ECONOMIC GROWTH CENTER

YALE UNIVERSITY

Box 1987, Yale Station
New Haven, Connecticut

CENTER DISCUSSION PAPER NO. 339

FLEXIBLE CONSUMER DEMAND SYSTEM WITH LINEAR ESTIMATION EQUATIONS

FOOD DEMAND IN INDIA

Gurushri Swamy and Hans P. Binswanger*

March 1980

Notes: Gurushri Swamy is an Economist at the World Bank and Hans Binswanger is an Associate of the Agricultural Development Council. Work on this study was done while both authors were visitors at the Economic Growth Center, Yale University under a project financed by the Agency for International Development, AID/DSAN-C-0020. However, the views expressed in this paper do not necessarily reflect those of AID.

The authors are especially grateful to Sidnie Feit for her excellent programming and computational assistance.

ABSTRACT

This study estimates income and price elasticities for foodgrains in India, using a cross-section of time series data across 10 states and for 20 years. Previous attempts to estimate demand elasticities for India have been largely restricted to single equation estimates of income or expenditure elasticities based on a limited number of observations. The few studies which have attempted to estimate a complete system of equations generally deal with broad groups of commodities, using the Linear Expenditure System, which makes restrictive assumptions about the underlying utility function.[1] This study attempts a more disaggregated analysis; demand elasticities are estimated for Rice, Wheat (i.e. the so-called superior cereals), Pulses and the so-called "inferior cereals" which form a substantial part of the cereal consumption of low income groups. The elasticities are estimated from a complete set of demand equations with cross-equation constraints on the price and income terms, using three new flexible functional forms.

INTRODUCTION

Any system of demand equations must satisfy the following conditions of consumer demand theory.

I homogeneity of degree zero in income and prices

II symmetry of the compensated cross-price terms

III adding-up constraint i.e., the weighted sum of income elasticities = 1

In addition, the nature of the data we are using, and the level of disaggregation we are attempting, make it desirable to have functional forms

---

1/ See for example, the work of R. Radhakrishna and K. N. Murthy (1978) which are simple to estimate and easy to interpret.

The data consists of time series covering 10 states in India and twenty years, and 4 foodgrain groups. We, therefore, need to deal with the problem of pooling cross-section and time-series data. Further, not only are these commodity groups close substitutes for each other, but one of them, i.e., "inferior" cereals may have a negative income elasticity. We also need functional forms which allow for decreasing or increasing income elasticities (as compared to constant elasticities) since, with foodgrain consumption, it is unrealistic to assume that the elasticity will be constant. We therefore need functional forms which:

I are linear in parameters

II are "flexible" in the sense that income and price terms are not constrained to be zero or unitary

III allow for positive, negative, increasing or decreasing income elasticities

IV allow for estimation of cross-price elasticities with a group of close substitutes or complements and do not assume different types of additivity.

In deriving a functional form for the demand equations which satisfy both the conditions of demand theory and our special needs, we use the results of duality theory as applied to consumer demand theory. Their usefulness can be best appreciated with a brief review of the historical development of consumer theory. (See also Barten, 1977)

REVIEW OF CONSUMER DEMAND THEORY

The first attempts to measure demand elasticities were in terms of single equations, specified directly and in an intuitive way, to include prices and income as explanatory variables. Whenever the prices of substitutes or complements were considered important, they were also included. But it is impossible to incorporate any of the restrictions of demand theory mentioned above (with the exception of homogeneity) to single equations since they refer to a system of demand equations which describe the allocation of a consumer's budget over an exhaustive set of commodities.

Systems of demand equations have, therefore, been derived from specifying a utility function $U = f(X_1, X_2, ..X_N)$ and maximizing it subject to a budget constraint. To derive estimation equations, this implies inversion of the bordered Hessian matrix, which may be quite large and difficult to handle without any (often unrealistic) constraints on the utility function itself. Different types of separability of the utility function were used such as "block-independence" which means that the marginal utility $\partial U / \partial X_i$ is independent of X_j, where the i^{th} and j^{th} commodities belong to different groups of commodities. This allows the Hessian matrix to be block-diagonal. The assumption of "preference independence", which means that the marginal utility of the i^{th} commodity depends only on X_i allows the Hessian matrix to be diagonal.[2] The underlying assumptions of separability have also been necessary to reduce the number of parameters to be estimated, given a small body of data.

Diewert also points out that the form specified for the utility function needs to be rather simple, to be able to obtain algebraic expressions for the demand functions. If we assume that the utility function is of a "flexible" form, such solutions are normally impossible.

Three systems of demand equations currently in use, which are derived from specific forms of a utility function are (i) the Linear Expenditure System (Stone, 1954) (and its extension, the Extended Linear Expenditure System (Lluch, et. al. 1977), (ii) the Quadratic Expenditure System (Pollack and Wales, 1978) and (iii) the Indirect Addilog System (Leser, 1941, Houthakker, 1960). The detailed features of these systems are discussed in Appendix D. It is important to note here that they all assume an additive form for the utility function. They therefore allow little flexibility in the price coefficients and assume that all goods are net substitutes. Further, the Linear Expenditure System (and its extensions) do not allow for inferior goods. Therefore, these systems are more suited to the analysis of broad aggregate groups and are unsuitable for our purpose.

In recent years, the development of duality relationships (and their application to consumer theory) has made it possible to avoid many of the problems associated with the traditional approaches. In particular, "it enables us to derive systems of demand equations which are consistent with maximizing or minimizing behavior on the part of an economic agent, consumer or producer, simply by differentiating a function, as opposed to solving explicitly a constrained maximization or minimization problem (Diewert, 1974, p.106). Specifically, the duality theorems have established a one-to-one correspondence between the direct utility functions, expenditure functions, indirect utility functions and the system of derived demand equations.

If F(X) is a consumer's utility function, X a vector of commodities, then C(U,P) is the minimum cost of achieving the utility level U, given that the consumer faces the commodity prices P. Under conditions discussed below there is duality between the consumer's utility function F(X) and the function C which can be called the expenditure function. If M is the budget constraint, then G(M,P) is the maximum utility that the consumer can attain given P and M, and G is called the consumer's indirect utility function which is again dual to F and C.

The basic optimization problem of maximizing a utility function F(X) subject to the budget constraint $P.X \leq M$ can be written as the indirect utility function

$$G(V) = \max_{X} \{ F(X) : V.X \leq 1, \; X \geq 0_N \} \tag{1}$$

M is given and is positive, and the constraint $P.X \leq M$ can be replaced by $V = P/M$ and $V \cdot X \leq 1$

In Diewert's (1978) notation the duality relationships state that if the utility function F(X) satisfies the following set of conditions

A: I F is a real valued function of N variables, defined over the non-negative orthant and is continuous on the domain

II F is increasing in X

III F is a quasiconcave function

and IV F is a positive function for $X \gg 0_N$.

Then, the indirect utility function also satisfies the following set of conditions B.

B: I G is a real valued function of N variables defined over the set of positive normalized prices V, and G is continuous

II G is decreasing in V

III G is quasiconvex over V

IV G has a continuous extension to the non-negative orthant[3]

It is important to note not only that the direct utility function completely determines the indirect utility function, but that the utility function F can be recovered from G since F(X) for $X \gg 0$ can be written as

$$F^*(X) = \min_{V} \{ G(V): V \cdot X \leq 1, V \geq 0_N \} \quad (2)$$

and G* (V) can be defined as

$$G^*(V) = \max_{X} \{ F(X) : V \cdot X \leq 1, X \geq 0_N \} \quad (3)$$

and therefore $G^*(V) = G(V)$ for all $V > 0_N$

If G(V) satisfies the relevant set of conditions B, and in addition is differentiable at $V \gg 0_N$ with a non-zero gradient vector $\nabla G(V) < 0_N$, then X*, which is the unique solution to the direct maximization problem, can also be found by applying Roy's identity to the indirect utility function, i.e.

$$X^* \equiv \nabla G(V)/V \cdot \nabla G(V) \quad (4)$$

These demand equations are homogeneous of degree zero in V and also satisfy symmetry and adding up constraints. This approach to specifying demand functions is particularly useful when we want to start out with "flexible" functional forms for the indirect utility function. The flexible functional forms are flexible in the sense that they do not *a priori* constrain the various income and price elasticities at a base point, *and* they provide a local second order approximation to an arbitrary differentiable direct or indirect utility function.

The two flexible functional forms (for the indirect utility function) used in demand studies have been the generalized Leontief and the Translog. The former is due to Diewert (1974) and has the form

$$G(V) = \sum_{i=1}^{N} \sum_{j=1}^{N} b_{ij} V_i^{1/2} V_j^{1/2} + 2 \sum_{j=1}^{N} b_{oj} V_j^{1/2} + b_{00} \quad (5)$$

where $V_i = P_i/M$, and $b_{ij} = b_{ji}$. The latter is due to Lau and Mitchell (1970) and has the form

$$\log G(V) = \alpha_0 + \sum_{i=1}^{N} \log V_i + 1/2 \sum_{i=1}^{N} \sum_{j=1}^{N} \gamma_{ij} \log V_i \log V_j \quad (6)$$

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

Application of Roy's identity to these functional forms (and generally to flexible functional forms) gives derived demand equations which are non-linear in the unknown parameters. Since linearity is a desirable property for ease of estimation these functional forms are not very helpful in our context.

For specifying demand equations which are linear in the unknown parameters we again use the results of duality theory. It is important to note that the one-to-one correspondence established between a utility function and an indirect utility function holds here as well, i.e., a system of demand equations specified to satisfy the constraints of consumer demand theory is consistent with a direct utility function which satisfies the set of conditions A, and with an indirect utility function which satisfies the set of conditions B, and where the indirect utility function is differentiable.

In effect, therefore, the problem of estimating a system of demand equations consistent with the theory of utility maximization can also be approached by first specifying a functional form for the demand equations, which satisfies the conditions of symmetry, adding-up and of homogeneity of degree zero in prices and income.[4]

One recently developed linear system of demand equations is the Linear Logarithmic Expenditure System (Lau, et. al. 1978) which is derived by specifying a homogeneous translog indirect utility function, and applying Roy's identity to this indirect function. However, the assumption of homogeneity (of degree -1) while making the demand equations linear, is very restrictive because it implies that all income elasticities are unitary.[5]

Another widely used functional form which is linear in the parameter is the Rotterdam model, developed by H. Theil (1971, pp 330-333 and 574-580).

There are two models, the first of which is in terms of share weighted logarithms of real income and absolute prices. The model is a modification of the familiar double-logarithmic functional form in order to incorporate the symmetry constraint. Specifically, since the double-logarithmic form estimates elasticities directly, it is difficult to impose the symmetry constraints on the compensated price terms. By multiplying through by the value share of the commodity, the symmetry constraint can be directly imposed, as can be seen below.

$$S_i d \log x_i = a_i + b_{i1} S_i d \log m + \sum_{j \neq k} S_i C_{ij} d \log p_j \ldots\ldots\ldots \quad (7)$$

where S_i is the share of commodity i in total expenditures M, m is real expenditure, and d stands for discrete changes in the variables.

It can be seen that in this model symmetry can be imposed globally since:

$$\frac{\partial X_i}{\partial P_j} = C_{ij} \cdot \frac{M}{P_i P_j} = \frac{\partial X_j}{\partial P_i} = C_{ji} \frac{M}{P_j P_i} \tag{8}$$

Therefore symmetry holds as long as $C_{ij} = C_{ji}$.

But it appears that homogeneity can only be imposed at sample means since homogeneity in this case means that

$$\sum_{j=1}^{n} C_{ij} \frac{M}{P_j P_i} = 0.^{6}$$

Theil has also developed a second model in real income and relative prices which attempts to reduce the number of unknown parameters by assuming either block-independence or preference independence for the utility function. As mentioned earlier, such assumptions are too restrictive for us. For details of this model, see Appendix D.

"Almost Ideal Demand System" (AIDS)

Finally, we considered the AIDS which is due to Deaton and Muellbauer (1977), which came to our attention after we had derived our functional forms. The AIDS satisfies all the axioms of consumer theory and is therefore consistent with an (unknown) indirect utility function satisfying conditions B.

It has the following form:

$$S_i = a_i + b_{i1} \log (M/\tilde{P}) + \sum_{j=1}^{N} C_{ij} \log P_j \tag{9}$$

where S_i = share of commodity i in total expenditures, M is money income and $\tilde{P}$ is a price index used to deflate money expenditures. This model provides a local approximation to any arbitrary demand system,(Deaton and

Muellbauer also claim that it aggregates perfectly over consumers; (see Deaton & Muellbauer, 1977).

