LSM - 11
OCT. 1981

Living Standards
Measurement Study
Working Paper No. 11

Three Essays on a Sri Lanka Household Survey

Angus Deaton

LSMS Working Papers

No. 1 *Living Standards Surveys in Developing Countries*

No. 2 *Poverty and Living Standards in Asia: An Overview of the Main Results and Lessons of Selected Household Surveys*

No. 3 *Measuring Levels of Living in Latin America: An Overview of Main Problems*

No. 4 *Towards More Effective Measurement of Levels of Living, and Review of Work of the United Nations Statistical Office (UNSO) Related to Statistics of Levels of Living*

No. 5 *Conducting Surveys in Developing Countries: Practical Problems and Experience in Brazil, Malaysia, and the Philippines*

No. 6 *Household Survey Experience in Africa*

No. 7 *Measurement of Welfare: Theory and Practical Guidelines*

No. 8 *Employment Data for the Measurement of Living Standards*

No. 9 *Income and Expenditure Surveys in Developing Countries: Sample Design and Execution*

No. 10 *Reflections on the LSMS Group Meeting*

No. 11 *Three Essays on a Sri Lanka Household Survey*

No. 12 *The ECIEL Study of Household Income and Consumption in Urban Latin America: An Analytical History*

No. 13 *Nutrition and Health Status Indicators: Suggestions for Surveys of the Standard of Living in Developing Countries*

No. 14 *Child Schooling and the Measurement of Living Standards*

No. 15 *Measuring Health as a Component of Living Standards*

No. 16 *Procedures for Collecting and Analyzing Mortality Data in LSMS*

No. 17 *The Labor Market and Social Accounting: A Framework of Data Presentation*

No. 18 *Time Use Data and the Living Standards Measurement Study*

No. 19 *The Conceptual Basis of Measures of Household Welfare and Their Implied Survey Data Requirements*

No. 20 *Statistical Experimentation for Household Surveys: Two Case Studies of Hong Kong*

No. 21 *The Collection of Price Data for the Measurement of Living Standards*

No. 22 *Household Expenditure Surveys: Some Methodological Issues*

No. 23 *Collecting Panel Data in Developing Countries: Does it Make Sense?*

No. 24 *Measuring and Analyzing Levels of Living in Developing Countries: An Annotated Questionnaire*

No. 25 *The Demand for Urban Housing in the Ivory Coast*

No. 26 *The Côte d'Ivoire Living Standards Survey: Design and Implementation*

No. 27 *The Role of Employment and Earnings in Analyzing Levels of Living: A General Methodology with Applications to Malaysia and Thailand*

(List continues on the inside back cover)

Three Essays on a Sri Lanka Household Survey

The Living Standards Measurement Study

The Living Standards Measurement Study (LSMS) was established by the World Bank in 1980 to explore ways of improving the type and quality of household data collected by Third World statistical offices. Its goal is to foster increased use of household data as a basis for policy decision making. Specifically, the LSMS is working to develop new methods to monitor progress in raising levels of living, to identify the consequences for households of past and proposed government policies, and to improve communications between survey statisticians, analysts, and policy makers.

The LSMS Working Paper series was started to disseminate intermediate products from the LSMS. Publications in the series include critical surveys covering different aspects of the LSMS data collection program and reports on improved methodologies for using Living Standards Survey (LSS) data. Future publications will recommend specific survey, questionnaire and data processing designs, and demonstrate the breadth of policy analysis that can be carried out using LSS data.

LSMS Working Papers
Number 11

Three Essays on a Sri Lanka Household Survey

Angus Deaton

The World Bank
Washington, D.C., U.S.A.

First printing October 1981
Second printing July 1985

When this paper was first published Angus Deaton was professor of Econometrics at the University of Bristol, England.

Library of Congress Cataloging in Publication Data

Deaton, Angus.
Three essays on a Sri Lanka household survey.

