat{C} = (1/2T)^{*} \ \Sigma_t \left[\Sigma_i \ e_{it}^{2} / \ w_{it} \right]$$

where $e_{it} = (\hat{w}_{it} - w_{it})$

The corrected information inaccuracy measures eliminate the averaging effect mentioned earlier and, therefore, can be used as a general indicator when evaluating demand systems. Naturally, the smaller the inaccuracy, the better the fit.

Also average information inaccuracy can be computed for each individual commodity. Denoting $I\hat{G}_i$ as the average information inaccuracy for group i, it is calculated by:

$$I\hat{G}_i = (\Sigma_t \ I_{it}) \ / \ T$$

where $I_{it} = w_{it} * \ln (w_{it}/\hat{w}_{it}) + (1-w_{it}) \ln [(1-w_{it})/(1-\hat{w}_{it})]$

To discriminate between different demand models, we may use a non-nested procedure such as the Cox test statistic proposed by Pesaran and Deaton(1978). The test was first applied by Pesaran(1974) to single equation linear regression models which was then extended to multivariate non-linear models by Pesaran and Deaton(1978), and applied to demand analysis by Deaton(1978). It is explained as follows:

Suppose that there are two competing hypotheses,

$$H_o : y_t = f_t(\beta) + e^0{}_t$$

$$H_1 : y_t = g_t(\gamma) + e^1{}_t$$

where $e^0{}_t$ and $e^1{}_t$ are assumed to be serially independent and to be multi-normally distributed as $N(0,\Omega_0)$ and $N(0,\Omega_1)$ under H_0 and H_1, respectively. The Cox test statistic N_0 is computed by first calculating

$$T_0 = (T/2) \ln (\det \hat{\Omega}_0 / \det \hat{\Omega}_{10})$$

where $\hat{\Omega}_{10}$ is the estimated expectation of Ω_1 under H_0. It can be computed by an auxiliary regression in which the predicted values from H_0 are used in H_1. The resulting covariance matrix is added to

$\hat{\Omega}_0$ which will give $\hat{\Omega}_{10}$. Then T_0 is normalized by dividing by its variance. The resulting Cox test statistic N_0 is asymptotically distributed as N(0,1), but the calculation of the variance of T_0 is complicated and interested readers should refer to the original papers. An example on the application of the Cox tests can be found in Murray(1984) who investigated the retail demand for meat in Australia.

Davidson and MacKinnon(1980) also proposed a simpler test statistic P_1 using an artificial compound model which is:

$$y_{it} - \hat{f}_{it} = \hat{F}_{it}'B + \alpha \hat{h}_{it} + e_{it}$$

where $\hat{h}_{it}$ is an element of

$$\hat{h} = \hat{\Omega}_0 \hat{\Omega}_1^{-1} (\hat{g} - \hat{f})$$

where $\hat{g}$ and $\hat{f}$ are the predicted values from the two competing models. $\hat{F}_{it}$ denotes the derivatives of $f_{it}(\beta)$ with respect to β, evaluated at $\hat{\beta}$. The test statistic P_1 is the t-statistic on $\hat{\alpha}$. Davidson and MacKinnon(1980) proved that the test statistic is asymptotically distributed as N(0,1) under H_0, and showed it to be asymptotically equivalent to the Cox test[13].

As well as the Cox test statistic being complicated to compute, the P_1 test is only discriminating a particular equation in the regression models and does not test the competing models as a whole. Also, they both require the dependent variables of the competing models to be the same. While the LES model can be expressed in either quantity or budget share form, the AIDS model is in budget share form and the DLOG model is a product of budget share and logarithm of quantity. Hence their usefulness for this particular demand analysis is doubtful.

Even though it is practically possible to test between the LES and AIDS, there is little theoretical justification to do so. The LES is obtained from maximizing an arbitrary Stone-Geary utility function with the demand equations automatically satisfying all the demand conditions. Further, it is additive in nature, implying an approximate proportional relationship between the own price and income elasticities. On the other hand AIDS is derived from a flexible functional form and the resulting demand equations are first order approximation to the unknown true model. Besides, it is indirectly nonadditive (Blanciforti and Green[1983]). While the LES model is not testable, we can test the AIDS model against each demand condition except adding-up which is satisfied by construction. The difference in the nature of the functional form makes it of little value for model discrimination. Also, the restrictiveness of the LES model as explained in section B of the chapter concludes that the model is inferior to the AIDS model.

Since we cannot practically discriminate between the AIDS and DLOG models using non-nested procedures, we can evaluate their performance in term of the goodness of fit using the crude measures described earlier in this section.

Comparisons of individual information inaccuracy measures for each model are made in Table 4.21. The LES model is the worst among the three models. The DLOG model performs better than the AIDS model for most of the equations except milk and alcohol. While the information inaccuracy for the unrestricted DLOG and AIDS models are presented in Table 4.14 and Table 4.18, the average and corrected information inaccuracy for the LES model are 100.3 and 533.4, respectively. Although the DLOG model has the lowest average information inaccuracy in absolute terms, after correction, the AIDS model is far superior to the other two models in terms of corrected inaccuracy measures.

The AIDS and DLOG models were again compared with the demand restrictions imposed. Based on the corrected information inaccuracies reported previously, the AIDS model once again shows superiority to the DLOG model even both restricted models rejected the demand theory.

Moreover, the AIDS model also possess advantages over the other two on theoretical issues. The LES model is too restrictive in the sense that no complementary or inferior goods can be included. The DLOG model belongs to a class of constant elasticity of substitution which implies the Engel curves will be straight lines from the origin. Also the double-log demand system is consistent with utility maximization only if the utility function is linear logarithmic. But the AIDS model does not have these limitations and its flexible form allows first-order approximation to any demand system. Also, it is a special case of the PIGLOG class which allows perfect aggregation over individuals. It is simple to estimate (in the linear approximation form) and provides a tool to test the validity of demand theory. For these reasons, the AIDS model will be chosen as the maintained model.

4.9 Size Correction

Homogeneity and symmetry conditions are frequently rejected in applied demand studies. Deaton(1972) argues that symmetry is "fundamentally a weak hypothesis" so that rejection of these restrictions is "intuitively implausible" (Bewley[1983]). It would appear that the reason for rejecting such restrictions would be because of the inappropriate nature of the test statistics. Laitinen(1978) demonstrates that the standard test for homogeneity is seriously biased towards rejection. He argues that the large sample criteria is highly misleading and the test statistic is in fact distributed as Hotelling's T^2, which is distributed as a multiple "[(n-1)(T-n-1)]/(T-2n+1)" of F[(n-1),(T-2n+1)] degrees of freedom. Meisner(1979), using a simulation experiment, shows that the standard test statistic for Slutsky symmetry is also biased towards rejection of the null hypothesis, particularly for large demand systems. The bias as illustrated by Meisner(1979) is due to the use of the estimator $\hat{\Omega}$ rather than the true yet unknown covariance matrix Ω.

As asymptotic tests are biased toward over-rejection, a size correction has been suggested to adjust the critical values of asymptotic χ^2 test. The idea is to center the distribution of the observed test statistic by multiplying a correction factor so that the test statistic coincides more closely with its hypothetical distribution. Arbitrarily chosen factors have been proposed by different people. For example, Bohm, Rieder and Tintner(1980) suggested the factor (T-k)/T where k is the number of parameters in each demand equation. The rationale is as follows :

Suppose that there is a linear system of n equations:

$$Y = X\beta + U \qquad (4.98)$$

Y is (T^*n), X is (T^*k), β is (k^*n) and U is (T^*n) where $U \sim N(0,\Omega \otimes I_t)$.

Imposing m linear restrictions

$$R\beta = g$$

where R is (m^*nk) and g is (m^*1).

As pointed out by Bewley(1983), if the covariance matrix Ω is known, the exact F test statistic is:

$$F = \{ \,tr[\Omega^{-1}(\tilde{U}'\tilde{U} - \hat{U}'\hat{U})]/m \,\} \,/\, \{ \,tr[\Omega^{-1}(\hat{U}'\hat{U})]/(T-k)n \,\} \qquad (4.99)$$

If Ω is unknown and replaced by the estimate $\hat{\Omega}$, the test statistic is:

$$W^* = (T-k)\ tr[\hat{\Omega}^{-1}(\tilde{\Omega} - \hat{\Omega})]$$

which is distributed as asymptotic χ^2 with m degrees of freedom. In fact, W^* can be defined with a small sample correction factor, (T-k)/T, times the Wald statistic W. Similarly, the correction factor can be applied to the LR and LM statistics and generate the corrected statistics LR^* and LM^*. It also follows that the inequality remain valid, ie,

$$W^* \geq LR^* \geq LM^*$$

Bera, Byron and Jarque(1981) argue that the size correction factor (T-k)/T is inadequate especially for large demand system.

They propose a factor which is equal to the ratio of the asymptotic 5% critical value to the empirical critical value. This method will be discussed in Chapter 6.

Later, Byron and Rosalsky(1985) proposed an Edgeworth correction method to approximate the true χ^2 tests. The suggested method appears to be fairly useful, but the computational cost is quite heavy especially for large SUR systems with a large number of restrictions. In addition, the reliability of Edgeworth corrections has yet to be established.

Deaton(1972) also suggests another test statistic which approximate the F statistic in (4.99) with a known covariance matrix by

$$\{ \operatorname{tr}[\tilde{\Omega}_{r-1}{}^{-1}(\tilde{\Omega}_r - \hat{\Omega}_0)]/m \} \, / \, \{ \operatorname{tr}[\tilde{\Omega}_{r-1}{}^{-1}\hat{\Omega}_0]/(T-k)n \}$$

and formed an asymptotic χ^2 statistic

$$D_r = \{ (T-k) \operatorname{tr}[\tilde{\Omega}_{r-1}{}^{-1}(\tilde{\Omega}_r - \hat{\Omega})] \} \, / \, \{ \operatorname{tr}[\tilde{\Omega}_{r-1}{}^{-1}\hat{\Omega}]/n \} \tag{4.100}$$

It is based upon the constrained Iterative Zellner Efficient (IZEF) method, a terminology by Kmenta and Gilbert(1968). Say for the model in (4.98), the constrained one-step GLS residuals $\tilde{U}_1 = Y - x\tilde{\beta}$, can be used to form a revised estimate of the covariance matrix $\tilde{\Omega}_1$

and subsequently a revised estimate of the constrained parameter matrix $\tilde{\beta}_2$. This r-step iterative procedure will generate a constrained estimator $\tilde{\beta}_r$ which is identical to the ML estimator, and the associated covariance matrix is (T-k)/T times the one from the ML estimator.

The limiting value of D_r is:

$$D^* = LM^* / \{ 1 - [LM^* / (T-k)n] \}$$

Bohm, Reider & Tintner(1980) noted an inequality of the form:

$$W^* \geq D_r \geq LM^* \qquad (4.101)$$

but no general statement can be made about the alternative magnitude of D_r and LR^*.

The size corrected test statistics from Table 4.14 and Table 4.18 are presented in Table 4.22. All W^*, LR^*, LM^* and D^* statistics strongly rejected the demand theory except for the homogeneity condition in the DLOG model. Homogeneïty restrictions are marginally accepted using the LM^* and D^* statistics in the DLOG model. The relationship of the test statistics in (4.101) is also satisfied.

4.10 Summary

Since the demand conditions are rejected by the maintained AIDS model, is it possible to conclude that consumers do suffer from money illusions and are irrational as they do not maximize their utility? Rejection of demand theory is not an uncommon finding in empirical research, but does it mean that the demand theory is in invalid and inappropriate in reality? Simmons and Weiserbs(1979) argue that given that the approximation of the true utility function is exact only at one point (the base point of approximation), it is only possible to test the true restrictions on the true utility at that point. But explicit restrictions on the approximated utility function can be tested over any subset of the data observed and will be considered at all points of the sample. Hence even though a restriction or a specification is rejected, this does not imply a conclusion can be reached concerning the validity of the theory. It does not mean that well behaved consumer preferences do not exist. Rather, it is probably the case that other specifications may perform better. This argument motivates the following analysis on diagnostic testing of the specification of the maintained model.

Footnotes

1. Pollak and Wales(1969) show that if using least squares method to minimize sum of square reisduals, the estimates will depend on which equation is omitted. That is n sets of parameters estimates are obtained by estimating the system with a different equation omitted each time. But if MLE is used, the procedure yields estimates which are independent of which equation is omitted. They also show that the least squares method yield ML estimates if the disturbance associated with all but the omitted expenditure equation are mutually independent, and the variance associated with the distribution in the retained equation are equal and constant over time.

2. K_{ij} can be expressed into: (from Phlips[1974])

$$K_{ij} = \lambda\mu^{ij} - [\ \lambda/(\partial\lambda/\partial y)\ ]\ (\partial q_i/\partial y)(\partial q_j/\partial y)$$

where the first term on the RHS is the specific substitution and the second term is the general substitution.

λ is the marginal utility of money and is equal to

$$\lambda = 1/p_i\ (\partial\mu/\partial q_i)$$

μ^{ij} denotes the ij element of the inverse of the Hessian matrix U^{-1}.

3. It was illustrated in p.337 of Amemiya's(1983) article, "Non-Linear Regression Model" in Handbook of Econometrics, Vol I, edited by Z. Griliches & M.D. Intrilgator.

4. Maximum Likelihood estimation method is not preferred because the condition of multi-normally distributed errors is too strong a priori assumption in a cross-sectional study. Nevertheless, the Least Squares estimators becomes MLE if normality is satisfied.

5. I am thankful to Dr. R. Byron for his assistance in computer programming.

6. I am indebted to Dr. Byron for allowing me to use his computer program.

7. Similiar exercises on complete demand model were carried out by Byron(1968) with reference to Australia and Byron(1970) using Barten's 16 commodity consumer expenditure data for Holland.

8. K_{ij} is invariant under monotonic increasing transformation.

Proof :

Expressed K_{ij} into (from footnote 2 above)

$$K_{ij} = \lambda\mu^{ij} - [\ \lambda/(\partial\lambda/\partial y)\ ]\ (\partial q_i/\partial y)(\partial q_j/\partial y)$$

If μ^{ij} is transformed into $A\mu$ where A is a positive constant, then λ is transformed into

$$A\lambda = A\ (\partial\mu/\partial q_i)\ \ (1/pi),$$

and μ^{ij} is transformed into $(1/A)\ \mu^{ij}$,
and $(\partial\lambda/\partial y)$ will be transformed into $A(\partial\lambda/\partial y)$
while $(\partial q_i/\partial p_j)$ and $(\partial q_i/\partial y)$ are invariant.

So after monotonic increasing transformation, K_{ij} becomes

$$\begin{aligned} K_{ij}^* &= [\ A\lambda\ \ (1/A)\ \mu^{ij}\ ] - [\ A\lambda\ /\ (A(\partial\lambda/\partial y))\ ]\ (\partial q_i/\partial y)\ (\partial q_j/\partial y) \\ &= \lambda\mu^{ij} - [\ \lambda/(\partial\lambda/\partial y)\ ]\ (\partial q_i/\partial y)\ (\partial q_j/\partial y) \\ &= K_{ij} \end{aligned}$$

9. Referenced from Beggs(1987).

10. I have to thank Dr. Ji-Chu Ryu for using his computer program.

11. The flexibility of the Translog functional form is well studied by Wales(1977), Guilkey(1978), and Simmons & Weiserbs(1979) and will not be discussed in this paper.

12. The original paper by Barten(1966) was in Dutch. The mathmatical derivation of the model can also be found from Barten's(1968) paper.

13. A brief review on testing non-nested alternatives was given by MacKinnon(1983).

Table 4.1
The Simple Linear Model

	q_1 Rice/Grain	q_2 Tubers	q_3 Fish	q_4 Dried Fish
Constant	1766.4 (11.5)	358.46 (2.46)	133.31 (2.31)	281.28 (2.89)
p_1	**-423.24** (8.69)	94.075 (2.03)	-25.105 (1.37)	-46.461 (1.50)
p_2	24.922 (1.00)	**-96.202** (4.07)	21.450 (2.30)	-28.004 (1.77)
p_3	14.118 (2.10)	-26.277 (4.11)	**-32.633** (12.9)	21.571 (5.04)
p_4	-34.783 (11.3)	20.018 (1.89)	3.0854 (0.72)	**-42.889**(5.95)
p_5	-5.0068 (2.13)	-2.1098 (0.94)	-.91073 (1.03)	-.13617 (0.09)
p_6	-.88774 (0.27)	6.8817 (2.21)	.95898 (0.78)	5.2278 (2.50)
p_7	-68.575 (0.77)	-39.096 (0.46)	59.492 (1.78)	93.121 (1.65)
p_8	-6.0762 (1.59)	-1.4387 (0.40)	-3.3401 (2.32)	-3.6687 (1.51)
p_9	-64.924 (1.31)	-68.582 (1.46)	22.190 (1.19)	32.861 (1.04)
p_{10}	17.695 (1.50)	-13.397 (1.19)	16.316 (3.67)	24.025 (3.20)
p_{11}	-62.976 (2.76)	-28.371 (1.31)	-15.555 (1.81)	-44.528 (3.07)
p_{12}	-334.34 (5.69)	-180.92 (3.24)	-63.615 (2.88)	-70.758 (1.89)
p_{13}	-3.1350 (0.19)	-4.4846 (0.28)	8.8620 (1.40)	-34.436 (3.22)
p_{14}	127.46 (3.18)	89.724 (2.36)	-60.428 (4.01)	-34.716 (1.36)
p_{15}	-17.960 (0.56)	88.717 (2.91)	40.422 (3.35)	14.648 (0.72)
p_{16}	9.9488 (2.29)	-2.8480 (0.69)	-.85534 (0.52)	-2.2405 (0.81)
p_{17}	93.849 (3.68)	26.263 (1.08)	11.658 (1.22)	57.934 (3.57)
y	.06969 (12.3)	.00288 (0.53)	.01995 (9.35)	.01480 (4.11)
T	424	424	424	424
F statistics	24.224	9.191	36.637	11.645
$\bar{R}^2$	0.4970	0.2585	0.6026	0.3118

Absolute t-ratios are in parenthesis

Table 4.1 (continued)
The Simple Linear Model

	q_5 Meat	q_6 Chicken	q_7 Egg	q_8 Milk
Constant	11.057 (0.78)	9.8532 (1.91)	86.413 (1.14)	-3.4822 (0.86)
p_1	-2.9611 (0.66)	-.54628 (0.33)	22.601 (0.94)	-1.1245 (0.87)
p_2	2.3476 (1.03)	1.9263 (2.30)	61.880 (5.03)	1.1923 (1.80)
p_3	.62861 (1.01)	.38786 (1.71)	14.495 (4.35)	-.10162 (0.57)
p_4	-.58249 (0.56)	.04026 (0.11)	2.2611 (0.40)	.71661 (2.38)
p_5	**-1.1137**(5.12)	-.11169 (1.41)	.85670 (0.73)	.12308 (1.96)
p_6	-.63521 (2.10)	**-.33106**(3.00)	-2.3818 (1.47)	-.10431 (1.20)
p_7	-14.461 (1.76)	-9.1629 (3.06)	**-477.28**(10.8)	-4.1791 (1.77)
p_8	.92751 (2.63)	.29612 (2.30)	5.5575 (2.94)	**-.06119**(0.60)
p_9	-3.3814 (0.74)	-4.2272 (2.53)	18.001 (0.73)	-.14143 (0.11)
p_{10}	-3.7687 (3.46)	-.44557 (1.12)	-4.3538 (0.74)	-.64560 (2.05)
p_{11}	2.9585 (1.41)	1.0681 (1.39)	25.200 (2.23)	1.1538 (1.90)
p_{12}	-10.069 (1.86)	-6.1626 (3.11)	1.7538 (0.06)	4.6252 (2.96)
p_{13}	-.49991 (0.32)	.22326 (0.39)	6.0554 (0.73)	-.39257 (0.88)
p_{14}	18.710 (5.07)	3.3388 (2.48)	-2.3192 (0.12)	.13504 (0.13)
p_{15}	-1.1475 (0.39)	1.2941 (1.20)	-5.2681 (0.33)	.28260 (0.33)
p_{16}	.55353 (1.38)	-.22010 (1.51)	-2.1986 (1.02)	-.06753 (0.59)
p_{17}	-2.3057 (0.98)	-2.9756 (3.47)	-69.846 (5.53)	-2.1558 (3.18)
y	.00386 (7.39)	.00145 (7.58)	.02297 (8.18)	.00157 (10.4)
T	424	424	424	424
F statistics	13.047	11.393	34.335	26.837
$\bar{R}^2$	0.3389	0.3066	0.5865	0.5237

Absolute t-ratios are in parenthesis

Table 4.1 (continued)

The Simple Linear Model

	q_9 Vegetable	q_{10} Legumes	q_{11} Fruit	q_{12} Condiment
Constant	676.86 (5.87)	114.11 (4.15)	169.54 (2.70)	369.56 (3.42)
p_1	99.085 (2.70)	27.736 (3.17)	67.989 (3.40)	16.008 (0.47)
p_2	106.30 (5.68)	5.1212 (1.15)	20.136 (1.98)	30.434 (1.74)
p_3	-4.1126 (0.81)	2.0294 (1.68)	-2.0001 (0.73)	-3.0591 (0.65)
p_4	-16.214 (1.90)	1.6201 (0.80)	4.9001 (1.05)	-15.556 (1.95)
p_5	-.23835 (0.13)	.76795 (1.81)	-3.2644 (3.38)	-.48304 (0.29)
p_6	.18693 (0.08)	.92306 (1.57)	.76945 (0.57)	1.7922 (0.77)
p_7	-68.136 (1.02)	-65.941 (4.13)	-26.879 (0.74)	-5.7412 (0.09)
p_8	-2.3342 (0.81)	.49707 (0.72)	-1.4393 (0.92)	-.47211 (0.18)
p_9	**-393.50** (10.5)	-26.446 (2.97)	-28.634 (1.41)	-16.504 (0.47)
p_{10}	-18.764 (8.90)	**-20.410** (9.61)	-5.4857 (1.13)	-7.6374 (0.92)
p_{11}	-2.1295 (0.12)	20.007 (4.88)	**-73.552** (7.85)	-16.148 (1.00)
p_{12}	10.426 (0.24)	61.646 (5.84)	-27.723 (1.15)	**-227.44**(5.49)
p_{13}	-37.598 (2.96)	-.74909 (0.25)	3.4519 (0.50)	-1.5600 (0.13)
p_{14}	-63.186 (2.10)	-1.7265 (0.24)	-1.1157 (0.07)	-51.236 (1.81)
p_{15}	-62.540 (2.59)	-14.800 (2.57)	7.6796 (0.58)	30.965 (1.37)
p_{16}	-1.1327 (0.35)	-.71820 (0.92)	-3.0279 (1.78)	.08582 (0.03)
p_{17}	42.485 (19.2)	-36.957 (8.07)	-17.916 (10.5)	18.013 (1.00)
y	.06367 (14.9)	-.00117 (1.15)	.01763 (7.57)	.02945 (7.36)
T	424	424	424	424
F statistics	33.914	34.639	15.687	12.334
$\overline{R}^2$	0.5834	0.5887	0.3846	0.3254

Absolute t-ratios are in parenthesis

Table 4.1 (continued)
The Simple Linear Model

	q_{13} Cooking oil	q_{14} Additive	q_{15} Pre. food	q_{16} Alcohol
Constant	57.712 (4.89)	899.46 (5.30)	-966.79 (1.74)	4.8929 (1.92)
p_1	16.778 (4.46)	.12774 (.002)	113.32 (0.64)	-3.0399 (3.76)
p_2	6.5423 (3.41)	-24.286 (0.88)	-58.946 (0.65)	-.92924 (2.25)
p_3	-.66845 (1.29)	-26.412 (3.54)	24.625 (1.01)	-.05062 (0.45)
p_4	-2.2480 (2.57)	8.1310 (0.65)	-47.036 (1.14)	.54619 (2.90)
p_5	.36850 (2.03)	-.65803 (0.25)	8.6729 (1.01)	-.06298 (1.61)
p_6	-.83579 (3.30)	2.4618 (0.68)	-24.353 (2.04)	.12992 (2.38)
p_7	-17.851 (2.60)	-215.56 (2.19)	-396.37 (1.23)	5.9782 (4.05)
p_8	.13861 (0.47)	4.4821 (1.06)	14.741 (1.06)	.07263 (1.14)
p_9	-4.0053 (1.05)	52.154 (0.95)	-464.04 (2.58)	-.98001 (1.19)
p_{10}	-1.1168 (1.23)	3.9546 (0.30)	-83.517 (1.94)	.20526 (1.05)
p_{11}	2.8677 (1.63)	-67.718 (2.68)	241.93 (2.91)	1.1038 (2.91)
p_{12}	-9.2640 (2.04)	187.70 (2.88)	199.01 (0.93)	-3.0088 (3.08)
p_{13}	**-8.5722** (6.60)	9.4294 (0.51)	114.03 (1.86)	-.61556 (2.20)
p_{14}	-6.8421 (2.22)	**-576.28** (13.0)	83.260 (0.57)	-.13869 (0.21)
p_{15}	.88711 (0.36)	25.995 (0.73)	**-299.94** (2.57)	-.43977 (0.83)
p_{16}	.57564 (1.72)	5.2388 (1.09)	13.038 (0.83)	**-.11837**(1.64)
p_{17}	-9.2197 (4.69)	-100.93 (3.57)	-30.450 (0.33)	.83180 (1.96)
y	.00537 (12.3)	.08502 (13.5)	.11148 (5.40)	.00020 (2.09)
T	424	424	424	424
F statistics	38.881	37.959	4.791	5.842
$\bar{R}^2$	0.6171	0.6113	0.1389	0.1709

Absolute t-ratios are in parenthesis

Table 4.1 (continued)

The Simple Linear Model

	q_{17} Tobacco
Constant	1270.4 (2.91)
p_1	-137.27 (0.99)
p_2	-220.93 (3.12)
p_3	-9.2747 (0.48)
p_4	-111.87 (3.46)
p_5	14.043 (2.09)
p_6	-2.8342 (0.30)
p_7	1311.1 (5.17)
p_8	-4.4816 (0.41)
p_9	175.95 (1.25)
p_{10}	29.136 (0.87)
p_{11}	-24.933 (0.38)
p_{12}	170.16 (1.02)
p_{13}	-24.644 (0.51)
p_{14}	27.180 (0.24)
p_{15}	17.909 (0.20)
p_{16}	-31.970 (12.4)
p_{17}	**-1636.7** (22.5)
y	.19299 (11.9)
T	424
F statistics	69.664
$\overline{R}^2$	0.745

Absolute t-ratios are in parenthesis

Table 4.2

Price and Income Elasticities for the Simple Linear Model

	q_1	q_2	q_3	q_4	q_5	q_6	q_7	q_8
p_1	**-.849***	2.907	-1.58	-.669	-1.52	-.254	.4050	-1.14
p_2	-.216	**-1.94**	.6318	-.209	.5311	.4214	.4865	.5293
p_3	.0838*	-2.49	**-6.59**	.9261	.9357	.5217	.7339	-.290
p_4	-.062	.6145	.1379	**-.588**	-.229	.0162	.0326	.5319
p_5	-.086*	-.567	-.513	-.016	**-4.97**	-.472	.1196	1.039
p_6	-.012*	1.396	.3909	.4779	-.218	**-.991**	-.283	-.652
p_7	-.045*	-.405	1.156	.4329	-2.52	-1.39	**-.284**	-1.37
p_8	-.068*	-.246	-1.30	-.316	2.623	.7788	.5548	**.3508**
p_9	-.074*	-1.28	.7045	.2707	-1.01	-1.10	.1811	-.078
p_{10}	.0631*	-.761	1.666	.5978	-3.62	-.381	-.142	-1.22
p_{11}	-.110*	-.803	-.876	-.588	1.211	.4132	.3527	.9351
p_{12}	-.216*	-1.74	-1.47	-.336	-1.76	-.969	.0097	1.500
p_{13}	-.016*	-.339	1.482	-1.29	-.675	.2674	.2794	-1.08
p_{14}	.1229*	1.325	-2.02	-.257	4.791	.7617	-.021	.0728
p_{15}	-.014*	1.039	1.010	.0895	-.248	.2487	-.039	.1190
p_{16}	.0676*	-.300	-2.04	-.123	1.063	-.376	-.141	-.266
p_{17}	.0847*	.3666	.3533	.3721	-.635	-.684	-.688	-1.15
y	.5882*	.4011	4.451	.8656	7.794	2.594	1.669	6.056

* Significant at 5%

Table 4.2 (continued)

Price and Income Elasticities for the Simple Linear Model

	q_9	q_{10}	q_{11}	q_{12}	q_{13}	q_{14}	q_{15}	q_{16}
p_1	.4164*	3.798	1.098	.0979*	.7892	.0004	8.562	-8.46
p_2	.2341	.3246	.0673	.0994	.1516	-.039	-2.06	-1.18
p_3	-.051	.7163	-.096	-.056*	-.091	-.236	5.199	-.365
p_4	-.061	.1564	.0668	-.083	-.087	.0204	-3.03	1.007
p_5	-.009	.7487	-.465	-.026*	.1399	-.017	5.262	-1.48
p_6	.0051	.8187	.0798	.0707*	-.256	.0479	-11.8	1.963
p_7	-.095	-2.97	-.141	-.012*	-.275	-.212	-.950	5.147
p_8	-.055	.3511	-.133	-.016*	.0369	.0754	6.552	1.083
p_9	**-.978**	-2.01	-.258	-.057*	-.105	.0885	-20.4	-1.47
p_{10}	-.146	**-5.80**	-.161	-.084*	-.095	.0212	-12.2	1.068
p_{11}	-.008	2.055	**-1.06**	-.086*	.1117	-.178	15.23	2.505
p_{12}	.0139*	2.279	-.144	**-.459**	-.139	.1723*	5.025	-2.55
p_{13}	-.402	-.255	.1409	-.024*	**-1.04**	.0712	23.11	-4.49
p_{14}	-.134	-.126	-.009	-.155*	-.169	**-.963**	3.374	-.188
p_{15}	-.111	-.870	.0502	.0752*	.0175	.0315	**-11.3**	-.393
p_{16}	-.017	-.372	-.176	.0018*	.0952	.0553	3.755	**-2.45**
p_{17}	.0857	-.298	-.138	.0508	-.231	-.150	-1.17	.9097
y	1.117*	-.663	1.133	.7506*	1.025	1.046	33.18	2.016

* Significant at 5%

Table 4.2 (continued)

Price and Income Elasticities for the Simple Linear Model

	q_{17}
p_1	-.334
p_2	-.269
p_3	-.064
p_4	-.220
p_5	.2609
p_6	-.043
p_7	.9999
p_8	-.057
p_9	.2270
p_{10}	.1220
p_{11}	-.048
p_{12}	.1262
p_{13}	-.152
p_{14}	.0337
p_{15}	.0205
p_{16}	-.294
p_{17}	**-2.49**
y	1.725

* Significant at 5%

Table 4.3

<u>Slutsky Substitution Coefficients and Income Effect for the Simple Linear Model</u>

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rice/Grain	**-338.***	97.59*	-.774	-28.4*	1.747	1.222*	50.62*	.7903
Tubers	37.62*	**-95.7***	25.08*	-25.3*	3.051*	2.191*	66.06*	1.478*
Fish	23.25*	-25.9*	**-30.0***	23.51*	1.134*	.5779*	17.50*	.1041
Dried fish	-15.7	20.81*	8.543*	**-38.8***	.4735	.4370	8.545*	1.146*
Meat	-3.86*	-2.06*	-.583	-1.12	**-1.05***	-.088*	1.234*	.1489*
Chicken	-.218	6.909*	1.151*	5.370*	-.598*	**-.317***	-2.16*	-.089*
Egg	-52.1*	-38.4*	64.22*	96.63*	-13.5*	-8.82*	**-471.***	-.381*
Milk	-5.58*	-1.42*	-3.20*	-3.56*	.9549*	.3064*	5.721*	**-.050***
Vegetable	-20.6	-66.8*	34.87*	42.27*	-.928	-3.31*	32.60*	.8564*
Legumes	22.92*	-13.2*	17.81*	25.14*	-3.48*	-.337*	-2.63	-.528*
Fruit	-49.8*	-27.8*	-11.8*	-41.7*	3.688*	1.342*	29.54*	1.450*
Condiment	-304.*	-179.*	-55.1*	-64.4*	-8.42*	-5.54*	11.57*	5.296*
Cooking oil	1.112	-4.31*	10.08*	-33.5*	-.265*	.3116*	7.455*	-.297*
Additive	192.1*	92.39*	-41.9*	-21.0*	22.29*	4.683*	18.98*	1.591*
Pre. food	3.346	89.60*	46.52*	19.17	.0326	1.737	1.755	.7626
Alcohol	10.03*	-2.85*	-.833*	-2.22*	.5578*	-.218*	-2.17*	-.066*
Tobacco	221.7*	31.55*	48.25*	85.08*	4.775	-.316	-27.7	.7241
Income Effect	-85.0*	-.525	-2.61	-4.05	-.063	-.014	-5.44	-.011

* Significant at 5%

Table 4.3 (continued)

Slutsky Substitution Coefficients and Income Effect
for the Simple Linear Model

	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Rice/Grain	176.7*	26.31*	89.49*	51.93*	23.33*	103.8*	249.3*	-2.80*
Tubers	117.9*	4.908*	23.35*	35.80*	7.521*	-8.80	-38.6	-.893*
Fish	4.230	1.876*	.3099	.7997	.0352	-15.3	39.23*	-.024
Dried fish	1.205	1.300*	9.723*	-7.50	-.779	31.39*	-16.5	.6009*
Meat	.8074	.7487*	-2.97*	.0006	.4567*	.7383	10.50*	-.060*
Chicken	.7990	.9118*	.9389*	2.075*	-.784*	3.279*	-23.3*	.1318*
Egg	-53.1*	-66.2*	-22.7*	1.237	-16.6*	-195.*	-369.*	6.026*
Milk	-1.88*	.4888*	-1.31*	-.263	.1768*	5.086*	15.53*	.0741*
Vegetable	**-353.***	-27.2*	-17.4*	2.212	-.593	106.2*	-393.*	-.853*
Legumes	-14.0*	**-20.5***	-4.16*	-5.43*	-.714*	10.33*	-75.2*	.2203*
Fruit	9.895	19.79*	**-70.2***	-10.6*	3.882*	-51.7*	263.0*	1.142*
Condiment	37.63*	61.15*	-20.2*	**-214.***	-6.97*	224.0*	246.6*	-2.92*
Cooking oil	-33.7*	-.820*	4.526*	.2349	**-8.24***	14.61*	120.8*	-.603*
Additive	-4.15	-2.81*	15.23*	-23.9*	-1.86	**-497.***	186.6*	.0467
Pre. food	-43.1	-15.2*	13.07	39.97	2.529	51.99	**-266.***	-.379*
Alcohol	-1.06*	-7.18*	-3.01*	.1185	.5816*	5.333*	13.17*	**-.118***
Tobacco	159.3*	-39.1*	14.42	72.03*	.6305	55.02	174.0	1.199*
Income Effect	-40.5*	.0878	-3.33	-12.6*	-.327*	-78.8*	-34.1	-.0002