It should benoted that the Rotterdam Model, the AIDS as well as the forms developed below are specified in "real income" and nominal prices. This is necessary to get compensated price terms on which symmetry constraints can be impoased.

FLEXIBLE LINEAR DEMAND SYSTEMS IN NOMINAL PRICES AND REAL INCOME

From the Review of Consumer Theory we retain two ideas: the first is that duality between direct, indirect utility functions as well as systems of demand equations allows us to specify linear demand equations which are consistent with consumer demand theory and be sure that the (unknown) underlying direct and indirect utility functions also satisfy the requirements of consumer demand theory.

Second we retain the idea of both the Rotterdam and the AIDS of writing the demand equations in real income and nominal prices. This has as its effect that the coefficients of the (nominal) prices in the consumer demand equations reflect (income-) compensated price effects. This is because the income effects of any price changes observed in the data are already reflected in changes of the price deflator $\tilde{P}$, and therefore in a change in real income. This "purges" the price coefficients of the effect of the price changes on real income.

The key advantage of this formulation is that the symmetry constraint $\frac{\partial X_i^c}{\partial P_j} = \frac{\partial X_j^c}{\partial P_i}$ which relates to the compensated price effects, can directly be imposed on the coefficients of the price terms as linear constraints, which would be impossible if the coefficients reflected uncompensated price effects.

The key problem of the nominal price - real income approach is the choice of a correct price deflator. In principle, to estimate real income of a consumer one needs to know the parameters of his utility function, and thus needs to know the consumer demand system which one is trying to estimate even before estimating it, which is impossible. However, recent advances in index number theory (which are again based on duality theory) show that "sufficiently good"approximation exist to true price deflators, to circumvent this awkward problem.

If the utility function were known it could be used to derive exact price indexes, which, if used to deflate nominal income, would estimate real income changes which correspond exactly to the changes in utility levels.

The following paragraphs depends heavily on Diewert (1976 and 1978b): An unknown linearily homogeneous utility function can be approximated to the second degree by a large number of'flexible' functional forms.[7] Diewert calls those index numbers which are exact for one of the flexible functional forms "superlative" index numbers. For arbitrary utility functions, superlative index numbers therefore approximate to the second degree the exact index numbers corresponding to the arbitrary utility function. Furthermore, Diewert shows that all superlative index numbers approximate each other closely for small changes in quantities and prices. Chaining of index numbers does lead to small changes. Therefore, any chained superlative index number can provide a good second order approximation to the exact price index corresponding to an unknown homogeneous utility function.[7]

However, these results apply strictly to utility functions which are linearly homogeneous. Since one does not want to constrain the utility function to be linearly homogeneous, Diewert makes use of the approximation results of Kloeck and Theil which do not require the utility function to be homogeneous. He shows that any quadratic mean of order r (quantity) index can approximate an arbitrary non-homogeneous utility function to the second order and any quadratic mean of order r price index can approximate an arbitrary indirect utility function.[8] This implies that:

(I) we can choose a chained quadratic mean of order r price index for deflating nominal income and obtain a second order approximation to real income, and

(II) we can therefore use functional forms in real income and nominal prices, even when the utility function us unknown.

Among the quadratic means of order r index numbers, the Fisher's ideal index numbers have certain features which make them the preferred pair. Firstly, there is the computational advantage that the price and implicit quantity indices (or vice versa) can be obtained by simply interchanging the quantities and prices in the same general formula, i.e.

$$\begin{aligned} P_f &= \left[P^1 \cdot X^0 P^1 \cdot X^1 \,/\, P^0 \cdot X^0 P^0 \cdot X^1 \right]^{\frac{1}{2}} \\ Q_f &= \left[P^1 \cdot X^1 P^0 \cdot X^1 \,/\, P^1 \cdot X^0 P^0 \cdot X^0 \right]^{\frac{1}{2}} \end{aligned} \tag{9}$$

Furthermore, the Fisher's index numbers are the only pair among the quadratic mean of order r index numbers which satisfy the factor reversal test[9], i.e.,

$$P_f\,(P^0,\ P^1,\ X^0,\ X^1)\ \ Q_f\,(P^0,\ P^1,\ X^0,\ X^1) \ = P^1 . X^1 / P^0 . X^0 \ . \tag{10}$$

We will therefore use these index numbers throughout.

The discussion above provides an alternative justification for the AIDS system. Furthermore, the above reasoning justifies the use of any linear-in-parameters demand system which can satisfy the homogeneity, symmetry, and adding-up constraint. Three such forms are presented below:

Demand Equations (Linear) in Relative Prices (DERP)[10]

$$X_i = a_i + b_{i1} m + b_{i2} m^2 + \sum_{j \neq N} C_{ij} \frac{P_j}{P_N} \qquad i=1,\ldots,N \tag{11}$$

where $m = M/\tilde{P}$ is real income.

Homogeneity of degree zero in prices and nominal income is automatically satisfied but cannot be tested. Symmetry implies that $C_{ij} = C_{ji}$ and can be imposed for all sample points. The adding-up constraint implies that:

$$\sum_i S_i \eta_{iM} = 1$$

$$\sum_i \frac{X_i P_i}{M} \frac{\partial X_i}{\partial M} \frac{M}{X_i} = \sum_i P_i \frac{\partial X_i}{\partial m} \frac{\partial m}{\partial M}$$

$$= \sum_i \frac{P_i}{\tilde{P}} \frac{\partial X_i}{\partial m} = \sum_i \frac{P_i}{\tilde{P}} (b_{i1} + 2b_{i2} m) = 1 \tag{12}$$

This constraint depends on the sample values of P_i, $\tilde{P}$ and m. Therefore, it cannot be imposed for all sample values. Instead we choose to impose it for the mean of the sample.

Additionally we know from Symmetry that for a commodity k

$$\frac{\partial X_k}{\partial P_N} = -\sum_{j=1}^{N-1} C_{kj} \frac{P_j}{P_N^2} = \frac{\partial X_N}{\partial P_k} = \frac{C_{Nk}}{P_N} \qquad k \neq N$$

Therefore, we have the additional symmetry constraint

$$C_{Nk} = -\sum_{j=1}^{N-1} C_{kj} \frac{P_j}{P_N}$$

It can only be imposed at sample means.

Demand Equations in Square Roots of Relative Prices (DESP)[11]

$$X_i = a_i + b_{i1} m + b_{i2} m^2 + \sum_{j \neq i} C_{ij} \left(\frac{P_j}{P_i} \right)^{\frac{1}{2}} \qquad i = 1,\ldots,N \qquad (13)$$

This is also homogeneous of degree zero in prices and nominal income and the constraint cannot be tested. Symmetry implies $C_{ij} = C_{ji}$ and can be imposed for all sample points. The adding-up constraint is the same as for DERP, i.e. equation No.(12) and can only be imposed at sample means.

Shares Equations in Logarithmic Prices (SELP)[12]

This is an extension of the Almost Ideal Demand System, i.e. AIDS

$$S_i = a_i + b_{i1}\log m + b_{i2}(\log m)^2 + \sum_j C_{ij}\log P_j \qquad i=1,\ldots,N-1 \qquad (14)$$

Since shares add up to one, only N-1 equations are linearily independent and for estimation purposes one equation has to be dropped. Adding-up constraints cannot therefore be tested.[13]

Homogeneity of degree zero implies that $\sum_j C_{ij} = 0$, for all j, and can be tested and imposed.

Symmetry implies that $C_{ij} = C_{ji}$ for all i,j

Table 1 gives the formulas for the price and income elasticities for each of these functional forms.

As can be seen, the elasticities not only depend on the estimated parameters but also on where in the sample space the formula is evaluated. The corresponding standard errors and t-values are calculated as linear combinations of the variances and covariances of the estimated coefficients. For the relevant formulas, see section on estimation procedures.

All the three systems contain square terms in incomes. This allows income elasticities to increase or decline which we wanted to achieve for the data sets considered. However, in all three systems this extra flexibility implies that one cannot impose the adding-up constraint for all sample points.

In each of these systems, real income could of course, be introduced in different functional transformations than the ones proposed here.

ESTIMATION PROCEDURES

For SELP, only N-1 equations need to be estimated, and the formulae given in Table 1 will give all the parameters of the N^{th} equation. Note also that from the formula for the N^{th} income elasticity, i.e.

$$\eta_{NM} = \frac{1 - \sum_{i=1}^{N-1} S_i \, \eta_{iM}}{S_N}$$

we get an estimate of η_{NM} only, and not of b_{N1} and b_{N2} separately.

All the systems form sets of "seemingly unrelated" regression equations (with cross-equation constraints) in the sense of Zellner(1962). As mentioned earlier, the coefficients estimated in the system do not make much sense themselves, and have to be converted into elasticities for interpretation. The formulae for these elasticities are given in Table 1. The corresponding standard errors are also linear combinations of the variances and covariance of the estimated parameters.

Table 1. Formulae for Price and Income Elasticities

	DERP	DESP	SELP
1. Compensated Own Price Elasticity	$\eta^c_{ii} = C_{ii}\frac{P_i}{X_i P_N}$ for i=1 ...,N-1. $\eta^c_{NN} = -\sum_{j=1}^{N-1} C_{Nj}\frac{P_j}{P_N X_N}$	$\eta^c_{ii} = \frac{-1}{2X_i}\sum_{j\neq i} C_{ij}\left(\frac{P_j}{P_i}\right)^{1/2}$ for i=1 ...,N	$\eta^c_{ii} = \frac{C_{ii}}{S_i} + S_i - 1$ for i=1...N-1 $\eta^c_{NN} = \frac{-\sum_{j=1}^{N-1} C_{nj}}{S_N} + S_N - 1$
2. Compensated Cross Price Elasticity	$\eta^c_{ij} = C_{ij}\frac{P_j}{X_i P_N}$ for i=1 j = 1...,N-1 i ≠ j	$\eta^c_{ij} = \frac{1}{2X_i}C_{ij}\left(\frac{P_j}{P_i}\right)^{1/2}$ for i = 1 j = 1..,N i ≠ j	$\eta^c_{ij} = \frac{C_{ij}}{S_i} + S_j$ for i = 1...N j = 1...N i ≠ j $\eta^c_{iN} = \frac{-\sum_{j=1}^{N-1} C_{ij}}{S_i} + S_N$
for j = N	$\eta^c_{iN} = -\sum_{j=1}^{N-1}\eta^c_{ij}$ $= -\sum_{j=1}^{N-1} C_{ij}\frac{P_j}{X_i P_N}$		
3. Income Elasticity	$\eta_{iM} = \frac{1}{X_i}\, b_{i1}m + 2b_{i2}m^2$ for i=1...,N		$\eta_{iM} = \frac{b_{i1}+2b_{i2}\log m}{S_i} + 1$ for i = 1, ...N-1 $\eta_{NM} = \frac{1-\sum_{i=1}^{N-1} S_i\,\eta_{iM}}{S_N}$
4. Uncompensated Own Price Elasticity		$\eta^\mu_{ii} = \eta^c_{ii} - S_i\eta_{iM}$	
5. Uncompensated Cross Price Elasticity		$\eta^\mu_{ij} = \eta^c_{ij} - S_j\,\eta_{iM}$	

As will be explained more fully in the section on data construction, our data consists of 200 observations over 20 years across 10 states. In general, we can expect that the classical assumptions of normally and independently distributed errors with zero mean and constant variance will not be satisfied when the observations stretch across two variational directions.

In a balanced sample, one way for accounting for these different effects is to transform the variables so that they are expressed as deviations from different means, (i.e. the covariance transformation) thus allowing for constant region and time effects. In recent years, the error-component model has been widely used to pool cross-section and time series data (Wallace and Hussain, 1969). The model assumes that the region and time effect are not fixed but random, and are independently distributed with zero means and (usually positive) variances. If the model is

$$Y_{it} = \alpha + \beta_i X_{1it} + \beta_2 X_{2it} + \varepsilon_{it}, \text{ and } \varepsilon_{it} = \mu_i + \nu_t + \eta_{it},$$

then μ_i, ν_t, and η_{it} have variances σ_μ^2, σ_ν^2, and σ_η^2 respectively.