(LSMS working paper, ISSN 0253-4517 ; no. 11)
"October 1981."
Bibliography: p.
1. Cost and standard of living—Sri Lanka. 2. Household surveys—Sri Lanka. 3. Quality of life—Sri Lanka. 4. Food consumption—Sri Lanka. 5. Children—Economic aspects—Sri Lanka. I. Title. II. Series.
HD049.8.D4 1985 339.4'1'095493 85-11535
ISBN 0-8213-0038-5

THREE ESSAYS ON A SRI LANKA HOUSEHOLD SURVEY

TABLE OF CONTENTS

LIST OF TABLES

Table No.		Page No.

ESSAY III

I. ANALYZING THE FOOD SHARES IN A HOUSEHOLD SURVEY

A. INTRODUCTION

Household surveys from developing countries can be used to illuminate welfare issues in a variety of different ways. This essay explores one possibility, the analysis of the share of food in household expenditure. Such studies are of limited interest unless they can be replicated in a number of different countries, even though the details of the analysis will inevitably vary from country to country and from survey to survey. The best approach seems to be to work illustratively, carrying out substantive studies designed in such a way as to be easily replicated elsewhere. Here we report results on one such study, based on the 1969-70 socio-economic survey of Sri Lanka and carried out as part of the World Bank's Living Standards Measurement Study.

The food share has been emphasized in studies of household welfare at least since the nineteenth century. Since food is seen as the first necessity, the demand for which rises much less rapidly than do resources, at least once subsistence needs are met, the share of food in total expenditure can be regarded as an (inverse) indicator of welfare. It is also a very convenient indicator, since its definition as a dimensionless ratio renders it comparable over time periods and between geographical locations, at least if the relative price of food does not vary too much. However, the real interest in the food share is that it may be capable of acting as a better indicator of welfare than measures based on income or expenditure alone. Once again, this is an old idea going back at least to Engel who noticed that larger households had larger food shares than smaller families with the same total outlay and suggested that households with the same food share be regarded as equally well-off irrespective of their size. This idea will be directly exploited later in the paper, but the principle is of wide potential applicability for measuring factors contributing to welfare.

In principle it would be more desirable to analyze the complete structure of the household budget rather than focussing on the foodshare alone. For LDC's, however, the food share is typically between 50% and 80%, ratios two to four times higher than those currently existing in the U.S. or the U.K. Hence, other individual categories of consumption, even if quite broadly defined, play a correspondingly limited role. It is also likely that their measurement is subject to much larger proportionate errors. Thus, relatively little of the <u>general</u> information about welfare is likely to be lost by looking at food alone. Of course, specific issues are likely to require attention to other items; housing is the obvious example. However, the increase in technical difficulty and presentational complexity in going from one commodity share to many is unlikely to be matched by a corresponding increase in information.

The remainder of the paper is divided as follows. Section B discusses the basic presentation of the data and illustrates the type of tabulation which can be used to descry the basic features of the food share in a large data set such as that provided by the Sri Lankan survey. Section C moves from the data to simple bivariate graphical and tabular analysis linking the food share with the most obvious indicator of welfare, household <u>per capita</u> expenditure. A transformation of the data is suggested which generates rough linearity in the relationship, simple bivariate regressions are estimated, and first attempts made at assessing the total expenditure elasticity of food demand as well as the shape of the Engel curve.

Section D turns to the relationship between food consumption and household composition, to the question of how household size and its age and sex composition affect the relationship between food consumption and total outlay. Again, several techniques are illustrated, from analyzing separately households with different compositions, to attempting to combine the various effects into a multivariate regression. In particular, attention is focussed on the sex, age, and numbers of adults, children and old people in the household. For Sri Lanka, and this particular data set, it seems that, in addition to per capita expenditure, only the separation of children and adults contributes very much to the explanation of the variance of the foodshare. Even so, it also appears that there are economies of scale in the costs of maintaining a household, particularly for adults and particularly for better-off families. Section D carries these descriptive results a stage further and uses the results together with Engel's method to calculate behavioral adult equivalence scales for different types of Sri Lankan families. Again, different methodologies are discussed together with their implications for the constructed scales. In this particular study, and in line with the economies of scale already mentioned, it appears that, for a given household composition, better-off families have lower scales. Hence, correction of per capita expenditures to an <u>equivalent</u> per capita basis would have two offsetting effects on measured inequality; on the one hand, the economies of scale would benefit larger families more and these typically have lower levels of per capita expenditure, while, on the other hand, such corrections are larger for the better-off households.