* Significant at 5%

Table 4.3 (continued)

Slutsky Substitution Coefficients and Income Effect for the Simple Linear Model

	(17)
Rice/Grain	98.10
Tubers	-185.*
Fish	16.01
Dried fish	-59.1
Meat	17.21*
Chicken	-.979
Egg	1356.*
Milk	-3.11
Vegetable	298.6*
Legumes	43.62*
Fruit	11.52
Condiment	252.6*
Cooking oil	-12.9*
Additive	206.1*
Pre. food	76.91
Alcohol	-31.8*
Tobacco	**-1282.***
Income Effect	-354.0

* Significant at 5%

Table 4.4

The Univariate Double Logarithmic Model

	$\ln q_1$ Rice/grain	$\ln q_2$ Tubers	$\ln q_3$ Fish	$\ln q_4$ Dried fish
Constant	2.0560 (4.47)	-.48864 (0.24)	-6.9637 (4.00)	.13787 (0.08)
$\ln p_1$	**-.95025** (10.7)	1.1442 (2.87)	-1.6895 (5.01)	-.39997 (1.26)
$\ln p_2$	-.01872 (0.77)	**-.82866** (7.58)	.23649 (2.56)	-.06498 (0.75)
$\ln p_3$	.10085 (2.93)	-.58283 (3.78)	**-1.5543** (11.9)	.45106 (3.67)
$\ln p_4$	-.03052 (1.40)	-.08559 (0.88)	.43577 (5.30)	**-.29855**(3.85)
$\ln p_5$	-.05706 (1.81)	.17451 (1.24)	-.02742 (0.23)	.20810 (1.85)
$\ln p_6$	-.00893 (0.24)	.24672 (1.49)	-.05555 (0.40)	.16661 (1.26)
$\ln p_7$	-.04515 (0.81)	-.13748 (0.55)	.79026 (3.72)	.36135 (1.81)
$\ln p_8$	-.06165 (1.62)	-.20747 (1.21)	-.33337 (2.31)	-.34527 (2.54)
$\ln p_9$	-.05855 (1.07)	-.72652 (2.97)	.83401 (4.04)	.16098 (0.83)
$\ln p_{10}$	.07761 (1.87)	-.17791 (0.96)	.78063 (4.97)	.56809 (3.84)
$\ln p_{11}$	-.05282 (1.48)	-.40287 (2.52)	-.52195 (3.86)	-.58483 (4.59)
$\ln p_{12}$	-.17678 (5.38)	-.13821 (0.94)	-.44944 (3.62)	-.00836 (0.07)
$\ln p_{13}$	-.12378 (1.69)	-.36046 (1.10)	-.32962 (1.19)	-1.1733 (4.48)
$\ln p_{14}$	.11694 (2.72)	.68265 (3.54)	-.94394 (5.80)	.08576 (0.56)
$\ln p_{15}$	-.01046 (0.42)	.34291 (3.07)	.30848 (3.26)	.01891 (0.21)
$\ln p_{16}$	.06924 (2.19)	.06486 (0.46)	-.05393 (0.45)	-.20578 (1.82)
$\ln p_{17}$	.07245 (3.70)	.02695 (0.31)	.13759 (1.86)	.22762 (3.26)
$\ln y$	.64485 (14.0)	.64827 (3.15)	1.7839 (10.2)	.74726 (4.55)
T	424	424	424	424
F statistics	29.581	15.213	61.804	11.709
$\bar{R}^2$	0.5488	0.3769	0.7212	0.3131

Absolute t-ratios are in parenthesis

Table 4.4 (continued)

The Univariate Double Logarithmic Model

	$\ln q_5$ Meat	$\ln q_6$ Chicken	$\ln q_7$ Egg	$\ln q_8$ Milk
Constant	-13.383 (5.25)	-12.243 (5.60)	-6.6454 (3.99)	-15.263 (7.30)
$\ln p_1$	.92863 (1.88)	-.19219 (0.45)	-.32556 (1.01)	.78680 (1.94)
$\ln p_2$	.39206 (2.90)	.38774 (3.35)	.26906 (3.05)	.29649 (2.67)
$\ln p_3$	.39068 (2.05)	.29619 (1.81)	.54513 (4.38)	.08431 (0.54)
$\ln p_4$	.17208 (1.43)	-.04903 (0.48)	-.10105 (1.29)	-.00445 (0.05)
$\ln p_5$	**-.95479** (5.47)	-.18173 (1.22)	.32238 (2.83)	.31449 (2.20)
$\ln p_6$	-.26657 (1.30)	**-.37084** (2.11)	-.23151 (1.73)	.07873 (0.47)
$\ln p_7$	-1.6253 (5.23)	-.86625 (3.25)	**-1.4298** (7.05)	-.26494 (1.04)
$\ln p_8$	.45068 (2.13)	.27399 (1.51)	.04431 (0.32)	**.15602**(0.90)
$\ln p_9$	-.74033 (2.45)	-.98848 (3.82)	.17559 (0.89)	.04417 (0.18)
$\ln p_{10}$	-.45217 (1.97)	.08433 (0.43)	-.08885 (0.59)	-.52484 (2.78)
$\ln p_{11}$	.66092 (3.34)	.45965 (2.71)	.60655 (4.70)	.50731 (3.13)
$\ln p_{12}$	-.44137 (2.43)	-.56751 (3.64)	.46354 (3.90)	.52817 (3.54)
$\ln p_{13}$	-.25656 (0.63)	-.09766 (0.28)	-.01245 (0.05)	-.34204 (1.03)
$\ln p_{14}$	.86665 (3.64)	.63180 (3.09)	.14509 (0.93)	-.17907 (0.92)
$\ln p_{15}$	-.16716 (1.21)	.14190 (1.20)	-.12162 (1.35)	.07717 (0.68)
$\ln p_{16}$	-.19607 (1.12)	-.31070 (2.07)	-.08039 (0.70)	-.09020 (0.63)
$\ln p_{17}$	-.28029 (2.59)	-.28289 (3.05)	-.36953 (5.23)	-.26911 (3.03)
$\ln y$	1.8206 (7.14)	1.5966 (7.31)	1.1153 (6.70)	1.6995 (8.13)
T	424	424	424	424
F statistics	19.103	13.433	29.173	33.634
$\overline{R}^2$	0.4351	0.3460	0.5452	0.5814

Absolute t-ratios are in parenthesis

Table 4.4 (continued)
The Univariate Double Logarithmic Model

	$\ln q_9$ Vegetable	$\ln q_{10}$ Legumes	$\ln q_{11}$ Fruit	$\ln q_{12}$ Condiment
Constant	-3.5669 (4.77)	5.3811 (2.39)	-3.7890 (3.56)	-.09224 (0.11)
$\ln p_1$	.62316 (4.30)	.48824 (1.12)	.02600 (0.13)	.14116 (0.88)
$\ln p_2$	.16007 (4.04)	.14599 (1.22)	.24762 (4.39)	.08197 (1.86)
$\ln p_3$	.03421 (0.61)	.13466 (0.80)	-.01103 (0.14)	-.04896 (0.79)
$\ln p_4$	-.15076 (4.27)	-.04352 (0.41)	-.01634 (0.33)	-.09243 (2.10)
$\ln p_5$	.02560 (0.50)	.51168 (3.32)	-.22476 (3.09)	.03011 (0.53)
$\ln p_6$	-.06487 (1.08)	-.04469 (0.25)	-.00980 (0.11)	.07083 (1.06)
$\ln p_7$	-.25432 (2.79)	-.99115 (3.61)	-.16769 (1.29)	-.13350 (1.32)
$\ln p_8$	.02403 (0.39)	.13466 (0.72)	-.04232 (0.48)	-.05462 (0.79)
$\ln p_9$	**-.97905** (11.0)	-.05224 (0.20)	-.02507 (0.20)	-.00551 (0.06)
$\ln p_{10}$	-.04072 (0.60)	**-1.6818** (8.28)	-.21582 (2.25)	-.01049 (0.14)
$\ln p_{11}$	.03671 (0.63)	.91065 (5.21)	**-.76809** (9.31)	-.11588 (1.80)
$\ln p_{12}$	.10032 (1.88)	.83162 (5.18)	.05852 (0.77)	**-.36137**(6.10)
$\ln p_{13}$	-.32576 (2.74)	-.66016 (1.84)	.03930 (0.23)	.01572 (0.12)
$\ln p_{14}$	.03037 (0.43)	.14883 (0.71)	-.06920 (0.70)	-.10646 (1.37)
$\ln p_{15}$	-.05580 (1.38)	-.26534 (2.17)	.02612 (0.45)	.07412 (1.65)
$\ln p_{16}$	-.02011 (0.39)	-.22339 (1.44)	-.11887 (1.63)	.06126 (1.07)
$\ln p_{17}$	.03719 (1.17)	-.56755 (5.93)	-.10276 (2.27)	.01341 (0.38)
$\ln y$	1.1240 (15.0)	-.06730 (0.30)	1.1586 (10.9)	.62668 (7.55)
T	424	424	424	424
F statistics	32.449	25.900	24.595	14.752
$\overline{R}^2$	0.5723	0.5145	0.5010	0.3692

Absolute t-ratios are in parenthesis

Table 4.4 (continued)

The Univariate Double Logarithmic Model

	$\ln q_{13}$ Cooking oil	$\ln q_{14}$ Additive	$\ln q_{15}$ Pre. food	$\ln q_{16}$ Alcohol
Constant	-.24783 (2.27)	-1.0810 (1.55)	-19.656 (5.19)	-4.3874 (1.99)
$\ln p_1$	.56917 (2.69)	.15482 (1.15)	.87330 (1.19)	-1.1415 (2.67)
$\ln p_2$	.09686 (1.67)	-.01620 (0.44)	-.21149 (1.05)	-.04779 (0.41)
$\ln p_3$	-.19264 (2.36)	-.18600 (3.56)	.33667 (1.19)	-.11492 (0.70)
$\ln p_4$	-.10875 (2.11)	-.02998 (0.91)	-.58439 (3.27)	.09636 (0.92)
$\ln p_5$	.27927 (3.74)	-.08836 (1.85)	.58050 (2.24)	-.41938 (2.77)
$\ln p_6$	-.30607 (3.48)	.02595 (0.46)	-.51336 (1.68)	.29192 (1.64)
$\ln p_7$	.06004 (0.45)	-.08368 (0.99)	-.18848 (0.41)	.94843 (3.52)
$\ln p_8$	-.00231 (0.03)	.05156 (0.89)	.21107 (0.67)	.30534 (1.67)
$\ln p_9$	-.05657 (0.44)	.16187 (1.96)	-.80212 (1.79)	-.51917 (1.99)
$\ln p_{10}$	.04342 (0.44)	-.01817 (0.29)	-.31961 (0.94)	.09806 (0.49)
$\ln p_{11}$	.30447 (3.60)	-.10432 (1.93)	1.2263 (4.17)	.38909 (2.27)
$\ln p_{12}$	.04402 (0.57)	.29857 (6.00)	.52966 (1.96)	-.40597 (2.58)
$\ln p_{13}$	**-.91563** (5.26)	.15608 (1.41)	.25421 (0.42)	-.61035 (1.74)
$\ln p_{14}$	-.35047 (3.43)	**-1.1948** (18.3)	.83155 (2.35)	-.47056 (2.28)
$\ln p_{15}$	.04301 (0.73)	.06494 (1.72)	**-1.1361** (5.53)	.08945 (0.75)
$\ln p_{16}$	.03814 (0.51)	.05151 (1.07)	.05040 (0.19)	**-.63901**(4.21)
$\ln p_{17}$	-.18117 (3.90)	-.16609 (5.61)	-.24235 (1.51)	.10106 (1.08)
$\ln y$	.84612 (7.75)	.87298 (12.5)	2.3328 (6.16)	.82094 (3.72)
T	424	424	424	424
F statistics	24.758	52.644	11.798	5.885
$\bar{R}^2$	0.5027	0.6873	0.3148	0.1721

Absolute t-ratios are in parenthesis

Table 4.4 (continued)
The Univariate Double Logarithmic Model

	$\ln q_{17}$ Tobacco
Constant	-5.2034 (4.89)
$\ln p_1$	-.33930 (1.65)
$\ln p_2$	-.33093 (5.86)
$\ln p_3$	-.07352 (0.92)
$\ln p_4$	-.07262 (1.45)
$\ln p_5$	.27480 (3.77)
$\ln p_6$	-.02520 (0.29)
$\ln p_7$	.54776 (4.22)
$\ln p_8$	.08480 (0.96)
$\ln p_9$	.10691 (0.85)
$\ln p_{10}$	.07648 (0.80)
$\ln p_{11}$	.00607 (0.07)
$\ln p_{12}$	.16635 (2.19)
$\ln p_{13}$	-.06899 (0.41)
$\ln p_{14}$	.01037 (0.10)
$\ln p_{15}$	-.18681 (3.23)
$\ln p_{16}$	-.23910 (3.27)
$\ln p_{17}$	**-.93237** (20.6)
$\ln y$	1.3819 (13.0)
T	424
F statistics	67.201
$\bar{R}^2$	0.7380

Absolute t-ratios are in parenthesis

Table 4.5
Slutsky Coefficients and Income Effect for the
Univariate Double-Logarithmic Model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rice/Grain	**-408.***	110.8	-67.4	-21.2	10.23	1.080	-2.02	3.541
Tubers	-7.03	**-194.**	33.02	-13.8	6.246	3.552	56.80	1.664
Fish	29.62	-18.4	**-30.9**	21.74	1.287	.5981	21.62	.2317
Dried Fish	-.484	-8.05	37.98	**-43.2**	2.316	.1069	-6.74	.2599
Meat	-2.81	2.375	.0799	3.591	**-.921**	-.065	4.519	.1353
Chicken	-.031	3.482	-.321	3.348	-.269	**-.232**	-3.53	.0491
Egg	-58.0*	-34.1	138.3	136.4	-34.9	-10.9	**-460.**	-2.07
Milk	-5.82*	-3.25	-3.52	-7.83	.6190	.2251	1.060	**.1028**
Vegetable	-5.60	-108.	92.05	47.04	-7.13	-6.33	49.19	.9631
Legumes	30.08*	-7.84	24.31	39.50	-1.69	.3276	-2.78	-.786
Fruit	-20.6	-44.4	-34.2	-85.4	6.083	2.550	75.56	1.915
Condiment	-294.*	-36.2	-81.8	5.332	-10.2	-7.64	167.5	5.643
Cooking oil	-22.3*	-11.1	-6.26	-56.5	-.574	-.073	1.173	-.356
Additive	204.8*	124.7	-92.2	40.42	14.59	6.823	56.94	-.048
Pre. food	3.986	75.25	49.47	10.87	-2.43	2.203	-27.1	1.111
Alcohol	11.54*	1.662	-.916	-7.65	-.439	-.420	-2.64	-.086
Tobacco	237.1	23.34	59.38	103.5	-.643	-.708	-63.6	-.178
Income Effect	-106.*	-7.42	-5.10	-7.57	-.124	-.024	-8.39	-.015

* Significant at 5%

Table 4.5 (continued)

Substitution Coefficient and Income Effect
for the Univariate Double-Logarithmic Model

	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Rice/Grain	261.6*	15.65	29.01	60.04*	21.04*	160.7*	183.8	-.438
Tubers	110.6	10.89	49.16	39.47	6.222	.8371	-44.3	-.017
Fish	11.83	1.515	2.842	.1738	-1.16	-16.5	20.00	-.006
Dried fish	-40.3	-2.31	4.114	-14.1	-2.50	5.152	-93.3	.0806
Meat	2.065	1.980	-2.24	1.183	1.028	-3.44	10.16	-.026
Chicken	-2.22	-.248	.0856	2.385*	-1.27*	2.447*	-10.3	.0207
Egg	-205.*	-108.	-37.8	-71.0*	6.231*	-86.3*	-57.9	1.225
Milk	1.758	.8139	-.516	-1.76	.0265	4.498*	5.559	.0276
Vegetable	**-453.***	-3.83	10.38	15.70	.7307	170.3*	-148.	-.342
Legumes	-.467	**-35.8**	-8.55	.9617	1.087	2.711	-18.5	.0293
Fruit	26.18*	35.15	**-79.4**	-21.9	10.70	-36.8	181.9	.2177
Condiment	126.3*	85.30	26.43	**-223.**	6.170	437.0*	248.5	-.751
Cooking oil	-32.9*	-9.02	2.658	2.818*	**-9.77***	30.87*	17.75	-.117
Additive	85.64	10.31	8.431	67.31*	-15.1*	**-987.***	294.7	-.383
Pre. food	-23.5	-28.6	11.60	46.14	4.773	97.37	**-377.**	.1365
Alcohol	-1.68	-2.37	-3.07	3.584	.3235	6.586*	2.192	**-.102**
Tobacco	152.3	-58.9	15.67	53.54	-3.82	-35.6	26.32	.2612
Income Effect	-51.3	.0762	-5.17	-13.0	-.346	-84.2	-144.	-.001

Table 4.5 (continued)
Substitution Coefficient and Income Effect
for the Univariate Double Logarithmic Model

	(17)
Rice/Grain	28.60
Tubers	-510.
Fish	10.55
Dried fish	-9.13
Meat	31.52
Chicken	-8.37
Egg	1359.
Milk	14.51
Vegetable	300.5
Legumes	54.88
Fruit	51.43
Condiment	545.2
Cooking oil	-6.84
Additive	254.0
Pre. food	-358.
Alcohol	-61.1
Tobacco	**-1797.**
Income Effect	-586.

Table 4.6
Test Homogeneity Condition on the Simple Models
(Univariate Double-Logarithmic Models)

Equation	Double-Log Model T	Double-Log Model F, Wald
(1) Rice	-4.36	18.98
(2) Tubers	-.601*	.3617*
(3) Fish	-1.46*	2.143*
(4) Dried fish	-.203*	0.041*
(5) Meat	.4630*	0.214*
(6) Chicken	-.063*	0.004*
(7) Egg	2.177***	4.738***
(8) Milk	5.424	29.42
(9) Vegi	1.592*	2.535*
(10) Legumes	-2.24***	5.027***
(11) Fruit	-.793*	0.629*
(12) Condiment	.9240*	0.853*
(13) Cook oil	.7560*	0.571*
(14) Additive	-.276*	0.076*
(15) Pre. food	3.336	11.13
(16) Alcohol	-2.18***	4.730***
(17) Tobacco	1.421*	2.018*

T is distributed as t (405)
F is distributed as F (1, 405)
Wald is distributed as $\chi^2(1)$

Critical Values

	0.1	0.05	0.01	α
T(2 sided)	1.645	1.96	2.576	
F	2.7	3.84	6.63	
Wald	2.7	3.84	6.63	

* significant at $\alpha = 0.1$
** significant at $\alpha = 0.05$
*** significant at $\alpha = 0.01$

Table 4.7

Stone's Linear Expenditure System

(using Non-linear and Step-wise Least Squares Procedures)

Budget Share w_i	Group	β_i	γ_i
1	Rice	.0946(10.2)	918.15(25.9)
2	Tubers	.0241(8.93)	-15.833(1.04)
3	Fish	.1419(21.8)	-26.618(3.24)
4	Dried fish	.0421(12.0)	117.04(9.70)
5	Meat	.0567(18.3)	-7.1785(5.74)
6	Chicken	.0195(10.3)	0.3184(0.32)
7	Egg	.0246(15.4)	-6.8754(0.40)
8	Milk	.0193(16.1)	-3.6364(5.19)
9	Vegetable	.1050(21.0)	47.938(1.40)
10	Legumes	.0114(3.80)	38.914(7.29)
11	Fruit	.0423(20.1)	26.662(3.12)
12	Condiment	.0232(16.6)	172.67(11.6)
13	Cooking oil	.0340(22.7)	14.599(6.67)
14	Additive	.1062(27.2)	170.80(6.01)
15	Pre. food	.0495(7.28)	-145.54(2.50)
16	Alcohol	.0019(6.33)	-0.5983(1.88)
17	Tobacco	.2037(---)	78.011(1.34)

Table 4.8
Price and Income Elasticities for the
Linear Expenditure System

	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8
Rice/Grain	**-.274**	-.565	-1.29	-.365	-1.59	-.631	-.568	-1.68
Tubers	.0007*	**-1.21***	.0098	.0029	.0117	.0049	.0042	.0120
Fish	.0065*	.0477	**-1.68**	.0315	.1291	.0526	0.466	.1343
Dried fish	-.008	-.066	-.116	**-.265**	-.151	-.071	-.056	-.150
Meat	.0052*	.0363	.0883	.0238	**-2.58**	.0456	.0349	.1057
Chicken	-.0002*	-.001	-.003	-.001	-.004	**-.934***	-.013	-.004
Egg	.0002*	.0014	.0030	.0009	.0040	.0015	**-1.05***	.0040
Milk	.0017*	.0125	.0320	.0087	.0343	.0143	.0128	**-2.72**
Vegetable	-.002*	-.017	-.034	-.010	-.049	-.018	-.016	-.047
Legumes	-.006*	-.043	-.089	-.026	-.126	-.049	-.043	-.135
Fruit	-.002*	-.014	-.033	-.010	-.035	-.015	-.013	-.039
Condiment	-.005*	-.034	-.091	-.023	-.010	-.041	-.034	-.100
Cooking oil	-.003*	-.023	-.056	-.016	-.067	-.026	-.024	-.073
Additive	-.007*	-.051	-.134	-.035	-.151	-.058	-.055	-.168
Pre. food	.0048*	.0340	.0841	.0234	.1121	.0418	.0377	.1064
Alcohol	.0002*	.0013	.0033	.0009	.0039	.0015	.0013	.0042
Tobacco	-.003*	-.023	-.055	-.015	-.077	-.027	-.027	-.075
Income	.3422*	2.668	4.886	1.524	6.836	2.572	2.426	6.958

The first column refers to prices and income

* Significant at 5%

Table 4.8 (continued)
Price and Income Elasticities for the
Linear Expenditure Syetem

	w_9	w_{10}	w_{11}	w_{12}	w_{13}	w_{14}	w_{15}	w_{16}
Rice/Grain	-.326	-.323	-.339	-.189*	-.261	-.257*	-4.02	-.454
Tubers	.0028*	.0026	.0028	.0017	.0022	.0024	0.320	.0038
Fish	.0277	.0245	.0284	.0162*	.0219	.0224	.3328	.0380
Dried fish	-.035	-.029	-.036	-.021	-.027	-.027	-.407	-.041
Meat	.0221	.0185	.0231	.0125*	.0165	.0175	.2612	.0314
Chicken	-.001	-.001	-.001	-.0004*	-.001	-.001*	-.009	-.001
Egg	.0008	.0008	.0008	.0005*	.0006	.0006*	.0093	.0011
Milk	.0074	.0067	.0077	.0042*	.0059	.0057*	.0936	.0097
Vegetable	**-.919***	-.009	-.010	-.006*	-.008	-.008*	-.118	-.013
Legumes	-.025	**1.269**	-.026	-.015*	-.020	-.020*	-.317	-.035
Fruit	-.008	-.007	**-.815***	-.005*	-.006	-.007*	-.098	-.011
Condiment	-.020	-.016	-.021	**-.543***	-.016	-.015*	-.254	-.027
Cooking oil	-.013	-.013	-.014	-.008*	**-.700***	-.010*	-.170	-.019
Additive	-.031	-.034	-.032	-.018	-.026	**-.796***	-.395	-.042
Pre. food	.0217	.0228	.0224	.0120	.0174	.0162	**-5.92**	.0252
Alcohol	.0008	.0008	.0008	.0004	.0006	.0006	.0099	**-.178**
Tobacco	-.014	-.017	-.014	-.008	-.012	-.011	-.165	-.017
Income	1.401*	1.444	1.423	.8566*	1.148	1.140*	16.62	1.842

The first column refers to prices and income

* Significant at 5%

Table 4.8 (continued)
Price and Income Elasticities for the
Linear Expenditure Expenditure

	w_{17}
Rice/Grain	-.390
Tubers	.0034
Fish	.0327
Dried fish	-.042
Meat	.0246
Chicken	-.001
Egg	.0009
Milk	.0084
Vegetable	-.011
Legumes	-.029
Fruit	-.009
Condiment	-.023
Cooking oil	-.016
Additive	-.036
Pre. food	.0272
Alcohol	.0009
Tobacco	**-.932***
Income	1.608

The first column refers to prices and income

* Significant at 5%

Table 4.9

Own Substitution Coefficients and Income Effect for the Linear Expenditure System

Group	Substitution Coefficient	Income Effect
Rice/Grain	-119.	-51.3*
Tubers	-240.	-5.86
Fish	-20.8	-3.28
Dried fish	-89.6	-7.16
Meat	-1.26	-.062
Chicken	-.635	-.014
Egg	-322.	-8.06
Milk	-.766	-.011
Vegetable	-422.*	-54.1*
Legumes	-10.9	-.242
Fruit	-89.5	-4.66
Condiment	-384.	-15.7
Cooking oil	-8.07	-.376*
Additive	-652.	-95.2
Pre. food	-273.	-22.7
Alcohol	-.187	-.0003
Tobacco	-1869.	-520.

Table 4.10

Estimates for the Linear Expenditure System
Using Non-Linear SUR Procedure from SAS/ETS

Group	β_i	γ_i
Rice/Grain	.1535(16.0)	633.4(16.4)
Tubers	.0238(8.61)	-16.6(.960)
Fish	.1346(20.2)	-58.2(7.01)
Dried Fish	.0391(12.7)	84.34(7.48)
Meat	.0409(15.7)	-4.58(4.79)
Chicken	.0147(9.17)	.4688(.560)
Egg	.0212(14.2)	-19.1(1.21)
Milk	.0104(10.8)	-.701(1.38)
Vegetable	.1004(20.8)	-41.9(1.21)
Legumes	.0195(5.67)	32.21(5.09)
Fruit	.0386(18.9)	8.366(1.05)
Condiment	.0172(14.0)	198.4(17.0)
Cooking oil	.0334(23.3)	7.798(4.13)
Additives	.1027(26.9)	86.86(3.33)
Pre. Food	.0371(7.49)	-117.(2.99)
Alcohol	.0017(6.50)	-.915(3.96)
Tobacco	.2109(---)	-229.(4.30)

Table 4.11
Own Price and Income Elasticities
for the Linear Expenditure System
Estimated using Non-Linear SUR Procedure in SAS/ETS

	Own Price Elasticity	Income Elasticity
Rice/Grain	-.532*	.5553*
Tubers	-1.22*	2.635
Fish	-2.50	4.635
Dried Fish	-.469	1.416
Meat	-2.03	4.931
Chicken	-.902*	1.939
Egg	-1.15*	2.091
Milk	-1.33*	3.750
Vegetable	-1.07*	1.340*
Legumes	.8624	2.471
Fruit	-.942*	1.299
Condiment	-.472*	.6350*
Cooking oil	-.840*	1.128
Additives	-.896*	1.102*
Pre. Food	-5.01	12.46
Alcohol	-2.19	1.648
Tobacco	-1.20*	1.665

* Significant at 5%

Table 4.12
Own Substitution Coefficients and Income Effect
For the Linear Expenditure System
Estimated using Non-Linear SUR from SAS/ETS Version 5

	Substitution Coefficient	Income Effect
Rice/Grain	-216.	-83.2*
Tubers	-241.	-5.78
Fish	-25.5	-3.11
Dried Fish	-110.	-6.65
Meat	-1.19	-.045
Chicken	-.633	-.010
Egg	-340.	-6.95
Milk	-.576	-.006
Vegetable	-491.*	-51.7*
Legumes	-12.4	-.415
Fruit	-99.7	-4.25
Condiment	-349.	-11.3
Cooking oil	-9.25	-.369*
Additives	-725.	-92.1
Pre. Food	-331.	-17.1
Alcohol	-.202	-.0003
Tobacco	-2130.	-539.

* Significant at 5%

Table 4.13
Unrestricted Double-Log Demand System

	Rice/Grain $w_1 \ln q_1$	Tubers $w_2 \ln q_2$	Fish $w_3 \ln q_3$	Dried Fish $w_4 \ln q_4$
w_i	7.979(355.)	7.055(249.)	6.044(165.)	6.613(264.)
$\ln p_1$	**-.199**(7.88)	-.002(.340)	.0591(3.48)	-8.8E-4(.12)
$\ln p_2$	.0576(8.49)	**-.014**(8.34)	.0107(2.42)	.0052(2.66)
$\ln p_3$	.0095(.950)	.0021(.820)	**-.076**(11.5)	.0033(1.11)
$\ln p_4$	.0173(2.74)	.0067(4.31)	-.015(3.62)	**-.045**(23.4)
$\ln p_5$	-.073(8.46)	-.007(3.16)	-.015(2.74)	-.017(6.89)
$\ln p_6$	.0022(.210)	-1.3E-4(.05)	.0054(.770)	.0082(2.64)
$\ln p_7$	.0814(5.30)	.0028(.740)	.0058(.580)	.0161(3.63)
$\ln p_8$	-.041(3.81)	-4.0E-4(.15)	-.002(.260)	-.002(.770)
$\ln p_9$	.1053(7.01)	.0152(4.12)	-.009(.960)	.0163(3.78)
$\ln p_{10}$	.0136(1.13)	.0008(.280)	-.013(1.63)	.0054(1.54)
$\ln p_{11}$	-.007(.690)	.0055(2.13)	.0198(2.91)	.0063(2.05)
$\ln p_{12}$	.0079(.830)	-.004(1.83)	.0134(2.20)	-.001(.510)
$\ln p_{13}$	-.094(4.57)	-.004(.770)	.0089(.680)	-.016(2.73)
$\ln p_{14}$	-.008(.660)	-.007(2.28)	.0055(.660)	-.007(1.94)
$\ln p_{15}$	-.005(.680)	-.001(.630)	-6.8E-4(.14)	-3.9E-4(.19)
$\ln p_{16}$	-.035(3.91)	-.006(2.64)	-.011(1.82)	-.006(2.19)
$\ln p_{17}$	.0115(1.99)	-5.4E-4(.38)	-.002(.410)	.0019(1.14)
Income Ratio	.0477(7.25)	-3.2E-4(.20)	.0097(2.27)	.0043(2.27)

Table 4.13 (continued)
Unrestricted Double-Log Demand System

	Meat $w_5 \ln q_5$	Chicken $w_6 \ln q_6$	Egg $w_7 \ln q_7$	Milk $w_8 \ln q_8$
w_i	4.393(98.0)	3.500(118.)	6.550(235.)	3.521(97.2)
$\ln p_1$	-.034(3.41)	.0029(.870)	.0054(2.00)	-.007(2.50)
$\ln p_2$	-.005(1.88)	.0003(.280)	.0029(3.93)	-9.8E-4(1.4)
$\ln p_3$	-.005(1.16)	-9.2E-4(.69)	-.001(1.10)	-.002(1.69)
$\ln p_4$	-.004(1.76)	.0011(1.30)	.0024(3.61)	.0023(3.55)
$\ln p_5$	**-.024**(7.01)	-.002(1.39)	-.006(6.73)	-9.1E-4(1.0)
$\ln p_6$	-.005(1.11)	**-.012**(8.35)	.0002(.190)	-5.2E-4(.46)
$\ln p_7$	.0322(5.18)	.0022(1.06)	**-.012**(7.19)	.0013(.800)
$\ln p_8$	-.001(.270)	-4.6E-4(.32)	-3.1E-4(.27)	**-.009**(8.39)
$\ln p_9$	.0225(3.79)	.0096(4.85)	.0057(3.58)	.0029(1.89)
$\ln p_{10}$	-.008(1.75)	-.003(1.94)	.0015(1.15)	.0024(1.95)
$\ln p_{11}$	-.004(.960)	-.002(1.46)	-.001(1.01)	-.003(2.81)
$\ln p_{12}$	-6.3E-4(.17)	.0022(1.78)	-4.7E-4(.47)	-.001(1.42)
$\ln p_{13}$	-.008(1.06)	.0017(.640)	-.006(2.54)	-9.1E-4(.43)
$\ln p_{14}$	.0021(.400)	-.002(1.34)	-3.6E-4(.27)	.0017(1.34)
$\ln p_{15}$	.0067(2.33)	.0003(.300)	.0004(.520)	-6.2E-4(.82)
$\ln p_{16}$	.0056(1.58)	-1.1E-4(.09)	-.003(2.77)	.0008(.910)
$\ln p_{17}$	.0074(3.21)	.0010(1.36)	.0015(2.31)	.0016(2.58)
Income Ratio	.0135(5.19)	.0030(3.43)	.0008(1.10)	.0028(4.05)