Originally, Wallace and Hussain derived the formulae for the error components from the residuals of the OLS regressions fitted to the above model.[14] However, the β's estimated by OLS are inefficient, and Amemiya (1971) uses the β's from the covariance transformed regression to calculate residuals for the error-components. These error components are then used to compute the second round generalized least square estimates.

Once the error components are estimated for single equations, they are used to transform the original data. The transformed data is then directly fed into the package for estimating the Zellner type of system of equations with cross-equation constraints of symmetry and adding up.

RESULTS

The three functional forms were tested to examine how well they conform to the restrictions of demand theory. The results of the F-tests are given in Table 2. SELP takes the restrictions of homogeneity and symmetry better than the other two functional forms. Homogeneity cannot be tested in DESP, but given homogeneity, the F-test would appear to reject the symmetry restriction. Homogeneity cannot be tested separately for DERP either, but the imposition of the homogeneity restriction through the functional form simultaneously constrains the cross-price terms of the n^{th} equation.[16] This is why the homogeneity and symmetry conditions are tested concurrently. The F-value indicates that this full set of constraints is just barely acceptable. In addition, the DERP and DESP reject the adding-up constraint which is imposed at the mean of the sample. Unfortunately it is not testable for SELP.

Note also that the test (for the functional form) that the second expenditure term equals zero is rejected for SELP. The test was not conducted for DERP or DESP in view of their relatively poorer performance with symmetry and adding up constraints. The regression results for SELP are reported in Table 8 at the end. Because the coefficients themselves are hard to interpret we will instead look at elasticities. But the SELP regression equations are what one would use in any projection work.

Before we look at these elasticities for SELP, it may be appropriate to report here the characteristic roots of the matrix of price derivatives and thus check for negative semidefiniteness. The matrix of price derivatives is derived at sample means since

$\frac{\partial X_i}{\partial P_j} = \eta^c_{ij} \frac{X_i}{P_j}$. The characteristic roots of this matrix are

-.000019, -.00165, -.00699, -.01966 and -.04366.

Table 2

Summary of Test Results

		F-value	P> F
SELP			
i)	Test for Homogeneity	1.337	0.254
ii)	Test for Symmetry given Homogeneity	0.236	0.975
iii)	Test for quadratic income term = 0, given homogeneity and symmetry	6.154	0.0001
DESP			
i)	Test for Symmetry (homogeneity is given and cannot be tested)	1.823	0.046
ii)	Test for adding up constraint, given homogeneity and symmetry	2.940	0.053
DERP			
i)	Test for homogeneity and symmetry (they cannot be tested separately)	1.762	0.056
ii)	Test for adding up constraint given symmetry and homogeneity	5.636	0.004

Table 3. Expenditure Elasticities at Predicted Means

	DERP	DESP	SELP
Rice	0.652 (4.84)	0.566 (4.31)	0.942 (6.58)
Wheat	0.645 (2.54)	0.822 (3.15)	1.077 (5.91)
Inferior Cereals	0.713 (3.45)	0.629 (3.24)	0.362 (1.89)
Pulses	0.058 (0.20)	0.058 (0.21)	- 0.097 (0.39)
Other Commodities	1.194 (25.13)	1.206 (25.71)	1.160

Note: t-values in parenthesis

Expenditure and Price Elasticities from SELP, DERP and DESP

The test results suggest that SELP is the preferred functional form. However, Table 3 gives the expenditure elasticities from DERP, DESP and SELP computed with the formulae in Table 1. The elasticities are estimated at mean values of the dependent variables (i.e., budget shares in SELP and quantities in DERP and DESP) which are themselves predicted by the estimated equations when mean values of the independent variables are plugged in. The three functional forms give, in general, different expenditure elasticities. However, the expenditure elasticity for pulses appears consistently too low and insignificant in all the functional forms.[17] The expenditure elasticity for other commodities is consistently high and interestingly of about the same value, regardless of functional form. Those for rice and wheat are higher in SELP than in the other two forms. The elasticity for inferior cereals is positive across functional forms but appears too high in DERP (where it is even higher than for rice or wheat) and DESP. The SELP elasticity for inferior cereals is positive and small.

Table 4 gives the compensated own and cross-price elasticities while Table 5 gives the uncompensated own and cross-price elasticities. Only the own price elasticities are reported for DERP and DESP since these functional forms reject symmetry and adding up constraints. (The reported elasticities are only for purposes of comparison with SELP).

Table 4 and 5 show that the SELP functional form is the only one that gives consistently negative and significant own price elasticities. (The magnitude of SELP elasticities are also more consistent with earlier estimates from single equations.) DESP estimates positive (although not very significant) own price elasticities for all commodities. While

Table 4. <u>Compensated own and cross-price elasticities at predicted means</u>

		Rice	Wheat	Inferior Cereals	Pulses	Other Commodities	own price elasticities DERP	DESP
Rice	SELP	-0.5273 (-6.93)	0.1031 (1.99)	0.1748 (3.25)	-0.0699 (-1.65)	0.3193 (4.70)	-0.1424 (2.36)	0.406 (0.65)
Wheat	SELP	0.1807 (1.99)	-0.2881 (-2.37)	-0.0536 (-0.56)	0.2579 (3.45)	-0.0969 (-1.09)	-0.0657 (0.48)	0.1774 (1.30)
Inferior Cereals	SELP	0.2939 (3.25)	-0.0514 (-0.56)	-0.6561 (-5.07)	0.0454 (0.56)	0.3682 (3.10)	-0.2167 (-1.84)	0.2147 (1.68)
Pulses	SELP	-0.1923 (-1.65)	0.4049 (3.45)	0.0743 (0.56)	-0.5553 (-3.74)	0.2684 (1.93)	-0.0822 (-0.51)	0.2128 (1.56)
Other Commodities	SELP	0.0643 (4.70)	-0.0111 (-1.09)	0.0441 (3.10)	0.0196 (1.92)	-0.1169	-0.0444 (-1.70)	0.0346 (2.93)

<u>Note.-</u> (i) t-values in paranthesis

(ii) symmetry conditions ensure that the price terms are symmetric. But the elasticities may be different because of the way they are computed.

Table 5. Uncompensated Own and Gross Price Elasticities at Predicted Means

		Rice	Wheat	Inferior Cereals	Pulses	Other Commodities	DERP	DESP
Rice	SELP	-0.6530 (-8.03)	0.0314 (0.59)	0.1001 (1.84)	-0.1156 (-2.60)	-0.3047 (-3.11)	-0.2285 (-3.56)	-0.037 (-0.57)
Wheat	SELP	0.0370 (0.39)	-0.3701 (-3.01)	-0.1391 (-1.46)	0.2057 (2.64)	-0.8107 (-6.53)	-0.1112 (-0.80)	0.1216 (0.87)
Inferior Cereals	SELP	0.2457 (2.58)	-0.0790 (-0.85)	-0.6848 (-5.35)	0.0279 (0.33)	0.1285 (0.77)	-0.2753 (-2.34)	0.161 (1.23)
Pulses	SELP	-0.1794 (-1.48)	0.4123 (3.47)	0.0820 (0.62)	-0.5506 (-3.61)	0.3326 (1.76)	-0.0853 (-0.52)	0.2097 (1.58)
Other Commodi-ties	SELP	-0.0904	-0.0994	-0.0479	-0.0366	-0.8852	-0.834 (-23.40)	-0.758 (-22.03)

Note: t-values in parenthesis

DERP estimates negative own-price elasticities, they are small in magnitude and not very significant in many cases.

Cross-price elasticities are more difficult to evaluate since there is no a priori reason to expect a particular sign. But in general, if there is a presumption that these commodities are substitutes, then clearly SELP as a functional form does better. Rice and pulses appear to be complements, so does wheat and other commodities but neither sign is statistically significant. All other cross price elasticities are either positive and significant or insignificant.

We have used SELP to estimate shares and expenditure elasticities across different expenditure levels. Aggregate real per capital expenditure presented in our data range from 55% (Bihar) to 182% (Punjab) of the mean value and we have therefore produced estimates ranging from 40% to roughly 200% of mean expenditure. Prices are kept constant at the level of the last year of the sample, i.e., 1975-76.[18]

Graph 1 plots the predicted shares. It can be seen that for Rice, Wheat and Pulses, the shares first rise and then fall with expenditure while the share of inferior cereals declines throughout the expenditure range. (Note that this pattern corresponds fairly well to the NSS data reported in Table 7 which is an entirely different data set.) The pulse share becomes negative at high expenditure levels, a further indication that the Pulse equation is not well estimated. Finally note that, as a consequence of the quadratic functional form chosen, the shares of Rice, Wheat and Pulses become negative if extrapolated very much outside the data range.

Fig. Predicted Expenditure Shares using SELP Coefficients

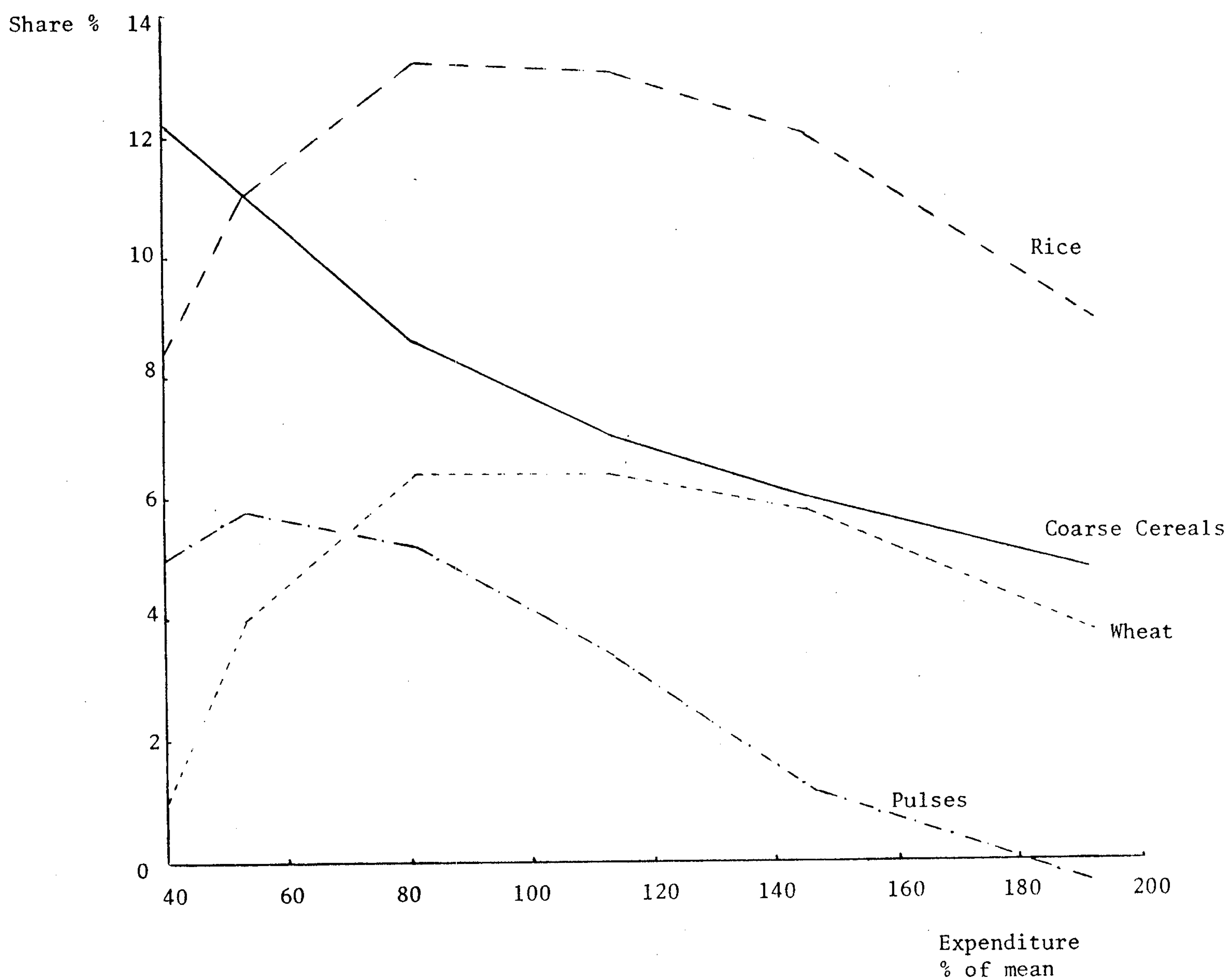

Table 6 reports the expenditure elasticities for the same expenditure levels, using predicted shares. These elasticities vary sharply across expenditure ranges except for inferior Cereals and other commodities. While one would expect the elasticities for all cereals to decline with expenditure, they should not become negative within the data range. One reason for such results could be that shares are estimated with large errors using the estimation equations, especially since the weighted R^2 for the entire system is only 0.203. In table 7 we have therefore estimated the elasticities using observed shares from 28th round of the National Sample Survey. But even here the wheat and rice elasticities turn negative within the data range, while the pulse elasticities remain erratic as before.