Finally, there are three appendices. Appendix 1 covers econometric issues, specifically the question of pooling samples from different sectors or regions and also the effects of residual heteroskedasticity on assessing the multivariate regressions reported in the text. Appendix 2 contains the technical mathematical theory supporting the equivalence scale methodology in Section D. It is unnecessary to an understanding of that section but is designed to show that the scales can be supported by the relevant consumer theory. Appendix 3 gives a brief list of the minimal data requirements needed to replicate this study.

B. PRESENTATION OF THE DATA

1. Background

The 1969-70 socio-economic survey of Sri Lanka is fully described in the two volumes published by the Sri Lanka Department of Census and Statistics (1973). This publication also contains the original questionnaire as well as numerous summary tables and cross-tabulations. A total of 9,694 households were covered by the survey and information was collected, not only on income and households' social, economic and demographic characteristics, including, essentially for present purposes, the age and sex of each member of the household. One-member households were excluded, as were boarding houses and institutions. For the remainder, a two-stage sampling design was adopted with census blocks as primary, and households as secondary sampling units. The island as a whole was stratified into three sectors, urban, rural and estate, with selection probabilities set to favor the inclusion of urban households. Rows 1 and 2 of Table 1 give the total number of households by sector together with those included in the sample. These disproportionate weights, particularly in favor of urban households, must be borne in mind when computing population figures from the sample. However, the three sectors differ so widely in type and in economic characterization that it is wise, for Sri Lanka, to analyze each stratum or sector separately. Whether or not valid pooling can take place is addressed as a separate issue later in the paper.

Table 1

Sectoral Composition of the Sample

	All Island	Urban	Rural	Estate
Total Households	2,372,197	375,042	1,695,610	301,545
In sample	9,694	4,037	3,657	2,000
In analysis	9,663	4,022	3,652	1,989

For the current analysis, some 31 households were excluded, leaving an effective data base of 9663 households, the sectoral composition of which is shown in the third row of Table 1. Exclusion was forced when either food or total expenditure data were unusable. Apart from cases of unavailability, the basic requirement that the share lay between zero and unity was checked for each household. More sophisticated editing techniques might well have been worthwhile but were not carried out in the present case. Even so, it is worth emphasizing the importance of this type of checking or editing prior to analysis. Even simple statistical analysis is computationally expensive when nearly 10,000 observations are involved and it is extremely costly to repeat calculations when errors are discovered at a late stage in the analysis. However, even the most sophisticated editing techniques cannot detect all errors, and an important second line of defense is the use of robust techniques at early stages of the analysis. Specifically, graphical analysis, such as scatter diagrams, will reveal outliers while not allowing them to lead to incorrect inference. Regression analysis without graphical or other display can be extremely sensitive and a single nonsensical observation can easily dominate the results even in a regression with many thousands of observations.

2. The Variables

The principle variables in the analysis, i.e., food, and total and per capital household expenditure, are largely self-explanatory; nonetheless, there are a number of definitional points. Food is "food and drink," drink covering tea, coffee, milk, etc. but excluding liquor. Tobacco and betel are also excluded. In the survey, food consumption was monitored over a seven-day period, but the figures are grossed up to a monthly basis to match other expenditure categories. Expenditure on durable goods was measured on an annual basis and divided by twelve; this at least partially avoids the "bumps" which otherwise would occur in total expenditure if the sample period happened to coincide with the purchase of an expensive durable. During 1969-70, every person over the age of one received a free rice ration (or its equivalent) and the imputed value of these is included in food consumption and hence in total expenditure. Other important imputed elements in total expenditure are the rental value of owner occupied housing as well as the value of free housing, the latter mainly on the estates. Production for own consumption is also valued and added both to food and total expenditure.