Table 4.13 (continued)
Unrestricted Double-Log Demand System

	Vegetable $w_9 \ln q_9$	Legumes $w_{10} \ln q_{10}$	Fruit $w_{11} \ln q_{11}$	Condiment $w_{12} \ln q_{12}$
w_i	7.419(383.)	5.324(176.)	6.319(252.)	7.150(302.)
$\ln p_1$	.0062(.750)	.0195(3.38)	.0096(2.19)	.0094(2.41)
$\ln p_2$	.0168(7.58)	.0051(3.34)	.0041(3.48)	.0075(7.17)
$\ln p_3$	-.001(.340)	.0012(.510)	-.001(.660)	.0013(.850)
$\ln p_4$	.0102(4.93)	.0025(1.75)	.0034(3.09)	.0034(3.51)
$\ln p_5$	-.016(5.90)	-.009(4.50)	-.008(5.69)	-.008(6.51)
$\ln p_6$	.0018(.530)	.0032(1.32)	.0028(1.51)	-.002(1.49)
$\ln p_7$	.0284(5.67)	.0137(3.83)	.0133(4.99)	.0179(7.58)
$\ln p_8$	-.015(4.25)	-.006(2.66)	-.005(2.77)	-.006(3.65)
$\ln p_9$	**-.042**(8.57)	-.002(.480)	.0103(3.97)	.0156(6.77)
$\ln p_{10}$	.0017(.450)	**-.020**(7.22)	.0060(2.87)	.0008(.460)
$\ln p_{11}$	.0004(.110)	-.002(.910)	**-.035**(19.3)	.0019(1.16)
$\ln p_{12}$	.0013(.430)	.0012(.550)	.0016(1.02)	**-.032**(21.3)
$\ln p_{13}$	-.032(4.84)	.0023(.490)	-.012(3.49)	-.014(4.48)
$\ln p_{14}$	-.011(2.66)	-.007(2.35)	-9.0E-4(.41)	-.002(.790)
$\ln p_{15}$	.0003(.110)	1.5E-5(.01)	-.001(.870)	-.001(1.04)
$\ln p_{16}$	-.010(3.42)	-.002(1.05)	-.003(1.90)	-.005(3.95)
$\ln p_{17}$	.0061(3.23)	-3.8E-4(.28)	.0017(1.69)	.0016(1.83)
Income Ratio	.0102(4.73)	.0031(2.05)	.0033(2.91)	.0026(2.57)

Table 4.13 (continued)
Unrestricted Double-Log Demand System

	Cooking Oil $w_{13}\ln q_{13}$	Additives $w_{14}\ln q_{14}$	Pre. Food $w_{15}\ln q_{15}$	Alcohol $w_{16}\ln q_{16}$
w_i	4.972(206.)	7.696(347.)	8.340(184.)	3.057(73.4)
$\ln p_1$	.0136(3.62)	.0205(1.91)	-.006(.280)	-8.4E-4(.89)
$\ln p_2$	.0076(7.64)	.0177(6.19)	.0043(.820)	-2.3E-4(.93)
$\ln p_3$	.0022(1.46)	.0021(.500)	-.007(.910)	.0002(.440)
$\ln p_4$	.0025(2.67)	.0087(3.28)	.0141(2.81)	.0001(.470)
$\ln p_5$	-.010(7.75)	-.025(6.85)	-.009(1.40)	.0008(2.68)
$\ln p_6$	.0007(.480)	.0013(.290)	.0015(.170)	.0004(.930)
$\ln p_7$	.0042(1.86)	.0359(5.53)	-.005(.430)	-7.4E-4(1.3)
$\ln p_8$	-.004(2.24)	-.016(3.46)	7.3E-5(.01)	-1.0E-4(.25)
$\ln p_9$	.0107(4.86)	.0407(6.43)	.0150(1.27)	.0009(1.60)
$\ln p_{10}$	.0016(.880)	.0068(1.34)	-.009(.980)	.0003(.660)
$\ln p_{11}$	-.003(1.67)	.0032(.710)	-.023(2.80)	-2.5E-4(.65)
$\ln p_{12}$	-.001(.750)	.0044(1.09)	-.002(.240)	.0002(.480)
$\ln p_{13}$	**-.041**(13.6)	-.039(4.52)	-.006(.390)	.0026(3.35)
$\ln p_{14}$	.0039(2.11)	**-.090**(17.0)	-.015(1.53)	.0006(1.37)
$\ln p_{15}$	.0006(.610)	-.003(.940)	**-.010**(1.71)	-3.6E-4(1.3)
$\ln p_{16}$	-.004(2.99)	-.014(3.59)	.0018(.250)	**-7.5E-4**(2.3)
$\ln p_{17}$	.0022(2.54)	.0035(1.41)	.0064(1.40)	-2.7E-6(.01)
Income Ratio	.0077(7.93)	.0101(3.62)	.0032(.620)	-8.6E-4(3.5)

Table 4.13 (continued)
Unrestricted Double-Log Demand System

	Tobacco $w_{17} \ln q_{17}$
w_i	8.301(250.)
$\ln p_1$	.0325(1.53)
$\ln p_2$	.0356(6.14)
$\ln p_3$	-.004(.490)
$\ln p_4$	.0072(1.36)
$\ln p_5$	-.048(6.78)
$\ln p_6$	-.008(.940)
$\ln p_7$	.0838(6.32)
$\ln p_8$	-.038(4.25)
$\ln p_9$	.0674(5.39)
$\ln p_{10}$	-.001(.140)
$\ln p_{11}$	-.004(.490)
$\ln p_{12}$	.0119(1.51)
$\ln p_{13}$	-.049(2.85)
$\ln p_{14}$	-.023(2.20)
$\ln p_{15}$	.0047(.760)
$\ln p_{16}$	-.018(2.35)
$\ln p_{17}$	**-.125**(25.6)
Income Ratio	.0276(4.88)

Table 4.14
<u>Hypothesis Testing of the Multivariate Double-Log Model</u>

	Unrestricted	Homogeneity	Symmetry	Adding-up
Parameter	323	306	187	322
Objective *T	6688.85	6700.99	5890.41	6372.51
Det $\|\Sigma\|$	2.3517E-68	2.5755E-68	6.7500E-67	2.2651E-66
Information Inaccuracy				
I * 1000	-3.88	-3.81	-4.65	-3.45
IC * 1000	9463.11	10227.31	8251.94	6039.93
Asy. χ^2 Tests for Restrictions				
Wald	---	43.34	3780.55	39960.7
LR	---	38.54	1423.36	1936.69
LM	---	34.14	774.95	422.54
GJ	---	57.38	977.1	1466.84
Degree of freedom	---	17	136	1

Critical values :

	α	$\chi^2_{(1)}$	$\chi^2_{(17)}$	$\chi^2_{(136)}$
*	0.1	2.705	24.769	157.52
**	0.05	3.841	27.587	164.22
***	0.01	5.024	33.409	177.28

Table 4.15

Uncompensated Own Price and Income Elasticities for the Multivariate Double-Log System

	Own Price	Income
Rice/Grain	-.770*	.1725*
Tubers	-1.54	-.0.35
Fish	-2.61	.3354
Dried Fish	-1.64	.1562
Meat	-2.93	1.632
Chicken	-1.58	.3943
Egg	-1.18	.0768
Milk	-3.49	1.000
Vegetable	-.567*	.1354*
Legumes	-2.54	.3913
Fruit	-1.19	.1124
Condiment	-1.19*	.0961*
Cooking oil	-1.40	.2601
Additives	-.981*	.1083*
Pre. Food	-3.33	1.089
Alcohol	-.726	-.839
Tobacco	-1.01	.2178

* Significant at 5%

Table 4.16

Substitution Coefficients and Income Effect for the Multivariate Double-Log System

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rice/Grain	**-377.***	-8.67*	39.80*	-2.27*	-8.24	.8448*	31.52*	-2.29*
Tubers	237.2*	**-154.**	16.44	32.26	-2.78	.1651	36.85*	-.725
Fish	6.403*	3.152	**-20.1**	3.127	-.411	-.097	-2.48*	-.217
Dried fish	44.54*	40.93	-14.3	**-189.**	-1.47	.4374	19.45*	1.098
Meat	-18.1*	-3.78	-1.42	-5.98	**-.897**	-.061	-4.72*	-.042
Chicken	.6865*	-.095	.5946	3.474	-.186	**-.610**	.2094*	-.029
Egg	474.7*	35.68*	12.18*	129.7*	23.84	1.964*	**-220.***	1.363*
Milk	-14.8*	-.313	-.234	-1.18	-.053	-.026	-.341*	**-.680**
Vegetable	352.2*	115.5	-11.4*	78.33	9.484	4.993*	59.92*	1.801*
Legumes	14.57*	1.953	-4.94*	8.062	-1.14	-.503*	4.944*	.4779*
Fruit	-16.7*	29.25	17.07	20.63	-1.22	-.721	-7.99*	-1.29
Condiment	51.14*	-60.4	31.53	-12.4	-.537	2.236	-9.33*	-1.67
Cooking oil	-71.9*	-6.36*	2.478*	-16.9*	-.843	.2048*	-13.2*	-.129*
Additive	-33.0*	-60.3*	7.902*	-38.1*	1.021	-1.37*	-4.43*	1.289*
Pre. food	-26.6*	-13.1	-1.28	-2.92	4.495	.2418	6.718	-.616
Alcohol	-20.7*	-7.42	-2.23	-4.53	.4299	-.010	-4.86*	.0928
Tobacco	57.82	-5.59	-2.72	12.65	4.657	.8255	22.88	1.452
Income Effect	-25.8*	.0778	-.225	-.734	-.015	-.002	-.255	-.002

* Significant at 5%

Table 4.16 (continued)
Substitution Coefficients and Income Effect for the Multivariate Double-Log System

	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Rice/Grain	20.71*	20.73*	22.29*	60.30*	10.45*	81.71*	-28.8*	-.177
Tubers	127.6	11.94	22.05	105.6	12.51*	152.5*	48.49	-.106
Fish	-1.34*	.4402*	-.987	3.080	.5969*	3.034*	-13.2	.0126
Dried fish	48.80	3.717	11.13	30.39	2.593*	47.72*	102.5	.0312
Meat	-7.27*	-1.26	-2.75	-7.63	-1.01	-12.9*	-6.33	.0250
Chicken	1.018*	.5506*	1.066	-2.59	.0950*	.8590*	1.254	.0128
Egg	298.8*	45.68*	95.83*	355.8*	10.05*	448.9*	-85.5	-.485
Milk	-9.54*	-1.32*	-2.32*	-7.58*	-.529*	-12.2*	-.073	-.004
Vegetable	**-261.***	-3.19*	43.05*	177.8*	14.64*	292.8*	143.1	.3335
Legumes	3.404*	**-12.9**	7.939*	3.123*	.6825*	15.58*	-28.4	.0357
Fruit	1.514*	-2.87*	**-110.**	15.16	-2.42*	15.59*	-147.	-.064
Condiment	15.01*	4.310*	13.60	**-772.**	-2.77*	59.69*	-31.9	.1304
Cooking oil	-44.2*	.9988*	-11.7*	-37.7*	**-13.3***	-65.1*	-13.6	.2296
Additive	-77.8*	-15.0*	-4.44*	-20.7*	6.453*	**-806.***	-171.	.2880
Pre. food	2.512	.0451	-7.12	-21.4	1.439	-33.8	**-180.**	-.211
Alcohol	-10.5*	-.715*	-2.15	-11.1	-.966*	-17.2*	2.950	**-.059**
Tobacco	54.20	-1.11	10.20	27.69	4.559	37.49	93.21	-.001
Income Effect	-5.23*	-.066	-.368	-1.70	-.085*	-9.05	-1.49	.0001

* Sigificant at 5%

Table 4.16 (continued)
Substitution Coefficients and Income Effect for the Multivariate Double-Log System

	(17)
Rice/Grain	163.1
Tubers	369.8
Fish	-7.26
Dried fish	47.09
Meat	-31.6
Chicken	-6.99
Egg	1298.
Milk	-36.7
Vegetable	600.9
Legumes	-4.06
Fruit	-25.5
Condiment	203.7
Cooking oil	-101.
Additive	-247.
Pre. food	68.97
Alcohol	-29.3
Tobacco	**-2177.**
Income Effect	-70.5

* Significant at 5%

Table 4.17
Unrestricted Almost Ideal Demand System

	Rice/Grain w_1	Tubers w_2	Fish w_3	Dried Fish w_4
Constant	1.249(9.97)	.0953(1.95)	.0963(1.01)	.1330(2.17)
ln p_1	-.002(.090)	.0215(2.19)	-.128(6.72)	-.015(1.23)
ln p_2	-.014(2.09)	.0006(.220)	.0234(4.47)	-.008(2.61)
ln p_3	.0166(1.66)	-.019(4.99)	-.032(4.16)	.0233(4.76)
ln p_4	-.018(2.89)	.0004(.160)	.0226(4.73)	.0244(7.94)
ln p_5	-.014(1.51)	-2.7E-4(.08)	-.018(2.64)	.0048(1.10)
ln p_6	-7.2E-4(.07)	.0036(.860)	.0139(1.71)	.0057(1.09)
ln p_7	-.005(.290)	.0014(.220)	.0435(3.48)	.0148(1.84)
ln p_8	-.016(1.42)	-.002(.410)	-.019(2.23)	-.013(2.44)
ln p_9	-.029(1.89)	-.011(1.91)	.0545(4.67)	.0137(1.84)
ln p_{10}	.0114(.960)	-.012(2.59)	.0501(5.52)	.0132(2.27)
ln p_{11}	-.018(1.80)	-.006(1.45)	-.027(3.44)	-.019(3.83)
ln p_{12}	-.052(5.46)	-.011(3.02)	-.022(3.02)	-.003(.570)
ln p_{13}	-.029(1.36)	.0059(.720)	.0167(1.03)	-.032(3.12)
ln p_{14}	.0335(2.67)	.0081(1.66)	-.061(6.36)	-.001(.210)
ln p_{15}	-.007(1.03)	.0103(3.64)	.0174(3.13)	.0023(.640)
ln p_{16}	.0201(2.19)	.0004(.100)	.0056(.800)	-.003(.570)
ln p_{17}	.0103(1.79)	.0024(1.08)	.0076(1.74)	.0095(3.35)
Income Ratio	-.106(8.71)	-.006(1.29)	.0118(1.26)	-.007(1.29)

Table 4.17 (continued)
Unrestricted Almost Ideal Demand System

	Meat w_5	Chicken w_6	Egg w_7	Milk w_8
Constant	.1151(2.20)	.0267(1.05)	-.002(.100)	-.031(1.84)
$\ln p_1$	.0285(2.72)	.0032(.620)	.0048(1.30)	.0092(2.73)
$\ln p_2$	.0121(4.22)	.0052(3.69)	.0052(5.12)	.0037(4.03)
$\ln p_3$	.0059(1.42)	.0046(2.28)	.0063(4.27)	.0004(.340)
$\ln p_4$	.0029(1.12)	.0012(.970)	-9.9E-4(1.06)	.0010(1.19)
$\ln p_5$	-.015(4.02)	-.004(2.11)	.0017(1.32)	.0011(.880)
$\ln p_6$	-.002(.400)	.0058(2.68)	-.001(.900)	-6.2E-4(.44)
$\ln p_7$	-.025(3.65)	-.008(2.40)	-.009(3.78)	-.005(2.39)
$\ln p_8$	.0091(1.97)	.0032(1.40)	.0029(1.79)	.0066(4.48)
$\ln p_9$	.0017(.270)	-.004(1.30)	.0019(.830)	.0032(1.57)
$\ln p_{10}$	-.009(1.97)	-.002(.850)	-.001(.580)	-.003(1.82)
$\ln p_{11}$	.0115(2.66)	.0029(1.36)	.0047(3.06)	.0017(1.22)
$\ln p_{12}$	-.006(2.66)	-.006(2.95)	.0007(.470)	.0031(2.47)
$\ln p_{13}$	-.005(.650)	.0011(.250)	.0032(1.01)	1.8E-5(.010)
$\ln p_{14}$	.0237(4.52)	.0060(2.36)	.0004(.230)	.0002(.140)
$\ln p_{15}$	-.003(1.00)	.0019(1.26)	-3.41E-4(.32)	-1.2E-4(.13)
$\ln p_{16}$	-.002(.440)	-.003(1.74)	-.002(1.40)	-.001(.980)
$\ln p_{17}$	-.007(3.13)	-.004(3.31)	-.006(7.21)	-.003(3.72)
Income Ratio	-.011(2.18)	-.004(1.71)	-.001(.790)	.0018(1.11)

Table 4.17 (continued)
Unrestricted Almost Ideal Demand System

	Vegetable w_9	Legumes w_{10}	Fruit w_{11}	Condiment w_{12}
Constant	.0684(1.25)	.2290(5.62)	.1256(3.94)	.0786(2.90)
$\ln p_1$	.0435(3.97)	.0252(3.08)	.0092(1.44)	-7.2E-4(.13)
$\ln p_2$	.0149(4.99)	-.002(.710)	.0079(4.51)	.0004(.250)
$\ln p_3$	.0014(.330)	-3.4E-4(.11)	-.002(.800)	-.002(.750)
$\ln p_4$	-.009(3.50)	-.005(2.31)	.0015(.980)	-.003(2.49)
$\ln p_5$	.0029(.730)	.0089(3.05)	-.012(5.34)	.0011(.590)
$\ln p_6$	-.003(.700)	.0028(.810)	.0022(.820)	.0022(.930)
$\ln p_7$	-.019(2.63)	-.020(3.76)	-.003(.810)	-.005(1.32)
$\ln p_8$	-6.1E-4(.13)	.0035(.980)	-.003(.970)	.0002(.080)
$\ln p_9$	.0076(1.15)	-.015(2.98)	.0019(.500)	-.003(1.05)
$\ln p_{10}$	-.004(.730)	-.021(5.30)	-.005(1.59)	-.002(.880)
$\ln p_{11}$	.0025(.550)	.0156(4.63)	.0080(3.05)	-.003(1.43)
$\ln p_{12}$	.0048(1.16)	.0192(6.19)	.0014(.590)	.0192(9.30)
$\ln p_{13}$	-.024(2.65)	.0014(.200)	-6.9E-4(.13)	.0019(.430)
$\ln p_{14}$	-1.9E-6(.00)	-.005(1.20)	-.005(1.60)	-.004(1.54)
$\ln p_{15}$	-.003(.850)	-.007(3.05)	.0004(.240)	.0018(1.15)
$\ln p_{16}$	-.002(.550)	-.004(1.43)	-.003(1.22)	.0013(.650)
$\ln p_{17}$	.0057(2.29)	-.016(8.71)	-.004(2.90)	6.3E-5(.05)
Income Ratio	.0030(.570)	-.027(6.78)	-.006(1.94)	-.005(2.00)

Table 4.17 (continued)
Unrestricted Almost Ideal Demand System

	Cooking Oil w_{13}	Additives w_{14}	Pre. Food w_{15}	Alcohol w_{16}
Constant	.0872(3.78)	.1993(3.24)	-.861(8.76)	-.005(.870)
$\ln p_1$	.0196(4.25)	.0138(1.12)	.0025(.130)	-.003(2.53)
$\ln p_2$	.0037(2.94)	-.003(.880)	-.003(.580)	-3.8E-4(1.3)
$\ln p_3$	-.003(1.44)	-.016(3.30)	.0123(1.57)	.0002(.360)
$\ln p_4$	-.004(3.69)	-.005(1.59)	-.007(1.38)	.0002(.910)
$\ln p_5$	.0052(3.15)	-.006(1.41)	.0151(2.15)	-7.7E-4(2.0)
$\ln p_6$	-.006(3.35)	.0035(.660)	-.017(2.10)	.0012(2.57)
$\ln p_7$	-.004(1.50)	-.010(1.24)	-.023(1.80)	.0021(3.01)
$\ln p_8$	.0007(.360)	.0057(1.05)	.0138(1.59)	.0006(1.29)
$\ln p_9$	3.9E-5(.01)	.0125(1.66)	-.034(2.83)	-9.3E-4(1.4)
$\ln p_{10}$	-5.6E-4(.26)	-.003(.460)	-.019(2.10)	.0010(2.02)
$\ln p_{11}$	.0027(1.39)	-.013(2.46)	.0329(4.05)	.0014(3.18)
$\ln p_{12}$	-.002(.930)	.0269(5.75)	.0053(.710)	-.002(3.78)
$\ln p_{13}$	.0048(1.25)	.0191(1.83)	.0280(1.68)	-.002(1.77)
$\ln p_{14}$	-.006(2.56)	-.017(2.73)	.0181(1.84)	-4.3E-4(.81)
$\ln p_{15}$	-1.0E-4(.08)	.0050(1.40)	-7.6E-5(.01)	-3.8E-6(.01)
$\ln p_{16}$	.0002(.120)	.0064(1.43)	.0099(1.38)	-2.4E-5(.06)
$\ln p_{17}$	-.006(5.50)	-.016(5.80)	.0037(.830)	.0004(1.54)
Income Ratio	-.008(3.82)	-.013(2.15)	.091(9.43)	.0007(1.27)

Table 4.17 (continued)
Unrestricted Almost Ideal Demand System

	Tobacco w_{17}
Constant	-.605
$\ln p_1$	-.002
$\ln p_2$	-.045
$\ln p_3$	-.008
$\ln p_4$	-.002
$\ln p_5$	-.007
$\ln p_6$	-.011
$\ln p_7$	.0742
$\ln p_8$	.0073
$\ln p_9$	-.0001
$\ln p_{10}$	.0059
$\ln p_{11}$	.0022
$\ln p_{12}$	.0234
$\ln p_{13}$	.0106
$\ln p_{14}$	.0094
$\ln p_{15}$	-.018
$\ln p_{16}$	-.023
$\ln p_{17}$	.0223
Income Ratio	.0857

Table 4.18
Hypothesis Testing of the AIDS

	Unrestricted	Homogeneity	Symmetry	Homogeneity & Symmetry
Parameter	304	288	184	168
Objective *T	6480.00	6496.00	6011.0	6023.59
Det $\|\Sigma\|$	2.1792E-60	2.4258E-60	1.2519E-59	1.8168E-59
Information Inaccuracy				
I * 1000	60.07	60.87	76.86	78.34
IC * 1000	404.44	409.64	436.72	446.53
Asy. χ^2 Tests for Restrictions				
Wald	---	50.12	898.99	1114.95
LR	---	45.47	741.28	899.19
LM	---	41.24	623.16	742.03
GJ	---	63.99	854.19	1033.46
Degree of freedom	---	16	120	136

Critical values :

	α	$\chi^2_{(16)}$	$\chi^2_{(120)}$	$\chi^2_{(136)}$
*	0.1	23.54	140.23	157.52
**	0.05	26.30	146.57	164.22
***	0.01	32.00	158.95	177.28

Table 4.19

Own Price and Income Elasticities for the Almost Ideal Demand System

	Own Price	Income
Rice/Grain	-.902*	.6139*
Tubers	-.929*	.3187
Fish	-2.11	1.406
Dried Fish	-.110	.7207
Meat	-2.80	-.344
Chicken	-.225	.4365
Egg	-1.91	.8581*
Milk	1.385	1.651
Vegetable	-.910*	1.041*
Legumes	-3.57	-2.42
Fruit	-.724*	.7973*
Condiment	-.285	.8037*
Cooking oil	-.827*	.7099*
Additive	-1.17*	.8610*
Pre. Food	-1.12*	31.41
Alcohol	-1.02*	1.643
Tobacco	-.904*	1.706

* Significant at 5%

Table 4.20

Substitution Coefficients and Income Effect for the Almost Ideal Demand System

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rice/Grain	**-381.***	105.4	-69.2*	-9.18	8.084	1.735*	53.41*	4.057*
Tubers	4522.	**-225.**	18.61	-23.1	7.688	3.549*	69.68*	2.817
Fish	271.8	-27.2	**-28.7**	25.51	.7570	.5688*	15.45*	.1607
Dried fish	1128.	8.049	21.38	**-57.4**	2.212	.6844	-2.26	.6668
Meat	356.2	.7976	-3.89	3.536	**-1.58**	-.129	1.759	.0624
Chicken	422.6	3.448	-.512	3.972	.0025	**-.377***	-1.05*	-.027
Egg	6751.	32.19	67.19	143.8	-16.9*	-.699	**-490.**	-5.32*
Milk	1898.	2.363	-7.94	-2.09	1.053	.2191	3.608	**-.091**
Vegetable	386.8	-76.7	72.47*	81.53*	1.446	-1.65*	34.01*	2.485*
Legumes	1702.	-23.2	9.514	26.69	-.909	-.258	-1.27	-.501*
Fruit	768.7	-26.6	-24.1	-56.6	3.842	1.172*	37.31*	.8637
Condiment	2144.	-148.	-58.0	-5.26	-4.31	-5.42*	22.49*	4.074
Cooking oil	246.2	10.91*	4.372	-31.5*	-.433	.1760*	9.013*	.0553
Additive	607.0	82.15	-74.7*	15.09	12.69*	4.305	25.92	.9529
Pre. food	23278.	144.5	-32.8	72.31	1.076	2.190	5.183	.1422
Alcohol	5333.	4.586	-10.9	18.26	.8107	-.185	-2.48	-.147
Tobacco	531.6	43.62	37.66	103.1	-2.98	-1.86	-53.9	-1.02
Income Effect	-106.	-9.95	-3.13	-8.82	-.056	-.012	-7.05	-.010

* Significant at 5%

Table 4.20 (continued)
Substitution Coefficients and Income Effect for the
Almost Ideal Demand System

	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Rice/Grain	228.3*	28.08*	40.40*	39.59	20.52*	155.3*	169.3	-.411
Tubers	110.7	5.546	49.32	22.25	8.813	184.5	-629.	-.133
Fish	8.677	.5890	.7479	1.093	-.073	-6.03	6.749	.0315
Dried fish	-32.8	-1.93	10.64	-18.3	-2.66	36.07	-219.	.1053
Meat	1.255	1.611	-3.08	2.167	.7417	10.46	-.229	-.019
Chicken	-2.65	1.431	1.403	3.358	-.590	16.97	-45.7	.0416
Egg	-205.*	-55.9*	-14.1	-72.9	-7.66*	55.01	-643.	1.399
Milk	-6.86	3.154	.3867	3.615	.9187	45.65	-26.0	.0260
Vegetable	**-421.***	-26.3*	18.61	-16.0	3.273	155.5*	-277.	-.258
Legumes	-7.20	**-33.3**	-3.61	-2.81	.5296	33.09	-99.8	.1314
Fruit	19.97	22.19*	**-80.3**	-18.0	3.697	-21.7	181.8	.3796
Condiment	74.16*	74.65*	21.43	**-170.**	-1.22	473.2*	-37.3	-1.12
Cooking oil	-30.5*	1.220	.6203	7.878*	**-9.06**	43.95*	46.30	-.136
Additive	60.33*	-8.35	-10.1	-23.5	-5.24	**-950.***	280.5	-.086
Pre. food	-85.3	63.33	24.22	83.69	9.567	738.2	**-398.**	.0394
Alcohol	-23.3	13.45	2.298	11.33	2.671	134.2	-484.	**-.162**
Tobacco	165.1	-42.4	2.383	62.62	-3.51	-15.8	262.9	.4432
Income Effect	-47.2	-.558	-3.80	-17.3	-.314	-84.8	-103.	-.001

* Significant at 5%

Table 4.20 (continued)

Substitution Coefficients and Income Effect for the Almost Ideal Demand System

	(17)
Rice/Grain	131.8
Tubers	-4061.
Fish	-74.5
Dried fish	-659.
Meat	-217.
Chicken	-309.
Egg	-1881.
Milk	-1108.
Vegetable	-114.
Legumes	-613.
Fruit	-462.
Condiment	-1233.
Cooking oil	-121.
Additive	93.74
Pre. food	-15430.
Alcohol	-2969.
Tobacco	**-1721.**
Income Effect	-653.

* Significant at 5%

Table 4.21

Average Information Inaccuracy for Individual Group[1]

	LES	DLOG[2]	AIDS[2]
Rice/Grain	8.825	.1169	5.387
Tubers	8.040	.2634	5.787
Fish	16.46	.7363	9.669
Dried Fish	8.248	.1423	5.605
Meat	8.781	3.995	6.749
Chicken	3.717	.7140	2.851
Egg	2.668	.0794	1.603
Milk	3.134	5.196	1.808
Vegetable	3.679	.0463	2.918
Legumes	10.72	.5929	4.404
Fruit	2.662	.1011	2.163
Condiment	1.584	.0272	1.581
Cooking Oil	1.671	.0528	1.191
Additive	4.630	.1805	3.083
Pre. Food	20.78	.8299	16.65
Alcohol	-.582	5.085	1.505
Tobacco	------	.1379	------

1. Average Information Inaccuracy * 1000
2. Unrestricted model

Table 4.22
Size Correction for the AIDS and DLOG Models

A. DLOG[a]

	Homogeneity	Symmetry	Adding-up
W^*	41.397877	3611.1386	38170.008
LR^*	36.812972	1359.5774	1849.9044
LM^*	32.610142***	740.22347	403.60542
D^*	32.765332***	829.39364	428.73850

a. n=17
*** insignificantly different from zero at α = .01

B. AIDS[b]

	Homogeneity	Symmetry	Homogeneity & Symmetry
W^*	47.874057	858.70507	1064.9876
LR^*	43.432429	708.06226	858.89611
LM^*	39.391981	595.23538	708.77866
D^*	39.632910	655.44257	795.82561

b. n=16

CHAPTER V

THE MAINTAINED MODEL - DIAGNOSTIC TESTING AND CORRECTIONS

Based on the information inaccuracy measures in Chapter 4, AIDS was chosen as the maintained model. Since all demand conditions are rejected, perhaps the next step is to determine whether an adequate representation of the data has been achieved in AIDS. Is it possible for any misspecification to be present in the model or in the error component? The test statistic used to check the correctness of the model is called a diagnostic. It is the main objective of this chapter to look at any possible source of error in the maintained model and to carry out appropriate correction. Different diagnostic procedures will be examined.

5.1 Diagnostic Tests

There are four important assumptions embodied in a classical linear regression model.

(1) The conditional mean of e_t is zero, reflecting the belief that both the functional form and the regressors are

correctly specified.

(2) Coefficient constancy. Both β and σ^2 are fixed over the sample period.

(3) Serial independence in e_t.

(4) e_t is assumed to be normally distributed.

As might be expected, violation of these assumptions generally has some deleterious effects upon model performance. Therefore it is important to apply diagnostic checking to determine the extent to which the assumptions are violated in any given context[1].

There are many well developed diagnostic checking procedures in the present econometric literature to detect the "health" of a regression model. Most of them are in the context of single equation analysis, such as Ramsey's(1969) RESET test for model misspecification, Chow's(1960) test for parameter constancy and structural change, the Breusch-Pagan's(1979) Test for heteroscedasticity, Jarque-Bera's(1980) test for normality. Extensive reviews on the topic can be found in Pagan and Hall(1983), Pagan(1983), McAleer and Deistler(1986) and Beggs(1987).

In general, little attention have been paid to the testing procedures in a system framework. Also, very little is known about it too. It is unusual to have diagnostic checking in demand studies. The reason are two-fold. First, there has been little research on testing a demand model, possibly due to its complexity. Secondly, most of applied demand research ignores the importance of diagnostic checking, concentrating mainly on discriminating among competing demand models and carrying out the usual elasticity analysis (e.g. Klevmarken[1981]). Accepting a model without going through the proper diagnostic checking is, however, dangerous. One reason is that the presence of heteroscedasticity generates inefficient estimates. As well, non-normal error distributions will produce misleading inference results.

The AIDS model under examination is of the SUR type. This means the error disturbance terms are contemporaneously correlated over different equations in the model but uncorrelated over different observations. While the OLS estimator is still unbiased, it is inefficient relative to the SUR estimator. The SUR estimator, however, is not best linear unbiased. In fact, its finite sample properties are not, in general, easy to derive. Therefore, only asymptotic results are justified. The SUR estimator is consistent and asymptotically normal. It is asymptotically more efficient than the least squares estimator. SUR is asymptotically superior, or at least not inferior, to OLS in the sense that there is a possible gain in efficiency through joint consideration of the equations, especially when the errors among different equations are highly correlated or when across-equation restrictions are imposed.

The SUR estimator reduces to OLS if the regressors in each equation are identical. Hence, if no across equation restrictions are imposed, SUR is identical to OLS and there will be no gain in efficiency from its use.

Since OLS estimates in an unrestricted SUR model are unbiased, it is valid to apply single equation diagnostic tests to the unrestricted AIDS model. Even though the OLS estimator is inefficient relative to SUR, it is still consistent. It may also be relatively inefficient due to multicollinearity of the price variables. However, inefficiency only reduces the power of the diagnostic tests and therefore does not invalidate their application[2]. Nevertheless, it is important to stress that the test statistics used here apply only to single equations. How the single equation results extend to a system of equations is not yet known. To what extent the poor equations will contaminate the whole system is also unknown. Obviously, no clear cut conclusions can be reached in the absence of a definitive theory. However, by applying the diagnostic checking procedure to each and every equation, we may at least have an insight into the causes of rejection.

Beggs(1987) suggests the undertaking of diagnostic evaluation of each equation to establish their reliability before applying systems estimation to improve the overall efficiency. He also argues that poor equations in a demand system contaminate good equations. Although diagnostic checking is necessary to evaluate the performance of each equation, we have to be careful that it should not be used to exclude poorly performing equations without any theoretical justification. From an economic viewpoint, it is generally accepted that a food demand system can be thought of as a second stage in Strotz's(1959) two-stage budgeting process (e.g. Blanciforti and Green[1983]). It is the weak separability assumption that made the two-stage budgeting process possible. Once the optimal income allocation to the food group is decided, it will be allocated to the commodities within the group so that optimal utility is achieved. Hence, the commodities should always be considered in a complete manner. In addition, the OLS estimates are poorly determined because of collinearity in the prices, and consequently any tests will have negligible power. Using the poorly determined OLS estimates for specification checking, for example, is therefore quite hazardous[3] and should be interpreted with some caution.

5.1.1 Functional Form Misspecification Test

In the standard regression framework

$$Y = X\beta + e$$

there is a stochastic specification which is referred as the orthogonality assumption. This is that the conditional mean of e is zero,

i.e. $E(e/X) = 0$

or in large sample,

$$\text{plim } (1/T)\ X'e = 0$$

implying X and e are independent. Failure of this condition means that the functional form is incorrectly specified. Misspecifying the functional form of an econometric model will, in general, lead to biased and inconsistent parameter estimates and consequently, inconsistent estimates of marginal effects and elasticities.