Table 6. Expenditure Elasticities at Different Income Levels
(Predicted Shares)

Expenditure Level as Proportion of Mean		Rice	Wheat	Inferior Cereals	Pulses	Other Commodities
0.40		0.487	2.315	0.485	1.845	0.688
0.53	Range of our data	1.650	3.023	0.444	1.126	0.843
0.81		1.105	1.407	0.386	0.285	1.075
1.00		0.942	1.077	0.361	-0.097	1.160
1.13		0.743	0.717	0.335	-1.135	1.246
1.45		0.412	0.083	0.296	-6.842	1.350
1.92		-0.182	-1.437	0.253	7.655	1.436

The pulses elasticities remain unreasonable partly because the data for pulses, i.e. consumption and prices is weak in our state data-set. The state-level disaggregation is unlikely to be reliable for pulses data, not only because the production statistics are poor, but also because inter-state movements of pulses occur substantially by road and are therefore not captured

Table 7. Expenditure Elasticities at Budget Shares and Expenditures Corresponding to the USS Survey Data - 28th Round

Income Level as proportion of Mean	Budget Shares corresponding to NSS Survey data					Income Elasticities				
	Rice	Wheat	Inferior Cereals	Pulses	Other Commod:	Rice	Wheat	Inferior Cereals	Pulses	Other Commod:
0.40	0.150	0.052	0.165	0.017	0.616	1.76	3.42	0.59	3.18	0.66
0.53	0.145	0.067	0.143	0.043	0.602	1.47	2.20	0.57	1.08	0.85
0.81	0.135	0.075	0.092	0.043	0.655	0.81	1.14	0.44	-0.15	1.18
1.13	0.115	0.069	0.062	0.034	0.720	0.67	0.57	0.35	-1.36	1.26
1.45	0.105	0.071	0.054	0.052	0.718	0.21	0.05	0.26	-1.09	1.42
1.92	0.070	0.066	0.041	0.043	0.780	-0.86	-0.74	0.19	-2.35	1.54

Table 8. Estimated SELP regressions

	Constant	Price terms: Rice	Wheat	Inf.cereals	Pulses	log m	$(\log m)^2$
Rice	- 0.117124 (- 2.23)	0.045261 (4.38)	0.003599 (0.51)	0.012736 (1.75)	-0.015793 (-2.74)	0.795855 (2.43)	-0.071172 (-2.51)
Wheat	- 0.147969 (- 3.17)	0.003599 (0.51)	0.048394 (5.14)	-0.010121 (-1.37)	0.015942 (2.75)	0.791344 (3.29)	-0.069564 (-3.33)
Inferior Cereals	0.054566 (0.98)	0.12736 (1.75)	-0.010121 (-1.37)	0.020991 (2.01)	0.0002478 (-0.04)	-0.164532 (-0.69)	0.010087 (0.49)
Pulses	-0.148049 (-2.42)	-0.015793 (-2.74)	0.015942 (2.75)	-0.000243 (-0.04)	0.019209 (2.62)	0.539340 (2.82)	-0.052476 (-3.17)

Number of observations = 200

Weighted R^2 for the system = 0.2031

at all. Unfortunately the NSS survey data are also weak on pulses since they only record quantities of chickpeas.

Conclusions

This study is an attempt to use results from duality theory to estimated demand systems using flexible functional forms. Although several such systems have been estimated, SELP appears to conform better to the restrictions of demand theory than DERP or DESP.

The system has been estimated using a cross-section of time series data on state-level average consumption. The results are interesting and appear convincing (except for the equation on pulses) at the mean values of the sample. The data for pulses are less reliable than for the other food commodities and this is probably the major reason for the poor results from the pulses equation. In view of this for further work we will include pulses into "other commodities", which is admittedly a catch-all residual variable.

The system appears to predict expenditure shares well over the range of expenditures covered by the sample. But expenditures elasticities are not so well estimated as soon as expenditures deviate substantially from their mean. Compared to other studies (see Appendix A) our expenditure elasticities for rice are considerably higher at the mean. On the other hand most other studies find similar expenditure elasticities for wheat (except for Radhakrishna and Murthy, 1978) and coarse cereals.

Few studies have estimated price elasticities which are comparable. For single equation studies the comparison needs to be made with the uncompensated price elasticities. Compared to table A-2 we find a much higher price elasticity for rice but a somewhat lower one for wheat. Our pulse price elasticity is very close to that of Chopra and Swamy, 1975 Comparisons with elasticities from linear expenditure systems is not appropriate because they are only reflections of income effects.

A further study is planned using the same techniques with two rounds of NSS data which contain quantity and price information, and a much wider range of incomes. We therefore expect a better fix on expenditure elasticities. However the NSS data set probably has poorer price data and may not give us a high quality price elasticities as those reported here.

FOOTNOTES

1. See for example, the work of R. Radhakrishna and K. N. Murthy (1978)

2. These assumptions are used by Theil (1971, pp 575-596) for estimating a demand system which uses relative prices rather than absolute prices,

3. This condition is necessary to derive a continuous utility function from the indirect utility function.

4. However, not all functional forms for consumer demand equations so chosen will allow us to find analytical solutions for the indirect or direct utility functions. Therefore, if the researcher is interested in questions of welfare economics where he does need an analytical expression for the Indirect Utility Function, he is probably better off specifying and starting from an indirect utility function which yields such information more readily. Our purpose here is more limited, i.e. to basically estimate the demand elasticities and we do not wish to be involved, at least this stage, with questions of welfare economics.

5. The only way of admitting the possibility that income elasticities are not unitary in the homogeneous generalized Leontif and homogeneous translog indirect utility functions is to introduce a vector of "committed" expenditures. However, this does not eliminate the problem of non-linearity and secondly, we then have to address ourselves to the whole host of questions on what these "committed" expenditures are -- a problem of interpretation that the Linear Expenditure System (and its extensions) also faces.

6. In general, homogeneity does not appear to be a sustainable hypothesis within the Rotterdam System, as its application to data has shown. See A. S. Deaton (1972)

7. The degree of approximation is very close. See W. E. Diewert, 1978b, Appendix 2.

8. In fact, any one of them can be exact for a non-homogeneous utility function which in turn approximates an arbitrary utility function to the second degree. W. E. Diewert, (1976) Theorems 2.16, and 2.17, pp 122-123 and for analogous results for quadratic means of order $r \neq 0$, see page 134.

9. The Fisher's index numbers, along with other quadratic mean of order r index numbers satisfy (I) the commodity reversal test, (II) the identity test, (III) the commensurability test, (IV) the determinateness test, (V) proportionality test, (VI) the time reversal test. But they do not satisfy the circularity test.

10. This system is analogous to the output supply and factor demand system derived from a quadratic normalized profit function.

11. In production theory, this system is derived from the Generalized Leontif Function.

12. In production theory this corresponds to the Translog System.

13. For SELP, the adding-up constraint reads :

$$\sum_i (b_{i1} + 2b_{i2} \log m) = 0.$$ Unless, as in the AIDS, all b_{i2} terms are zero, this constraint can also hold only for sample means.

14. M. Nerlove (1971) gives a more elegant derivation.

15. For the sequencing of these tests, see L. R. Christensen, D. W. Jorgensen and L. J. Lau, (1973).

16. For details, see Appendix E.

17. In a single-equation estimate made by G. Swamy for cross-section

data from family budget data, the elasticity was 0.58 with a t-value of 1.98. The estimate was however estimated from national data. For details, see K. Chopra and G. Swamy (1975).

18. Per capita expenditure ranges for individuals are of course much larger. According to National Sample Survey (India) data of the 28th round, 1973/74, the means of the lowest and highest expenditure class are 20% and 486% of overall mean per capita expenditure. We feel, however, that it is inappropriate to extrapolate our results so far out of the estimation range. Also note that elasticities can be computed for each state.

APPENDIX A

Income and Price Elasticities for India - Review of Literature

The earliest studies on demand elasticities in India depended heavily on the NSS household consumption surveys to estimate expenditure elasticities generally from only one of the several rounds of data that have become available in the past 25 years. Much of the earliest work is presented in A Gangulee (ed) Studies in Consumer Behaviour, Asia Publishing House, Bombay 1960. Subsequently, a large number of studies have been made using different Engel curve forms. For a careful and comprehensive review of this literature, see N. Bhattacharya (1978). The number of these studies and the variety of functional forms tried has been large and Table A-1 below gives the results from only a few of these studies for illustrative purposes.

These NSS based studies have also explored some relationships between consumption on the one hand, and household composition, land ownership and other non-price, household characteristics.

Single equation estimates of price <u>and</u> income elasticities have also been obtained by a few authors who have used aggregate average consumption data and/or NSS survey data. One of the earliest attempts seems to have been made by the National Council for Applied Economic Research (NCAER 1962). Using annual national data from 1948-49 to 1957-58, they estimate income and price elasticities.

Again using national data, A.K. Chakravarty (1961) has estimated the following income and price elasticities for <u>wheat</u>, B.K. Barpiyari & K. Chandra (1961) have used data from 1950-51 to 1957-78 on cereal consumption to estimate income and own price elasticity. These elasticities are given

Table A.1 <u>Some Expenditure Elasticities from NSS data</u> (Rural)

Rice			0.63-0.50	0.562-0.511
Wheat			1.54-1.20	0.926-1.100
Other Cereals			0.31-0.10	
Pulses			0.77-0.70	
Cereals				0.436-0.426
Foodgrains	0.53	0.49		0.514-0.484 (Cereals & Cereal Substitutes)
Milk	1.65	1.78		
Other Foods	1.05	0.99		
Clothing	1.74	1.84		
Fuel	0.63	0.59		
Other Non-Fuel	1.75	1.91		
Year	1954-55	1960-61		1961-62
Source	Pushpam Joseph, Economic & Political Weekly, April 1968		Lydall & Ahmad ISI mimeo,1961	T. Maitra+ ISI technical report Econ. 2/69

\+ The first estimate is a weighted average of regional averages while the second one is a direct all-India estimate.

Table A.2 Income & Price Elasticities from annual time-series data

	Income Elasticity			Own Price-Elasticity		
Rice	0.16			-0.19		
Wheat	1.25	0.91		-0.73	-0.35	
Major Cereals	0.46			-0.34		
All Cereals			0.50			- 0.19
Groundnut Oil	1.72					
Mustard Oil	1.02					
Clothing	1.26					
Time Period	1948-49 to 1957-58	1924-25 1941-42	1950-51 1957-58			
Source	NCAER	A.Chakra-varty †	B.K. Barpi-yari †	NCAER	A.Chakra-varty †	B.K. Barpi-yari †

† These results are published in V.K.R.V. Rao et. al. eds (1961)

in Table 2.

There are two studies which estimate expenditure and price elasticities using dynamic demand models of the stock-flow variety. The methodology is as developed by Houthakker and Taylor (1966). The basic model is

$$q(t) = \alpha + \beta s(t) + \gamma x(t) + \lambda p(t) + U_t \qquad \text{A-1}$$

where q(t) and x(t) are instantaneous flows of consumption and income respectively, P(t) is price and s(t) is the level of the "state variable". The sign of β determines whether q(t) is subject to inventory adjustment ($\beta < 0$) or habit formation ($\beta > 0$). The generally unobservable state variable is eliminated by using the relationship

$$\dot{S}(t) = q(t) - \delta S(t) \qquad \text{A-2}$$

where $\dot{S}(t)$ is the time derivative of S and δ is a depreciation rate assumed to be a constant proportion of the level of the state variable.

C. C. Maji et. al. (1971) use a time series of a changing cross-section of 46 households covering the period 1949-50 tc 1963-64 on Punjab households collected by the Board of Economic Enquiry, Punjab. They estimate price and income elasticities for Rice, Wheat and Maize. The results are mixed and not very convincing.