The use of imputed values in the analysis raises potentially difficult theoretical and practical questions. Pricing non-marketed commodities at market prices implicitly assumes that consumers would have bought those commodities at the prevailing prices had they not obtained them by other means, for example, through the free rice ration or free housing. If indeed it is possible for the consumer to sell unwanted rations or home produce, the assumption is a reasonable one. For the rice ration, few problems arise. The ration was about half of average consumption in 1969-70 and even very poor households bought rice outside the ration. However, for very poor consumers on the estates (and possibly elsewhere) the imputation of value for free housing may cause distortions in the behavior of the food share. Intuitively, it is reasonable to expect that, for very poor households, a very large fraction of spendable resources will be spent on food and that this fraction will tend to increase as resources decrease. The relationship between the food share and outlay would thus be as illustrated by AB in Figure 1. If households near A are allocated free housing, they are likely to wish to sell it or exchange it

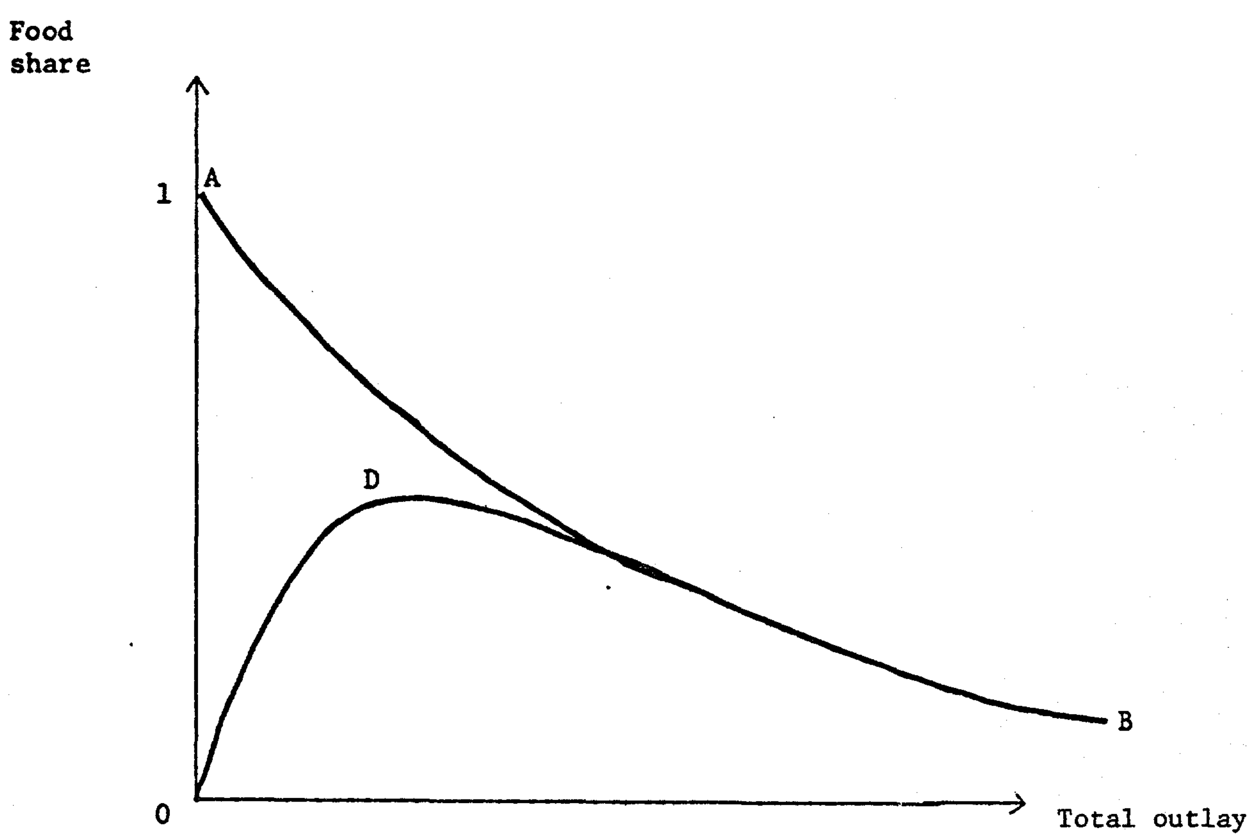


Figure 1: The effect of imputing value to free housing for very poor consumers

for food. In practice, they cannot do so, so that the shadow price to the recipients is in fact much lower than the market price which is assigned by the survey accountants. Algebraically, if x is spendable resources, $r(x)$ is expenditure on non-food items, and a is the imputed value of housing, the food share, w, is given by

$$w = (x - r(x))/(x + a). \tag{1}$$

For the very poor, $r(x)$ and its rate of change with x are very small, so that $w \simeq x/(x + a)$ which tends to zero with x so that the food share, at least initially, is a rising function of outlay. This is illustrated by the curve ODB in Figure 1. Of course, the extreme points near 0 are not observed, but, as will be shown, there is a distinct tendency in the Sri Lankan data for the food share to stop rising and flatten out at low levels of total outlay. Since, on the estates, average imputed income from housing is about 5% of average total income, this phenomenon may be of considerable importance lower in the income distribution and clearly deserves further investigation.