The functional form testing literature generally deals with two types of comparison (Beggs[1987]). One is with a non-specific alternative model, and other is to a specific non-nested alternative model. Generally, as pointed out by Godfrey, McAleer and McKenzie(1986), if there is information about the likely nature of misspecification, then the general tests for a non-specific alternative are likely to be less powerful than tests for the specific alternative which use this information. Since the latter type of non-nested testing procedures have already been discussed in Chapter 4, it will not be repeated here. There are several tests designed to detect functional form misspecification; they are discussed below.

Basically, misspecification tests are based upon a technique known as "Variable Addition". Detailed and extensive discussion on the topic is given by Pagan(1983) and Godfrey, McAleer and McKenzie(1986).

Suppose the true linear model is

$$Y_t = X_t' \beta + u_t \tag{5.1}$$

and assume the disturbance $u_t \sim NID(0,\sigma^2)$ (the term "NID" stands for normally and independently distributed). Also, assume a potentially misspecified model of

$$Y_t = X_t'\beta + Z_t'\alpha + v_t$$

A natural way to test for correct specification would be to regress u against Z and test if $\alpha = 0$. Unfortunately, the disturbances u are unknown and subsequently the least squares residuals $\hat{e}$ must be used. As explained by Pagan and Hall(1983), the shift from disturbance to residuals means the estimator $\hat{\alpha}$ will generally be an inconsistent estimator of α unless $\alpha = 0$. As well, even if the model is known to be misspecified the variable Z may not be known. In this case, a proxy of Z has to be used. The disadvantage of using proxy $\hat{Z}$

rather than Z lies in the fact that the power of the F or t statistic for $\alpha = 0$ depends upon the correlation of $\hat{Z}$ and Z. When $\hat{Z}$ is uncorrelated with Z poor power performance is likely (Pagan[1983]). Generally, cases where $\hat{Z}$ are totally unrelated to Z are rare, and adding $\hat{Z}$ to the model will at least yield some information concerning model adequacy.

The most simple and commonly applied diagnostic test for omitted variables and/or incorrect functional form is Ramsey's(1969) Regression Specification Error Test (RESET)[4], in which powers of the predictions are added to the regression equation and tested for their significance using the usual t or F statistic[5]. $\hat{Y}_t^2$, $\hat{Y}_t^3$ and higher powers are proxies for the actual misspecification. The rationale for such a procedure is obvious in that we should expect all of the additional terms to be insignificant if the model is correctly specified. We denote the tests as RESET2, RESET3 and RESET23 (with both $\hat{Y}_t^2$ and $\hat{Y}_t^3$ included).

A different specification test was proposed by Hausman(1978). The test statistic is constructed by comparing the OLS estimator, which is consistent and efficient under the null hypothesis of no misspecification, with an instrumental variables (IV) estimator, which is consistent under the alternative. A substantial difference between the two estimators implies the functional form is highly likely to be misspecified.

Hausman's(1978) testing procedure can be generalized into a single regression if the explanatory variables correlated with e can be identified. Suppose X_1 are possibly correlated with e while X_2 are uncorrelated, then the procedure is to construct an augmented model by adding an extra variable $\tilde{X}_1$ such that

$$Y = X_1\beta_1 + X_2\beta_2 + \tilde{X}_1\alpha + v$$

where $\tilde{X}_1 = P_zX_1$ and $P_z = Z(Z'Z)^{-1}Z'$. Z are instrumental variables which should include X_2. Testing $H_0: \alpha = 0$ would indicate whether the model is misspecified.

White(1980b) also suggested a functional form test which is, in principle, similar to Hausman's(1978) specification test. Instead of analysing the OLS and IV estimators, it compares the OLS with the Weighted Least Squares (WLS) estimator. The test can be generalized into an artificial regression

$$Y = X\beta + (\Omega^{-1}X)\,\alpha + v$$

White(1980b) suggests using weights that are the reciprocal of the squared prediction under the null hypothesis. That is, run

$$Y = X\beta + (X/\hat{Y}^2)\,\alpha + v$$

and test for the additional term by seeing if $\alpha = 0$.

The three testing procedures, Ramsey's(1969) RESET test, Hausman's(1978) specification test and White's(1980b) functional form tests are all very similar and it is therefore difficult to select among them. A Monte Carlo study conducted by Godfrey, McAleer and McKenzie(1986) shows that RESET and White's(1980) test are well behaved in terms of size, and are similar in terms of power. However they found that RESET is more robust against non-normal error distribution. Together with its simplicity, RESET is therefore more useful and the results of RESET tests are reported in Table 5.1. Only 6 of the total 17 equations passed all RESET tests easily. This implies the existence of serious functional form misspecification in the maintained model. Although the form of misspecification is unknown, it may be related to the omission of relevant explanatory variables. The possible variables omitted are urban/rural dummy and 5 island dummies[6]. In addition, misspecification of the error component (e.g. heteroscedasticity in the error variance) is another possible problem. Both possibilities will be examined later in this chapter.

5.1.2 <u>Parameter Constancy Test</u>

As reliable inference in econometric models generally depends upon certain parameters remaining constant over a set of observations, parameter constancy tests have an important role in modelling. Assuming there is a known switch point for all parameters in the model, parameter inconstancy tests can be applied. They are considered below.

(i) <u>In the context of the univariate model</u>

Chow(1960) developed two well known tests for parameter constancy and structural stability for the general model in (5.1). Since the stability test is only associated with a time series study, we will look at the parameter constancy test and refer it as the "Chow test".

The Chow test determines whether the whole structure of a regression model has changed between one set of data and another. The test itself looks at the entire set of coefficients and determines whether or not the regression surface has changed. It may be explained as follows.

Suppose we have g observations in one group of data and h observations in the other. The total number of observations is (g+h) = T. Given k regressors, including the intercept term, Chow's(1960) F-ratio test statistic can be computed as follows:

(1) fit a regression using the same k regressors to each set of data separately and calculate SSE(g) and SSE(h), which are

the sum of squares error for the models with g and h observations respectively.

(2) Fit another regression with the same regressors to the pooled g+h=T observations and calculate SSE(T).

(3) Chow's test statistic is calculated as

$$C = \{ [SSE(T) - [SSE(g)+SSE(h)]] / [SSE(g) + SSE(h)] \} * \{ (T-2k) / k \}$$

which is distributed as F(k, T-2k).

A significantly large value of F leads us to reject the null hypothesis and to conclude that there is parameter inconstancy in the set of regression coefficients considered as a whole. Also, to ensure positive degrees of freedom, we must have $T > 2k$, $g > k$ and $h > k$.

(ii) In the context of multivariate model

Based upon the LR principle, Anderson and Mizon(1983) considered an extension of the univariate Chow test to dynamic, non-linear simultaneous equations. Denote the simultaneous equations analogues of Chow test statistic as AM. By definition

$$AM = -2 (L_T - L_g - L_h)$$

which is distributed as $\chi^2(p + n(n+1)/2)$ where p and n are respectively the total number of parameters and the number of equations to be estimated in the model. L_T, L_g and L_h are the log likelihood functions with reference to T, g and h observations respectively.

Harvey(1981) shows that the joint density function of the SUR system with n equations at t is given by

$$P(Y_t) = (2\pi)^{-n/2} * |\Omega|^{-1/2} * \exp\{ (Y_t-X_t\beta)' \Omega^{-1} (Y_t-X_t\beta) \}$$

The log likelihood function for all T observations is therefore

$$L_T = (-Tn/2 * \log 2\pi) - (T/2 * \log |\Omega|) - [1/2 *$$

$$\Sigma_t (Y_t-X_t\beta)' \Omega^{-1} (Y_t-X_t\beta)]$$

Note that the last summation term on the right hand side is exactly the term " objective * T " generated by the SAS/ETS.

Both univariate and multivariate testing procedures have been applied to AIDS. The results of the univariate Chow test for urban and rural samples are presented in Table 5.2. This shows that 10 out of the 17 equations significantly rejected the null hypothesis, implying the problem of parameter inconstancy exists between the urban and rural samples.

Two estimations of the multivariate AIDS were carried out with respect to the urban and rural samples. The results are given in Table 5.3. All demand conditions were strongly rejected in both samples by the Wald, LR and LM statistics. The Anderson and Mizon's(1983) AM test statistics were calculated for the unrestricted and restricted models and the results were reported in Table 5.4. As the statistics are significantly different from the critical values, for the demand system as a whole there is inconstancy in parameters between urban and rural regions in all unrestricted and restricted models. The results support the earlier findings from the univariate models.

The immediate and conventional correction for parameter inconstancy is to include a dummy variable in the model. In this case, a urban/rural dummy is suggested. Also it is possible that the six groups of islands themselves may contribute to the misspecification. Unfortunately we are unable to test for parameter constancy among islands because one of the six main islands, Makulu and Irian Jaya, has only 11 observations, violating the condition of positive degrees of freedom. Consequently, we re-estimated the equations with an urban/rural dummy and a set of 5 island dummies. Notice that the dummies are just shift effects and have no effect on slope coefficients.

The augmented models with added dummy variables were again tested using RESET2 and RESET3 and the results are reported in Table 5.5. Ten out of the 17 equations are misspecified even with both urban and island dummies included. Comparing the results with Table 5.1, no striking improvement is observed. In order to test for the usefulness of adding dummies in the model, we applied a simple F-ratio statistic to test for their joint significance (Hebden [1983]) in the univariate models. The test statistic is calculated as follows

$$F_{(p,q)} = [(R^2 \text{ with dummies} - R^2 \text{ without dummies}) / p\ ] \ *$$
$$[\ q / (1 - R^2 \text{ with dummies})\ ]$$

where p is the number of dummies in the model and

q = (no. of observations - no. of dummies - no. of parameters)

The R^2 statistics for each model are presented in Table 5.6 and the F-ratio test statistics are reported in Table 5.7. Significantly large F values imply the dummies are statistically significant in the model. Surprisingly, this shows that in most of the situations the dummies are significantly different from zero. Including both urban and island dummies, 16 of the 17 equations show statistical improvement.

As well, the unrestricted and restricted SUR systems were estimated with the dummies included and tested for demand conditions. The results are given in Table 5.8, Table 5.9 and Table 5.10 for the model with urban dummy, island dummies, and both kinds of dummies respectively. Although the test statistics strongly rejected the demand conditions, there is significant improvement after incorporating dummies. Based on the calculated information inaccuracy measures, I and IC, inclusion of an urban dummy reduced the inaccuracy by 3% to 6% compared to the maintained model in Table 4.18. The island dummies, on the other hand, improved the model by 20% to 24%. By incorporating both urban and island dummies, there is an average improvement of 21% to 25%. Examining the Wald, LR, LM and GJ statistics, inclusion of dummies has generally reduced the test statistics to an extent that varied between restrictions and the type of dummies included. The urban dummy alone reduced the statistics for the symmetry condition by 7% to 10% but, on the contrary, increased the statistics for homogeneity by 7% to 13%. The island dummies have reduced the statistics by 7% to 10% for the homogeneity condition, and a large extent of 40% to 47% for the symmetry condition. Inclusion of both kinds of dummies decreased the homogeneity statistics by 10% to 12% and a striking decrease of 46% to 57% for the symmetry condition. Even though there is such a large reduction, the demand conditions are still significantly rejected by the statistics. Nevertheless, the inclusion of dummies is encouraging both in terms of the fitness of the models and the inference statistics.

Another LR statistic is constructed to test for the joint significance of those dummy variables in the SUR system. The LR formula is the same as the one reported in Chapter 4, except the unrestricted and restricted regressions are now referring to the models with and without dummy variables. The results are also given in Table 5.8, Table 5.9 and Table 5.10. The dummies are found to be significantly different from zero in every situation. Once again, this reinforced the importance of the dummy variables. Therefore it is apparent that it is necessary to include both urban and island dummies in AIDS.

5.1.3 <u>Heteroscedasticity Test</u>

Another possible source of misspecification is from the error component. As stated, one of the stochastic specifications of a classical model is that the error disturbance is distributed with constant but unknown variance,

i.e. $\quad Var(e/X) = \sigma^2 I$

sometimes referred to as the sphericality assumption. If the disturbances do not have constant variance for each observation, they are said to be heteroscedastic. Applying OLS when the sphericality condition is violated will lead to inefficient but still unbiased parameter estimates. In addition, we are likely to obtain a biased estimator of the covariance matrix of the estimates which leads to misleading statistical inferences. Goldfeld and

Quandt(1972) provide a clear exposition of the likely effects and the possible detecting methods of heteroscedasticity (e.g. the Goldfeld and Quandt Test, the Ramsey Test and the Peak Test). Assuming a linear model in (5.1), there are two simple diagnostic procedures to test for the presence and source of heteroscedasticity and are discussed below.

5.1.3.a <u>A General LM Test for Heteroscedasticity</u>

Assume the variance of the error term is has the relationship

$$\sigma_t^2 = \sigma^2 + Z_t'\gamma$$

It follows that the LM test for an unspecified form of heteroscedasticity as illustrated by McAleer and Deistler(1986) is calculated by regressing the squared OLS residual, $\hat{e}_t^2$, on a constant and the source of error variance Z_t[7]. If $\hat{Y}_t$ is used as a proxy to Z_t, the auxiliary regression becomes

$$\hat{e}_t^2 = \alpha + \gamma\hat{Y}_t + v_t$$

which is similar to the variable addition approach. The LM test statistic to test for $\gamma = 0$ is obtained by multiplying the R^2 from the auxiliary regression by the number of observations T and is distributed as $\chi^2(1)$ under the null hypothesis of homoscedasticity. Besides, the test statistic is, generally, robust to non-normality (Goldfeld, McAleer & McKenzie[1986]).

The test has been applied to each commodity equation in four different AIDS models, the maintained AIDS without any dummy variable, the AIDS with an urban dummy, AIDS with island dummies, and AIDS with both urban and island dummies. The results are given in Table 5.11. They show that a majority of the equations do have an heteroscedastic problem given that 11 of the 17 equations in the maintained model significantly rejected the null hypothesis of homoscedasticity. The situation is slightly worse after including the dummy variables. Thirteen equations imply heteroscedasticity when incorporating both urban and island dummies. In order to determine its source, another test statistic has to be used.

5.1.3.b <u>Breusch-Pagan Test: a Specific Test for Heteroscedasticity</u>

The LM test above is simple and general enough to compute, but the source of heteroscedasticity is still undetermined. If we have some idea about the nature of the variance, that is, if Z_t is known, we can then apply the Breusch-Pagan(1979) test.

To calculate the test statistic

- run OLS on the original model to obtain the estimated variance of the model $\hat{\sigma}^2$, and the squared residual $\hat{e}_t^2$.

- then run an auxiliary regression of $(\hat{e}_t^2/\hat{\sigma}^2)$ on the suspected source of variance Z_t (with a constant term included).
- the Breusch-Pagan LM test statistic is obtained by multiplying the explained sum of squares from the auxiliary regression by 1/2 with the statistic being distributed as $\chi^2(q)$ under the null hypothesis where q is the number of Z_t (excluding the constant term).

The suspected source in this case may come from some exogenous factors, such as the number of members in the household, the number of adults, total household income, total household expenditure, and the total expenditure on food items. The logarithm of each of the variables, signified respectively as LKTHH, LKTAH, LKTY, LKTE and LEXP, are tested separately in the four AIDS models. The results are given in Table 5.12, Table 5.13, Table 5.14 and Table 5.15. For the maintained model (Table 5.12), heteroscedasticity present in the majority of the 17 equations in relating to LKTY, LKTE and LEXP. similar results are observed in Table 5.13, Table 5.14 and Table 5.15, reinforcing the finding that the possible sources of error variance are LKTY, LKTE and LEXP.

Another source of heteroscedasticity is originated from the nature of grouped data (Johnston[1983] and Kmenta[1971]). Since each observation in the study is a representative household in each geographical location (as was explained in Chapter 3), it is a grouped average of individuals' information. Suppose a regression model

$$y_t = \alpha + \beta x_t + e_t \qquad t = 1, 2, \ldots n$$

where e_t is homoscedastic. If the data have been averaged within G groups, where n_g indicates the number of observations in the g'th group and g = 1, 2, ... G, then the appropriate model is

$$\overline{y}_g = \alpha + \beta\overline{x}_g + \overline{e}_g$$

and clearly

$$E(\overline{e}_g) = E[(1/n_g)(e_{1g}+e_{2g}+\ldots+e_{n_g g})] = 0$$

and $$Var(\overline{e}_g) = (1/n_g)(\sigma^2+\sigma^2+\ldots+\sigma^2) = (n_g\sigma^2)/(n_g{}^2) = \sigma^2/n_g$$

It means that unless the number of observations is the same in every group, the disturbance for grouped data is heteroscedastic. Kmenta(1971) also demonstrates that by grouping the observations and estimating the regression coefficients from group means rather than from the individual observations, we are losing some information contained in the sample, namely the information about the variation of observations within each group. Therefore, there would be some loss of efficiency in going from estimation based on all individual observations to estimation based on group means unless there is no variation of the values of X within each group. Further, the loss of efficiency will be small if the variation in X within group is small compared with the variation of the group means of X around the overall mean (Kmenta[1971]). This conclusion holds whether the groups contain the same number of observations or not. In other word, having groups of equal size would make e_g homoscedastic but would not prevent a loss of efficiency as a result of grouping.

Detecting the existence of heteroscedasticity in a system of equations is very difficult. Kelejian(1982) has proposed a large sample test if heteroscedasticity is associated with one or more equations in a linear simultaneous equations system (assuming the exact specification is unknown). The testing procedure is simple only if heteroscedasticity is presented in only one equation of the system. It involves two auxiliary regressions of the squared residuals on the source elements (which are assumed to be contained in the regressors), and on a single constant term. Suppose the concern for heteroscedasticity relates only to the i'th equation and the remaining equations are assumed to be trouble-free. Specifically, suppose that $var(e_{ti})$ is a bound function of some or all of the elements of the non-stochastic exogenous variables, say $G(X_t)$. Suppose also that the exact specification of the function is not known, but that the approximation

$$G(X_t) \approx d_{ii} + P_{ti}\alpha_i$$

is considered, where P_{ti} is a (1^*q_i) vector of observable functions of the relevant elements of X_t. Let $RSS_e{}^i$ and $RSS_R{}^i$ be, respectively, the error sums of squares from the least squares regression of $\hat{e}_{ti}{}^2$ on $(1, P_{ti})$, and of $\hat{e}_{ti}{}^2$ on only the constant term. The large sample test for heteroscedasticity is

$$K_i = [\ T^{-1}\Sigma\hat{e}_{ti}{}^4 - (T^{-1}\Sigma\hat{e}_t{}^2)\ ]\ (RSS_R{}^i - RSS_e{}^i)$$

which is distributed as $\chi^2(q_i)$.

But if heteroscedasticity is suspected to be existed in more than one equation, the computation of the test statistic is very complicated. Generalizing, suppose only the disturbance terms in the first r equations are heteroscedastic. Suppose also that

$$\sigma_{ijt} \approx d_{ij} + P_{ijt}\alpha_{ij} \qquad\qquad i, j \in\ S$$

where S denotes the set of r combinations considered and P_{ijt} is a (1^*q_{ij}) vector of observations on q_{ij} functions of the elements of

X_t. Basically, the procedure will be to jointly estimate the parameter vector, α_{ij}, say $\alpha' = (\alpha'_{i1j1}, \ldots, \alpha'_{irjr})$, by a GLS procedure based on the estimated values of the squares and cross products of the corresponding disturbance terms, and then establish a large sample test for H_0 by testing for $\alpha = 0$. Let

$$\hat{\psi}_t = [\ (\hat{e}_{ti1}\hat{e}_{tj1}), \ldots, (\hat{e}_{tir}\hat{e}_{tjr})\ ]$$

be an estimate of the covariances ψ_t. Under H_0, let

$$E(\hat{\psi}_t) = \hat{\eta}$$

and
$$E[(\hat{\psi}_t-\hat{\eta})'(\hat{\psi}_t-\hat{\eta})] = \hat{V}_\psi$$

Also, let $\hat{V}_\psi$ be any consistent estimator of V_ψ, which is the covariance of ψ_t. Under H_0, one such estimator is

$$\hat{V}_\psi = [\ \Sigma(\hat{\psi}_t-\hat{\eta})'(\hat{\psi}_t-\hat{\eta})\ ] / T$$

where $\hat{\eta} = (\hat{\psi}_1 + \ldots + \hat{\psi}_T) / T$

The large sample test for heteroscedasticity is

$$K = \hat{\alpha}'\ (P\hat{V}_\varepsilon{}^{-1}P)\ \hat{\alpha}$$

where $\hat{\alpha} = (P'\hat{V}_\varepsilon{}^{-1}P)^{-1}P'\hat{V}_\varepsilon{}^{-1}\ \mathrm{vec}(\hat{\psi})$,

$$\hat{V}_\varepsilon = \hat{V}_\psi \otimes I_T,$$

$$vec(\hat{\psi})' = [(\hat{e}_{1i1}\,\hat{e}_{1j1}),\ldots,(\hat{e}_{Ti1}\,\hat{e}_{Tj1}),\ldots,(\hat{e}_{1ir}\hat{e}_{1jr}),\ \ldots,(\hat{e}_{Tir}\hat{e}_{Tjr})],$$

and $P = diag(P_{iSjS})$. The test statistic K is distributed as $\chi^2(h)$ where $h = (q_{i1j1} + \ldots + q_{irjr})$.

The calculation of the test statistic K is so inefficient and complicated that its feasibility is doubtful.

5.1.3.c Weighted Least Square Procedure

White(1980a) points out an important aspect of cross-sectional surveys, which is that the stochastic regressors X_t and e_t are independent but not necessary identically distributed (i.n.i.d.). While the non-linear least square estimator is proved by Jennrich(1969) to be strongly consistent and asymptotically normally distributed, it is based on the condition that the stochastic regressors X_t and e_t are independent and identically distributed (i.i.d.). It is therefore important to ensure that these properties can be retained even under the condition of i.n.i.d.. White(1980a) suggests an estimated weighted least squares method and found that the estimators are both strongly consistent and asymptotically normal. The difficult question is how to correctly specify the weights.

White(1980a), implicitly assuming the source of error variance is known, suggests that partitioning the full data set into fixed finite number of cells, say c_j where $j = 1, 2, ..., J$, for which the error variances differ, and follows.

First, obtain a consistent estimate $\hat{\theta}_n$ by solving

$$\min \quad T^{-1} \Sigma_t^T (Y_t - f(X_t,\theta))^2$$

Secondly, consistently estimate $\hat{\sigma}_{jn}$

$$\hat{\sigma}_{jn} = T_j^{-1} \Sigma_{t \in cj} (Y_t - f(X_t,\hat{\theta}_n))^2$$

and let $\hat{w}_{tn} = 1 / \hat{\sigma}_{jn}^2 , \quad t \in c_j, \quad j = 1, 2, ..., J$

And finally, obtain the consistent estimated WLS estimator, $\tilde{\theta}_n$ by solving

$$\min \quad T^{-1} \Sigma_t (Y_t - f(X_t,\theta))^2 \hat{w}_{tn}$$

Generally, as pointed out by White(1980a), the non-linear WLS will not be asymptotically efficient. But if the errors are normally distributed with zero mean and are independent of the regressors, the WLS becomes the Maximum Likelihood Estimator and is asymptotically efficient.

The question still remains of what criteria should one subset the sample observations. This means the source of the error variance has to be identified before subsampling the full samples. In fact, this is the problem encountered given correction for heteroscedasticity using the standard weighted least square procedure.

Basically, the presence of heteroscedasticity in the error component is one of the phenomena of i.n.i.d.. If errors exhibit different variances across samples, their distributions are not identical. Hence, the suggestion by White(1980a) of using the weighted least squares method is not surprising.

The simplest form of WLS estimation is as follows:

Assume heteroscedasticity is presented in a linear model,

$$Y_t = X_t\beta + e_t$$

with the relationship

$$E(e_t^2) = \sigma^2 Z_t$$

where Z_t is known and exogenous. Then divide every dependent and independent variables in the equation by the square root of Z_t. Applying OLS to the transformed equation produces efficient estimates if the assumed form of heteroscedasticity is correct. It can be demonstrated as follows:

Suppose $Y_t = X_t\beta + e_t$ where $E(e_t^2) = \sigma^2 Z_t$

Then $Y_t / (Z_t^{1/2}) = (X_t / (Z_t^{1/2})) * \beta + \varepsilon_t$ where $\varepsilon_t = e_t / (Z_t^{1/2})$

Hence $E(\varepsilon_t^2) = E(e_t^2 / Z_t) = (1/Z_t) * E(e_t^2) = (1/Z_t) * \sigma^2 Z_t = \sigma^2$

If heteroscedasticity results from several explanatory factors, then the correction of the procedure is as follows:

(1) assume

$$E(e_t^2) = \sigma_t^2 = \alpha_0 + \alpha_1 Z_{t1} + \alpha_2 Z_{t2}$$

(2) calculate $\hat{e}_t = Y_t - X_t \hat{\beta}$

(3) regress $\hat{e}_t^2 = \alpha_0 + \alpha_1 Z_{t1} + \alpha_2 Z_{t2}$

(4) calculate $\hat{\sigma}_t^2 = \hat{\alpha}_0 + \hat{\alpha}_1 Z_{t1} + \hat{\alpha}_2 Z_{t2}$

(5) divide each observation by the weight $\hat{\sigma}_t$, which is the square root of $\hat{\sigma}_t^2$, and then apply the usual least squares procedure to the transformed model.

From the earlier Breusch-Pagan Test results and the illustration by Johnston(1983) and Kmenta(1971), there is evidence to show that the possible explanatory factors of heteroscedasticity are LEXP, LKTY, LKTE and the inverse of the number of observations in each group, GHH. The reciprocal of their square roots are separately used to weight the two AIDS models. The first model is the maintained AIDS model without any dummy variables. The second AIDS model includes both urban and island dummies. The weighted models are tested against heteroscedasticity using the general LM test statistics, and the results are presented in Table 5.16 and 5.17. From the results, it is obvious that the inverse of GHH is not the major cause of heteroscedasticity relative to income factors, LEXP, LKTY or LKTE. Among the three income factors, eleven equations in the weighted maintained model were characterised by rejection of the null hypothesis of homoscedasticity. Inclusion of dummies has only marginally improved the situation. Comparing the results with those in column two and column five of Table 5.10 (the unweighted models), it is clear that both weighted and unweighted results are almost identical. Generally, there is no major improvement in correcting heteroscedasticity in the use of the weighted approach. Also the results are invariant to the choice between the three weighting factors.

The same weightings were then applied to two multivariate AIDS models, the maintained model, with and without the dummies. The inference results are recorded in Tables 5.18 and 5.19. Yet the demand conditions are strongly rejected. But it is important to compare the two tables with Tables 4.18 and 5.10 (the unweighted AIDS model excluding and including dummies), since this illustrates the usefulness of the weighted approach and the dummies.

Comparing Table 4.18 with Table 5.18, and Table 5.10 with Table 5.19, the effects of incorporating the weighting factors with reference to the multivariate models with and without dummies. There is a dramatic improvement in Theil's(1971) inaccuracy measures given that the corrected information inaccuracy values, IC, are sharply reduced by 67% to 70%. This means the performance of the models are significantly improved by using weighting factors, and the choice of the weights have been found to be unimportant. On the other hand, there is no improvement in the test statistics themselves. The test statistics for the within equation restriction, homogeneity, are only reduced by about 1% by excluding dummies and about 3% by including dummies, with the other statistics being increased by a small proportion (about 0.8%). This implies the weighting does not help when testing demand conditions.

On the other hand, comparing Tables 4.18 with 5.10 (unweighted AIDS model with and without dummies), and Tables 5.18 with 5.19 (weighted AIDS model with and without dummies) revealed a different story. While the corrected inaccuracy measures are improved by about 20% to 26%, the test statistics are also reduced. The homogeneity test statistics decreased by 12% to 14% and the others are reduced by a large magnitude, about 48% to 57%. This leads us to the conclusion that while weighted factors have significantly improved the overall fit of the model, the inclusion of dummies has reduced the test statistics by a large proportion. Even so, all the demand conditions are rejected. Hence, the weighted AIDS model with urban and island dummy variables is preferred. The inaccuracy measures show that the weighting factor, the inverse of the square root of LKTY, is marginally superior to the other, and is therefore the preferred weighting factor. Subsequently, the weighted AIDS model with urban and island dummies and weighting factor, the inverse of the square root of LKTY, is chosen as the new maintained model.

The parameter estimates for the new maintained model are presented in Table 5.20. The associated own price and income elasticities of the model are shown in Table 5.21. Except chicken, milk and prepared food have (insignificant) positive own price elasticities, all own price elasticities are negative. Comparing the Table 5.21 with the results from the unweighted AIDS model in Table 4.19, we find that the results are fairly similar. The only difference is the signs of the own price elasticities of chicken and prepared food which are now positive (but insignificantly different from zero). Meat and legumes remained as inferior items but statistically insignificant. Rice, dried fish, egg, vegetable, fruit, condiment, cooking oil and additives are necessities (statistically significant). The only luxury item which is statistically significant is milk. The substitution and income effects are given in Table 5.22. The negativity condition is satisfied but only three of the terms are significant. Among the 136 pairs of substitution coefficients, 68 are consistent in sign, of which 52 are substitutes (only 3 are significant) and 16 are complements (all insignificant).

The averaged information inaccuracies are calculated and reported in Table 5.23. Comparing with the results of the unweighted AIDS model in Table 4.21, there is obvious improvement after incorporating dummies and weighting factor. The joint significance tests for the urban and island dummies using the LR principle are recorded in Table 5.24. The test statistics, once again, confirm the joint significance of dummies in every situation. Size corrected statistics for the new maintained model are presented in Table 5.25. Even the size corrected test statistics are reduced, they are still far too large to accept the restrictions.

5.1.4 Normality Test

In a classical linear regression model

$$Y = X\beta + e$$

Suppose this satisfies the four basic assumptions described early in this chapter. Given normality, the OLS estimator is BLUE since it attains the Cramer-Rao bound. Unfortunately, the effects of departures from normality are not easily or clearly understood (Gnanadesikan[1977]). But, generally, the estimator will have poor efficiency and is no longer BLUE (White and MacDonald[1980]).

Unlike the Maximum Likelihood Method, the Non-linear Least Squares Estimator does not require the errors to be normally distributed. But the Wald, LR, LM and GJ test statistics all require normality. Otherwise, the tests are, strictly speaking, invalid.

As well, the diagnostic tests presented earlier assume the normality condition. Although RESET tests are found to be robust against a non-normal distribution, the LM tests, in general, are not particular useful since they can be very unreliable when disturbances are not normally distributed (Godfrey, McAleer and McKenzie[1986]).

Basically, the classical method of evaluating univariate normality is by measuring skewness and kurtosis. The measures may be utilized individually or can be combined into an omnibus test statistics. Gnanadesikan(1977) has briefly discussed several univariate normality test statistics while their properties are examined by White and MacDonald(1980) using simulation techniques. Jarque and Bera(1980)'s normality test, based on the

Lagrange Multiplier principle, will now be examined.

Normal disturbances have the property that the third moment (m_3) is zero and the forth moment (m_4) about the mean is three times the square of the second moment, the variance (σ^2). The normality hypothesis can be tested by considering the joint hypothesis

$$m_3 = 0,$$

and $$m_4 - 3\sigma^4 = 0$$

in the system of SUR equations.

The test statistic for normality derived by Jarque and Bera(1980) is given by

$$LM_N = (\hat{\gamma}_3{}^2 / 6\hat{\sigma}^6 + \hat{\gamma}_4{}^2 / 24\hat{\sigma}^8)$$

which is distributed as $\chi^2(2)$,

where $\hat{\gamma}_3 = T^{-1} \Sigma_t \hat{e}_t{}^3$, and

$$\hat{\gamma}_4 = T^{-1} \Sigma_t (\hat{e}_t{}^4 - 3\hat{\sigma}^2 \hat{e}_t{}^2)$$

Jarque and Bera's(1980) normality test is performed on each single equation of four different AIDS models, the unweighted and weighted models with and without dummies. The test results are reported in Table 5.26. Only one of the 17 equations, the budget share of tobacco, has satisfied totally the normality condition in every model.

Jarque and McKenzie(1983) followed Mardia's(1970) multivariate measures of skewness and kurtosis and suggested a multivariate normality test statistic which can be applied to simultaneous equations models. The statistic, denoted as JM, is calculated as follow,

$$JM = T \{ b_1/6 + [b_2-n(n+2)]^2/[8n^*(n+2)] \}$$

$$b_1 = 1/T^2 * \Sigma_i\Sigma_j(e_i'\Omega^{-1}e_j)^3$$

$$b_2 = 1/T * \Sigma_i(e_i'\Omega^{-1}e_i)^2$$

where n is the number of equations in the model, e_i and e_j are mean adjusted (n*1) residuals vector with i,j = 1, 2, ..., T, and $\Omega = \Sigma_i(e_ie_i'/T)$. The JM test statistic is asymptotically distributed as $\chi^2(r)$ where $r = [n^*(n+1)^*(n+2)/6]+1$. If n=1, JM would reduce to the univariate Jarque and Bera's LM test statistic. As pointed out by Jarque and McKenzie(1983), the JM test has not been shown to satisfy any optimality condition. Therefore, its power properties are unknown. However, univariate techniques can also be used to evaluate multivariate non-normality. Gnanadesikan(1977), Mardia(1980) and Jarque and McKenzie(1983) all claim that although marginal normality does not necessary imply joint normality, the presence of multivariate non-normality will be reflected by the marginal distribution and is therefore likely to be detected by univariate techniques. Naturally when there is a need for tests that explicitly exploit the multivariate nature of the data so as to yield more conclusive analysis, statistics such as JM test statistic are required. But in this case, statistical results from the univariate Jarque and Bera's(1980) normality test suggest that the errors are multivariate non-normal. It is well known that violation of the normality assumption may lead to inaccurate inferencial statements, and hence, the inference results obtained should be exercised with care. In general, the RESET test and the LM test with $T^{*}R^{2}$ principle are robust to non-normality (Godfrey, McAleer & McKenzie[1986]).