The dynamic demand function has also been used by S. Tendulkar (1969) who attempts to construct a theoritical model of consumer behaviour of rural households in a semi-monetized economy where the household decides on how much of its produce to retain for self-consumption and how much of it to sell. His model therefore differentiates between cash and non-cash components of total expenditures. The results are given in Table A.3 but it is interesting to note that cash expenditures appear to have an inventory adjustment process while non-cash expenditures indicates a habit-formation process.

Table A.3: Price and Expenditure Elasticities for food from Dynamic Demand Equations

		Expenditure Elasticities		Price Elasticities		Substitution Effect	
		Short-run	Long-run	Short-run	Long-run	Short-run	Long-run
1a	Cash Expenditure	1.0187	0.5933				
1b	Non-Cash Expenditure	0.2998	0.9497				
1c	Total expenditure	0.5503	0.7702				
2a	Cash Expenditure	1.0589	0.7449	-1.2703	-0.8940	-0.9811	-0.6906
2b	Non-Cash Expenditure	0.4480	0.9977	-0.2394	-0.5331	-0.0218	-0.3155
2c	Total Expenditure	0.7440	0.8889	-0.6741	-0.8054	-0.1095	-0.1308
3a	Cash Expenditure	1.2321	0.7344	0.3682	0.2194	0.7025	0.4187
3b	Non-Cash Expenditure	0.1879	0.7043	-0.0814	-0.3052	0.0009	0.0369
3c	Total Food Expenditures	0.6198	0.3910	0.2247	0.1418		

The substitution effect is given by $(\eta_p^i + \alpha_i \eta_e^i)$ where η_p^i and η_e^i are elasticities with respect to price and total expenditure, evaluated at mean, and α_i denotes the proportion of the i^{th} expenditure group in total expenditure. Source: S. D. Tendulkar, 1969

Note that two alternative price variables are used. They are (i) the Rural Retail Price Index of Food Articles (2a-2c) and (ii) the Parity Index representing the ratio of the agricultural wholesale prices received by farmers to the Rural retail price index of non-agricultural commodities bought by them (3a-3c).

IV. K. Chopra and G. Swamy, 1975, estimated a demand function for pulses using NSS survey data from the 14th round, and national annual data on per capita availability. The expenditure elasticity is computed from the cross-section data and own price and cross-price elasticities are computed from a pooled regression using annual time series data. The resulting equation is

$$\log X = 2.51263 + 0.58 \log (E/P_b) - 0.63307 \log (P_p/P_b) + 0.64516 \log (P_c/P$$

t-values (1.98) (-5.62) (1.83)

--- A-3

$R^2 = 0.80$

$DW = 2.21$

where X = per capita consumption of pulses, E = total per capita expenditure, P_p = price of pulses, P_c = price of cereals (the substitute) and P_b = a general wholesale price index.

It can be seen that the expenditure elasticities of pulses measured by SELP differ markedly from the one in equation A-3, which appears more reasonable given our knowledge of consumption habits.

Several rounds of the NSS survey data have also been used to estimate price and expenditure elasticities using the Linear Expenditure system. These provide probably the only consistent set of demand elasticities for India in the sense that they are derived from a system of demand equations that satisfies all the restrictions of demand theory.

Before discussing these, however, we may briefly review the work of N. S. Iyengar and L. R. Jain (1974) who extend Honthakker's Indirect Addilog model to estimate income and price elasticities for food and non-food commodities (The model is not extended to more than two commodities) although this can be done by appropriately restricting the coefficients - see Appendix D)

The reported elasticities are given below.

Table A-4: Price and Income Elasticities from Extension of Indirect Addilog System

	Food	Non-Food	Food	Non-Food
Income	0.4729	1.8548	0.6669	2.0488
Own-Price	-0.4180	-0.9110	-0.6322	-0.8908
Cross-Price	-0.0549	-0.9438	-0.0347	-1.1580

The Linear Expenditure System is a convenient system for the estimation of elasticities for broad groups of commodities. (For the merits and limitations of the system, see Appendix D). Although Paul Pushpam and Ashok Rudra (1964) had experimented with the functional form and NSS survey data, the really comprehensive work using the linear expenditure system has been done by R. Radhakrishna and K. N. Murthy (1978).

In the study, the NSS data from 25 rounds covering the period April 1951 to June 1971 have been used. The different expenditure groups in different surveys have been reconciled by deflating them and regrouping them into five consistent expenditure classes, separately for rural and urban areas. The linear expenditure system is then fitted to each expenditure group.

The commodities are classified into nine groups in the first instance and into a larger number of more specific food groups in subsequent estimations (depending on the availability of data).

The expenditure elasticities are reported for each expenditure class and for rural and urban areas separately. All the results are not reproduced here, but in general the expenditure elasticities for Cereals, Edible Oil and Other Food show a clearly declining trend as income rise. The trend is not as clear for Milk and Milk Products, Meat and Fish or Sugar (and Gur). The expenditure elasticities each for Fuel, Clothing and other non-food clearly rise with incomes. The average (Group III) elasticities are reported below.

Table A.4. Expenditure Elasticities from LES

	Rural	Urban	Rural	Urban
1. Cereals	0.583	0.461	-0.528	-0.348
2. Milk & Milk Products	2.222	2.055	-1.234	-0.951
3. Edible Oil	0.968	1.067	-0.568	-0.515
4. Meat,Fish & Eggs	1.569	1.589	-0.982	-0.744
5. Sugar & Gur	1.537	1.302	-0.885	-0.616
6. Other Food	1.121	1.069	-0.693	-0.584
7. Clothing	1.468	0.979	-0.853	-0.470
8. Fuel & Light	0.814	0.792	-0.495	-0.401
9. Other non-food	1.763	1.601	-1.007	-0.804

Note: Pulses are included in Other Food.

NSS data on 6 rounds which give the requisite data for a finer 11-commodity classification have been used to estimate (for each expenditure group) a 11-commodity LES, which breaks up Cereals into Rice Wheat and other Cereals.

As mentioned in the text, the LES does not admit inferior goods and it may therefore be inappropriate to disaggregate cereals when one of them is indeed inferior for higher income groups, a limitation, which the authors recognize. They indeed show that the estimated 'γ' parameters see equation D-1 from the first three equations (Rice, Wheat and Inferior Cereals) in the 11-commodity model do not add up to the 'γ' estimate of the first equation in the 9-commodity model, i.e. consistency is not maintained in a number of cases.

In general the expenditure elasticities at the mean of the sample for Rice is lower than that for inferior cereals in both rural and urban areas. In the urban areas, even the elasticity for wheat is lower than for inferior cereals as the following table shows.

Table A.5 Expenditure Elasticities from LES (11-Commodity model) average group (Group III)

	Rural	Urban
Rice ..	0.399	0.738
Wheat ..	0.589	0.349
Inferior Cereals	0.511	0.962
Milk & Products	1.065	1.690
Edible Oil ..	0.385	0.574
Meat, Fish & Eggs	0.977	0.972
Sugar & Gur ..	0.090	1.181
Other Food ..	2.333	1.141
Clothing ..	1.770	1.584
Fuel & Light ..	0.563	0.753
Other non-food	3.023	1.389

APPENDIX B

DATA

The data set consists of time series data on aggregate foodgrain availability of 10 Indian States for the period 1956-57 to 1975-76 and is called "State Data".

Period covered: 1956-57 to 1975-76 (agricultural years).

The data on consumption is derived as follows. Production figures by States are published by the Government of India, in Estimates of Area, Production and Yield, Directorate of Economics and Statistics, India.

1. Rice
2. Wheat
3. Jowar (Sorghum)
4. Bajra (Pearl Millet)
5. Maize
6. Ragi (Finger Millet)
7. Barley
8. Bengal Gram (Chickpea)
9. Tur (Pigeonpea or Redgram)
10. Urad (Black Gram)
11. Green Gram (Mung Bean)
12. Masur (Lentil)
13. "Other" Pulses (or Lentils)

In Estimates of Area, Production and Yield, production of the pulses 10, 11 and 12 is aggregated into "Other Pulses". However, production data on these pulses have been put together by K. Chopra and G. Swamy (1975) for most of the years. For years not covered by their study, the data was gathered from Agricultural Situation in India, various issues.

The production data is then adjusted for:

(I) seed feed (animal) and wastage

(II) changes in government stocks

(III) inter-state movements by rail of these grains.

I. The requirements of seed and feed and the loss from wastage is assumed to be as follows for the different crops:

Rice	:	7.6%
Wheat	:	12.1%
Other Cereals	:	12.5%
Chickpea	:	22.1%
Other Pulses	:	12.5%

II. Changes in Government Stocks.

In the Bulletin on Food Statistics, Directorate of Economics and Statistics, the Government of India publishes data (on a calender year basis) on Internal procurement, Total Public distribution (by State and Central Governments) and Closing Stocks (with State and Central Governments). This data is available for Rice, Wheat and "Other Grains", and for all States.

Changes in government stocks are taken to be the difference between closing stocks at the end of two years. When the difference is positive, the figure is substracted from consumption since this means a decline in what is available for consumption. When this difference is negative, the figure is added to consumption.

It should be noted that the differences are measured from December of one year to December of the previous year, while production figures relate to the agricultural year (July to June). Therefore, the time periods are not coincident though they overlap for six months. We have, therefore, assumed that the change in stock from any year 't' to 't+1' relates to the agricultural year 't' to 't+1'. It should also be noted that the category "other grains" is assumed to refer to the most important of the coarse cereals produced in that state.

III. Inter-state movements of foodgrains.

The Government of India also publishes (Bulletin on Food Statistics) data on inter-state movements of foodgrains (and other commodities) by rail, (and up to 1968-69 by river). Although movement of foodgrains also takes place by road, there is no data on this. Further, rail-traffic is still the most important mode of movement, and the Food Corporation of India also transfers its stocks by and large, by rail.

The data is available for each State, and for each cereal from 1956-57 till current time, and for pulses (chickpea and other pulses) from 1963-64.

Although most of the States we consider were reorganized by 1956-57 (so that their geographical boundaries have not altered through the period) the data on inter-state movements is not tabulated to correspond to the reorganized boundaries till 1960-61. Thus for the years 1956-57 to 1959-60, data which is available for the pre-reorganization States have been added together to approximate to the new boundary lines. Thus, for example, for these years, the state of Madhya Pradesh is defined to include the old Madhya Pradesh, Madhya Bharat, Bhopal, Vindhya and Vidarbha.

Two of the states were reorganized after 1956-57, i.e. the old Bombay State was divided into Gujarat and Maharashtra, in 1960-61, and the data on inter-state movements tabulated according to the new states appears only in 1962-63. The old state of Punjab was divided into Punjab and Haryana in 1967-68. The data for these states were adjusted to make them conform to constant state boundaries. The states which are included in the sample are:

1. Andhra Pradesh
2. Bihar
3. Gujarat
4. Karnataka
5. Madhya Pradesh
6. Maharashtra
7. Punjab (and Haryana)
8. Rajasthan
9. Uttar Pradesh
10. West Bengal

It should be noted that data on inter-state movements are given separately for exports and imports from the mainland and exports and imports from the ports which belong to these states. However, exports from and imports into ports do not either decrease or increase the availability of grains in the mainland. The reason is that ports are not producing units and exports from the ports can only occur if in fact the mainland had previously exported grains to the ports, and the latter is already accounted for the in the export figures of the state's mainland. On the other hand, since the ports are not consuming units either, any imports into the ports must be considered as commodities in transit which will reach the mainland as the state's imports. Thus the relevant export and import figures for each state are those for the mainland, i.e. excluding the ports. The net of these exports and imports constitute the net addition on to or subtraction from what is available for consumption in these states.

Therefore, availability in each state (which is a measure of consumption) is estimated as:

Availability = Production - Seed, Feed, Wastage + Change in Government Stocks + Net Imports or Exports from States

This is divided by population to arrive at per capita availability or consumption.

Nominal Income and Total Expenditure

The per capita income figures for these states was obtained directly from the Central Statistical Organization, India for the period 1960-61 to 1975-76. For the years prior to this, figures of per capita income estimated and published by the West Bengal State Statistical Bureau (Estimates of State Income, 1965) were used. These were adjusted to conform to the first series by taking the ratio of the two estimates of per capita income for the year 1960-61 and splicing the series backwards.

From these figures of per capita income, "savings" are deducted by assuming that the national net saving rate (for the different years) as published by the CSO applies to all states. This is necessary since estimates of savings in the states are not generally available. The resulting series is per capita expenditures.