Per capita total household expenditure (PCE) is the basic variable used in the study to explain variations in the food share. For reasons of accuracy of measurement, as well as the other issues discussed, for example in Deaton (1980), expenditure is preferred to income as the first approximation to welfare. Similarly, the deflation from total household expenditure to per capita expenditure is a first attempt to measure expenditure in relation to needs. More sophisticated corrections will be attempted in Sections C and D below.

3. Univariate Analysis

Table 2 lists the means and standard deviations of the food share, per capita expenditure and the logarithm of per capita expenditure. The All-Island figures are weighted averages using the population weights of the three sectors, from Table 1. Note the relative poverty of the rural and estate sectors, particularly the latter, whether assessed on the basis of PCE or of the food share. Dispersion is also less away from the urban sector, both in PCE and in the food share. Much of the positive skewness in the PCE distribution is removed by taking logarithms so that the standard deviations of lnPCE are more useful indicators of dispersion than those of PCE itself.

Table 2

Means and Standard Deviations of Food Shares PCE and lnPCE by Sector (Rupees per month)

	Foodshare		PCE		lnPCE	
	mean	s.d.	mean	s.d.	mean	s.d.
Urban	0.614	0.138	78.72	54.91	4.193	0.565
Rural	0.662	0.118	55.57	30.40	3.899	0.474
Estate	0.676	0.115	53.64	27.66	3.889	0.414
All Island	0.656	0.122	58.98	36.17	3.944	0.494

Table 3 provides a greater disaggregation of the food share, again by sectors. The model class is a food share from 70-75% of total expenditure in all three sectors. Also in all three sectors, the distributions are slightly negatively skewed. This corresponds to a "tail" of very well-off consumers with low budget shares corresponding to the long upper tail of the PCE distribution. This negative skewness is particularly marked in the urban sector corresponding to the greater inequality of PCE as compared with the country and the estates. The higher level of PCE in the urban sector is also apparent from the distribution: 21% of urban households have food expenditures of less than 50% of total expenditure while the corresponding figures for rural and estate are

10% and 8%, respectively. At the other end of the distribution, amongst the very poor, 14% of households on the estates (43,000 households) and 10% in the rural sector (176,000 households) have food shares greater than 80% compared with only 6 1/2% in the urban sector (24,000 households). Even this probably understates relative deprivation on the estates where, with largely free accommodation, the imputation phenomenon discussed above prevents the food share attaining even higher levels.