5.2 Additivity of Diagnostics

Strictly speaking, the univariate heteroscedastic tests (the general LM and the Breusch-Pagan tests) discussed previously are incorrect in this case because they are, naturally, "one-directional tests". They have optimal properties only when all other standard underlying assumptions are satisfied. For example, the heteroscedastic residual tests explicitly assume the error to be normally distributed. Similarly, the Jarque-Bera's normality test assumes the presence of homoscedastic residuals. Most of the uni-directional tests are designed to verify the validity of one particular specification at a time and are not, in general, robust in the presence of other misspecification. As noted by Bera and Jarque(1982), it is analytically difficult to examine the robustness of uni-directional tests when the required assumptions are violated. They examined the problem extensively using a simulation study and found that when fewer-directional tests are applied than actually required, inferences will not be reliable due to the lack of robustness of the smaller directional tests (Bera and Jarque named this phenomenon "undertesting"). Similarly, "overtesting" occurs when applying higher-directional tests than is required and may affect the power and the significance level of the tests. Simulation results reported by Bera and Jarque(1981) show that the consequences of overtesting are not very serious, whereas those of undertesting can lead to highly misleading results. Also, for one-directional tests, violation of maintained assumptions can lead to a loss of power, incorrect conclusions and inefficient testing procedures.

The question now is: how are the uni-directional tests coordinated to reflect the multi-directional departures possible from the standard regression assumptions? One of the possible solution is to construct an omnibus test (e.g. a joint test) for all specification errors deemed likely. The potential for simplification is that by assuming independence under the null hypothesis, uni-directional tests may be added to produce an omnibus test. Jarque and Bera(1980) and Bera and Jarque(1982) propose a LM procedure to derive efficient joint test for normality, homoscedasticity and serial independence. The independence property of diagnostics is derived from the block diagonality of the inverse information matrix[8] (Pagan[1983] and Pagan and Hall[1983]). It is a multi-directional test and is simple to compute because of its "additive" nature. The joint test itself, in fact, is the combination of the uni-directional tests and is distributed as χ^2 with k degree of freedom where k is the sum of degrees of freedom from the uni-directional tests. Jarque and Bera(1980) claim that the test is asymptotically equivalent to the LR test and has maximum local power for large samples, and proposed a multiple comparison procedure (Bera and Jarque[1982]) to identify the sources of departure from the null. This test relies on the repetition of a random sampling process to derive empirical critical values which are then used to evaluate the statistics computed from the original data. This multiple comparison process is found to perform reasonably well in their study.

But the simple solution of adding the diagnostic tests to form a joint test on misspecification has not been utilized extensively, possibly because of the belief that the tests would not be independently distributed. As argued by Pagan and Hall(1983), " ... the major disadvantage to an examination of additivity properties through the information matrix resides in the fact that this quantity is based upon a set of specific distributional assumptions, and independence may therefore be reflecting nothing more than the chosen distribution". Pagan and Hall(1983) developed the conditions required for independence of diagnostic test statistics based on a residuals approach. Essentially, the tests can be separated into two blocks. The first one includes tests for misspecification, serial correlation and heteroscedasticity, while the second contains the normality test. Their finding is that the two blocks are always additive but additivity within the first block does not always hold.

In a cross-sectional survey, serial independence must be satisfied. But if following Bera and Jarque's approach to test heteroscedasticity and normality, we should adopt a "two-directional" joint test, LM_{HN}, which is,

$$LM_{HN} = LM_H + LM_N$$

and is distributed as $\chi^2(3)$. LM_H and LM_N refer to the Breusch-Pagan test of $\chi^2(1)$, and Bera and Jarque Test of $\chi^2(2)$, respectively.

Due to the fact that both one-directional tests all rejected the null hypothesis of homoscedasticity and normality, we expect the same findings from the joint tests.

Since the normality condition is violated, the question arises is: to what extent is the rejection of demand theory due to the non-normal nature of the residuals? In order to answer the question, we have to construct a Monte-Carlo experiment, "Parametric Bootstrapping", a methodology to be discussed in the following chapter.

Footnotes

1. Assumption (3) is naturally satisfied in a cross-sectional study.

2. I am grateful to Dr. M. McAleer for his helpful comment.

3. I am thankful to Dr. R. Byron for this argument.

4. The test was modified later by Ramsey and Schmidt(1976).

5. As distinquished by Pagan(1983), three options can be used to test that $\alpha = 0$. Two of them are now discussed.

   The first option is to fit the augmented model
   $$Y = X\beta + Z\alpha + v$$
   and test $\alpha = 0$ using the t or F statistic.

   The second option involves the subtraction of $X\beta$, the predicted value, from both sides and yields a model
   $$Y - X\hat{\beta} = \hat{e} = X(\beta-\hat{\beta}) + Z\alpha + v$$
   Then regress the above model and test if $\alpha = 0$ using the t or F statistic.

   Since the transition from the first to the second option only involves the subtraction of the same quantity from both sides, the estimator of α and the associated t or F statistic from the models must be identical. Accordingly, both give identical answers to the hypothesis of $\alpha = 0$. In this study, the second option is chosen when applying RESET tests.

6. Indonesia is composed of six main groups of islands, but to avoid singularity, one of the six island dummies is redundant and is therefore excluded.

7. The procedure is originated by White(1980) which involves regressing the squared OLS residuals from the estimation of the null model on the cross-products of the regressors.

8. The block diagonality and subsequently the independence will fail in the presence of lagged dependent variable.

Table 5.1
RESET Tests on the Maintained AIDS

Equation	RESET2	RESET3	RESET23[a]
Rice/Grain	-.362*	-.358*	.0660*
Tubers	3.914	2.532	20.25
Fish	.2025*	-.552*	5.976
Dried fish	-2.25	-2.86	9.274
Meat	1.481*	1.022*	3.264
Chicken	-.134*	-.711*	3.663
Egg	2.308	2.052*	3.200
Milk	5.537	5.554	15.85
Vegetable	1.880*	1.867*	1.769*
Legumes	5.967	4.107	31.86
Fruit	1.812*	1.676*	2.636*
Condiment	1.564*	1.648*	1.567*
Cooking oil	-2.63	-2.72	3.963
Additives	-1.29*	-1.38*	1.888*
Pre. food	14.22	14.50	114.4
Alcohol	10.44	14.85	117.8
Tobacco	-.137*	-.196*	.1568*

a. Distributed as $F(2,403)$

* Insignificantly different from zero at $\alpha = .05$
critical value for $F(2,403) = 3.0181$

Table 5.2

Chow Test for Parameter Constancy for the Maintained AIDS

Equation	SSE(n+m)	SSE(n)	SSE(m)	F(19,386)
Rice/Grain	.9672	.2159	.6259	3.027
Tubers	.1475	.0146	.1175	2.383
Fish	.5624	.1897	.3118	2.466
Dried fish	.2314	.0672	.1398	2.372
Meat	.1683	.0083	.0544	4.541
Chicken	.0403	.0215	.0155	1.864***
Egg	.0213	.0103	.0082	3.089
Milk	.0172	.0116	.0027	4.232
Vegetable	.1838	.0745	.0884	2.606
Legumes	.1022	.0367	.0513	3.280
Fruit	.0625	.0205	.0357	2.243
Condiment	.0453	.0201	.0227	1.204**
Cooking oil	.0327	.0078	.0219	1.953
Additive	.2335	.0820	.1381	1.238**
Pre. food	.5945	.3579	.1855	1.910***
Alcohol	.0017	.0006	.0011	.7815*
Tobacco	.9959	.3416	.5803	1.631***

n = samples in urban regions = 185

m = samples in rural regions = 239

number of parameters including constant = 19

		critical values
*	insignificantly different from zero at 10%	1.4526
**	insignificantly different from zero at 5%	1.6154
***	insignificantly different from zero at 1%	1.9529

Table 5.3
Unrestricted and Restricted Estimations for the AIDS with respect to Urban and Rural samples

A. Urban Regions (Sample = 185)

	Unrestricted	Homogeneity	Symmetry	Homogeneity & Symmetry
Parameter	304	288	184	168
Objective*T	2656	2672	2507.29	2519.6
Det $\|\Sigma\|$	2.4817E-61	2.8664E-61	1.6078E-60	1.8143E-60
Wald	---	73.8201	1022.46	1086.49
LR	---	61.1116	792.244	843.474
LM	---	50.2237	628.099	670.421
Log Likelihood Function	8859.61	8838.28	8761.12	8743.79

B. Rural Regions (Sample = 239)

	Unrestricted	Homogeneity	Symmetry	Homogeneity & Symmetry
Parameter	304	288	184	168
Objective*T	3520	3536	3232.56	3234.34
Det $\|\Sigma\|$	1.2605E-61	1.4046E-61	8.7128E-61	1.3845E-60
Wald	---	52.9343	1037.84	1322.03
LR	---	45.8800	819.710	1016.09
LM	---	39.6004	663.946	804.452
Log Likelihood Function	11482.24	11394.93	11461.30	11338.70

Critical values :		α	$\chi^2(16)$	$\chi^2(120)$	$\chi^2(136)$
	*	.1	23.54	140.23	157.52
	**	.05	26.29	146.57	164.22
	***	.01	32.00	158.95	177.28

Table 5.4

Multivariate Parameter Constancy Test for the Maintained AIDS

	Unrestricted	Homogeneity	Symmetry	Homogeneity & Symmetry
Log Likelihood				
L_T (Total)[a]	19648.30	19617.57	19512.11	19426.92
L_g (Urban)	8859.61	8838.28	8761.12	8743.79
L_h (Rural)	11482.24	11461.30	11394.93	11338.70
AM Statistic				
$-2(L_T-L_g-L_h)$	1387.09	1364.01	1287.88	1311.15
degrees of freedom	440	424	320	304

a. Calculated from Table 4.18

Table 5.5
RESET Tests for the AIDS with Dummies

Equation	Urban		Island		Urban & Island	
	RESET2	RESET3	RESET2	RESET3	RESET2	RESET3
Rice/Grain	.458	.485	1.07	.861	2.38*	2.20*
Tubers	4.75*	3.37*	2.31*	.641	3.10*	1.38
Fish	.212	-.55	-.10	-.53	-.13	-.56
Dried fish	-2.1*	-2.7*	-.36	-.82	-.05	-.53
Meat	3.29*	3.40*	3.68*	3.09*	4.81*	4.56*
Chicken	-.02	-.05	1.77	.844	1.78	.968
Egg	2.39*	2.24*	3.01*	2.59*	3.24*	2.90*
Milk	5.82*	5.94*	6.45*	6.21*	7.07*	7.28*
Vegetable	1.98	1.97	-2.1*	-1.9	-2.0	-1.8
Legumes	5.85*	3.87*	3.84*	3.13*	4.26*	3.40*
Fruit	1.77	1.65	2.83*	2.86*	2.68*	2.61*
Condiment	1.68	1.74	1.91	1.98	1.78	1.84
Cooking oil	-2.7*	-2.9*	-4.4*	-4.3*	-4.4*	-4.4*
Additive	-1.8	-1.9	.298	.139	-.14	-.31
Pre. food	14.6*	15.0*	14.6*	14.3*	15.3*	15.3*
Alcohol	10.2*	14.3*	11.7*	12.2*	11.6*	12.2*
Tobacco	-.28	-.35	-.61	-.74	-.62	-.76

* significantly different from zero at 5%

Table 5.6

R-Square Statistics for Models With and Without Dummies

Equation	No Dummy	Urban Dummy	Island Dummies	Urban & Island Dummies
1	.5316	.5610	.6177	.6431
2	.2507	.2654	.3654	.3775
3	.5409	.5409	.6981	.6983
4	.3653	.3849	.4326	.4430
5	.3043	.3934	.4676	.5218
6	.2197	.2410	.3095	.3202
7	.4561	.4793	.4769	.4970
8	.3934	.4522	.4202	.4732
9	.2362	.2377	.4377	.4382
10	.6364	.6572	.7488	.7637
11	.2046	.2103	.2680	.2696
12	.3252	.3313	.3441	.3483
13	.2698	.2779	.3242	.3317
14	.2701	.2801	.3119	.3239
15	.2690	.2705	.3045	.3082
16	.1508	.1513	.2094	.2100
17	.3092	.3099	.3939	.3943

Table 5.7

Joint Significance Tests for the Inclusion of Dummies in the AIDS

Equation	Urban F(1,404)	Island F(5,400)	Urban & Island F(6,399)
Rice/Grain	27.06	18.02	20.78
Tubers	8.043	14.46	13.55
Fish	0.000*	41.66	34.69
Dried fish	12.87	9.489	9.277
Meat	59.34	24.54	30.25
Chicken	11.34	10.40	9.831
Egg	18.00	3.181	5.407
Milk	43.36	3.698	10.07
Vegetable	.7950*	28.67	23.91
Legumes	26.63	36.37	36.33
Fruit	2.916**	6.929	5.918
Condiment	3.685**	2.305**	2.357***
Cooking oil	4.532***	6.440	6.159
Additive	5.612***	4.860	5.292
Pre. food	.8307*	4.083	3.768
Alcohol	.2380*	5.930	4.983
Tobacco	.4098*	11.18	9.343

* insignificantly different from zero at 10%

** insignificantly different from zero at 5%

*** insignificantly different from zero at 1%

Critical values :

	F(1,404)	F(5,400)	F(6,399)
α = 0.1	2.7181	1.8942	1.8096
0.05	3.8645	2.2677	2.1407
0.01	6.7002	3.0803	2.8556

Table 5.8

Estimations of the AIDS with Urban Dummy

	Unrestricted	Homogeneity	Symmetry	Homogeneity & Symmetry
Parameter	320	304	200	184
Objective *T	6460.00	6480.00	6015.53	6027.27
Det $\|\Sigma\|$	1.5933E-60	1.7897E-60	7.8269E-60	1.0795E-59
Information Inaccuracy				
I * 1000	59.21	60.02	73.48	72.25
IC * 1000	391.48	397.06	416.43	420.65
Asy. χ^2 Test for restriction				
Wald	---	54.56	806.53	979.87
LR	---	49.29	674.90	811.23
LM	---	44.53	573.92	682.78
GJ	---	72.11	796.90	949.19
Degrees of freedom	---	16	120	136
Joint Test for dummies				
LR	132.78	128.94	199.14	220.73
Degrees of freedom	16	16	16	16

Critical values :

	α	$\chi^2_{(16)}$	$\chi^2_{(120)}$	$\chi^2_{(136)}$
*	0.1	23.54	140.23	157.52
**	0.05	26.30	146.57	164.22
***	0.01	32.00	158.95	177.28

Table 5.9

Estimations of the AIDS with Island Dummies

	Unrestricted	Homogeneity	Symmetry	Homogeneity & Symmetry
Parameter	384	368	264	248
Objective *T	6400.00	6416.00	6052.72	6047.89
Det $\lvert\Sigma\rvert$	3.6671E-61	4.0448E-61	9.6244E-61	1.2463E-60
Information Inaccuracy				
I * 1000	50.59	51.70	57.85	56.71
IC * 1000	319.09	329.46	332.24	339.35
Asy. χ^2 Test for restriction				
Wald	---	45.67	478.32	613.54
LR	---	41.57	409.12	518.72
LM	---	37.83	355.71	446.14
GJ	---	59.19	522.23	623.57
Degrees of freedom	---	16	120	136
Joint Test for dummies				
LR	755.63	759.52	1087.79	1136.10
Degrees of freedom	80	80	80	80

Critical values :

	α	$\chi^2_{(80)}$	$\chi^2_{(16)}$	$\chi^2_{(120)}$	$\chi^2_{(136)}$
*	0.1	96.58	23.54	140.23	157.52
**	0.05	101.88	26.30	146.57	164.22
***	0.01	112.33	32.00	158.95	177.28

Table 5.10

Estimations of the AIDS with Urban & Island Dummies

	Unrestricted	Homogeneity	Symmetry	Homogeneity & Symmetry
Parameter	400	384	280	264
Objective *T	6384.00	6400.00	6058.25	6044.62
Det $\|\Sigma\|$	2.8380E-61	3.1199E-61	6.3767E-61	7.7416E-61
Information Inaccuracy				
I * 1000	49.72	50.48	55.93	54.67
IC * 1000	319.54	323.74	326.90	333.46
Asy. χ^2 Test for restriction				
Wald	---	44.04	382.84	477.77
LR	---	40.15	343.25	425.48
LM	---	36.58	309.90	381.35
GJ	---	57.55	461.14	548.55
Degrees of freedom	---	16	120	136
Joint Test for dummies				
LR	864.30	869.60	1262.33	1337.99
Degrees of freedom	96	96	96	96

Critical values :

	α	$\chi^2_{(96)}$	$\chi^2_{(16)}$	$\chi^2_{(120)}$	$\chi^2_{(136)}$
*	0.1	114.13	23.54	140.23	157.52
**	0.05	119.81	26.30	146.57	164.22
***	0.01	131.14	32.00	158.95	177.28

Table 5.11
The General LM Test for Heteroscedasticity
for the Unrestricted AIDS

Equation	No Dummy	Urban Dummy	Island Dummies	Urban & Island Dummies
Rice/Grain	18.57	18.06	15.90	12.89
Tubers	28.96	32.56	65.38	59.87
Fish	23.24	23.24	32.44	32.56
Dried fish	12.68	13.48	7.844	8.353
Meat	27.31	29.55	44.94	42.19
Chicken	15.39	18.02	28.87	32.01
Egg	4.706***	5.639***	6.233***	6.954
Milk	11.74	12.68	12.72	13.06
Vegetable	0.000*	.0424*	5.003***	4.876***
Legumes	12.04	13.78	25.95	26.12
Fruit	.1224*	3.731**	8.098	8.353
Condiment	1.611*	1.738*	1.314*	1.569*
Cooking oil	1.654*	1.738*	0.000*	0.000*
Additive	.0424*	.2968*	.1272*	0.000*
Pre. food	80.26	79.50	88.36	87.22
Alcohol	35.70	35.62	44.73	45.33
Tobacco	15.31	15.56	26.08	25.91

		$\chi^2(1)$ critical values
*	insignificantly different from zero at 10%	2.706
**	insignificantly different from zero at 5%	3.841
***	insignificantly different from zero at 1%	6.635

Table 5.12

Breusch-Pagan Test for Specific Heteroscedasticity for the Maintained AIDS

	LKTHH	LKTAH	LKTY	LKTE	LEXP
Rice/Grain	1.626*	9.039	9.630	22.50	12.39
Tubers	.0351*	.0416*	131.8	83.94	34.29
Fish	9.045	2.371*	.1856*	.6461*	1.649*
Dried fish	4.626***	7.771	6.476***	9.088	2.639*
Meat	23.32	25.08	21.67	22.70	28.24
Chicken	.0965*	1.059*	5.976***	9.813	7.899
Egg	.0003*	1.098*	4.759***	6.756	2.032*
Milk	18.64	7.863	62.77	97.61	21.29
Vegetable	.2188*	.7337*	.2292*	.2623*	.9857*
Legumes	2.263*	.2596*	22.56	18.66	24.84
Fruit	.3273*	.1600*	.4203*	.0529*	3.099**
Condiment	6.621***	7.669	1.846*	1.140*	1.401*
Cooking oil	3.117**	.1312*	5.305***	7.314	1.395*
Additive	2.133*	2.536*	2.902**	6.766	2.138*
Pre. food	2.102*	2.542*	21.35	36.05	48.16
Alcohol	78.99	5.704***	19.89	37.36	94.30
Tobacco	2.103*	2.356*	1.448*	1.632*	.0877*

		$\chi^2(1)$ critical values
*	insignificantly different from zero at 10%	2.706
**	insignificantly different from zero at 5%	3.841
***	insignificantly different from zero at 1%	6.635

Table 5.13

Breusch-Pagan Test for Specific Heteroscedasticity for the original AIDS with Urban Dummy

	LKTHH	LKTAH	LKTY	LKTE	LEXP
Rice/Grain	.6927*	7.477	8.197	21.19	11.94
Tubers	.0877*	.0959*	133.2	88.48	36.48
Fish	8.976	2.335*	.1856*	.6445*	1.639*
Dried fish	4.124***	8.409	7.340	10.65	2.905**
Meat	21.27	25.98	17.79	22.44	25.81
Chicken	.1620*	.9181*	5.574***	9.238	6.892
Egg	.0008*	1.167*	8.228	10.15	4.303***
Milk	21.21	9.184	79.34	112.3	19.96
Vegetable	.2451*	.6444*	.3795*	.4164*	1.235*
Legumes	2.929**	1.508*	22.44	21.08	26.81
Fruit	.5981*	.0417*	.6205*	.0576*	3.286**
Condiment	6.911	6.985	2.098*	1.282*	1.711*
Cooking oil	3.196**	.1449*	5.012***	7.793	1.590*
Additive	2.958**	2.913**	2.821**	6.657	2.019*
Pre. food	2.505*	2.894**	20.18	34.21	46.81
Alcohol	79.01	5.553***	19.36	36.88	93.45
Tobacco	2.210*	2.566*	1.082*	1.651*	.0994*

		$\chi^2(1)$ critical values
*	insignificantly different from zero at 10%	2.706
**	insignificantly different from zero at 5%	3.841
***	insignificantly different from zero at 1%	6.635

Table 5.14

Breusch-Pagan Test for Specific Heteroscedasticity for the original AIDS with Island Dummies

	LKTHH	LKTAH	LKTY	LKTE	LEXP
Rice/Grain	2.486*	4.819***	10.13	21.98	14.81
Tubers	.5885*	.1393*	150.5	108.5	63.79
Fish	1.154*	.6370*	.7164*	1.302*	1.540*
Dried fish	6.456***	9.937	13.40	20.68	12.12
Meat	12.27	7.197	10.34	11.59	18.66
Chicken	.0083*	2.849**	6.320***	11.12	9.457
Egg	.1479*	.2068*	4.196***	6.017***	2.348*
Milk	19.96	9.294	60.19	86.78	16.18
Vegetable	1.328*	.0180*	.5098*	.0821*	.1537*
Legumes	6.371***	1.769*	43.76	32.72	43.04
Fruit	.3697*	.2151*	.8635*	.0018*	1.264*
Condiment	8.458	9.673	1.879*	1.063*	.9495*
Cooking oil	2.857**	1.071*	4.440***	6.554***	1.554*
Additive	1.704*	2.240*	2.862**	5.469***	1.747*
Pre. food	1.341*	.8824*	19.92	34.60	45.53
Alcohol	86.88	10.72	13.47	28.07	76.34
Tobacco	.7227*	1.318*	1.817*	1.946*	.0224*

		$\chi^2(1)$ critical values
*	insignificantly different from zero at 10%	2.706
**	insignificantly different from zero at 5%	3.841
***	insignificantly different from zero at 1%	6.635

Table 5.15

Breusch-Pagan Test for Specific Heteroscedasticity for the original AIDS with Urban and Island Dummies

	LKTHH	LKTAH	LKTY	LKTE	LEXP
Rice/Grain	1.598*	4.397***	8.976	20.04	13.79
Tubers	1.696*	.4160*	157.4	117.2	69.87
Fish	1.146*	.5566*	.6718*	1.296*	1.526*
Dried fish	5.669***	9.804	12.91	20.42	10.94
Meat	13.18	10.97	10.28	14.38	19.85
Chicken	.0003*	2.361*	6.299***	10.91	8.744
Egg	.0360*	.2898*	6.784	8.575	4.320***
Milk	19.48	9.024	74.34	99.52	15.90
Vegetable	1.304*	.0174*	.6076*	.1131*	.1211*
Legumes	8.242	3.966***	48.91	38.16	47.16
Fruit	.4775*	.0038*	1.012*	.0013*	1.338*
Condiment	8.883	9.293	2.057*	1.214*	1.241*
Cooking oil	2.767**	1.243*	4.062***	6.681	1.572*
Additive	2.619*	2.799**	2.587*	5.357***	1.655*
Pre. food	1.943*	1.269*	18.39	31.61	43.34
Alcohol	86.76	10.54	13.13	27.85	75.76
Tobacco	.7722*	1.456*	1.836*	2.017*	.0209*

		$\chi^2(1)$ critical values
*	insignificantly different from zero at 10%	2.706
**	insignificantly different from zero at 5%	3.841
***	insignificantly different from zero at 1%	6.635

Table 5.16

General LM Test for Heteroscedasticity
for the Weighted AIDS Model without Dummy Variables

Equation	Wgt = Sqrt(GHH)	Wgt = Inv (Sqrt(LEXP))	Wgt = Inv (Sqrt(LKTY))	Wgt = Inv (Sqrt(LKTE))
Rice/Grain	52.15	20.77	21.33	21.45
Tubers	38.80	29.43	30.74	29.93
Fish	30.49	23.19	23.74	23.66
Dried fish	39.47	13.14	13.48	13.53
Meat	49.52	26.71	27.60	27.43
Chicken	33.67	15.35	15.22	15.31
Egg	10.43	4.664***	4.325***	4.070***
Milk	5.470***	11.62	11.28	11.07
Vegetable	35.79	.0424*	.0424*	.0424*
Legumes	50.29	13.69	13.40	13.40
Fruit	12.97	3.774**	4.579***	4.409***
Condiment	12.97	2.205*	2.035*	2.035*
Cooking oil	14.71	1.145*	1.272*	1.145*
Additive	11.41	.2968*	.2544*	.3816*
Pre. food	109.6	78.48	79.25	78.86
Alcohol	43.42	34.34	35.49	34.98
Tobacco	44.52	14.59	15.14	15.56

		$\chi^2(1)$ critical values
*	insignificantly different from zero at 10%	2.706
**	insignificantly different from zero at 5%	3.841
***	insignificantly different from zero at 1%	6.635

Table 5.17

General LM Test for Heteroscedasticity
for the Weighted AIDS Model with Urban and Islands Dummy Variables

Equation	Wgt= Sqrt(GHH)	Wgt = Inv (Sqrt(LEXP))	Wgt = Inv (Sqrt(LKTY))	Wgt = Inv (Sqrt(LKTE))
Rice/Grain	39.26	20.73	21.07	21.07
Tubers	44.44	60.21	63.35	61.65
Fish	29.47	33.16	33.75	33.84
Dried fish	25.82	10.90	11.02	11.28
Meat	51.52	40.28	41.00	40.83
Chicken	45.79	29.68	29.04	29.55
Egg	7.166	6.700	6.275***	6.106***
Milk	6.106***	12.25	12.08	11.91
Vegetable	39.56	3.816**	4.070***	4.155***
Legumes	80.05	28.70	28.24	28.45
Fruit	28.92	6.911	8.056	7.590
Condiment	13.61	2.374*	2.247*	2.290*
Cooking oil	12.68	.0424*	.0424*	.0424*
Additive	12.97	.0848*	.0424*	.1272*
Pre. food	119.74	44.31	43.71	43.59
Alcohol	47.45	44.22	45.54	45.07
Tobacco	55.08	13.36	14.08	14.33

		$\chi^2(1)$ critical values
*	insignificantly different from zero at 10%	2.706
**	insignificantly different from zero at 5%	3.841
***	insignificantly different from zero at 1%	6.635

Table 5.18

Estimations of the Weighted AIDS without Dummies

	Unrestricted	Homogeneity	Symmetry	Homogeneity & Symmetry
A. Weight=Inv(Sqrt(LEXP))				
Information Inaccuracy				
I * 1000	20.08	20.27	25.53	25.24
IC * 1000	133.36	135.01	144.01	147.26
Asy. χ^2 Tests				
Wald	---	48.97	907.36	1123.07
LR	---	44.47	746.18	903.20
LM	---	40.38	625.90	743.66
GJ	---	62.89	858.39	1037.65
B. Weight=Inv(Sqrt(LKTY))				
Information Inaccuracy				
I * 1000	18.11	18.28	23.07	22.92
IC * 1000	120.80	122.33	130.55	133.47
Asy. χ^2 Tests				
Wald	---	50.03	907.22	1123.88
LR	---	45.39	746.63	904.36
LM	---	41.18	626.72	744.98
GJ	---	63.90	859.63	1039.53
C. Weight=Inv(Sqrt(LKTE))				
Information Inaccuracy				
I * 1000	18.38	18.10	23.33	23.12
IC * 1000	121.91	123.44	131.67	134.61
Asy. χ^2 Tests				
Wald	---	49.44	907.88	1124.82
LR	---	44.89	746.83	904.86
LM	---	40.74	626.67	745.28
GJ	---	63.35	859.22	1038.85

Table 5.19

Estimations of the Weighted AIDS with Urban and Island Dummies

	Unrestricted	Homogeneity	Symmetry	Homogeneity & Symmetry
A. Weight=Inv(Sqrt(LEXP))				
Information Inaccuracy				
I * 1000	16.49	16.80	18.51	18.09
IC * 1000	104.90	106.22	107.25	109.39
Asy. χ^2 Tests				
Wald	---	42.59	383.89	477.75
LR	---	38.88	343.96	425.23
LM	---	35.46	310.35	380.94
GJ	---	56.18	462.13	548.92
B. Weight=Inv(Sqrt(LKTY))				
Information Inaccuracy				
I * 1000	14.88	15.22	16.82	16.35
IC * 1000	95.12	96.34	97.42	99.33
Asy. χ^2 Tests				
Wald	---	43.76	387.85	482.06
LR	---	39.90	347.19	428.61
LM	---	36.36	313.02	383.63
GJ	---	57.28	465.32	552.49
C. Weight=Inv(Sqrt(LKTE))				
Information Inaccuracy				
I * 1000	15.01	15.35	16.93	16.48
IC * 1000	95.82	97.03	98.09	100.16
Asy. χ^2 Tests				
Wald	---	42.83	385.26	479.24
LR	---	39.08	345.14	426.46
LM	---	35.64	311.38	381.99
GJ	---	56.41	462.99	549.95

Table 5.20

Unrestricted Almost Ideal Demand System

With Weighting Factor Inv(Sqrt(LKTY)), Urban and Island Dummies

	Rice/Grain w_1	Tubers w_2	Fish w_3	Dried Fish w_4
Constant	1.239(10.3)	.1560(3.22)	.3096(3.65)	.1180(1.87)
Urban	-.032(5.34)	-.007(2.73)	.0023(.540)	-.009(2.75)
Island 1	.0911(5.80)	-.051(7.80)	-.067(6.06)	.0355(4.31)
Island 2	.0464(2.73)	-.050(7.06)	-.107(8.90)	.0296(3.31)
Island 3	.1003(5.75)	-.053(7.32)	-.084(6.78)	.0176(1.92)
Island 4	.0480(3.09)	-.048(7.52)	-.048(4.42)	.0456(5.60)
Island 5	.0652(4.13)	-.047(7.22)	-.013(1.18)	.0314(3.78)
$\ln p_1$	.0050(.210)	.0194(1.98)	-.089(5.32)	-.017(1.37)
$\ln p_2$	-.013(1.90)	.0037(1.26)	.0071(1.43)	-.004(1.24)
$\ln p_3$	.0153(1.67)	-.013(3.42)	-.013(1.98)	.0190(3.95)
$\ln p_4$	-.017(3.07)	-7.8E-4(.34)	.0187(4.73)	.0254(8.62)
$\ln p_5$	-.012(1.23)	-.003(.680)	.0042(.620)	-.002(.380)
$\ln p_6$	.0133(1.34)	-2.2E-4(.05)	.0054(.770)	.0014(.260)
$\ln p_7$	.0201(1.31)	-.007(1.10)	.0156(1.44)	.0144(1.79)
$\ln p_8$	-.015(1.49)	.0014(.330)	-8.9E-4(.13)	-.012(2.31)
$\ln p_9$	-.038(2.65)	-.010(1.79)	.0329(3.26)	.0044(.580)
$\ln p_{10}$	-.006(.490)	-.014(3.07)	.0251(3.17)	.0071(1.20)
$\ln p_{11}$	-.001(.130)	-.003(.800)	-.002(.350)	-.008(1.52)
$\ln p_{12}$	-.036(3.94)	-.011(2.95)	-.012(1.85)	-.010(2.19)
$\ln p_{13}$	-.003(.160)	.0016(.200)	-.012(.860)	-.027(2.59)
$\ln p_{14}$	.0234(2.01)	.0055(1.14)	-.048(5.84)	-.001(.170)
$\ln p_{15}$	-.009(1.31)	.0086(3.14)	.0021(.450)	.0032(.900)
$\ln p_{16}$	.0170(2.07)	.0005(.160)	-.003(.510)	-.006(1.41)
$\ln p_{17}$	-.010(1.80)	.0011(.480)	.0040(.980)	.0064(2.13)
Income Ratio	-.118(10.2)	-.007(1.49)	-.013(1.59)	-.005(.820)

Table 5.20 (continued)
Unrestricted Almost Ideal Demand System
With Weighting Factor Inv(Sqrt(LKTY)), Urban and Island Dummies

	Meat w_5	Chicken w_6	Egg w_7	Milk w_8
Constant	.0887(1.88)	.0211(.810)	-.002(.080)	-.023(1.39)
Urban	.0160(6.73)	.0033(2.51)	.0040(4.07)	.0054(6.42)
Island 1	.0021(.340)	-.002(.480)	.0028(1.13)	-.001(.650)
Island 2	.0049(.740)	-.001(.390)	.0030(1.09)	-.002(.700)
Island 3	.0402(5.86)	.0127(3.35)	.0065(2.29)	-.004(1.67)
Island 4	.0024(.390)	.0024(.700)	.0052(2.07)	-.005(2.19)
Island 5	.0048(.770)	.0014(.410)	.0020(.800)	-5.1E-4(.23)
$\ln p_1$	.0343(3.69)	.0065(1.27)	.0043(1.12)	.0095(2.88)
$\ln p_2$	.0060(2.19)	.0041(2.73)	.0041(3.65)	.0007(.690)
$\ln p_3$	.0053(1.48)	.0051(2.57)	.0055(3.69)	.0006(.450)
$\ln p_4$	.0032(1.47)	.0013(1.07)	-.001(1.11)	.0007(.860)
$\ln p_5$	.0016(.420)	.0025(1.22)	.0037(2.39)	.0016(1.19)
$\ln p_6$	.0059(1.53)	.0078(3.64)	-.001(.690)	-3.3E-4(.24)
$\ln p_7$	-.017(2.77)	-.007(2.12)	-.008(3.11)	-.005(2.10)
$\ln p_8$	.0023(.590)	.0011(.490)	.0017(1.06)	.0072(5.12)
$\ln p_9$	.0038(.670)	-.005(1.70)	.0012(.520)	.0048(2.40)
$\ln p_{10}$	-.006(1.34)	-.003(1.13)	-3.5E-4(.20)	-7.1E-4(.45)
$\ln p_{11}$	-.002(.410)	.0006(.260)	.0031(1.89)	-7.4E-4(.52)
$\ln p_{12}$	.0044(1.24)	-.002(1.07)	.0013(.900)	.0036(2.83)
$\ln p_{13}$	-.016(2.10)	-.005(1.06)	.0011(.360)	-1.1E-4(.04)
$\ln p_{14}$	.0194(4.25)	.0038(1.50)	.0002(.100)	.0031(1.89)
$\ln p_{15}$	-.001(.400)	.0023(1.60)	.0003(.320)	-5.8E-4(.63)
$\ln p_{16}$	.0032(.990)	-.002(1.12)	-.001(.800)	-6.8E-4(.59)
$\ln p_{17}$	-.002(.700)	-.002(1.64)	-.005(5.16)	-.002(2.27)
Income Ratio	-.014(3.10)	-.0051(2.02)	-.002(1.00)	.0001(.090)