Prices

Month-end wholesale price quotations for many agricultural commodities are recorded for a large number of "centers" or markets in India. This data has been published in the following yearly publications:

> Agricultural Prices in India, 1952-62, Directorate of Economics and Statistics, Government of India
>
> Agricultural Prices in India, 1963-74, Directorate of Economics and Statistics, Government of India, Comprehensive Volume.

This forms the basis of our price-data. For each state and for each commodity, the price quotations for the different centers are averaged for each month and then averaged over the agricultural years.

The number of centers varies, depending on the commodity, and its importance in the production pattern of the state. Further, although the publications prior to 1963 publish price data for a very large number of centres, the number of centres covered is much smaller in the comprehensive volume which publishes data for the period 1963-74. Therefore, we have worked backwards, i.e. we have included those centres for which data is available for 1963-74 and extended the series back to 1956-57, using the data on the same centres from the earlier publications.

For the years 1974 to mid-1976, the data comes from the Bulletin on Food Statistics, Directorate of Economics and Statistics, Government of India, which again, reduces the number of centres covered to a slightly smaller set. We adjust for this by multiplying the average (over the centres quoted in the Bulletin on Food Statistics) prices for 1974-75 and 1975-76 by a ratio of the average price (over the centres quoted in agricultural prices in India) to the average price (over the centres quoted in the Bulletin on Food Statistics) for the years 1973-74, 1972-73 and 1971-72, and splicing the series.

It is important to note that the price quotations in each centre refers to a particular variety of a commodity. In most cases, the price quotation is for the same variety throughout the period. However, in the more recent years, starting from the late sixties till 1975-76, price quotations often refer to the new high-yielding varieties of Rice and (particularly) Wheat. Thus, although the average over centres represents, in general, an average over different varieties of the commodity, this average tends to represent the newer varieties in the later years of the sample.

Another deficiency in the price data cannot be as easily ignored, and that is, that for Rice in particular, but also for Wheat and Jowar to some extent, price quotations for some years after 1964-65 refer to "controlled" rates fixed by the State governments. These are not free-market prices in the sense that they do not represent the price at which consumers could buy as much as they wished, but represent instead the price prevailing in a smaller government market where limited amounts could be bought at the controlled rate.

For those years when only controlled rates are quoted for the centres, we have made use of Farm Harvest Prices, adjusted to conform to the wholesale price series by a ratio of wholesale price to farm-harvest price in the previous years in that state. We have also used Farm Harvest Prices as a complete price series in itself when there is no wholesale price data for that commodity in that state (as for example wheat in Andhra Pradesh and Karnataka), or when the wholesale price has been controlled practically throughout the time period (as for Jowar in Maharashtra). Only when we have neither a full series on wholesale prices, nor farm harvest prices, have we used the All-India wholesale price index to fill in the missing values -- but this only happens in a few cases for Pulses. The sources for the Farm Harvest Prices are:

(I) Farm Harvest Prices in India, Directorate of Economics and Statistics, Government of India, 1962-63.

(II) Agriculture in Brief -- various issues, Directorate of Economics and Statistics, Government of India

(III) Agricultural Situation in India -- various issues, Directorate of Economics and Statistics, Government of India

Price and Quantity Indices

As discussed in the paper and Appendix C, chained Fisher price and Quantity indices have been constructed. The Fisher's index is the geometric mean of the Laepayers and Paasche index numbers. The quantity index numbers use the same formulas with the price and quantity variables interchanged.

Note that all quantity variables are in per capita terms, i.e. the quantities derived in the previous section are divided by population.

Adjustment for Base-year Differences in Quantities and Prices Across States

The quantity and price indices estimated in the manner are all equal to 1 in the base year for all the states and therefore do not take into account the differences in base-year prices and quantities. To account for this, the following procedure is used.

A reference quantity for commodity i is defined for the base year as an average over all the states, i.e.

$$\bar{Q}_{0i} = \frac{\sum_k N_0^k Q_{0i}^k}{\sum_k N_0^k} \qquad \text{(B-1)}$$

where 0 = base year, k = the states and N = population, and Q_{0i}^k = base year per capita quantity consumed.

Then a reference price is defined as: $\bar{P}_{0i} = \sum_k w_{0i}^k P_{0i}^k$, where:

$$w_{0i}^k = \frac{N_0^k Q_{0i}^k}{\sum_k N_0^k Q_0^k} \qquad \text{(B-2)}$$

Then reference expenditures are calculated as the expenditures in State k, evaluated at the reference prices.

$$\text{REFEXPS}_0^k = \sum_i Q_{0i}^k \bar{P}_{0i} \qquad \text{(B-3)}$$

The ratio of observed expenditures to reference expenditures is a measure of the proportion by which price differences across states account for observed expenditure differences. This ratio is called Reference Ratio (REFRATIO) and is an index of the prices in state k relative to the average prices in the base year.

$$\mathrm{REFRATIO}_0^k = \frac{\sum_i Q_{0i}^k \; P_{0i}^k}{\sum_i Q_{0i} \; \bar{P}_{0i}} \qquad \text{(B-4)}$$

If $\bar{P}_{kt}$ is the (Fisher's) Price Index in year t for state k and $\bar{Q}_{kt}$ is the (Fisher's) Quantity index of state k in period t, then we define the adjusted price indices P^*_{kt} and the adjusted quantities Q^*_{kt} as

$$P^*_{kt} = \mathrm{REFRATIO}_0^k \mathrm{x} \quad \bar{P}_{kt} \qquad \text{(B-5)}$$

$$Q^*_{kt} = \mathrm{REFEXPS}_0^k \mathrm{\ x} \quad \bar{Q}_{kt} \qquad \text{(B-6)}$$

Note that P^*_{kt} is an index number which is dimensionless, but Q^*_{kt} has the dimension of the expenditure series, i.e. it is of the same order of magnitude than the expenditures. It thus is a quantity series, and not a quantity index.

The Fisher's chained price indices and quantities are calculated for the following sets of commodities for each of the ten states.

1. Rice
2. Wheat
3. Superior Cereals (Rice and Wheat)
4. Inferior Cereals (Jowar (Sorghum), Bajra (Pearl Millet), Maize and Ragi, and Barley)
5. Pulses (Chickpea, Tur, Mung, Urad and Masur)

Note that the system of demand equations is necessarily a system that considers the allocation of total expenditure on an exhaustive bundle of commodities, and cannot be used to estimate demand functions for a subset of these commodities only, without assuming separability. In our study therefore, although the primary interest is to estimate demand elasticities for the above mentioned sets of commodities, the system would not be complete without including "other commodities" not mentioned above.

Admittedly, this category of "other commodities" is a mixed bundle which includes non-grain food commodities as well as non-food commodities. We do not have specific data on prices and quantities of these commodities, but the expenditure on them is derived simply as total expenditure minus expenditure on the foodgrain commodities listed above. If we have a quantity or price index for these commodities, the other index can be derived because the following equation must hold:

$$P\,(P^0, P^1, X^0, X^1)\quad Q(P^0, P^1, X^0, X^1) = \frac{P^1X^1}{P^0X^0}$$

We have used a proxy for the price index of other commodities and this is weighted average of the consumer price index number (for non-food commodities) for industrial workers and the consumer price index number (for non-food commodities) for agricultural laborers, where the weights are the proportion of the population that is urban and rural respectively.

These price indices for all commodities and for food commodities are estimated for each state and published with 1960 as base year for the Industrial Worker's Index, and with 1960-61 as the base year for the agricultural worker's index. The former indices are estimated for several urban centres in each state, while the latter is estimated for the rural sector as a whole. The non-food index is derived from the all commodities index and the food-index, using the weights as given by the Labor Bureau, Ministry of Labour, which computes these indices.

For the period 1960-61 to 1975-76, therefore, this data is published in a directly usable form. However, for the period 1956-57 to 1959-60, we do not have these indices for either the industrial workers or the Agricultural Workers. However, we do have the "working class" Price Index with base 1950-51, and this index is adjusted to correspond to the Industrial Workers Index from 1960-61 to 1975-76, and the whole series is then arithmetically adjusted to shift the base to 1956-57. To extend the agricultural workers price index back to 1956-57, we have used some information published in the Pocket Book of Labor Statistics, Ministry of Labour, 1968. This gives the agricultural workers' price index (with base year 1960-61) for most of the states for the year 1956-57 only. The data for the years 1957-58, 1958-59 and 1959-60 are filled in by assuming that the agricultural worker's non-food price index was proportional to the industrial worker's index, the ratio referring to the year 1960-61.

The weighted average of these two indices is our price index for "all commodities" not included in the list above, and is used to derive the quantity index for other commodities, which is one of the dependent variables in the specifications DERP and DESP. In specification SELP, since only N-1 equations need to be estimated, the equation for "other commodities" is left out, although the price index for other commodities enters as the deflator to ensure homogeneity.

Expenditure Deflator

To deflate the expenditure figures, we use the Fisher's chained price index of all commodities, including other commodities. (The quantity of "other commodities" is assumed to equal 1 in the base period).

Appendix C: Index Numbers

This section draws heavily on Diewert (1976, 1978a, 1978b). One of the important questions facing econometricians who construct data series is the choice of functional form for an index number. If the functional form for the utility function is known, then an "exact" index number exists in the sense that, if the utility function is $U = f(X)$, then the change in utility levels between periods 0 and 1 can be precisely measured by a quantity index:

$$U^1/U^0 = f(X^1)/f(X^0) = Q(X^0, X^1; P^0, P^1) \tag{C-5}$$

where the term on the right hand side is a quantity index which is some function of prices and quantities. Given a quantity index (or price index) the other function can be defined by the Fisher's weak factor reversal requirement, that the product of a quantity and a price index be equal to the ratio of expenditure in the two periods, i.e.,

$$P(P^0, P^1, X^0, X^1)\ R(P^0, X^1; P^0, P^1) = P^1X^1/P^0X^0$$

The utility function may be unknown, but a flexible functional form may exist which provides a second order approximation to the unknown, arbitrary, twice differentiable linearly homogeneous function. If an index number exists which is exact for the flexible (approximate) utility function, then that index number has been called "superlative" by Diewert. Since the flexible functional form is a second order approximation to an arbitrary homogeneous utility function, the superlative index number which is exact for the flexible form is a second order approximation to the exact index number of the true utility function.

Diewert then shows that index numbers of quadratic means of order r are superlative for linearly homogeneous utility functions or indirect utility functions which are also quadratic means of order r. In particular, the Divisia Index is superlative for the linearly homogeneous translog utility function, and the Fisher's index is superlative for the linearly homogeneous quadratic mean of order 1 function. Similarly the quadratic mean of order r = 2 index is superlative for the linearly homogeneous generalized Leontif utility function. Thus a wide choice of superlative index numbers exist.

However, in another paper Diewert shows that for small changes in quantities and prices, all superlative index numbers approximate each other to the second degree. Since generally changes in price and quantities between successive periods are smaller than changes relative to a fixed base period, chaining of indices can bring about extra-ordinarily close approximation between all the superlative index numbers. For examples see Diewert (1978). Thus, once superlative index numbers are chained one can choose any one of them.

However, all the results discussed so far about superlative index numbers refer to utility functions which are linearly homogeneous. Since we do not want to constrain our utility function (or indirect utility function) to be homogeneous, these results do not apply strictly. Instead, Diewert used the results of Kloeck and Theil which do not require the aggregator function to be linearly homogeneous. Diewert (by extending their results to have a global character) has shown that the Divisia price index is exact for the general translog cost function, as well as for functional forms other than the translog if the assumption of homogeneity is dropped. A similar result is proved for the Divisia quantity index in the context of an utility function which is not necessarily linearly homogeneous (Diewert, 1976). Diewert has further shown that similar results can be proved for all quadratic means of order r index numbers.

These extraordinary results imply that chained superlative quantity indices can measure utility changes to a very close extent even if the utility function is unknown and that chained superlative price indices can approximate "true" cost of living changes to an equally close extent. The approximation of chained superlative index numbers to each other appears to be so close that any error which could be introduced by not choosing the "correct" one must pale into insignificance compared to other error in data problems commonly encountered in econometric analysis.

APPENDIX D

Other Demand Systems

Systems of Demand Equations Derived from Specified Utility Functions

We review briefly two systems of demand equations which have been used widely. The Linear Expenditure System (and its extension, the Extended Linear Expenditure System) and the Indirect Addilog Systems satisfy all the conditions imposed by consumer theory, but are nevertheless inappropriate for our needs.