C. THE RELATIONSHIP BETWEEN THE FOOD SHARE AND PCE

Figure 2 is a (somewhat rough) sketch of the empirical joint distribution of the food share and per capita expenditure over all the Sri Lankan households in the sample. Ideally, the corresponding diagrams for each sector should be examined separately, but as will be seen, the relationship between the share and PCE is sufficiently stable across the three sectors for a single diagram not to be misleading. This sketch is prepared from a standard scatter diagram and the contour lines indicate the density of households falling in the various areas with the inner contours corresponding to the higher densities. The isolated crosses marked are intended to give an idea of the occurrence of such isolated observations without faithfully reproducing them on a one for one basis.

The steepness of the contours close to the vertical axis reflects the sharp drop in numbers of households with PCE below the mode, but note the very wide range of food shares associated with any given level of PCE, even when the latter takes on extremely low values. Clearly, if the food share turns out to be a good welfare-ranking device, it will give results rather different from those which would be obtained using PCE. Even so, the empirical relationship

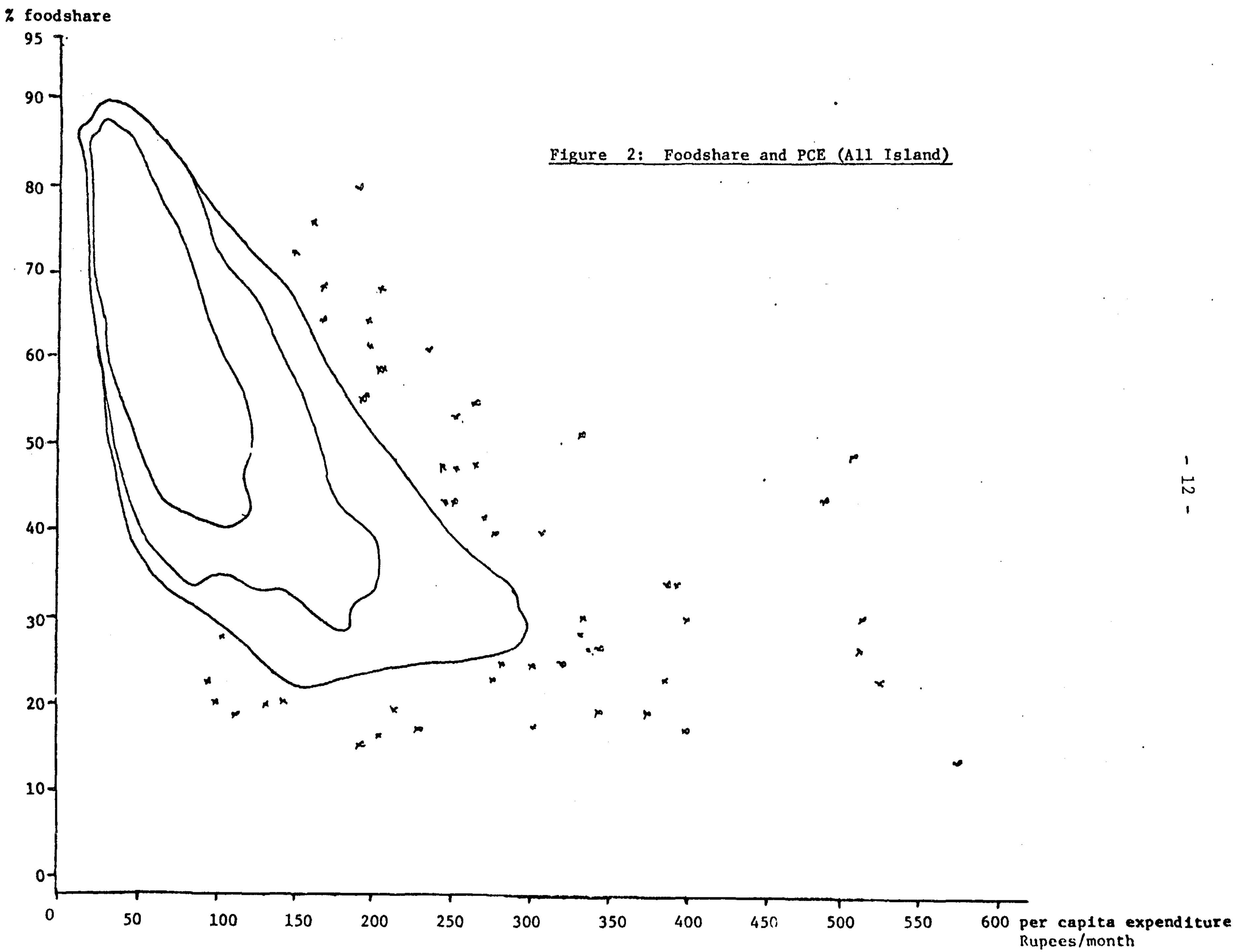
Table 3

Frequency Distribution (%) of Households

Classified by Foodshare and Sector

Foodshare	Frequency Urban	Frequency Rural	Frequency Estate	Cumulative Frequency Urban	Cumulative Frequency Rural	Cumulative Frequency Estate
≤ .15	0	0	0	0	0	0
≤ .20	0.2	0	0.1	0.2	0	0.1
≤ .25	0.4	0.1	0.2	0.6	0.1	0.2
≤ .30	1.3	0.2	0.2	2.0	0.4	0.4
≤ .35	2.2	0.8	0.4	4.2	1.1	0.7
≤ .40	3.7	1.6	0.9	7.8	2.7	1.6
≤ .45	6.0	2.8	2.4	13.8	5.5	4.0
≤ .50	7.2	4.8	3.9	21.0	10.3	7.9
≤ .55	9.8	6.5	6.1	30.8	16.8	14.0
≤ .60	11.1	10.8	10.0	41.9	27.6	24.0
≤ .65	13.3	13.4	13.7	55.2	41.0	37.7
≤ .70	14.2	17.6	15.7	69.4	58.6	53.4
≤ .75	14.2	16.5	17.9	83.6	75.0	71.4
≤ .80	9.9	14.6	14.5	93.5	89.6	85.8