Table 5.20 (continued)

Unrestricted Almost Ideal Demand System

With Weighting Factor Inv(Sqrt(LKTY)), Urban and Island Dummies

	Vegetable w_9	Legumes w_{10}	Fruit w_{11}	Condiment w_{12}
Constant	.2114(4.12)	.1467(4.06)	.1299(3.89)	.0729(2.50)
Urban	.0013(.520)	.0091(4.98)	.0015(.890)	-.002(1.58)
Island 1	-.003(.400)	-8.7E-4(.18)	-.017(3.67)	.0023(.590)
Island 2	-.034(4.70)	.0263(5.14)	-.015(3.21)	.0038(.930)
Island 3	-.024(3.16)	.0002(.050)	-.009(1.87)	-.002(.380)
Island 4	-.037(5.57)	.0004(.100)	-.021(4.89)	-.001(.360)
Island 5	-.034(5.05)	-.008(1.62)	-.009(2.13)	-.004(.920)
$\ln p_1$	.0263(2.60)	.0134(1.89)	.0147(2.23)	-.005(.950)
$\ln p_2$	.0077(2.58)	.0020(.970)	.0058(2.98)	.0021(1.26)
$\ln p_3$	.0051(1.31)	-.006(2.28)	.0009(.350)	-.002(1.12)
$\ln p_4$	-.013(5.29)	-.004(2.14)	.0009(.610)	-.003(2.43)
$\ln p_5$	-.005(1.17)	.0012(.430)	-.008(3.07)	-.003(1.29)
$\ln p_6$	-.003(.600)	.0029(.980)	.0042(1.52)	.0020(.830)
$\ln p_7$	-.004(.610)	-.019(4.03)	-.004(1.05)	-.003(.880)
$\ln p_8$	.0042(.990)	-.002(.580)	-.002(.700)	-8.2E-5(.03)
$\ln p_9$	.0070(1.15)	-.002(.500)	.0052(1.32)	-.002(.470)
$\ln p_{10}$	-.004(.810)	-.006(1.86)	-.004(1.32)	-6.9E-4(.25)
$\ln p_{11}$	.0050(1.17)	.0015(.490)	.0051(1.80)	-.004(1.55)
$\ln p_{12}$	.0136(3.50)	.0116(4.25)	.0054(2.15)	.0178(8.07)
$\ln p_{13}$	-.011(1.34)	.0054(.900)	-.003(.600)	.0061(1.27)
$\ln p_{14}$	.0015(.310)	-.004(1.10)	-.004(1.39)	-.005(1.69)
$\ln p_{15}$	-.005(1.63)	-.003(1.35)	-9.5E-4(.52)	.0026(1.60)
$\ln p_{16}$	.0004(.120)	-.001(.440)	-.002(1.03)	.0015(.750)
$\ln p_{17}$	-.004(1.75)	-.009(5.34)	-.004(2.28)	-8.2E-4(.59)
Income Ratio	-.011(2.31)	-.016(4.47)	-.0078(2.40)	-.004(1.28)

Table 5.20 (continued)

Unrestricted Almost Ideal Demand System

With Weighting Factor Inv(Sqrt(LKTY)), Urban and Island Dummies

	Cooking Oil w_{13}	Additives w_{14}	Pre. Food w_{15}	Alcohol w_{16}
Constant	.0834(3.45)	.1952(3.01)	-.947(9.12)	.0005(.080)
Urban	.0026(2.15)	.0088(2.68)	-.007(1.38)	.0001(.530)
Island 1	-.004(1.30)	.0088(1.04)	-.025(1.87)	-.003(3.48)
Island 2	-8.8E-5(.03)	.0146(1.59)	.0049(.330)	-.003(3.82)
Island 3	-.009(2.67)	-.001(.120)	-.009(.590)	-.003(3.37)
Island 4	-.008(2.66)	.0266(3.18)	-.012(.920)	-.004(5.21)
Island 5	-.006(2.01)	.0128(1.49)	-.020(1.44)	-.003(3.97)
ln p_1	.0158(3.31)	.0120(.940)	-4.4E-4(.02)	-.003(2.72)
ln p_2	.0032(2.25)	-.003(.830)	.0077(1.28)	-6.5E-4(2.0)
ln p_3	-.003(1.70)	-.019(3.98)	.0109(1.37)	.0006(1.31)
ln p_4	-.004(3.90)	-.004(1.48)	-.005(1.11)	.0001(.390)
ln p_5	.0021(1.10)	-.008(1.54)	.0118(1.43)	-8.5E-4(1.9)
ln p_6	-.007(3.39)	-.002(.350)	-.017(2.04)	.0012(2.66)
ln p_7	-.004(1.29)	-.014(1.74)	-.030(2.28)	.0022(3.11)
ln p_8	.0009(.470)	.0054(.990)	.0080(.930)	.0008(1.72)
ln p_9	.0040(1.38)	.0081(1.05)	-.025(2.00)	-5.3E-4(.81)
ln p_{10}	.0034(1.51)	-.002(.300)	-.013(1.34)	.0012(2.36)
ln p_{11}	-3.2E-4(.16)	-.012(2.22)	.0252(2.88)	.0012(2.46)
ln p_{12}	-.003(1.69)	.0196(4.00)	-.001(.160)	-.001(2.56)
ln p_{13}	.0071(1.79)	.0145(1.36)	.0251(1.47)	-.002(1.63)
ln p_{14}	-.004(1.68)	-.013(2.12)	.0113(1.12)	-3.9E-4(.73)
ln p_{15}	6.04E-5(.05)	.0060(1.68)	.0035(.610)	-1.6E-4(.53)
ln p_{16}	.0008(.470)	.0053(1.20)	.0111(1.56)	.0001(.370)
ln p_{17}	-.005(4.46)	-.012(3.75)	.0093(1.87)	.0002(.720)
Income Ratio	-.007(3.19)	-.010(1.65)	.1046(10.4)	.0002(.420)

Table 5.20 (continued)

Unrestricted Almost Ideal Demand System

With Weighting Factor Inv(Sqrt(LKTY)), Urban and Island Dummies

	Tobacco w_{17}
Constant	-.800
Urban	.0026
Island 1	.0313
Island 2	.0786
Island 3	.0205
Island 4	.0534
Island 5	.0269
$\ln p_1$	-.047
$\ln p_2$	-.033
$\ln p_3$	-.011
$\ln p_4$	.0024
$\ln p_5$	.0123
$\ln p_6$	-.014
$\ln p_7$	.0696
$\ln p_8$	-.001
$\ln p_9$	.0113
$\ln p_{10}$	.0224
$\ln p_{11}$	-.008
$\ln p_{12}$	-.001
$\ln p_{13}$	.0181
$\ln p_{14}$	.0115
$\ln p_{15}$	-.010
$\ln p_{16}$	-.024
$\ln p_{17}$	.0343
Income Ratio	.1161

Table 5.21

Own Price and Income Elasticities for the
New Maintained Almost Ideal Demand System
With Weighting Factor Inv(Sqrt(LKTY)), Urban and Island Dummies

	Own Price	Income
Rice/Grain	-.863*	.5720*
Tubers	-.589	.2069
Fish	-1.43	.5493
Dried Fish	-.076	.8192*
Meat	-.797	-.707
Chicken	.0353	.3288
Egg	-1.76	.8145*
Milk	1.592	1.053*
Vegetable	-.895*	.8467*
Legumes	-1.78	-.983
Fruit	-.822*	.7382*
Condiment	-.339	.8662*
Cooking oil	-.752*	.7480*
Additive	-1.13*	.8885*
Pre. Food	.0796	36.12
Alcohol	-.863	1.219
Tobacco	-.845*	1.917

* Significant at 5%

Table 5.22

Substitution Coefficients and Income Effect for the
New Naintained Almost Ideal Demand System
With Weighting Factor Inv(Sqrt(LKTY)), Urban and Island Dummies

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rice/Grain	**-367.***	95.69	-46.6*	-12.0	9.270	2.638	49.65*	4.029*
Tubers	5022.	**-191.**	35.89	-6.12	4.618	2.892	56.26*	.5845
Fish	297.5	-17.2	**-23.7**	21.47	.7182	.6095	13.62*	.1584
Dried fish	1255.	.8241	25.30	**-53.1**	2.555	.6964	-2.46	.4818
Meat	395.5	-.457	3.197	.6637	**-.968**	.1179	3.236	.0853
Chicken	472.7	.9708	3.112	1.679	.3224	**-.277**	-.737	-.011
Egg	7630.	-72.3	64.61	134.2*	-10.5	-6.09*	**-463.**	-4.59*
Milk	2206.	5.418	6.166	-3.06	.9096	.1101	2.335	**-.051**
Vegetable	400.9	-71.0	48.06	37.16*	2.234	-2.32	26.56	3.378*
Legumes	1866.	-27.6	21.53	15.68	-.292	-.367	.9486	-.081
Fruit	894.3	-13.1	4.469	-20.1	-.142	.3418	25.93*	-.181
Condiment	2513.	-145.	-8.81	-76.9	5.113	-1.78	35.31*	4.566
Cooking oil	294.3	3.828	-1.43	-25.9*	-1.45	-.493*	4.143*	.0299
Additive	606.1	58.49	-56.1*	17.60	10.42	2.857*	22.36	2.932
Pre. food	25802.	129.2	81.72	61.38	3.100	2.604	17.11	-.316
Alcohol	5910.	5.498	12.84	8.309	1.438	-.049	-.659	-.074
Tobacco	456.9	28.75	27.38	85.69	.5039	-.489	-33.8	-.305
Income Effect	-100.	-9.70	-2.56	-9.28	-.053	-.012	-6.91	-.009

* Significantly different from zero at 5%

Table 5.22 (continued)

Substitution Coefficients and Income Effect for the

New Maintained Almost Ideal Demand System

With Weighting Factor Inv(Sqrt(LKTY)), Urban and Island Dummies

	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Rice/Grain	157.8*	19.04*	51.99*	12.20	17.80*	150.7*	173.4	-.499
Tubers	129.1	10.77	39.18	44.95	7.697	144.9	-602.	-.260
Fish	15.90	-1.74	3.192	-.739	-.192	-12.0	1.180	.0601
Dried fish	-28.2	-1.67	8.775	-17.6	-2.81	30.98	-235.	.0600
Meat	2.815	.4297	-1.63	-1.65	.4120	7.148	-4.12	-.022
Chicken	4.883	1.085	2.253	2.978	-.635	10.75	-50.3	.0434
Egg	58.46	-54.8*	-20.3	-48.2	-6.51*	-31.4	-802.	1.471
Milk	29.33	1.089	1.163	2.285	.8496	37.36	-38.0	.0341
Vegetable	**-425.***	-.499	31.93*	6.426	8.741*	123.2*	-184.	-.125
Legumes	17.80	**-24.2**	-2.28	2.265	2.242*	28.93	-86.4	.1547
Fruit	42.85*	3.569	**-89.7**	-22.7	.8878	-24.0	128.5	.3139
Condiment	217.1*	46.48*	54.48	**-204.**	-5.17	360.6*	-176.	-.779
Cooking oil	-8.18*	2.930*	-1.79	18.86*	**-8.33**	35.10*	37.55	-.128
Additive	69.50*	-3.79	-7.76	-29.3	-1.81	**-918.***	210.9	-.088
Pre. food	271.7	42.55	19.97	85.34	8.870	620.4	**-333.**	-.060
Alcohol	80.30	8.265	3.933	9.008	2.468	107.8	-560.	**-.149**
Tobacco	62.91	-17.6	4.503	51.86	-1.76	37.92	368.0	.3050
Income Effect	-39.8	-.799	-3.61	-18.4	-.327	-87.1	-110.	-.001

* Significantly different from zero at 5%

Table 5.22 (continued)
Substitution Coefficients and Income Effect for the
New Maintained Almost Ideal Demand System
With Weighting Factor Inv(Sqrt(LKTY)), Urban and Island Dummies

	(17)
Rice/Grain	77.59
Tubers	-5002.
Fish	-129.
Dried fish	-841.
Meat	-299.
Chicken	-404.
Egg	-2884.
Milk	-1448.
Vegetable	-86.3
Legumes	-759.
Fruit	-679.
Condiment	-2138.
Cooking oil	-154.
Additive	62.87
Pre. food	-19840.
Alcohol	-3842.
Tobacco	**-1533.**
Income Effect	-721.

* Significantly different from zero at 5%

Table 5.23

Average Information Inaccuracy[1]

For the New Maintained Almost Ideal Demand System

With Weighting Factor Inv(Sqrt(LKTY)), Urban and Island Dummies

	AIDS[2]
Rice/Grain	.9343
Tubers	1.535
Fish	1.846
Dried Fish	1.454
Meat	1.370
Chicken	.7117
Egg	.4431
Milk	.4758
Vegetable	.6031
Legumes	.8757
Fruit	.5854
Condiment	.4460
Cooking Oil	.3183
Additive	.7980
Pre. Food	4.429
Alcohol	.3988

1. Average Information Inaccuracy * 1000
2. Unrestricted model

Table 5.24

Size Correction for the New Maintained AIDS

With Weighting Factor Inv(Sqrt(LKTY)), Urban and Island Dummies

	Homogeneity	Symmetry	Homogeneity & Symmetry
W^*	41.179811	364.98149	453.63665
LR^*	37.547406	326.71889	403.33818
LM^*	34.216132	294.56363	361.01031
D^*	34.400507	308.81252	382.64881

a. n=16

Table 5.25
Joint Significance Tests for Dummies
in the Weighted AIDS Model
Weighting Factor = Inv(Sqrt(LKTY))

	Unrestricted	Homogeneity	Symmetry	Homogeneity & Symmetry
With Dummies				
Det $\|\Sigma\|$	4.8651E-78	5.3451E-78	1.1033E-77	1.3369E-77
No Dummies				
Det $\|\Sigma\|$	3.6677E-77	4.0821E-77	2.1338E-76	3.0954E-76
T	424	424	424	424
Joint Test For Dummies				
LR Statistics	856.50654	861.99895	1255.9659	1332.2711
Degree of Freedom	96	96	96	96

Critical values :

	α	$\chi^2_{(96)}$
*	0.1	114.13
**	0.05	119.81
***	0.01	131.14

Table 5.26
<u>Bera & Jarque Normality Test for the Univariate AIDS</u>

Equation	Unweighted No Dummy	Unweighted Urban & Island	Weight = Inv(Sqrt(LKTY)) No dummy	Weight = Inv(Sqrt(LKTY)) Urban & Island
Rice/Grain	164.8	330.9	157.1	288.9
Tubers	3731.	2170.	3970.	2151.
Fish	78.89	70.39	79.59	77.31
Dried fish	77.93	86.47	90.03	83.58
Meat	594.9	550.5	560.6	440.0
Chicken	364.4	250.1	361.6	276.2
Egg	40.16	31.59	38.31	30.64
Milk	13142.	19298.	12629.	20015.
Vegetable	12.29	4.122*	10.97	4.139*
Legumes	9.205***	19.91	10.53	25.51
Fruit	54.95	31.75	50.87	33.33
Condiment	954.8	1060.	928.5	991.1
Cooking oil	124.6	48.46	127.5	41.53
Additive	28.64	26.44	30.65	23.68
Pre. food	6231.	5711.	6222.	11824.
Alcohol	47428.	41830.	50725.	45026.
Tobacco	5.111**	3.942*	5.080**	3.968*

		$\chi^2(1)$ critical values
*	insignificantly different from zero at 10%	2.706
**	insignificantly different from zero at 5%	3.841
***	insignificantly different from zero at 1%	6.635

CHAPTER VI

BOOTSTRAPPING

So far, the theory has not been supported. An attempt was made to correct for heteroscedasticity and parameter inconstancy using dummy variables and weighted least squares procedure. However, these corrections did not completely resolve the problem.

The question arises, are the rejections of theory due to the misspecification of the error distribution or to the inappropriateness of the asymptotic test statistics in finite sample? To answer this question, it is necessary to set up a "bootstrapping" experiment, described below. Before any discussion, it is important to note that such a simulation experiment is computationally very expensive. Therefore a fast algorithm is needed. Byron(1982) suggested an algorithm using the Lyapunov equation, an estimating technique which is examined below.

As well, to test the appropriateness of the test statistics, it is necessary to construct Kolmogorov-Smirnov Two-sample tests which are considered below.

6.1 Estimation using the Lyapunov Equation

In order to perform "bootstrapping" on a large demand system with a total of 17 commodities, a fast algorithm is necessary to minimize computing cost. Byron(1982) has demonstrated a computationally efficient method for handling large symmetric demand systems using Lyapunov equations. The method was demonstrated on Theil's(1975) 14 sector Rotterdam model and revealed a 60 fold increase in speed over the conventional SUR solution. As well, significant storage savings were achieved. The new algorithm is encouraging and will be particularly useful in a simulation experiment such as bootstrapping.

Consider the linear system

$$Y = X\beta + U$$

where β satisfies the symmetric condition that $\beta = \beta'$. There are n linear equations with n regressors and T observations.

Subject to the symmetry condition, the first order conditions on the objective function $\theta = 1/2 \text{ (tr } \Omega^{-1} U'U \text{)}$ are

$$\partial\theta/\partial\beta = \Omega^{-1} U'X + X'U\Omega^{-1} = 0$$

Substituting $U = Y - X\beta$ into the function, we get

$$\Omega^{-1}(Y-X\beta)'X + X'(Y-X\beta)\Omega^{-1} = 0$$

$$\Rightarrow \quad \Omega^{-1}Y'X + X'Y\Omega^{-1} = \Omega^{-1}\beta X'X + X'X\beta\Omega^{-1}$$

$$\Rightarrow \quad \Omega^{-1}Y'X + X'Y\Omega^{-1} = \Omega^{-1}X'X\beta + \beta X'X\Omega^{-1} \qquad (6.1)$$

Denote $C = \Omega^{-1}Y'X + X'Y\Omega^{-1}$ and $A = X'X\Omega^{-1}$, then equation (6.1) becomes

$$C = A'B + BA$$

which is a system of Lyapunov equations and can be solved using characteristic equations.

Proposition (Simultaneous Decomposition of semi-definite and definite matrices)[1] : Let Ω^{-1} be a positive definite matrix and X'X be positive semi-definite matrix. Then there exists a non-singular matrix W such that

$$\Omega^{-1} = W'W \qquad \text{and} \qquad X'X = W'\wedge W$$

where $\wedge$ is a diagonal matrix of the eigen values (characteristics roots) of X'X in the metric of Ω^{-1}, and W is a non-singular matrix.

Then (6.1) becomes

$$W'WBW'\wedge W + W'\wedge WBW'W = C$$

Pre- and post-multiplying by the terms $(W')^{-1}$ and W^{-1} respectively, we get

$$WBW'\wedge + \wedge WBW' = (W')^{-1}CW^{-1}$$

Denote WBW' as G, and $(W')^{-1}CW^{-1}$ as H, we have

$$G\wedge + \wedge G = H \qquad (6.2)$$

From (6.2), the term G can be solved directly. For example, suppose n=2 for simplicity, then (6.2) becomes

$$\begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \begin{bmatrix} \wedge_{11} & 0 \\ 0 & \wedge_{22} \end{bmatrix} + \begin{bmatrix} \wedge_{11} & 0 \\ 0 & \wedge_{22} \end{bmatrix} \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}$$

$$\begin{bmatrix} G_{11}\wedge_{11} & G_{12}\wedge_{22} \\ G_{21}\wedge_{22} & G_{22}\wedge_{22} \end{bmatrix} + \begin{bmatrix} G_{11}\wedge_{11} & G_{12}\wedge_{11} \\ G_{21}\wedge_{22} & G_{22}\wedge_{22} \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}$$

$$\begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} = \begin{bmatrix} H_{11}/(\wedge_{11}+\wedge_{11}) & H_{12}/(\wedge_{22}+\wedge_{11}) \\ H_{21}/(\wedge_{11}+\wedge_{22}) & H_{22}/(\wedge_{22}+\wedge_{22}) \end{bmatrix}$$

Since WBW' = G, we can solve for β simply with the equation

$$\tilde{B} = W^{-1}GW'^{-1} \qquad (6.3)$$

Now the question is how to find the non-singular matrix W that satisfies the conditions?

Let $E = \Omega^{-1}$ and A = X'X.

Proposition : For a positive definite matrix E, there exists a non-singular matrix P such that

$$E = P'^{-1}P^{-1}$$

or

$$E^{-1} = PP' = \Omega$$

Consequently, the eigen values of A in the metric of E are simply the usual eigen values of P'AP.

Let Q be the (orthogonal) matrix of (ordinary) eigen vectors (characteristics vectors) of P'AP. Thus, we have

$$P'APQ = Q\Lambda$$

From that, it follows

$$A = P'^{-1}Q\Lambda Q'P^{-1}$$

Hence,

$$W = Q'P^{-1} \qquad (6.4)$$

Proposition : Since matrix Ω is symmetric positive (semi) definite, let R denote the diagonal matrix of its (real) eigen values and V the associated (orthogonal) matrix of eigen vectors. We have

$$\Omega V = VR$$

$$\Rightarrow \quad \Omega = VRV'$$

Since Ω = PP', then

$$PP' = VRV'$$

$$\Rightarrow \quad P = V\,(R^{1/2}) \qquad (6.5)$$

Having found P, the non-singular matrix from (6.5), we can then solve for W in (6.4). Then we form the matrix H, which is the right hand side component of (6.2). Subsequently, we can solve for $\tilde{B}$ using (6.3). The process simply involves the manipulation of the characteristic equations.

generalizing, suppose the linear symmetric demand system of n equations is of the form

$$Y = XB + Z\Gamma + U$$

where B is (n^*n) symmetric matrix and Z collects any other variables. Nuisance parameters, Γ, which are not part of the symmetric set may be concentrated out.

The first order conditions on the objective function yield an equation of the form

$$Z'U = 0$$

Substituting U into the equation, we get

$$Z'(Y - XB - Z\Gamma) = 0$$

$$\Rightarrow \quad Z'Y - Z'XB - Z'Z\Gamma = 0$$

$$\Rightarrow \quad \hat{\Gamma} = (Z'Z)^{-1}Z'Y - (Z'Z)^{-1}Z'XB \tag{6.6}$$

Substituting $\hat{\Gamma}$ into the original model, we have

$$Y = XB + Z(Z'Z)^{-1}Z'Y - Z(Z'Z)^{-1}Z'XB + U$$

$$\Rightarrow \quad Y - Z(Z'Z)^{-1}Z'Y = (X - Z(Z'Z)^{-1}Z'X)B + U$$

The concentrated model becomes

$$Y^* = X^* B + U$$

where $Y^* = (Y - Z(Z'Z)^{-1}Z'Y)$ and $X^* = (X - Z(Z'Z)^{-1}Z'X)$.

Once B is solved $\hat{\Gamma}$ can be recovered from equation (6.6). A minor problem is the imposition of the within-equations restriction, homogeneity ($\Sigma_j B_{ij} = 0$). To introduce homogeneity by substitution with the elimination of the redundant variable is appropriate. Thus X is redefined as $(X_1-X_n, X_2-X_n, ..., X_{n-1}-X_n)$ and is a $T^*(n-1)$ matrix. Y is $T^*(n-1)$ and B is symmetric $(n-1)^*(n-1)$.

6.2 Kolmogorov-Smirnov Two-Sample Test

The test is used to compare two empirical distribution functions[2]. For samples drawn independently from two populations with cumulative probability distribution $F_x(z)$ and $F_y(z)$, the hypothesis is:

$$H_0 : F_x(z) = F_y(z) \quad \text{for all } z$$

$$H_1 : F_x(z) \neq F_y(z) \quad \text{for some } z$$

The data consist of two mutually independent sets of random observations of size m and n, which are both ordinal. They are ordered in such a way that

$$x_1 < x_2 < \ldots < x_m, \text{ and}$$

$$y_1 < y_2 < \ldots < y_n.$$

Pooling, combined ordered arrangement of the (m+n) random samples may be formed, and is denoted as z. Then define two respective samples or empirical distribution functions as

$S_x(z)$ = (number of observations in the first sample, x, that are less than or equal to z) / m

$S_y(z)$ = (number of observations in the second sample, y, that are less than or equal to z) / n

The population distributions of x and y are the same, corresponding values of the empirical distribution functions, $S_x(z)$ and $S_y(z)$, should be small and there should be reasonable agreement between them for all values of z. The absolute values of the differences of $S_x(z)$ and $S_y(z)$ are a measure of this disagreement. If the maximum absolute difference is small, then so are all differences. Hence, the largest absolute difference is the value of the Kolmogorov-Smirnov test statistic which is denoted as D.

D is completely distribution-free for any continuous common population distribution. If m and n are large, the critical values, P, based on the asymptotic probability distribution, can be approximated by :

$\underline{\alpha}$	$\underline{P}$
0.1	1.07 * sqrt[(m+n)/(m*n)]
0.05	1.22 * sqrt[(m+n)/(m*n)]
0.01	1.36 * sqrt[(m+n)/(m*n)]

(6.7)

If the test statistic, D, is less than the critical value, P, we shall accept the null hypothesis that the two distributions are identical.

6.3 Non-Parametric Bootstrap

So far, none of the models examined accept the demand theory even after correcting for parameters inconstancy and heteroscedasticity. It is important to understand why the rejections are so strong. Did those test statistics over-reject the hypothesis? Are they testing with the same empirical size as the actual nominal values? In order to answer these questions, it is necessary to construct a non-parametric bootstrap.

The classical statistical techniques, whose justification in probability is based on specific assumptions about the population samples, are called parametric methods. A non-parametric method is a distribution-free procedure in which the analysis is based on some functions of the sample observations whose sampling distribution can be determined without knowledge of the specific distribution function of the underlying population. As explained by Efron and Gong(1983), a good parametric analysis, when appropriate, can be far more efficient than its non-parametric counterpart. Often, though, parametric assumptions are difficult to justify, in which case it is reassuring to have available the comparatively crude but trustworthy non-parametric techniques.

Basically bootstrapping is a general technique for appraising the properties of an estimator. It is distribution-free, that is, non-parametric, and it is intended to reproduce the appropriate finite sample behaviour for the estimator. Since the standard error is the traditional measure of accuracy, the bootstrap is a simulation procedure to estimate standard errors using Monte-Carlo experiments but with a non-parametric estimate of the underlying error distribution. The idea is to resample the original observations in a suitable way that we can construct "pseudo-data" on which the estimator of interest is exercised.

Freedman and Peters(1984) presented a detailed description of the construction of the bootstrap with reference to an econometric equation. Also, Efron and Tibshirani(1986) review the bootstrap methods and its basic ideas and applications. For a simple exposition, see Efron and Gong(1983).

For our purposes, we are particularly interested in investigating the possibility of over-rejection. In other words, we are going to analyse the empirical sizes of the tests (or type I error, which is the probability of rejecting the true null hypothesis) using non-parametric bootstraps without knowledge of the specific distribution function of the underlying population.

Generally, the construction involved continuous resampling of the estimated residual with replacement to find out the rejection percentage. The procedure is as follows:

First, use the actual data to estimate the demand model subject to the restrictions. Get estimates, $\tilde{\beta}$, and store the residuals, $\tilde{e}$.

Secondly, resample $\tilde{e}$ with replacement and inflate the residuals by multiplying the term "sqrt[T/(T-p)]"[3] to form e^*. (p is the number of parameters in each demand equation)

Thirdly, generate "pseudo-data" for Y,

i.e. $Y^* = X\tilde{\beta} + e^*$

and re-estimate the model with and without the restrictions imposed.

Finally, repeat the resampling process at least 100 times and record the inference results. Efron and Tibshirani(1986) argue that 50 to 200 replications are quite adequate for most situations.

The questions of importance are:

(1) Are the empirical test statistics distributed as χ^2?

(2) Are the empirical sizes of the test statistics the same as the actual (nominal) values? If they are different, what are the empirical critical values of the tests? Is the theory still rejected?

We estimated a system of Lyapunov Equations for the original AIDS model (unweighted and no dummies included) and experimented with the non-parametric bootstrap described above. With each demand restriction imposed separately and jointly, we followed the approach by Freedman and Peters(1984) and repeated the iteration process 100 times. Wald, LR and LM statistics were calculated. An illustration of estimating the Lyapunov Equations, with homogeneity and symmetry imposed, is given in Appendix 5[4].

The simulated test statistics were sorted in ascending order. Next, instead of dealing with 100 observations, a group of 20 observations were drawn, each corresponding to a 5% cumulative frequency distribution beginning from the 5% point. The 100% point was replaced by the 99% point to avoid infinity.

A true χ^2 distribution with the appropriate degrees of freedom was generated corresponding to the simulated probability distribution selected earlier. Then the two distributions were compared using the Kolmogorov-Smirnov two-sample test described previously. The results are reported in Table 6.1.

Since the Kolmogorov-Smirnov test statistics are significantly smaller than the critical values in all cases, we accept the null hypothesis that the two distributions are identical and conclude that the empirical distributions are distributed as χ^2. Notice that the mean of a χ^2 distribution should be equal to the number of degrees of freedom. Comparing the sample means of the empirical distributions (which are presented in Table 6.2) with the actual number of degrees of freedom, it is clear that the empirical distributions are not centred correctly. The problem is more serious when testing the homogeneity condition. Also, Wald statistics have the largest deviation from the actual means. LR statistics come second and LM statistics are the smallest.

The number of rejections was counted for each restriction and compared with the true sizes, the results being presented in Table 6.3. They reveal that the empirical sizes of the tests are, obviously, larger than the nominal values. This implies that the problem of over-rejection exists in the empirical testing procedures and is consistent with the general belief that the test statistics are biased towards rejection (Laitinen[1978) and Meisner[1979]). Since the statistics are actually asymptotic test criteria, the over-rejection may be due to the fact that we are dealing with a small finite sample, which may not be large enough to justify asymptotic theory. This is the small sample bias as examined by Laitinen(1978). Another possible explanation is the consequence of replacing the true but unknown population covariance matrix by the unbiased least squares estimate. The empirical sizes of the three statistics still satisfy the inequality developed by Berndt and Savin(1977).

The empirical critical values are reported in Table 6.4, with the empirical figures being larger than the nominal values. The differences in percentages vary with the significant level α and the restrictions imposed[5]. Interestingly for the within equation restriction of homogeneity, the differences increase as α increases. On the other hand, testing symmetry, the differences increase as α decreases. Generally, Wald statistics have the largest deviations from 11% to 25% for homogeneity, and 6% to 13% including symmetry. The LR statistics are second with 11% to 20% and 3% to 8% respectively. The LM statistics have the smallest which are about 3% to 16% and 0.1% to 5% respectively for the homogeneity and symmetry conditions. If evaluating the actual statistics from Table 4.18 using the empirical critical values, it is clear that the demand conditions are still significantly rejected.

There is an important implication from the non-parametric exercise. This is that the asymptotic test statistics all assume the satisfaction of normality condition, but in this study, the errors are non-normal. There is considerable uncertainty about the behaviour of the asymptotic test statistics. By using a simple and computer intensive technique, non-parametric bootstrapping, we observed that the test statistics are empirically asymptotic χ^2 distributed, and found that the empirical means are mis-positioned by about 1% to 18% (evaluated from the actual means). Also, the problem of over-rejection (over-size) is noted. A question of importance is : to what extent is this due to the non-normal and heteroscedastic errors? This motivates the following experiment on parametric bootstrapping.

6.4 Parametric Bootstrap (Monte-Carlo Experiments)

As the demand conditions have been consistently rejected by the data, perhaps another question is the appropriateness of asymptotic testing procedures in finite samples. Byron(1987) has suggested a method to test such a hypothesis in a demand model using "parametric bootstrapping". It is similar to the non-parametric bootstrap discussed earlier except in its resemblance to a Monte-Carlo simulation technique, which is based on the condition that the assumed specific underlying probability distribution is satisfied. For example, the Wald, LR and LM statistics all require the errors to be normally distributed with zero mean and constant variance. Otherwise, the statistics are misleading and, strictly speaking, invalid. It is highly possible that this problem exists in this particular study. Univariate

Jarque-Bera's normality tests on each single equation reported in Table 5.20 revealed that non-normal residuals exist in most of the equations in the original maintained model. This may explain the presence of over-rejection in the model. Therefore, we are concerned with the extent to which non-normality influences results.

The process of setting up a parametric bootstrap is as follows:

First, estimate the restricted demand model with the actual data, and store the estimated parameters $\tilde{\beta}$ and the covariance matrix of residuals $\tilde{\Omega}$. These estimates are treated as the population parameters.

Secondly, use the matrix $\tilde{\Omega}$ to generate a new set of disturbances V^* (which is a set of randomly generated numbers), in such a way that it satisfies the condition $V^* \sim N(0, \tilde{\Omega})$. This can be done by performing Cholesky's decomposition of the symmetric positive definite matrix.

Thirdly, generate a new set of "pseudo-data" for Y

$$Y^* = X\beta + V^*$$

and re-estimate both unrestricted and restricted models.

Finally, repeat the generating process of V^* and re-estimate the model, at least 100 times.

Since the parameters of the generating function obey the restrictions and the residuals are normally distributed by construction, a test of the conventional assumptions of homogeneity and symmetry imply the question : Assuming normality is satisfied, do the (Wald, LR and LM) tests over-reject the hypothesis in the finite sample? Also, are the parametric results similar to the non-parametric results generated earlier?