The Linear Expenditure System (Stone, 1954) - This is derived from a directly additive utility function

$$U(X) = \Sigma f_i(X_i)$$

and gives the following functional form for the demand equations

$$P_i X_i = \gamma_i P_i + \beta_i (M - \sum_{j=1}^{N} \gamma_j P_j) \quad i = 1, \ldots N \tag{D-1}$$

where M = total expenditure, and $\sum_{i=1}^{N} \beta_i = 1$

The assumptions further needed are that $M - \sum_{j=1}^{N} \gamma_j P_j > 0$

and $0 < \beta_i < 1$ hold for all i. The β_i are the marginal budget shares, and since they cannot be negative, inferior goods are ruled out. Further, the additive form of the utility function allows little flexibility in the adjustment of price coefficients, and all goods are net substitutes. There is rigidity in the income responses as well since marginal budget shares are constant. Therefore this system is more suited to the analysis of broad aggregate groups.

The Extended Linear Expenditure System (Lluch et.al. 1977) allows for the endogenous determination of savings, and hence of total consumption expenditure. It has the same properties as the LES in all other ways, and is therefore unsuitable for our purpose. Similarly, the Quadratic Expenditure System (Pollack and Wales, 1978) has the following functional forum of the demand equation

$$P_i X_i = \gamma_i P_i + \beta_i \left(M - \sum_{j=1}^{N} \gamma_j P_j\right) + (C_i - \beta_i)\, \lambda \Pi P_j^{-c_j} \left(M - \sum_{j=1}^{N} \gamma_j P_j\right)^2 \qquad \text{(D-2)}$$

$\Sigma\beta_j = 1$, and $\Sigma C_j = 1$. If $C_i = \beta_i$ for all i then this reduces to the LES

<u>The Indirect Addilog System (Houthakker, 1960)</u> - This system generates demand equations for pairs of commodities of the following kind

$$(\log X_i P_i - \log X_j P_j) = a_i + b_i \log\left(\frac{M}{P_i}\right) - b_j \log\left(\frac{M}{P_j}\right) \qquad \text{(D-3)}$$

If these equations are estimated separately, then N-1 distinct estimates for each of the b_i are obtained and hence the restriction that b_i take the same value in every equation is used in estimation. The test for the equality of the b_i can be used to test the compatibility of the addilog model and the data. However, like the LES (and its extensions) this also assume additivity of the utility function (as does the Rotterdam model with block-independence or preference independence). Therefore they impose the same kind of constraints on the price coefficients.

<u>Systems of Demand Equations derived from differentiating a general non-homogeneous flexible functional form for the indirect utility function.</u>

(i) Generalized Leontief Reciprocal Indirect Utility Function. This formulation is due to W.E. Diewert (1974) and has this form

$$\frac{1}{G(V)} = h(V) = \sum_{i=1}^{N} \sum_{j=1}^{N} b_{ij} V_i^{1/2} V_j^{1/2} + 2 \sum_{j=1}^{N} b_{0j} V_j^{1/2} + b_{00} \qquad \text{(D-4)}$$

where $b_{ij} = b_{ji}$. The i^{th} demand equation is

$$X_i(V_1,\ldots V_N) = \frac{\sum_{j=1}^{N} b_{ij} V_i^{-1/2} V_j^{1/2} + b_{0i} V_i^{-1/2}}{\sum_{k=1}^{N} \sum_{m=1}^{N} b_{km} V_k^{1/2} V_m^{1/2} + \sum_{m=1}^{N} b_{0m} V_m^{1/2}} \qquad \text{(D-5)}$$

(ii) The translog reciprocal indirect utility function has the form

$$\frac{1}{G(V)} = h(V) = \alpha_o + \sum_{i=1}^{N} \alpha_i \log V_i + 1/2 \sum_{i=1}^{N} \sum_{j=1}^{N} \gamma_{ij} \log V_i \log V_j \qquad \text{(D-6)}$$

$\gamma_{ij} = \gamma_{ji}$. Application of Roy's identity yields,

$$X_i(V) = \frac{V_i^{-1} (\alpha_i + \sum_{j=1}^{N} \gamma_{ij} \log V_j)}{\sum_{k=1}^{N} \alpha_K + \sum_{k=1}^{N} \sum_{m=1}^{N} \gamma_{km} \log V_m} \qquad \text{(D-7)}$$

Both these systems of equations are non-linear.

The translog indirect utility function can be made to yield demand equations which are linear in the unknown parameters if we make the indirect function linearly homogeneous (of degree -1) (and the direct function of 1) as in Lau, Lin and Yotopoulos (1978). The indirect translog utility function

$$\log W^* = \alpha_o + \sum_{j=1}^{N} \alpha_j \log P_j^* + 1/2 \sum_{j=1}^{N} \sum_{k=1}^{N} \beta_{jk} \log P_j^* \log P_k^* \qquad \text{(D-8)}$$

(The P* variable are normalized by dividing prices by income M) is homogeneous if $\sum_{i=1}^{N} \alpha_j = -1$, $\beta_{jk} = \beta_{kj}$ for all j and k, and $\sum_{i=1}^{N} \beta_{jk} = 0$ for all j, and $\alpha_o = 0$. However, homogeneity of the indirect utility function is a very restrictive assumption since it implies that all income elasticities are unitary. As mentioned in the text, the only way of admitting non-unitary income elasticities into the homogeneous version of either the generalized Leontif or the translog is to introduce a vector of committed expenditures. Diewert has done this for the generalized Leontief form, and Lau and Mitchell (1970) have done this for the Translog form. However, this does not eliminate non-linearity, and in addition, it is difficult to justify the existence of "committed"expenditure a problem that the Linear Expenditure System also faces.

Systems of Equations Linear in Parameters, but only partially consistent with consumer demand theory.

(1) Linear in logarithms of real income and prices.

$$\log X_i = a_i + b_{i1} \log m + b_{i2} (\log m)^2 + \sum_{j \neq k} c_{ij} (\log P_j - \log P_k) \qquad \text{(D-9)}$$

This is homogeneous but again, symmetry can be imposed only at sample means, since

$$\frac{\partial X_i}{\partial P_j} = C_{ij} \frac{X_i}{P_j} \neq \frac{\partial X_j}{\partial P_i} = C_{ji} \frac{X_j}{P_i} \qquad \text{(D-10)}$$

(2) The Rotterdam model discussed in the text is an improvement over (1) in the sense that symmetry can be imposed globally. However, as pointed out in the text, homogeneity does not seem to be a sustainable hypothesis in this system.

(3) The Rotterdam model in real income and relative prices (Theil, , pp 574-580) is an attempt to reduce the number of unknown parameters in the system by assuming block-independence or preference independence for the utility function. The compensated price effect, i.e., the substitution effect is written as a sum of two terms

$$\frac{\partial X_i}{\partial P_j} = \lambda U^{ij} - \frac{\lambda}{\partial\lambda/\partial m} \quad \frac{\partial X_i}{\partial m} \; \frac{\partial X_j}{\partial m} \qquad \text{(D-11)}$$

where λ = marginal utility of income, and U^{ij} is the element i, j of the inverse of the utility function

$$U_{ij} = \frac{\partial^2 U}{\partial X_i \, \partial X_j}$$

The first term is known as the specific substitution effect of the change in the j^{th} price on the i^{th} quantity, the second as the general substitution effect, and the sum as the total substitution effect in the sense that all commodities compete for the consumer's money. The demand function is then written as

$$S_i \log X_i = a_0 + a_1 \log m + \sum_{j=1}^{N} C_{ij} \left(\log P_j - \sum_{k=1}^{N} \mu_k \log P_k\right) \qquad \text{(D-12)}$$

i.e. the equation is in terms of relative prices in the sense that depending on whether the utility function is assumed to be block-independent or preference independent, the price term P_j is deflated by a weighted average (where the weights μ_k are <u>marginal</u> budget-shares) of those prices which are assumed not to affect the demand for the i^{th} commodity, i.e. the general substitution effect is substracted from the specific one. In the case of preference independence, each demand equation contains only one deflated price, i.e. the own price, deflated by all other prices; with the assumption of block-independence the number of price terms depends on the number of sets and the number of commodities in each set. This system is therefore estimable only if, in addition to symmetry, either preference independence or block-independence is imposed in the estimation.

As mentioned earlier, neither of these assumptions are tenable in our context. (It may also be noted that the estimation of this set of demand equations leads to non-linear procedures since the second term on the right hand side is not linear in the unknown parameters. However Theil suggests a way of linearising the relationship by using an estimator for the marginal value shares).

REFERENCES

Anton B. Barten, The System of Consumer Demand Function Approach : A Review, Econometrica, Vol 45, January 1977, pp 23-52

Amemiya, T. "The Estimation of the Variances in a Variance Component Model", IER, 1971, Vol 12, pp 1-13

Chopra, K. and G. Swamy, Pulses - An Analysis of Demand and Supply in India, Institute for Social and Economic Change, India, Monograph 2, 1975, New Delhi, Sterling Publishers

Christensen, L. R., D.W. Jorgensen and L. J. Lau, "Transcendental Logarithmic Production Frontiers", The Review of Economics and Statistics, Vol. 55, February 1973, pp 28-45

Deaton, A. and J. Muellbauer, "An Almost Ideal Demand System", University of Bristol, Department of Economics, Discussion Paper No.62/77, Novr.1977

Deaton, A. and A. Brown, "Models of Consumer Behavior : A Survey", Economic Journal, Vol.82, December 1972 : 1114-1236

Diewert, W. E., "Duality Approaches to Microeconomic Theory", Discussion Paper, 78-09, Department of Economics, University of British Columbia, March 1978a, pp .90

____________, "Applications of Duality Theory", in M. D. Intrilligator and D. A. Kendrick (Eds), Frontiers of Quantitative Economics, Vol.II Amsterdam:North Holland, 1974, p.106

____________, "Exact and Superlative Index Numbers", Journal of Econometrics Vol.4(1976) pp 115-145

____________, "Superlative Index Numbers and Consistency in Aggregation", Econometrica, Vol.4, July 1978c, pp 883-900

Directorate of Economics and Statistics, Ministry of Agriculture, Government of India

-- Estimates of Area, Production,and Yield, Various issues

-- Agricultural Situation in India, various issues

-- Bulletin of Food Statistics, various issues

-- Agricultural Prices in India, 1952-62

-- Agricultural Prices in India, 1963-74

-- Farm Harvest Prices in India, various issues

Houthakker, H. S., "The influence of Price and Income on Household Expenditures", Bulletin de l' Institute International de Statistique, XXXVII, 1960

Houthakker, H. S. "Additive Preferences", Econometrica, Vol.28, 1960, pp 244-257

______________, "On a class of Dynamic Demand Functions (Abstract), Econometrica, July 1962, Vol.30, pp 588-589

______________, and L. D. Taylor, "Consumer Demand in the U. S. 1929-70" Cambridge, Harvard University Press, 1966

Iyengar, N. S. and L. R. Jain, "Income and Price Effects on Households Expenditure", Indian Economic Journal, Vol.22, October-December 1974 pp. 128-141

Lau, L.J., N. Lin and P. A. Yotopoulus, "The Linear Logarithmic Expenditure System : An Application to Consumption-Leisure Choice", Econometrica, Vol.46, July 1978 : pp 843-868

________ and Mitchell B.M., 1970, "A Linear Logarithmic Expenditure System : An Application to U. S. Data", paper presented at the Second World Congress of the Econometric Society, Cambridge, England

Leser, C.E.V., "Family Budget Data and Price Elasticities of Demand" Review of Economic Studies, Vol.IX, 1941, pp 40-57

Lluch, C., A. A. Powell, Alan, A., and R. A. Williams, "Patterns in Household Demand and Saving, World Bank Research Publication, 1977, Oxford University Press for the World Bank, 1977, pp. xxxi, 280

Maji, C. C., D. Jha and L. S. Venkataraman, "Dynamic Supply and Demand Models for Better Estimations and Projections : An Econometric Study for Major Foodgrains in the Punjab Region, Indian Journal of Agricultural Economics, January-March 1971, Volume XXVI, pp 21-34

National Council of Applied Economic Research (NCAER) : Demand Forecast for Consumer Goods, New Delhi, 1959

______________, "Demand Projections for Foodgrains, Tobacco, Vegetable Oils and Cotton Textiles in India", 1960-61 to 1975-76, New Delhi 1960

______________, "Long Term Projections of Demand and Supply of selected Agricultural Commodities", New Delhi, 1962

Pollack, R.A., and T.J. Wales, "Estimation of Complete Demand System from Household Budget Data : The Linear & Quadratic expenditure systems" AER, June 1978, Volume 68 No.3, pp 348-359

Pushpam & A. Rudra, "Demand Elasticity for Foodgrains" Economic and Political Weekly, India, Vol.16, November 28, 1964, pp 1893-94

Radhakrishna, R. and K. N. Murthy, "Food Demand Model for India" Sardar Patel Institute of Economics and Social Research, Ahmedabad, 1978

Stone, R., "Linear Expenditure Systems and Demand Analysis : An Application to the Pattern of British Demand", Economic Journal, Vol.64, 1954, pp 511-527

Tendulkar, S. D. "Econometric Study of Monthly Consumption Expenditure in Rural Uttar Pradesh", American Journal of Agricultural Economics, Vol.51, 1969, pp 119-136

Theil, H., "Principles of Econometrics", New York, John Wiley andSons. Inc. 1971, pp 575-596 and pp. 328-330

Wallace, T. D. and A. Hussain, "The Use of Error Component Models in Combining Cross-Section with Time Series Data", Econometrica, 1969, vol.37, pp 55-72

Zellner, A., "An Efficient Method of Estimating Seemingly Unrelated Regression and Tests for Aggregation Bias", Journal of American Statistical Association Vol.57, 1962, pp 348-368

V. K. R. V. Rao et. al. eds. "Papers on National Income and Allied Topics" Vol.II, Asia Publishing House, Delhi, 1961

N. Bhattacharya for Indian Council for Social Science Research: A Survey of Research in Economics, Vol.7, Allied Publishers, New Delhi, 1978. pp. 176-232.