The Lyapunov Equations are estimated again with 100 repetitions. Wald, LR and LM test statistics were also calculated. Note that by assuming the true variance-covariance matrix Ω is known, the Wald and LM statistics are identical. They are calculated by

$$\text{Wald} = \text{LM} = T * \text{tr}\,[\Omega^{-1} * (\tilde{\Omega} - \hat{\Omega})]$$

where the true matrix Ω is the constrained estimates from the original model. As the information of the assumed true variance-covariance matrix cannot be incorporated into the LR statistics, the statistics are different from the Wald and LM values.

The Kolmogorov-Smirnov Two-sample tests are performed and the results are presented in Table 6.5. Again the Kolmogorov-Smirnov tests significantly accepted the hypothesis that the empirical distributions are χ^2. The calculated sample means of the statistics are given in Table 6.6. As with the previous finding, the empirical means are mis-centred and slightly skewed to the right. The skewness is more serious if homogeneity is imposed. Deviations from the actual means (in percentage terms) are 12 to 31 for homogeneity, and are less than 6% for symmetry.

The number of rejections in each case was counted to check if the empirical sizes are equal to the actual values. Table 6.7 notes the results and reveals that the problem of over-rejection still remains[5]. This implies that the problem experienced previously in the non-parametric bootstrap is not simply caused by non-normality.

Table 6.8 shows the calculated empirical critical values in each situation. Interestingly, evaluation of the actual statistics from the original maintained AIDS model given in Table 4.18 reveals that the homogeneity condition although rejected by the Wald and LR statistics, is marginally accepted by the LM criteria at $\alpha=0.01$. However the rejection of the joint validity of all demand conditions is once again confirmed.

In Chapter 4, we mentioned a size correcting factor suggested by Bera, Byron and Jarque(1981) which is a ratio of the nominal (actual asymptotic) critical values and the empirical χ^2 critical values from Monte Carlo simulation. The ratios are calculated with respect to each statistic at $\alpha=0.05$. The calculated ratios (the adjustment factors) and the adjusted test statistics are presented in Tables 6.9 and 6.10 respectively. Only the homogeneity condition is marginally accepted by the adjusted LM criteria.

6.5 Bootstrapping the New Maintained AIDS Model

Both non-parametric and parametric bootstraps were constructed with reference to the new maintained AIDS derived in chapter 5, which includes both urban and island dummies, and a weighting factor, the inverse of "sqrt(LKTY)". From the non-parametric bootstrapping, it was found that the empirical distributions of the test statistics are still χ^2. As reported in Table 6.11, they passed the Kolmogorov-Smirnov test easily even though the means are mis-positioned. Checking Table 6.12, the empirical distributions are positively skewed. The Wald statistics are about 8% to 16% different from the actual means. The LR and LM statistics are respectively about 6% to 13% and 3% to 10% larger than the actual means. Over-rejections are again observed in Table 6.13. Evaluating the actual test statistics in Table 5.19 based on the empirical critical values reported in Table 6.14 reveals that all demand conditions are still rejected.

Parametric bootstraps were also applied to the model. Kolmogorov-Smirnov test statistics recorded in Table 6.15

indicated that the empirical statistics are distributed as χ^2 in every case. Mis-positions of the means are again observed in Table 6.16 but these are negligible. Even assuming the population covariance is known, over-rejection still remains (Table 6.17). When testing symmetry with Wald and LM principles, almost exact sizes are recorded. Surprisingly, under-rejection is observed if testing homogeneity using the LR criteria. Empirical critical values are calculated and reported in Table 6.18. Rejection of all restrictions are confirmed, even based on the empirical sizes. Adjusting the statistics with the size correcting factors presented in Table 6.19, which are the ratios of the nominal and empirical critical values suggested by Bera, Byron and Jarque(1981), the resulting size corrected statistics reported in Table 6.20 still significantly rejected all restrictions.

6.6 Summary

Examining the bootstrapping results of the original and new maintained AIDS models, one may observe the following:

First, ignoring the error distribution for the moment, the asymptotic test statistics are appropriate in the analysis. Although the means of the empirical distributions are not positioned correctly and over-rejections are observed, they are empirically distributed as χ^2. The reliability of the results may be questioned because of the number of replications of 100 was too small. The problem can be examined using the binomial theorem. Suppose α = 0.05 (i.e. the population proportion of rejection is 5%) then the standard error of α is:

$$SE_{\alpha} = \text{Sqrt} [(\alpha^*(1-\alpha))/R]$$

where R is the number of replications. If R equals 100, then the value for SE_{α} is 0.02. This means that the proportion of rejections varies from 0.03 to 0.07. If R equals 500, then SE_{α} becomes 0.01, and the percentage of rejections will fall in the range of 0.04 to 0.06. Similarly, if R is set at 1000, SE_{α} becomes 0.007 and the empirical α will vary from 0.043 to 0.057. Since R is set at 100 in this study, the variability of the outcomes will be misleading. In order to reach a firm conclusion, we should at least experiment with 500 or more bootstrap replications[7].

Secondly, small sample corrected statistics (either evaluated with empirical critical values or adjusting the statistics with some arbitrary factors) rejected the joint validity of demand conditions (although homogeneity condition is sometimes marginally accepted). This implies the problem of small sample bias is not the major reason for the rejection of theory.

Thirdly, non-normal errors, strictly speaking, invalidate the testing procedures, but the problem may not be serious. Comparing the results from non-parametric and parametric bootstraps indicated that the difference in inference is only small. Therefore, in this particular case, one may expect to obtain similar results even if the errors satisfy normality.

Finally, although the new maintained AIDS model significantly improved the model performance (both in fit and inference measures), it still rejected the postulated demand theory. The results are the same with distribution-free approaches. This means the model may be inadequately specified, or may be due to the fact that the theory is inappropriate to this study.

Footnotes

1. The propositions in this chapter can be referenced from Dhrymes(1978).

2. The test appears in most of the texts relating to non-parametric techniques (e.g. Gibbons[1971] & [1976]).

3. As explained by Freedman and Peters(1984), some inflation of the residuals may be desirable to compensate for the deflation of the residuals in fitting.

4. The algorithm was first tested against Theil's(1975) Rotterdam model using Barten's(1966) original Dutch data and yielded identical results.

5. The empirical critical values should be exact. The discrepancy may be because of the number of repetition is too small. Perhaps 200 or even 500 repetitions would have been more appropriate.

6. In fact, as the population covariance matrix is assumed to be known, it is expected that the problem will be eliminated if the number of repetitions is increased, say to 500 or more.

7. I have to thank Dr. R. Byron for pointing this argument to me.

Table 6.1

Non-Parametric Bootstrap (Original Maintained AIDS)

Kolmogorov-Smirnov Two-sample Tests[a]

	Homogeneity	Symmetry	Homogeneity & Symmetry
Wald	.20*	.25*	.25*
LR	.20*	.20*	.15*
LM	.20*	.15*	.10*

* accept null at 10%

	Critical values
$\alpha = 0.1$	.3383637
$= 0.05$	.3857978
$= 0.01$	.4300697

a. sample size = 20

Table 6.2

Non-Parametric Bootstrap (Original Maintained AIDS)

Sample Means of the Statistics[b]

	Homogeneity	Symmetry	Homogeneity & Symmetry
Wald	18.87(17.93)[c]	130.46(8.72)	145.06(6.66)
LR	18.41(15.06)	127.25(6.04)	141.28(3.88)
LM	17.97(12.31)	124.18(3.48)	137.66(1.22)
Actual D.F.	16	120	136

b. sample size = 100

c. parethesis are percentage deviations from actual (hypothetical) means

Table 6.3

Non-Parametric Bootstrap (Original Maintained AIDS)

Number of Rejection[a]

	$\chi^2(16)$ Homogeneity	$\chi^2(120)$ Symmetry	$\chi^2(136)$ Homogeneity & Symmetry
Wald			
α = 0.1	21	29	18
0.05	16	17	16
0.01	4	6	5
LR			
α = 0.1	19	20	17
0.05	15	10	8
0.01	4	3	4
LM			
α = 0.1	17	15	11
0.05	14	8	6
0.01	3	3	3

a. sample size = 100

Critical values	$\chi^2(16)$	$\chi^2(120)$	$\chi^2(136)$
α = 0.1	23.54	140.23	157.52
0.05	26.30	146.57	164.22
0.01	32.00	158.95	177.28

Table 6.4

Non-parametric Bootstrap (Original Maintained AIDS)

Empirical Critical Values[a]

	Homogeneity	Symmetry	Homogeneity & Symmetry
Wald			
α = 0.1	29.2164(24.11)	150.2617(7.15)	167.9464(6.62)
0.05	30.8628(17.35)	160.1772(9.28)	176.8345(7.68)
0.01	35.6269(11.33)	177.1979(11.48)	200.0271(12.83)
LR			
α = 0.1	28.2538(20.02)	146.4746(4.45)	162.2443(2.99)
0.05	29.7913(13.27)	154.9380(5.71)	171.9220(4.69)
0.01	35.5175(10.99)	172.1386(8.30)	192.4435(8.55)
LM			
α = 0.1	27.3330(16.11)	142.5858(1.68)	157.6962(0.11)
0.05	28.7687(9.39)	149.7432(2.16)	167.2298(1.83)
0.01	32.8654(2.70)	167.3076(5.26)	185.3305(4.54)

a. Parathesis are the percentage deviations from the actual critical values

Table 6.5
Parametric Bootstrap (Original Maintained AIDS)
Kolmogorov-Smirnov Two-sample Tests[a]

	Homogeneity	Symmetry	Homogeneity & Symmetry
Wald	.25*	.10*	.10*
LR	.20*	.15*	.15*
LM	.25*	.10*	.10*

* accept null at 10%

	Critical values
α = 0.1	.3383637
= 0.05	.3857978
= 0.01	.4300697

a. sample size = 20

Table 6.6
Parametric Bootstrap (Original Maintained AIDS)
Sample Means of the Statistics[b]

	Homogeneity	Symmetry	Homogeneity & Symmetry
Wald	21.00(31.25)[c]	123.41(2.84)	137.22(0.89)
LR	17.90(11.88)	126.64(5.53)	141.16(3.79)
LM	21.00(31.25)	123.41(2.84)	137.22(0.89)
Actual degrees of freedom	16	120	136

b. sample size = 100
c. parathesis are percentage deviations from actual means

Table 6.7

Parametric Bootstrap (Original Maintained AIDS)

Number of Rejection[a]

	$\chi^2(16)$ Homogeneity	$\chi^2(120)$ Symmetry	$\chi^2(136)$ Homogeneity & Symmetry
Wald			
$\alpha = 0.1$	34	18	13
0.05	25	12	4
0.01	11	3	2
LR			
$\alpha = 0.1$	22	21	24
0.05	9	15	11
0.01	2	3	3
LM			
$\alpha = 0.1$	34	18	13
0.05	25	12	4
0.01	11	3	2

a. sample size = 100

Critical values	$\chi^2(16)$	$\chi^2(120)$	$\chi^2(136)$
$\alpha = 0.1$	23.54	140.23	157.52
0.05	26.30	146.57	164.22
0.01	32.00	158.95	177.28

Table 6.8

Parametric Bootstrap (Original Maintained AIDS)

Empirical Critical Values[a]

	Homogeneity	Symmetry	Homogeneity & Symmetry
Wald			
α = 0.1	33.19996(41.04)	147.7294(5.35)	160.7608(2.06)
0.05	37.56457(42.83)	148.9947(1.65)	163.6727(-.33)
0.01	49.64409(55.14)	169.2074(6.45)	182.1315(2.74)
LR			
α = 0.1	26.20929(11.34)	150.0668(7.01)	164.6537(4.53)
0.05	27.10980(3.08)	154.1945(5.20)	170.7362(3.97)
0.01	40.79705(27.49)	169.8056(6.83)	186.8452(5.39)
LM			
α = 0.1	33.19996(41.04)	147.7294(5.35)	160.7608(2.06)
0.05	37.56457(42.83)	148.9947(1.65)	163.6727(-.33)
0.01	49.64409(55.14)	169.2074(6.45)	182.1315(2.74)

a. Parathesis are the percentage deviations from the actual critical values

Table 6.9

Size Correcting Factors (Original Maintained AIDS)[a]

	Homogeneity	Symmetry	Homogeneity & Symmetry
Wald	.7001278	.9837262	1.003344
LR	.9701288	.9505527	.9618346
LM	.7001278	.9837262	1.003344

a. Bera, Byron & Jarque's(1981) correcting factors at $\alpha=0.05$

Table 6.10

Size Corrected Test Statistics (Original Maintained AIDS)[b]

	Homogeneity	Symmetry	Homogeneity & Symmetry
Wald	35.090405	884.36002	1118.6783
LR	44.111757	704.62571	864.87205
LM	28.873270***	613.01882	744.51127

b. Bera, Byron & Jarque's(1981) correcting factor at $\alpha=0.05$

*** insignificantly different from zero at $\alpha=0.01$

Table 6.11

Non-Parametric Bootstrap (New Maintained AIDS)

Kolmogorov-Smirnov Two-sample Tests[a]

	Homogeneity	Symmetry	Homogeneity & Symmetry
Wald	.20*	.30*	.35**
LR	.15*	.25*	.25*
LM	.15*	.20*	.20*

* accept null at $\alpha=0.1$

** accept null at $\alpha=0.05$

	Critical values
$\alpha = 0.1$	.3383637
$= 0.05$	.3857978
$= 0.01$	.4300697

a. sample size = 20

Table 6.12

Non-Parametric Bootstrap (New Maintained AIDS)

Sample Means of the Statistics[b]

	Homogeneity	Symmetry	Homogeneity & Symmetry
Wald	18.597(16.23)[c]	133.02(10.85)	148.12(8.91)
LR	18.139(13.37)	129.66(8.05)	144.19(6.02)
LM	17.697(10.61)	126.43(5.36)	140.43(3.26)

b. sample size = 100

c. parethesis are percentage deviations from actual (hypothetical) means

Table 6.13

Non-Parametric Bootstrap (New Maintained AIDS)

Number of Rejection[a]

	$\chi^2(16)$ Homogeneity	$\chi^2(120)$ Symmetry	$\chi^2(136)$ Homogeneity & Symmetry
Wald			
$\alpha = 0.1$	20	33	28
0.05	15	24	20
0.01	8	5	6
LR			
$\alpha = 0.1$	19	26	21
0.05	15	20	10
0.01	5	4	4
LM			
$\alpha = 0.1$	18	23	13
0.05	14	11	7
0.01	3	2	2

a. sample size = 100

Critical values	$\chi^2(16)$	$\chi^2(120)$	$\chi^2(136)$
$\alpha = 0.1$	23.54	140.23	157.52
0.05	26.30	146.57	164.22
0.01	32.00	158.95	177.28

Table 6.14

Non-parametric Bootstrap (New Maintained AIDS)

Empirical Critical Values[a]

	Homogeneity	Symmetry	Homogeneity & Symmetry
Wald			
α = 0.1	29.82857(26.71)	156.5965(11.67)	169.4211(7.56)
0.05	33.10423(25.87)	158.8700(8.39)	179.9863(9.60)
0.01	38.20424(19.39)	169.5186(6.65)	191.0089(7.74)
LR			
α = 0.1	28.82609(22.46)	151.9464(8.35)	163.7404(3.95)
0.05	31.87546(21.20)	154.1127(5.15)	174.4353(6.22)
0.01	36.57993(14.31)	163.3727(2.78)	184.3018(3.96)
LM			
α = 0.1	27.86804(18.38)	147.2853(5.03)	158.4910(0.62)
0.05	30.70677(16.76)	149.6220(2.08)	169.1717(3.02)
0.01	35.04641(9.52)	159.3641(0.26)	177.9823(0.39)

a. Parathesis are percentage deviations from the actual critical values

Table 6.15
Parametric Bootstrap (New Maintained AIDS)
Kolmogorov-Smirnov Two-sample Tests[a]

	Homogeneity	Symmetry	Homogeneity & Symmetry
Wald	.20*	.05*	.10*
LR	.15*	.25*	.25*
LM	.20*	.05*	.10*

* accept null at 10%

	Critical values
α = 0.1	.3383637
= 0.05	.3857978
= 0.01	.4300697

a. sample size = 20

Table 6.16
Parametric Bootstrap (New Maintained AIDS)
Sample Means of the Statistics[b]

	Homogeneity	Symmetry	Homogeneity & Symmetry
Wald	18.1425(13.39)[c]	121.43(1.19)	137.245(0.92)
LR	16.7370(4.61)	127.96(6.63)	144.140(5.99)
LM	18.1425(13.39)	121.43(1.19)	137.245(0.92)
Actual degrees of freedom	16	120	136

b. sample size = 100
c. parathesis are percentage deviations from actual means

Table 6.17

Parametric Bootstrap (New Maintained AIDS)

Number of Rejection[a]

	$\chi^2(16)$ Homogeneity	$\chi^2(120)$ Symmetry	$\chi^2(136)$ Homogeneity & Symmetry
Wald			
α = 0.1	19	11	16
0.05	9	5	6
0.01	2	1	0
LR			
α = 0.1	8	26	25
0.05	3	13	15
0.01	1	4	1
LM			
α = 0.1	19	11	16
0.05	9	5	6
0.01	2	1	0

a. sample size = 100

Critical values	$\chi^2(16)$	$\chi^2(120)$	$\chi^2(136)$
α = 0.1	23.54	140.23	157.52
0.05	26.30	146.57	164.22
0.01	32.00	158.95	177.28

Table 6.18
Parametric Bootstrap (New Maintained AIDS)
Empirical Critical Values[a]

	Homogeneity	Symmetry	Homogeneity & Symmetry
Wald			
α = 0.1	26.10350(10.89)	140.2548(0.02)	162.4874(3.15)
0.05	28.67176(9.02)	146.5270(-.03)	165.3816(.707)
0.01	32.51500(1.61)	156.9398(-1.26)	170.6417(-3.74)
LR			
α = 0.1	22.80033(-3.14)	147.1203(4.91)	168.8939(7.22)
0.05	24.37482(-7.32)	156.5786(6.83)	172.2473(4.89)
0.01	31.48814(-1.60)	162.9452(2.51)	176.5087(-.44)
LM			
α = 0.1	26.10350(10.89)	140.2548(0.02)	162.4874(3.15)
0.05	28.67176(9.02)	146.5270(-.03)	165.3816(.707)
0.01	32.51545(1.61)	156.8559(-1.26)	170.6417(-3.74)

Table 6.19

Size Correcting Factors (New Maintained AIDS)[a]

	Homogeneity	Symmetry	Homogeneity & Symmetry
Wald	.9172788	1.0002935	.9929762
LR	1.078982	.93607930	.9533966
LM	.9172788	1.0002935	.9929762

a. Bera, Byron & Jarque's(1981) correcting factors at α=0.05

Table 6.20

Size Corrected Test Statistics (New Maintained AIDS)[b]

	Homogeneity	Symmetry	Homogeneity & Symmetry
Wald	40.14012	387.96383	478.67411
LR	43.05139	324.99737	408.63532
LM	33.35226	313.11187	380.93546

b. Bera, Byron & Jarque's(1981) correcting factor at α=0.05

CHAPTER VII

CONCLUSION

In Chapter 4, we estimated several popular demand systems such as the LES, the Double-log system and the AIDS models. Based on the information inaccuracy measures, the AIDS model was chosen as the preferred model. After subjecting the system to a series of diagnostic tests in Chapter 5, a modified AIDS model was developed. However, the maintained model rejected demand theory. The causes of the rejection, to some extent, are directly related to certain statistical weaknesses such as the non-normal error distribution which invalidates the testing process (Jarque and McKenzie[1983]). Also, due to the fact that we are dealing with a small finite sample, the test statistics could be biased towards rejection (Laitinen[1978] and Meisner[1979]). But as illustrated in Chapter 6, the simulated results using the bootstrap indicate the effect of non-normality on inference was not particularly serious. Furthermore, the size corrected test statistics still rejected the conditions of demand theory significantly.

If the demand conditions are indeed a "fundamentally weak hypotheses" and it is "intuitively implausible" to reject these restrictions (Deaton[1972] & Bewley[1983]), it is important to examine the reasons for the rejection. These will be discussed below.

A general belief is that rejection of consumer theory can be the result of aggregation (Kiefer and MacKinnon[1976], Deaton and Muellbauer[1980a] and Mattei[1986]). Econometric analyses cannot deal with each individual item and instead focus on a limited groups of commodities. The assumption necessitated (i.e. the weak separability condition), as argued by Deaton and Muellbauer(1980a), is by no means trivial. Secondly, the transition from the microeconomics of consumer behaviour to the analysis of market demand is a serious problem. Deaton and Muellbauer(1980a) argue that there are few grounds to justify the argument by Hicks(1956) and Houthakker and Taylor(1970) that "the aggregation error" is negligible by assuming the variations in circumstances of individual household are averaged out in aggregate and only the systematic effects of variations in prices and budgets remain.

It is probably the case that the aggregation problem is an inevitable dilemma between theoretical issues and empirical analysis in applied demand studies. To make the abstract theory operational with current consumer theory, it may be necessary to make non-trivial assumptions. However, this is unavoidable.

There are other possible reasons for rejecting the demand conditions (Deaton and Muellbauer[1980a]). For example, total expenditure or income is defined to be exogenous. But they are in fact choice variables affected by the prices of commodities. Furthermore, quantities supplied are assumed to be elastic so that the suppliers can meet whatever demand emerges at a predetermined price (set by the manufacturers themselves). But in the real world, there may be difficulties in satisfying demand in some places due to the non-availability of some goods. Price expectations are also important in determining the current consumption behaviour. In addition, measurement errors could be another major problem in empirical studies.

Since the reasons for rejecting demand restrictions are largely unresolved, a conclusion cannot be reached concerning the validity of the theory. Clearly, there are different attitudes among researchers towards this issue. Some economists, such as Christensen, Jorgenson and Lau(1975), demonstrate the unambiguous rejection of demand conditions and conclude that the theory of demand is inconsistent with the data. On the other hand, other take a different view on this matter. For instance, as noted by Phlips(1974):

> "We find it difficult to take the results of these tests very seriously . . . Given that the demand equations have

to be specified in some way, a valid testing against unrestricted data is probably impossible. We therefore think that, if we want measurement to be meaningful, we must impose the general restrictions whatever the results of the sort of tests just referred to."

Simons and Weiserbs(1979) also take an optimistic view. They argue that the chosen model may not be adequately specified and therefore rejection of restrictions does not mean that well behaved consumer preferences do not exist or that the theory is invalid. Clearly, it is always possible for other specifications to perform better. Mattei's(1986) opinion is that a more "realistic" model of consumer behaviour is needed if we want a firm conclusion.

If the rejection of demand theory is something to do with the specification of the model, then there are at least two aspects worth investigating in future studies. These are: the treatment of demographic variables; and limited dependent variables. These are discussed briefly below.

When modelling a demand system using family budget data, we should account for demographic factors such as: household size and composition; race and religion; age and education. One easy way to do this is to model per capita demand as a function of prices and per capita income. But this approach is too restrictive given that a child is assumed to be equivalent to an adult in consumption, and this equivalence is assumed to be the same for each commodity. Alternatively, without any additional assumptions, the demand system can be applied to sub-samples of households with identical demographic profiles (this is referred to "unpooled" specification" later in this chapter). Such an approach does not require any specific form of relationship between the original model and the demographic variables. Other than these two methods, Pollak and Wales(1981) provide a detailed description of several general procedures for incorporating demographic variables into complete demand systems. The idea is to replace the original class of demand system by a related class which involves, for example, r demographic variables (S_1, S_2, ... S_r). Two of the most common procedures, translating and scaling, are briefly examined.

Demographic translating was first employed by Pollak and Wales(1978) as follows:

Denote the original class of demand system as

$$q_i = h_i (p, y) \qquad i = 1, 2, ..., n$$

Replacing the original model, h_i, in such a way that

$$h_i^* (p, y) = d_i + h_i (p, y - \Sigma_k p_k d_k)$$

where $d_i = D_i(S)$. For linear demographic translating,

$$D_i = \Sigma_r \delta_{ir} S_r$$

The second method of demographic scaling was first proposed by Barten(1964). The idea is to replace the original demand, h_i, by

$$h_i^*(p, y) = m_i + h_i(p_1 m_1, p_2 m_2, \ldots, p_n m_n, y)$$

where $m_i = M_i(S)$. For linear demographic scaling,

$$M_i(S) = 1 + \Sigma_r \delta_{ir} S_r$$

Both translating and scaling add at most (n^*r) independent parameters to the original system. If the original demand system is theoretically plausible, then so is the modified system[1]. As suggested by Pollak and Wales(1981), the results from the two procedures can be compared with the "pooled" and "unpooled" specifications to test for the significance of incorporating demographic variables. Pooled specification combines data from different demographic variables and estimates a single demand system implicitly assuming that consumption patterns are independent of demographic variables. The unpooled specification estimates $(\Sigma_r G_r)$ separate demand systems (where G_r is the number of the subset type in the demographic variable S_r), there being one for each demographic variable type which implicitly assumes that demographic variables affect all demand system parameters. Comparison with the pooled and unpooled results will indicate whether the demographic variables significantly affect consumption patterns, and whether they affect all demand system parameters.

Based on British household budget data from 1966 to 1972, Pollak and Wales(1981) applied the Generalized CES demand system and demonstrated that family size, for example, significantly affects consumption patterns. However, this variable affects only some of the demand system parameters. Similar results are confirmed by Barnes and Gillingham(1978) when estimating the QES using micro data from the 1972-73 Consumer Expenditure Survey in U.S.

If applying these methodologies to the current study, it is sensible to reduce the dimension of the problem since the number of extra independent parameters depends on n and r. Thus, instead of dealing with 17 commodities, it may be necessary to aggregate the commodities into a more manageable number. Note that one has to be prudent when determining the appropriate demographic variables. One may follow Strauss's(1982) approach to determine which characteristic variables should enter the system. His suggestion is to run single equation demand regressions using all of the potential variables and all the possible subsets of independent variables are examined and ranked by the adjusted R^2. Equations having the highest adjusted R^2 and including those variables should be chosen.

The second issue worthy of attention is the treatment of limited dependent variables. It refers to situations when a model's dependent variable can take only a certain range of values (truncated) or when some range of responses is unobservable (censored). Suppose that a random variable, Y is $N(\mu,\sigma^2)$ and that all the observations are for $Y \geq S$. We do not have any observations for $Y < S$. Then the density of Y is truncated and normally distributed. On the other hand, suppose we have a sample of size T, of which T_2 observations are less than S and $T_1 = T - T_2$ observations are equal to or greater than S, and only for the T_1 observations are the exact values known. This is the case of a censored distribution. To distinguish the two concepts in the regression context, we may follow the simple distinction by Maddala(1983). That is, in the case of the truncated regression model, we do not have any observations on either the explained variable Y or the explanatory variables X if the value of Y is above (or below) a threshold. In the case of the censored regression model we have data on the explanatory variables X for all the observations. As for the explained variable Y, we have actual observations for some, but for others we know only whether or not they are above (or below) a certain threshold.

In the analysis of household consumption data, zero responses for individual consumption items are not uncommon but very little attention has been paid to the problem. If we exclude the zero responses, the data is truncated. But if we included those zero responses, the data is censored. Regardless of whether or not the complete sample is used, applying OLS estimator in either case will generate biased and inconsistent estimates because the random disturbances have expectations which are not zeros and which depend upon the exogenous variables. In the case of a (single equation) linear truncated regression model, one may use the MLE (conditional on positive responses) which is consistent and asymptotically efficient. For the censored regression model, since the dependent variable has finite probability mass concentrated at some limit point, say zero, one should adopt the Tobit approach proposed by Tobin(1958). Olsen(1980) suggests a simple and easy way to approximate the maximum likelihood estimator of a truncated normal regression model with the results from the OLS regression. Based upon the known point of truncation, say zero (i.e to the left of zero are truncated), the mean and variance of the resulting incomplete normal distribution, can be used to calculate the mean and variance of the complete distribution using the method of moments by Pearson and Lee(1908).

After adjusting with a correcting factor, the OLS estimates can provide an approximation to the likelihood estimator. Greene(1981) considers a simple estimation of Tobin's(1958) limited dependent variable (the Tobit model is essentially a linear censored regression model in which non-positive observations of the dependent variable are replaced by zeroes) in the following way:

By assuming X_t and e_t are normally distributed, Greene(1981) shows that the bias of the OLS slope estimator can be corrected by multiplying it by the ratio T/T_1, that is the reciprocal of the sample proportion of non-limit observation. The results are shown to provide good approximations to the the Maximum Likelihood estimates proposed by Tobin(1958).

Deaton and Irish(1984) consider, in the context of single equation, a so called P-Tobit model in which the standard Tobit specification is supplemented by the operation of a single binary censor which randomly replaces a fraction of the observations generated by the Tobit model by zeroes. Their basic idea is to statistically deal with the zero responses resulting from misreporting[2] by using a specifically defined log likelihood function which captures the standard regression likelihood function and the binary censoring effect.

In the context of a complete system of equations, Wales and Woodland(1983) consider the problem of estimating a consumer demand system with binding non-negativity constraints. In their first approach they suggest the traditional method of directly maximizing a specific utility function subject to a set of non-negativity constraints using the Kuhn-Tucker conditions. The second method, the Amemiya-Tobin approach, is also considered to estimate systems of budget share or expenditure equations with the assumption that the observed shares or expenditures follow a truncated multivariate normal distribution. For more detailed explanation, one should refer to their original article. Generally speaking, in the Kuhn-Tucker model, the consumption vector for an individual is obtained by constrained maximization of a utility function, and may involve zero consumption of one or more goods. Randomness is incorporated by supposing that the parameters of the utility function are randomly distributed over the population. In the Amemiya-Tobin model, individuals have the same utility function. An individual's observed consumption vector is the sum of the utility maximizing consumption vector plus a vector of random disturbances which has a truncated distribution. This truncation also allows the observed consumption vector to involve zero expenditures on one or more goods.

In addition, there are several implications which can be found in this study. First, our results reinforce the findings of Park(1969) and Klevmarken(1981) that the parameters and estimated elasticities are very sensitive to, and highly dependent on, the model specified. It becomes clear simply by comparing the elasticities generated from the multivariate DLOG and AIDS models (reported in Table 4.15 and 4.19 respectively). Although the elasticity estimates from the two AIDS models (the old and new maintained models) are similar, they are inconsistent with those derived from the multivariate DLOG model. The differences are in terms of sign, magnitude and significance. The different estimation methods are also responsible for the conflicting results. This can be seen in the results on the LES model described in Chapter 4. Therefore, it is important to have the model correctly specified with an appropriate estimator.

An interesting finding from this study is that the elasticities of rice are similar whatever the underlying demand model. The income and own price elasticities for rice are significant and inelastic. This is consistent with the fact that rice, the major food crop and source of most nutrients, is a basic necessity to the people in Indonesia.

Another interesting question was whether or not luxuries are more price elastic than necessities. Following the approach of Clements, Kappelle and Roberts(1985), in which it was demonstrated that luxuries were relatively price elastic, we determined the relationship between income elasticities and own price elasticities. Each pair of elasticities was weighted by the reciprocal of the standard error of the price elasticity, a weighting which gives less weight to those commodities with price elasticities estimated imprecisely. Certain offending observations, such as those commodities with positive price elasticities or negative income elasticities, were excluded. The weighted price elasticities were regressed against the weighted income elasticities for each demand system examined before. The results are reported in Table 7.1.

Table 7.1
Relationships Between Weighted Own Price and Income Elasticities

Model	With Constant: Constant	With Constant: Slope	No Constant: Slope
1. Simple Linear	-.153	-.746	-.834
	(.682)	(4.48)	(7.74)
2. Univariate DLOG	-1.77	-.624	-.747
	(.918)	(4.06)	(9.73)
3. LES (NLLS)	.041	-.521	-.519
	(.068)	(11.5)	(15.4)
4. LES (SAS)	.858	-.776	-.732
	(1.73)	(20.5)	(23.6)
5. Multivariate DLOG	-.352	-4.20	-4.92
	(1.55)	(6.34)	(9.49)

6.	AIDS (No Dummies)	-12.8 (.435)	-.408 (2.01)	-.444 (2.38)
7.	AIDS - Weighted (With Dummies)	2.679 (2.27)	-1.51 (48.5)	-1.48 (42.7)

* Absolute t-ratios are in parentheses

From the results of Table 7.1, the intercepts are always insignificant except for the weighted AIDS model. The slope coefficients are, on the other hand, always significant. Generally speaking, there is a distinct tendency for the points to scatter around a negatively sloped line coming from the origin (the minus signs of the slope coefficients are a result of the opposite signs of price and income elasticities), which supports the proportionality hypothesis that luxuries are more price elastic than necessities. The result is consistent with those of Clements, Kappelle and Roberts(1985).

From the basic food demand analysis, one may further investigate nutrition. Indonesia is generally thought to have a malnutrition problem among its population because of its poverty and lack of food education. In order to have a closer look at the seriousness of nutrient deficiencies in the country, one may follow the approach by Chernichnovsky and Meesook(1983). Basically, the food demand vector, Q, can be transformed linearly into a nutrient vector, N, using a matrix of the nutrient factor, F, where the element F_{ij} can be interpreted as the amount of nutrient i contained in one unit of food j. That is

$$N = F * Q$$

where N is (m*1) vector of nutrient types, F is (m*n) matrix of nutrient content in each kind of food, and Q is (n*1) vector of quantity demanded for food. The nutrient vector N should reflect the general nutrient condition of the population. As pointed out by Klumper(1985), the food problem in Indonesia is essentially one of distribution rather than of non-availability. With the help of this kind of exercise, policy makers can understand not only the nutrient consumption level on average, but also are able to identify those population groups with nutrient deficiencies. Accordingly, effective intervention policies may be formulated.

There are other areas in this study worthy of further examination. These are urban/rural and rich/poor differentials. As illustrated by Klumper(1985), the poorer households in Indonesia are much more sensitive to changes in prices or income than other population groups. Therefore, when formulating policies in regard to food consumption or malnutrition, it is important to analyse not just the average impact of a change in government policy, but also its effect on particular individual socio-economic groups. In the current study, we incorporated an urban dummy variable into the weighted model which proved to be statistically significant. But this simple method only changes the intercept term of the equations and has no direct effect on the slope coefficients. Perhaps unpooling the data and estimating the subset samples separately may give a better picture. Alternatively, interacting the dummy with the price and income variables can reveal the impact of the subject dummy variable.

As a final brief conclusion, it is clear from this study that there is a well developed economic and statistical theory underlying the estimation of systems of demand functions. However, many empirical problems remained unresolved. In this study, we significantly rejected the demand restrictions, but we are not prepared, at this stage, to conclude that the theory is invalid or inconsistent. We need, at very least, an adequately or realistically specified model before we could make such a firm conclusion.

Footnotes

1. As explained by Pollak and Wales(1978), a complete system of demand equations is said to be "theoretically plausible" if it is derived from a "well-behavioured" utility function, or equivalently, if the demand equations are homogeneous of degree zero in prices and total expenditure, and the implied Slutsky matrix is symmetric and negative semi-definite.

2. Zero responses can result from false reporting by either the respondent or the enumerator when conducting the survey. They may also arise if purchases are made infrequently so that no purchase is recorded for some households over the limited period of the survey. This is especially the situation for durable goods.

Appendix 1

Description of the Data types

Type 10 Household Identifier

1	10	type identifier
2		urban/rural area
3		subround
4		province
5		Kabupaten
6		code number of SUSENAS sample
7		code number of 1980 SP sample
8		household serial number
9		total number of household members
10		total number of household members age 10 and older
11		social and economic classification
12		household income
13		household expenses
14		protein per household
15		calories per household

Type 21 Food Purchases and Consumption

1	21	value & amount of food consumption during last week
2		code of item (refer to appendix 4)
3		amount purchased
4		value of amount purchased
5		amount of own production
6		value of own production
7		amount received as gift
8		value of amount received as gift
9		amount of total consumption
10		value of total consumption
11		protein per items per day
12		calories per item per day

Appendix 2

Provinces in Indonesia, by Island

Island	Province	Code*
Sumatra	Daerah Istimewa Aceh	11
	Sumatera Utara	12
	Sumatera Barat	13
	Riau	14
	Jambi	15
	Sumatera Selatan	16
	Bengkulu	17
	Lampung	18
Java	D.K.I. Jakarta	31
	Jawa Barat	32
	Jawa Tengah	33
	D.I. Yogyakarta	34
	Jawa Timur	35
Nusa Tenggara	Bali	51
	Nusatenggara Barat	52
	Nusatenggara Timur	53
	Timor Timur	54
Kalimantan	Kalimantan Barat	61
	Kalimantan Tengah	62
	Kalimantan Selatan	63
	Kalimantan Timur	64
Sulawesi	Sulawesi Utara	71
	Sulawesi Tengah	72
	Sulawesi Selantan	73
	Sulawesi Tenggara	74
Maluku & Irian Jaya	Maluku	81
	Irian Jaya	82

* SUSENAS code

Appendix 3

Map of Indonesia, by Province

Keterangan/Note:

1. Daerah Istimewa Aceh
2. Sumatera Utara
3. Sumatera Barat
4. Riau
5. Jambi
6. Sumatera Selatan
7. Bengkulu
8. Lampung
9. DKI Jakarta
10. Jawa Barat
11. Jawa Tengah
12. Daerah Istimewa Yogyakarta
13. Jawa Timur
14. Bali
15. Nusa Tenggara Barat
16. Nusa Tenggara Timur
17. Timor Timur
18. Kalimantan Barat
19. Kalimantan Tengah
20. Kalimantan Selatan
21. Kalimantan Timur
22. Sulawesi Utara
23. Sulawesi Tengah
24. Sulawesi Selatan
25. Sulawesi Tenggara
26. Maluku
27. Irian Jaya

Appendix 4

Classification of Food Commodities (with SUSENAS code)

Group 1	Rice and Grain
	001 Imported rice
	002 Top-quality rice
	003 Local rice
	004 Glutinous rice
	005 Rice byproducts
	006 Fresh corn in husk
	007 Dried corn in husk
	008 Shelled corn
	009 Corn meal
	010 Wheat flour
Group 2	Tubers
	013 Cassava
	014 Dried cassava
	015 Tapioca
	016 Cassava flour
	017 Sweet potatoes
	018 Potatoes
	019 Taro
	020 Sago
Group 3	Fish
	023 Milkfish
	024 Yellowtail
	025 Tuna
	026 Skipjack
	027 Selar
	028 Anchovies
	030 Pike
	031 Mujair
	032 Carp
Group 4	Dried Fish
	034 Salted fish
	035 Anchovies
	036 Selar
	037 Shrimp
	038 Squid
	040 Canned fish

041 Canned shrimp
042 Canned squid
043 Canned crabs

Group 5 Meat

046 Beef
047 Carabao meat
048 Mutton
049 Pork
050 Horse meat
051 Dried seasoned meat
052 Corned beef
053 Fried shredded meat
054 Liver
055 Entrails
056 Bones

Group 6 Chicken

058 Chicken

Group 7 Eggs

061 Chicken eggs
062 Duck eggs
063 Salted eggs

Group 8 Milk

065 Fresh milk
066 Milk from dairy
067 Evaporated milk
068 Powdered milk in cans
069 Powdered milk in bulk

Group 9 Vegetables

072 Spinach
073 Kangkung spinach
074 Cabbage
075 Mustard greens
076 Peas
077 String beans
078 Stew tomatoes
079 Carrots

080 Cucumbers
081 Cassava leaves
082 Egg plant
083 Bean sprouts
084 Squash
085 Radish
086 Soup greens
087 Pickled vegetable
088 Young jackfruit
089 Young pawpaws
090 Tree beans
091 Fruit
092 Shallots
093 Garlic
094 Red pepper
095 Cayenne pepper
096 Green chilli
097 Canned vegetables

Group 10 Legumes
100 Peanuts
101 Green beans
102 Red beans
103 Soybeans
104 Black-eyed peas
105 Bean curd
106 Soybean cake
107 Salted soybeans
108 Peanut cake
109 Soybean flour
110 Guava

Group 11 Fruits
113 Citrus fruits
114 Mangos
115 Apples
116 Avocados
117 Rambutan
118 Dukuh
119 Durian
120 Salak

121 Pineapples
122 Ambon bananas
123 Raja bananas
125 Papayas
126 Sapodilla plums
127 Belimbing
128 Spanish plums
129 Watermelons
130 Jackfruit
131 Sweet tomatoes
132 Canned fruit

Group 12 Condiments
135 Salt
136 Candlenuts
137 Coriander and caraway seeds
138 White and black pepper
139 Tamarind
140 Nutmeg
141 Cloves
142 Vinegar
143 Fish paste
144 Soy sauce
145 Coconuts

Group 13 Cooking Oil
146 Coconut oil
147 Cooking oil
148 Corn oil
149 Butter

Group 14 Additives
150 Mixed spices
151 Granulated sugar
152 Refined sugar
153 Tea
154 Coffee
155 Cocoa
156 Fried shrimp chips
157 Fried vegetable chips
158 Wheat noodles

159 Rice noodles
160 Macaroni
161 Condiments and additives
(sasa, miwon brands of monosodium glutamate,etc)
162 Cordial

Group 15 Prepared Food

165 Iced cordial
166 Lemonade and other canned beverages
167 Biscuits
168 Unsweetened bread
169 Porridge of green beans
170 Special vegetables
171 Rice and side dishes

Group 16 Alcohol

174 Beer
175 Wine

Group 17 Tobacco

178 Commodore cigarettes
179 Kansas cigarettes
181 Gudung garam clove cigarettes
182 Bentoel clove cigarettes
184 Cigars
185 Tobacco
186 Betel

Appendix 5

An Illustration of Non-Parametric Bootstrapping with Lyapunov Solutions using Shazam Version 5 Software

```
* this program is to construct non-parametric boostrap
* total number of commodities is 17
* total number of observations is 424

par 900

* read data from unit 14 and 15
* shares.dat contains the dependent variables, budget shares
* inc.dat contains the independent variables, prices and real income

file 14 shares.dat
file 15 inc.dat
smpl 1 424
dim y 424 16 x 424 19
read(14) y / cols=16
read(15) x / cols=18

* generate the constant term

genr x:19=1

* budget shares in y, prices, income and constant in x

dim bhat 19 16 u 424 16 uuhat 16 16

* m-1 equations system, find unrestricted ols solutions

mat bhat=inv(x'*x)*(x'*y)
mat uuhat=((y'*y)-(x'*y)'*bhat)/424

* this gives the unconstrainted residuals var-covar matrix
* now set up the constrainted estimation with homogeneity and symmetry
* imposed
* first seperate nuisance variables, income and constant, place in z

dim z 424 2 one 16
mat one=diag(iden(16))
genr z:1=x:18
```

```
genr z:2=x:19

* imposed homogeneity by substitution

dim xs 424 16 ys 424 16 x2 424 16
copy x xs / fcol=1;16
mat xs=xs-(x:17*one')
copy xs x2
mat ys=y

* concentrate out the nuisance variables and to solve the Lyapunov
* equations using characteristic equations

mat xs=xs-z*(inv(z'*z)*(z'*xs))
mat ys=ys-z*(inv(z'*z)*(z'*y))
mat c=inv(uuhat)*(xs'*ys)'+(xs'*ys)*inv(uuhat)
mat v=eigvec(uuhat)
mat r=eigval(uuhat)
mat p=v*(diag(sqrt(r)))
mat pap=p*((xs'*xs)*p)
mat q=eigvec(pap)
mat qval=eigval(pap)
mat w=q'*(inv(p))
mat rhs=inv(w')*c*(inv(w))
mat g=rhs/(qval*one'+one*qval')
mat btit=inv(w)*g*(inv(w'))
mat gam=(inv(z'*z))*(z'*Y-z'*x2*btit)
mat us=y-xs*btit-z*gam
mat uutit=(ys'*ys+btit'*xs'*ys-ys'*xs*btit+btit'*(xs'*xs)*btit)/424

* btit and gam are the constrainted estimates
* uutit is the constrainted residuals covariance matrix
* now calculate the test statistics

mat wald=424*(trace(inv(uuhat)*(uutit-uuhat)))
mat lr=424*log(det(uutit)/det(uuhat))
mat lm=424*(trace(inv(uutit)*(uutit-uuhat)))

* to set up the bootstrap experiment for 100 repititions

dim tv 1 3 tv2 100 3  newu 424 16 abhat 19 16 abtit 16 16 agam 2 16
mat tv2=tv2'
set nodoecho

do $=1,100
```

```
mat newu=samp(us,424)

* to inflate the residuals

mat newu=newu*sqrt(424/(424-19))
mat y=x2*btit+z*gam+newu

* repeat the unconstrainted and constrainted estimations desribed above
.
.
.
mat abhat=abhat+bhat
mat abtit=abtit+btit
mat agam=agam+gam
mat tv:1=wald
mat tv:2=lr
mat tv:3=lm
mat tv=tv'
mat tv2:$=tv
mat tv=tv'

endo

mat abhat=abhat/100
mat abtit=abtit/100
mat agam=agam/100
smpl 1 100
genr waldt=tv2:1
genr lrt=tv2:2
genr lmt=tv2:3

* abhat, abtit and agam are the averaged unconstrainted and constrainted
* estimates
* waldt, lrt and lmt are the simulated test statistics in each repetition

stop
```

BIBLIOGRAPHY

Aitcheson J. & Silvey S.D. (1958) " Maximum Likelihood Estimation of Parameters Subjected to Restraints", ***Annuals of Mathematical Statistics***, p.813-828

Amemiya T. (1983) "Non-Linear Regression Models", ***Handbook of Econometrics***, Vol. I, edited by Z. Griliches & M.D. Intriligator, North-Holland Publishing Company

Anderson G.J. & Mizon G.E. (1983) "Parameter Constancy Tests : Old and New", Discussion paper in Economics and Econometrics no. 8325, University of Southhampton

Barnes R. & Gillingham R. (1984) "Demographic Effects in Demand Analysis: Estimation of the QES using Micro Data", ***Review of Economics and Statistics***, p.591-601

Barnett W.A. (1983) "Definitions of 'Second Order Approximation' and of 'Flexible Functional Form'", ***Economics Letters***, p.31-35

Barten A.P. (1964) "Family Composition, Prices and Expenditure Patterns" in ***Economic Analysis for National Economic Planning***, edited by P.E. Hart, G. Mills & J.K. Whitaker, London: Butterworth

------------------ (1966a) "Theorie en empirie van een volledig stelsel van viaigvergelijkingen" (Dutch), Doctoral dissertation, Rotterdam

------------------ (1966b) "Het verbruik doer gezinshuishoudingen in Netherland 1921-1939 en 1948-1962", Netherlands School of Economics, Report 6604

------------------ (1968) "Estimating Demand Equations", ***Econometrica***, p.213-251

------------------ (1977) "The Systems of Consumer Demand Functions Approach : A Review", ***Frontiers of Quantitative Economics***, Vol A, edited by M.D. Intriligator, N. Holland

Barten A.P. & Bohm V. (1982) "Consumer Theory", ***Handbook of Mathematical Economics***, Vol II, edited by K.J. Arrow and M.D. Intriligator, North Holland

Barten A.P. & Turnovsky S.J. (1966) "Some Aspects of the Aggregation Problem for Complete Demand Equations", ***International Economic Review***, p.231-257

Beggs J.J. (1987) "Diagnostic Testing in Econometrics", invited address to the Agricultural Economics Society Meetings, Feb.,1987. Working paper in Economics and Econometrics series, ANU

Bera A.K. (1982) "A Note on Testing Demand Homogeneity", ***Journal of Econometrics***, p.291-294

Bera A.K. & Byron R.P. (1983) "A Note on the Effects of Linear Approximation on Hypothesis Testing", ***Economics Letters***, p.251-254

Bera A.K., Byron R.P. & Jarque C.M. (1981) "Further Evidence on Asytotic Tests for Homogeneity and Symmetry in Large Demand System", ***Economics Letters***, p.101-105

Bera A.K. & Jarque C.M. (1981) "Efficient Tests for Normality, Homoscedasticity and Serial Regression Residuals : Monte Carlo Evidence", ***Economics Letters***, p.313-318

------------------ (1982) "Model Specification Tests: A Simultaneous Approach", ***Journal of Econometrics***, p.59-82

Berndt E.R. & Savin N.E. (1977) "Conflict Among Criteria for Testing Hypotheses in the Multivariate Linear Regression Model", ***Econometrica***, p.1263-1277

Bewley R.A. (1983) "Test of Restrictions in Large Demand Systems", ***European Economic Review***, p.257-269

----------------- (1986) ***Allocation Models: Specification, Estimation and Applications***, Ballinger Publishing Company, Cambridge, Mass.

Blanciforti L. & Green R. (1983) "An AIDS Incorporating Habits, an Analysis of Expenditure on Food and Aggregate Commodity Groups", ***Review of Economics and***

Statistics, p.511-515

Bohm B., Reider B. & Tintner G. (1980) " A System of Demand Equations for Austria", ***Empirical Economics***, p.129-142

Brown A. & Deaton A. (1972) "Surveys in Applied Economics: Models of Consumer Behaviour", the ***Economic Journal***, p.1145-1236

Breusch T.S. & Pagan A.R. (1979) "A Simple Test for Heteroscedasticity and Random Coefficient Variation", ***Econometrica***, p.1287-1294

Byron R.P. (1968) "Methods for Estimating Demand Equations Using Prior Information : A Series of Experiments with Australian Data", ***Australian Economic Paper***, p.227-248

-------------- (1970) "The Restricted Aitken Estimation of Sets of Demand Relations", ***Econometrica***, p.816-830

-------------- (1982) "Notes and Comments : A Note on the Estimation of Symmetric System", ***Econometrica***, p.1573-1575

-------------- (1983) "A Compact Version of the SUSENAS Household Expenditure Survey(1981) for Java", Department of Statistics, ANU, mimeo

-------------- (1984) "On the Flexibility of the Rotterdam Model", ***European Economic Review***, p.273-283

------------- (1985) "An Aggregate Model of Food Consumption in Java", ANU, mimeo

Byron R.P. & Bera A.K. (1983) "Least Squares Approximations to Unknown Regression Functions: A Comment", ***International Economic Review***, p.255-260

Byron R.P. & Rosalsky M.C. (1985) "Hypothesis Testing in Demand Systems : Some Examples of Size Corrections Using Egdeworth Approximations", ***Econometric Theory***, p.403-408

Chernichovsky D. & Meesook O.A. (1983) "Patterns of Food Consumption and Nutrition in Indonesia", Discussion paper, the World Bank

Chow G.C. (1960) "Test of Equality Between Sets of Coefficients in Two Linear Regressions", ***Econometrica***, p.591-605

------------- (1983) ***Econometrics***, McGraw-Hill

Christensen L.R., Jorgensen D.W. & Lau L.J. (1973) "Transcendental Logarithmic Production Frontiers", ***Review of Economics and Statistics***, p.28-45

-------------------- (1975) "Transcental Logarithmic Utility Functions", ***American Economic Review***, p.367-383

Clements K.W., Kappelle S. & Roberts E.J.(1985) "Are Luxuries More Price Elastic Than Necessities", Discussion paper no. 85-05, Department of Economics, University of Western Australia

Davidson R. & MacKinnon J.G. (1980) "Some Non-Nested Hypothesis Tests and the Relations among them", Discussion paper no. 409, Queen's Institute for Economic Research, Queen's University, Kingston, Ontario, Canada

Deaton A. (1972) "The Estimation and Testing of Systems of Demand Equations : A Note", ***European Economic Review***, p.399-411

------------ (1974) " The Analysis of Consumer Demand in the United Kingdom, 1900-1970", ***Econometrica***, p.341-367

------------- (1978) "Specification and Testing in Applied Demand Analysis", ***Economic Journal***, p.524-536

------------- (1986a) "Demand Analysis", ***Handbook of Econometrics***, vol III. Edited by Z. Griliches and M.D. Intrilgator, p.1767-1839

------------ (1980b) "Quality, Quantity and Spatial Variation inPrice", Discussion paper, Living Standards Unit, the World Bank, Research Program in Development Studies, Princeton University

Deaton A. & Irish M. (1984) "Statistical Models for Zero Expenditures in Household Budgets", ***Journal of Public Economics***, p.59-80

Deaton A. & Muellbauer J. (1980a) ***Economics and Consumer Behaviour***, Cambridge University Press

--------------------------- (1980b) "An Almost Ideal Demand System", ***American Economic Review***, p.312-326

Denny M. & Fuss M. (1977) "The Use of Approximation Analysis to Test for Separability and the Existence of Consistent Aggregates", ***American Economic Review***, p.404-418

Dhrymes P.J. (1978) ***Introductory Econometrics***, Springer-Verlag

Diewert W.E. (1971) "An Application of the Shepard Duality Theroem: A Generalized Leontief Production Function", ***Journal of Political Economy***, p.481-507

Edgeworth F.Y. (1881) ***Mathematical Psychics***, London: Kegan Paul

Efon B. & Gong G. (1983) " A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation", the ***Amercian Statistican***, p.36-48

Efon B. & Tibshiraui R. (1986) "Bootstrap Methods for Standard Errors, Confidence Intervals and other Measures of Statistical Accuracy", ***Statistical Science***, p.54-77

Engle R.F. (1984) "Wald, Likelihood Ratio, and Lagrange Multiplier Tests in Econometrics", ***Handbook of Econometrics***, Vol 2, Edited by Z. Griliches and M.D. Intriligator, p.775-826

Freedman D.A. & Peters S.C. (1984) "Bootstrapping a Regression Equation: Some Empirical Results", ***Journal of the American Statistical Association***, p. 97-106

Frisch R. (1959) "A Complete Scheme for Computing All Direct & Cross Demand Elasticities in a Model with Many Sectors", ***Econometrica***, p.177-196

Gallant A.R. (1975) "Seemingly Unrelated Non-Linear Regressions", ***Journal of Econometrics***, p.35-50

------------------ (1981) "On the Bias in Flexible Functional Forms and An Essentially Unbiased Form : The Fourier Flexible Form", ***Journal of Econometrics***, p.211-245

Gallant A.R. & Jorgenson D.W. (1979) "Statistical Inference for a System of Simultaneous, Non-Linear, Implicit Equations in the Context of Instrumental Variable Estimation", ***Journal of Econometrics***, p.275-302

Gibbons J.D. (1971) ***Nonparametric Statistical Inference***, 2nd edition, Revised and Expanded, Marcel Dekker Inc.

Gibbons J.D. (1976) ***Nonparametric Methods for Quantitative Analysis***, International Series in Decision Processes

Gnanadesikan R. (1977) ***Methods for Statistical Data Analysis of Multivariate Observations***, John Wiley & Sons

Godfrey L.G., McAleer M. & McKenzie C.R. (1986) "Variable Addition and Lagrange Multiplier Tests for Linear and Logarithmic Regression Models: Theory and Monte Carlo Evidence", Australian National University working paper

------------------ (1985) ***Nonparameteric Statistical Inference***, 2nd edition, Revised and Expanded, Marcel Debber Inc.

Goldfeld S.M. & Quandt R.E. (1972) ***Non-Linear Methods in Econometrics***, North Holland

Gorman W.M. (1959) "Separable Utility and Aggregation", ***Econometrica***, p.469-481

------------------ (1976) "Tricks with Utility Functions", ***Essays in Economic Analysis***, Edited by Artis M. and Nobay R., Cambridge University Press

Greene W.H. (1981) "On the Asymptotic Bias of the OLS estimator of the Tobit Model", ***Econometrica***, p.505-513

Guilkey D.K. & Lovell C.A. (1978) "On the Flexibility of the Translog Approximation", Department of Economics, University of Northern Carolina, working paper series

Harvey A.C. (1981) ***The Econometric Analysis of Time Series***, Philip Allan

Hausman J.A. (1978) "Specification Tests in Econometrics", ***Econometrica***, p.1251-1271

Hebden J. (1983) ***Applications of Econometrics***, Philip Allan

Hendry D.F. (1984) "Monte Carlo Experimentation", ***Handbook of Econometrics***, Vol II, edited by Z. Griliches & M.D. Intriligator, N. Holland

Hicks J.R. (1936) ***Value and Capital***, London: Oxford University Press

------------- (1946) ***Value and Capital***, 2nd Edition, Oxford Clarendon Press

------------- (1956) ***A Review of Demand Theory***, Oxford University Press, Oxford

Houthakker H.S. (1960) "Additive Preferences", ***Econometrica***, p.244-257

Houthakker H.S. & Taylor L.D. (1970) ***Consumer Demand in the United States: Analysis and Projections***, 2nd and enlarged edition, Harvard University Press, Cambridge

Jarque C.M. & Bera A.K. (1980) "Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals", ***Economics Letters***, p.255-259

Jarque C.M. & McKenzie C.R. (1983) "Testing for Multivariate Normality in Simultaneous Equations Models", Australian National University working paper no. 82

Jennrich R.J. (1969) "Asymptotic Properties of Non-Linear Least Squares Estimators", ***Annuals of Mathematical Statistics***, p.633-643

Johnston J. (1984) ***Econometric Methods***, 3rd Edition, Internatinal Student Edition

Jorgenson D.W., Slesnick D.T. & Stoker T.M. (1983) "Exact Aggregation over Individuals and Commodities", Discussion paper 1005, Harvard Institute for Economic Research, Harvard University, Cambridge, Massachusetts

Judge G.G., Griffiths W.E., Hill R.C. & Lee T.C. (1980) ***The Theory and Practice of Econometrics***, John Wiley & Sons

Judge G.G., Hill R.C., Griffiths W.E., Lutkepohl H. & Lee T.C. (1982) ***Introduction to the Theory and Practice of Econometrics***, John Wiley & Sons

Katzner D.W. (1970) ***Static Demand Theory***, Macmillan

Kelejian H.H. (1982) "An Extension of a Standard Test for Heteroscedasticity to a System Framework", ***Journal of Econometrics***, p.325-333

Kiefer N.M. (1982) "Microeconometric Evidence on the Neoclassical Model of Demand", working paper no. 282, Department of Economics, Cornell University

Kiefer N.M. & MacKinnon J.G. (1976) "Small Sample Properties of Demand System Estimates",***Studies in Nonlinear Estimation***, edited by S.M. Goldfeld & R.E. Quandt, Cambridge

King G. (1696) "Natural and Political Observations and Conclusions Upon the State and Condition of England" in Two Tracts by G. King, edited by G.E. Barnett, Johns Hopkins, Baltimore Press 1936

Klevmarken N.A. (1981) ***On the Complete Systems Approach to Demand Analysis***, The Industrial Institute for Economic and Social Research, Almqvist and Wiksell Internal, Stockholm, Sweden

Klein L.R. & Rubin H. (1947) "A Constant Utility Index of the Cost of Living", ***Review of Economic Studies***, p.84-87.

Klumper S.A. (1985) "Food Strategies in Indonesia: A Contribution to the Problem of Optimal Factor Allocation in Agriculture", Research Proposal, Research Centre for International Agrarian Development

Kmenta J. (1971) ***Elements of Econometrics***, Macmillan Company

Kmenta J. & Gilbert R.F. (1968) "Small Sample Properties of Alternative Estimates of Seemingly Unrelated Regressions", ***Journal of the American Statistical Association***, p.1180-1200

Knudsen O.K. & Scandizzo P.L. (1982) "The Demand for Calories in Developing Countries", ***American Journal of***

Agricultural Economics, p.80-86

Laitinen K. (1978) "Why is Demand Homogeneity so often Rejected", ***Economics Letters***, p.187-191

Leontief W. (1947) "Introduction to a Theory of the Internal Structure of Functional Relationships", ***Econometrica***, p.361-373

Malinvaud E. (1980) ***Statistical Methods of Econometrics***, Third revised edition, North-Holland

MacKinnon J.G. (1983) "Model Specification Tests Against Non-Nested Alternatives", ***Econometric Reviews***, p.85-110

Maddala G.S. (1977) ***Econometrics***, McGraw-Hill, New York

----------------- (1983) ***Limited Dependent and Qualitative Variables in Econometrics***, Cambridge University Press

Mardia K.V. (1970) "Measures of Multivariate Skewness and Kurtosis with Applications", ***Biometrika***, p.519-530

----------------- (1980) "Tests of Univariate and Multivariate Normality", p.279-320, ***Handbook of Statistics Vol I***, edited by P.R. Krishnaiah, North Holland

Marshall A. (1890) ***Principles of Economics***, Macmillan, London

Martin W. & Porter D. (1985) "Testing for Changes in the Structure of the Demand for Meat in Australia", ***Australian Journal of Agricultural Economics***, p.16-31

Mattei A. (1986) "Consumer Behaviour and Commodity Aggregation, A Simulation Study", Discussion paper, University of Lausanne, Switzerland

McAleer M. & Deistler M. (1986) "Some Recent Developments in Econometrics", ANU discussion paper, forthcoming in Bittanti S. (ed) ***Stochastic Models and Linear Systems***, Springer-Verlag

Meisner J.F. (1979) "The Sad Fate of the Asymptotic Slutsky Symmetry Test for Large System",***Economics Letters***, p.231-233

Muellbauer J. (1975) "Aggregation, Income Distribution and Consumer

Demand", ***Review of Economic Studies***, p,525-543

------------------ (1976) "Community Preference and the Representative Consumer", ***Econometrica***, p.979-999

Murray J. (1984) "Retail Demand for Meat in Australia: A Utility Theory Approach", ***Economic Record***, p.45-66

Olsen R.J. (1980) "Approximating a Truncated Normal Regression with the Method of Moments", ***Econometrica***, p.1099-1105

Pagan A.R. (1983) "Model Evaluation by Variable Addition", ***Econometrics and Quantitative Economics***, Edited by D.F. Hendry & K.F. Wallis, Basil Blackwell

Pagan A.R. & Hall A.D. (1983) "Diagnostic Tests as Residual Analysis", ***Econometric Reviews***, p.159-218

Park R.W. (1969) "Systems of Demand Equations : A Empirical Comparsion of Alternative Functional Forms", ***Econometrica***, p.629-650

Pearson K. & Lee A. (1908) "Generalized Probable Error in Multiple Normal Correlation", ***Biometrika***, p.59-68

Pesaran M.H. (1974) "On the General Problem of Model Selection", ***Review of Economic Studies***, p.153-171

Pesaran M.H. & Deaton A. (1978) "Testing Non-Nested Non-Linear Regression Models", ***Econometrica***, p.677-694

Phlips L. (1974) ***Applied Consumption Analysis***, North Holland

Pollak R.A. & Wales T.J. (1969) "Estimation of the Linear Expenditure System", ***Econometrica***, p.611-628

------------------ (1978) "Estimation of Complete Demand System from Household Budget Data: the Linear and Quadratic Expenditure System", ***American Economic Review***, p.348-359

------------------ (1980) "Comparison of the QES and Translog Demand System with an Alternative Specification of Demographic Effects", ***Econometrica***, p.595-612

------------------- (1981) "Demographic Variables in Demand Analysis", ***Econometrica***, p.1533-1551

Powell A.A. (1974) ***Empirical Analytics of Demand Systems***, Lexington Books

Ramsey J.B. (1969) "Test for Specification Errors in Classical Linear Least Squares Regression Analysis", ***Journal of the Royal Statistical Society B***, p.350-371

Ramsey J.B. & Schimdt P. (1976) "Some Further Results on the Use of OLS and BLUS Residuals in Specification Error Tests", ***Journal of the Amercian Statistics Association***, p.389-390

Rao L.R. (1948) "Large Sample Tests of Statistical Hypotheses Concerning Several Parameters with Application to Problems of Estimation", ***Proceeding of the Cambridge Philosophical Society***, p.50-57

Ray R. (1980) "Analysis of a Time Series of Household Expenditure Surveys for India", ***Review of Economics and Statistics***, p.595-602

Ryu J.C. (1986) "Static and Dynamic Aspects of the Demand for Goods in Korea", Ph.D. Dissertation, Department of Statistics, ANU

SAS/ETS User Guide, Version 5 Edition (1984) SAS Institute Inc. NC, USA

Savin N.E. (1976) "Conflicts Among Testing Procedures in a Linear Regression Model with Autoregressive Disturbances", ***Econometrica***, p.1303-1313

SHAZAM, The Econometrics and Computer Program, Version 5.1, User's Reference Mannual 1986, by K.J. White & N.G. Horsman

Silvey S.D. (1959) :The Langrange Multiplier Test", ***Annuals of Mathematical Statistics***, p.389-407

Simmons P. & Weiserbs D. (1979) "Translog Flexible Functional Forms Associated Demand System", ***American Economic Review***, p.404-418 lutsky E.E. (1915) "On the Theory of the Budget of the Consumer", Giornale degli economisti p.1-26. English translation in G. Stigler & K. Boulding (eds) ***Reading in Price Theory***, p.27-56, Homewood: Irwin, 1952

ono M. (1961) "The Effect of Price Changes on the Demand and Supply of Separable goods", ***International Economic Review***, p.239-271

tapleton D.C. (1984) "Errors-in-Variables in Demand Systems", ***Journal of Econometrics***, p.255-270

tone R. (1953) ***The Measurement of Consumers' Expenditure and Behaviour in the United Kingdom 1920-1938***, Vol I, Cambridge

---------- (1954) "Linear Expenditure System and Demand Analysis: An Application to the Patterns of British Demand", ***Economic Journal***, p.511-527.

trauss J. (1982) "Determinants of Food Consumption in Rural Sierro Leone", ***Journal of Development Economics***, p.327-353

-------------- (1984) "An Overview of Agricultural Household Model: Theory", Yale University, Economic Growth Centre, discussion paper no. 450

trotz R.H. (1957) "The Empirical Implications of a Utility Tree", ***Econometrica***, p.269-280

-------------- (1959) "The Utility Tree - A Correction and Further Appraisal", ***Econometrica***, p.482-488

heil H. (1965) The Information Approach to Demand Analysis, ***Econometrica***, p.67-87

------------- (1971) ***Principles of Econometrics*** , John Wiley & Sons

------------ (1975) ***Theory and Measurement of Consumer Demand***, Vol I, Amsterdam, North Holland

------------ (1976) ***Theory and Measurement of Consumer Demand***, Vol II, Amsterdam, North Holland

-------------- (1978) ***Introduction to Econometrics***, Prentice-Hall

Theil H. & Clements K.W. (1987) ***Applied Demand Analysis: Results from System-Wide Approach***, forthcoming, Ballinger Publishing Co., Cambridge, Mass.

Van de Walle D. (1986) "Welfare Analysis of Tax Reforms Using SUSENAS Data", Discussion paper, Department of Economics, Research School of Pacific Studies, Australian National University

Wald A. (1943) "Tests of Statistical Hypotheses Concerning Several Parameters when the Number of Observations is large", ***Transactions of the Amercian Mathematical Society***, p.426-482

Wales T.J. (1977) "On the Flexiblity of Flexible Functional Forms - An Empirical Approach", ***Journal of Econometrics***, p.183-193

Wales T.J. & Woodland A.D. (1983) "Estimation of Consumer Demand Systems with Binding Non-Negativity Constraints", ***Journal of Econometrics***, p.263-285

Walras L. (1874) Elements d'economie politique pure, Lausanne: Carbay; English Translations, ***Elements of Pure Economics***, George Allen & Unwin, London 1954

White H. (1980a) "Non-Linear Regression on Cross-Section Data", ***Econometrica***, p.721-746

-------------- (1980b) "Using Least Squares to Approximate Unknown Regression Functions", ***International Economic Review***, p.149-170

-------------- (1980c) "A Heteroscedasticity Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity", ***Econometrica***, p.817-838

White H. & MacDonald G.M. (1980) "Some Large Sample Tests for Non-normality in the Linear Regression Model", ***Journal of the Amercian Statistical Association***, p.16-31

Wilks S.S. (1938) " The Large Sample Distribution of the Likelihood Ratio for Testing Composite Hypothesis", ***Annuals of***

Mathematical Statistics, p.60-62

Yoshihara K. (1969) "Demand Functins: An Application to the Japanese Expenditure Pattern", ***Econometrica***, p.257-274

Zellner A. (1962) "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias", ***Journal of the Amercian Statistical Association***, p.348-368