APPENDIX E

ESTIMATION PROCEDURES

The demand equations (i.e. DERP, DESP and SELP) are specified to incorporate all the restrictions of demand theory. It is worthwhile recalling that

1. The specifications of DERP, DESP and SELP satisfy the homogeneity constraint (i.e. that the demand equations are homogeneous of degree zero in income and prices) although homogeneity cannot be tested in DERP and DESP.

2. They satisfy the adding up constraint at sample means. In specifications SELP, if we did not have the squared income term, then the adding up property would be satisfied globally since the equation is in terms of shares which must sum to unity. However, in DERP and DESP, even without the squared income term, we can satisfy the adding up constraint only at sample means.

3. Symmetry conditions are imposed in the estimation procedure, and hold globally for all the specifications. However, in DERP, the symmetry constraint on the cross-price terms of the n^{th} equation are derived from the symmetry condition

$$\frac{\delta X_i}{\delta P_N} = \frac{\delta X_N}{\delta P_i} \qquad \text{E-1}$$

which leads to

$$C_{Ni} = - \sum_{j=1}^{n-1} C_{ij} \frac{\bar{P}_j}{P_N} \qquad \text{E-2}$$

However this constraint is imposed only at the sample mean.

4. Conditions (2) and (3) imply that in the specification SELP only n-1 equations need to be estimated, the adding-up (at sample means) and

symmetry constraint will ensure that all the parameters of the n'th equation are then derivable from the estimated parameters of the n-1 equations. Symmetry constraints give all the cross-price terms for the n'th equation, homogeneity constraints give the cross-price terms for any i, with respect to the n'th commodity, and the own-price term for the n'th commodity. Of course, the income elasticity for the n'th commodity will be derivable only at some given value of log m since (as in Table 1),

$$\eta_{nM} = - \sum_{i=1}^{n-1} \left(\frac{b_{i1} + 2\, b_{i2} \log m}{S_i} + 1 \right) \qquad \text{E.3}$$

Also note that we do not have estimates of b_{n1} and b_{n2} separately, but only of η_{nM} at some value of log m.

5. In specifications DERP and DESP, however, since the demand equations have quantities rather than shares as the dependent variable, we can estimate all the n equations with the symmetry and adding up constraints imposed on the estimation procedure. The specifications impose homogeneity in each equation by using relative prices. Since in DESP each equation uses its own price as a deflator this means that for each commodity, the own price term (and elasticity) will have to be derived according to the formula given in Table 1.

All the systems form sets of "seemingly unrelated" regression equations in the sense of Zellner and are estimated using the SAS package "SYSREG". The system estimated without the symmetry constraints, (and the adding up constraint in specification DERP and DESP) will be equivalent to estimation with single equation OLS, if single time series are used.

As mentioned earlier, the coefficients estimated in this system do not make much sense themselves, and have to be converted into elasticities for interpretation. The formulae for these elasticities are given in Table 1. It is worth noting here that the corresponding standard errors and t-values are also linear combinations of the variances and covariances of the estimated parameters. In general, if we have a row vector of regression coefficients for each equation

$$b_i = (b_{i1}, b_{i2}, \ldots, b_{iH}) \qquad \text{E.4}$$

where H = total number of independent variables, and a row vector of constraints

$$\delta_i = (\delta_{i1}, \delta_{i2}, \ldots \delta_{iH}) \qquad \text{E.5}$$

the elasticities are computed as the scalar

$$P_i = \delta_i b_i^T + D \qquad \text{E.6}$$

where D is a constant. The variance of P_i is

$$\text{Var } P_i = \delta_i \, \Sigma_{\hat{b}_i \hat{b}_i} \, \delta_i^T \qquad \text{E.7}$$

where $\Sigma_{\hat{b}_i \hat{b}_i}$ is the variance covariance matrix of the regression coefficients. The t-value is given by

$$t_i = \frac{P_i}{\sqrt{\text{Var } P_i}} \qquad \text{E.8}$$

Pooling of Cross-Sections of Time-Series

The data set consists of 200 observations across 10 states in India and over twenty years. In general, we can expect that the classical assumptions of normal and independently distributed error, with zero mean and constant variance will not be satisfied when the observation stretch across two variational directions. We can expect for example, that the variation in the time direction is more or less than in the regional direction, and hence, the error term may also display the same variational characteristics as the dependent variable.

One way of accounting for these different effects and hence cleaning up the errors has been to introduce dummy variables for the regions and for time. This allows for constant time and region effects. The same thing can be accomplished by transforming the variables so that they are expressed as deviations from the means, provided the sample is balanced. Specifically, any variable Y_{it} (whether dependent or independent) is transformed so that the transformed variable is

$$Y^*_{it} = Y_{it} - \bar{Y}_{\cdot t} - Y_{i\cdot} + \bar{Y}_{\cdot\cdot} \qquad \text{E.9}$$

where the dot-bar notation indicates average over the suppressed subscript. The transformed variables can then be used in estimating the system of demand equations, using the Zellner procedures. The covariance transformation yields estimates of the regression parameters without estimating the coefficients of the dummy variables, and thus saving on degrees of freedom.

The error-component model has been used in recent years for pooling cross-section and time series data. This model assumes that the region and time effects are not fixed but random, are independently distributed, with zero means and (usually) positive variances. If the estimated model is

$$Y_{it} = \alpha + \beta_1 X_{1it} + \beta_2 X_{2it} + \varepsilon_{it}, \qquad \begin{array}{l} i = 1, n \\ t = 1, T \end{array} \qquad \text{E.10}$$

and $\varepsilon_{it} = \mu_i + \nu_t + \eta_{it}$

then μ_i , ν_t and η_{it} have variances σ^2_μ , σ^2_ν and σ^2_η respectively

Basically, therefore, this treats the intercept terms as random and is an intermediate solution to treating them all as different (Least-Squares with dummy variables) or treating them all as equal (OLS). The advantage of using the error-component model over the covariance transformed regressions is that, while the latter are consistent, we can get more efficient estimates from generalized least square estimates of the error-component model.

Originally, Wallace and Hussein derived the formulae for the error components from the residuals of the OLS regressions.

$$\tilde{\sigma}^2_\eta = \frac{S.\ S.\ E}{(N-1)\ (T-1)} \qquad \text{E.11}$$

$$\tilde{\sigma}^2_\mu = \frac{1}{N} \left(\sum_t \frac{\varepsilon._t{}^2}{N(T-1)} - \tilde{\sigma}^2_\eta \right) \qquad \text{E.12}$$

$$\tilde{\sigma}_{\nu}^{2} = \frac{1}{T} \left(\sum_{i} \frac{\hat{\varepsilon}_{i\cdot}}{T(N-1)} - \tilde{\sigma}_{\eta}^{2} \right) \qquad \text{E.13}$$

However, Amemiya uses estimates of the β's from the covariance transformed regression to calculate residuals. But in the covariance transformed equations, by construction, $\hat{\eta}$ and $\hat{\eta}_{\cdot t}$ are zero for all i and t, where $\hat{\eta}$ indicates the residuals from the equation. Hence to calculate the residuals in the formulas for the error components above, Amemiya uses the β's from the covariance transformed equations, but on the original X, Y data, not the transformed data. (According to a simulation study by Maddala and Mount, the two different estimators have virtually the same small sample properties and do equally well as several others that they examine).

These error components are used to compute the second round generalized least square estimates. Nerlove obtains the inverse of the error covariance matrix in a way that allows for transforming the original data. He defines the four distinct characteristics roots of the residual variance covariance matrix in terms of the ratios of the estimated error components.

First, if the total error is

$$\hat{\sigma}^2 = \hat{\sigma}_{\mu}^2 + \hat{\sigma}_{\nu}^2 + \hat{\sigma}_{\eta} \qquad \text{E.14}$$

Nerlove defines as

$\rho = \hat{\sigma}_{\nu}^2 / \sigma^2$, and $\omega = \hat{\sigma}_{\mu} / \sigma^2$

The four distinct characteristic roots are

$$
\begin{aligned}
&\text{(i)} && \lambda_1 = 1-\rho-\omega+\omega N + \rho T \\
&\text{(ii)} && \lambda_2 = 1-\rho-\omega+\omega N \\
&\text{(iii)} && \lambda_3 = 1-\rho-\omega + \rho T \\
&\text{(iv)} && \lambda_4 = 1-\rho-\omega
\end{aligned}
\qquad \text{E.15}
$$

The GLS estimates are given by regressing the transformed dependent variable on the transformed independent variables, where the variable Y_{it} is transformed in the following manner.

$$
Y^{**}_{it} = Y_{it} - \left(1-\frac{\sqrt{\lambda_4}}{\sqrt{\lambda_3}}\right)\bar{Y}_{i.} - \left(1-\frac{\sqrt{\lambda_4}}{\sqrt{\lambda_2}}\right)\bar{Y}_{.t}
$$

$$
+ \left(1 - \frac{\sqrt{\lambda_4}}{\sqrt{\lambda_2}} - \frac{\sqrt{\lambda_4}}{\sqrt{\lambda_3}} + \frac{\sqrt{\lambda_4}}{\sqrt{\lambda_1}}\right)\bar{Y}_{..} \qquad \text{E.16}
$$

The β estimates are asymptotically normal and consistent, and the estimated variances are consistent estimators of the variances in the limiting distribution. Therefore, all tests are valid asymptotically.

In general, the variance components are assumed to be positive. Wallace suggests that if one of the estimates is negative, then it can be assumed to be zero. Further, the gain in efficiency is of course dependent on the value of the components. If they are zero, or close to zero, then either OLS or the covariance transformed regressions techniques are adequate.

Fuller and Battese use a "fitting-of-constants" method for estimating the variance components. Given the original data, the total sum of squares of errors in E.11 above is computed as the residual sum of squares from the regressions of the covariance transformed variables Y^{*}_{it} as before. The sum of squares $\hat{\varepsilon}^2_{\cdot t}$ in E.12 is obtained by regressing the variable transformed as follows:

$$Y^{+}_{it} = Y_{it} - \bar{Y}_{\cdot t} \qquad \text{E.17}$$

and the sum of squares $\hat{\varepsilon}^2_{i\cdot}$ is obtained by regressing the variables transformed as follows

$$Y^{++}_{it} = Y_{it} - \bar{Y}_{i\cdot} \qquad \text{E.18}$$

The SAS Institute does have a package programme for estimating this error-component model. However, the package is usable only for the estimation of single equations, and at this stage, it is not possible to use this for estimating the system of equations. Therefore, we have used the formulas suggested by Amemiya to estimate the variance components, and the procedure outlined above to transform the original variables to get GLS estimates of the β's

Avery has developed a procedure for extending the error-component model to the case where the error components are allowed to be correlated across equations (as in the Zellner's procedure). The procedure is very complex and given the nature of our data, and our needs, we have decided that it is probably not worthwhile to go through the procedure.