≤ .85	4.9	8.0	10.3	98.4	97.6	96.1
≤ .90	1.5	2.1	3.7	99.9	99.7	99.7
≤ .95	0.1	0.2	0.3	100.0	100.0	100.0

Figure 2: Foodshare and PCE (All Island)



between the food share and PCE is strong and inverse. It also has a considerable (convex) curvature as might be expected from the strong positive skewness in the marginal distribution of PCE and the slight negative skewness in the marginal distribution of the food share.

Much of the non-linearity in Figure 2 can be removed by the use of a simple logarithmic transformation. Figure 3 represents the same joint distribution as in Figure 2 but with the horizontal axis on a logarithmic scale. Clearly, the relationship is now much closer to a linear one, although, as will be demonstrated later, there is still some curvature remaining. However, if linearity is taken as a first approximation, Figure 3 suggests a relationship of the form

$$w = \beta_o + \beta_1 \ln PCE + u \tag{2}$$

where w is the food share, u is an error, lnPCE is the (natural) logarithm of PCE, and β_o and β_1 are intercept and slope parameters respectively.

This relationship, although here suggested entirely on empirical grounds, is both extremely convenient and of rather distinguished ancestry. It was first used for empirical Engel curve analysis by Holbrook Working (1943) and later recommended by Leser (1963) as a simple functional form which performed well in competition with other specifications. Also, unlike most of the traditional Engel curves, such as those examined by Prais and Houthakker (1955), it is possible for all goods to conform to the Working-Leser specification without violating the constraint that the sum of all budget shares be unity. This is sufficient for the model to be made consistent with the standard theory of consumer behavior with all its attendant apparatus for welfare analysis. Less essentially for present purposes, the model also possesses extremely elegant properties when applied to aggregates

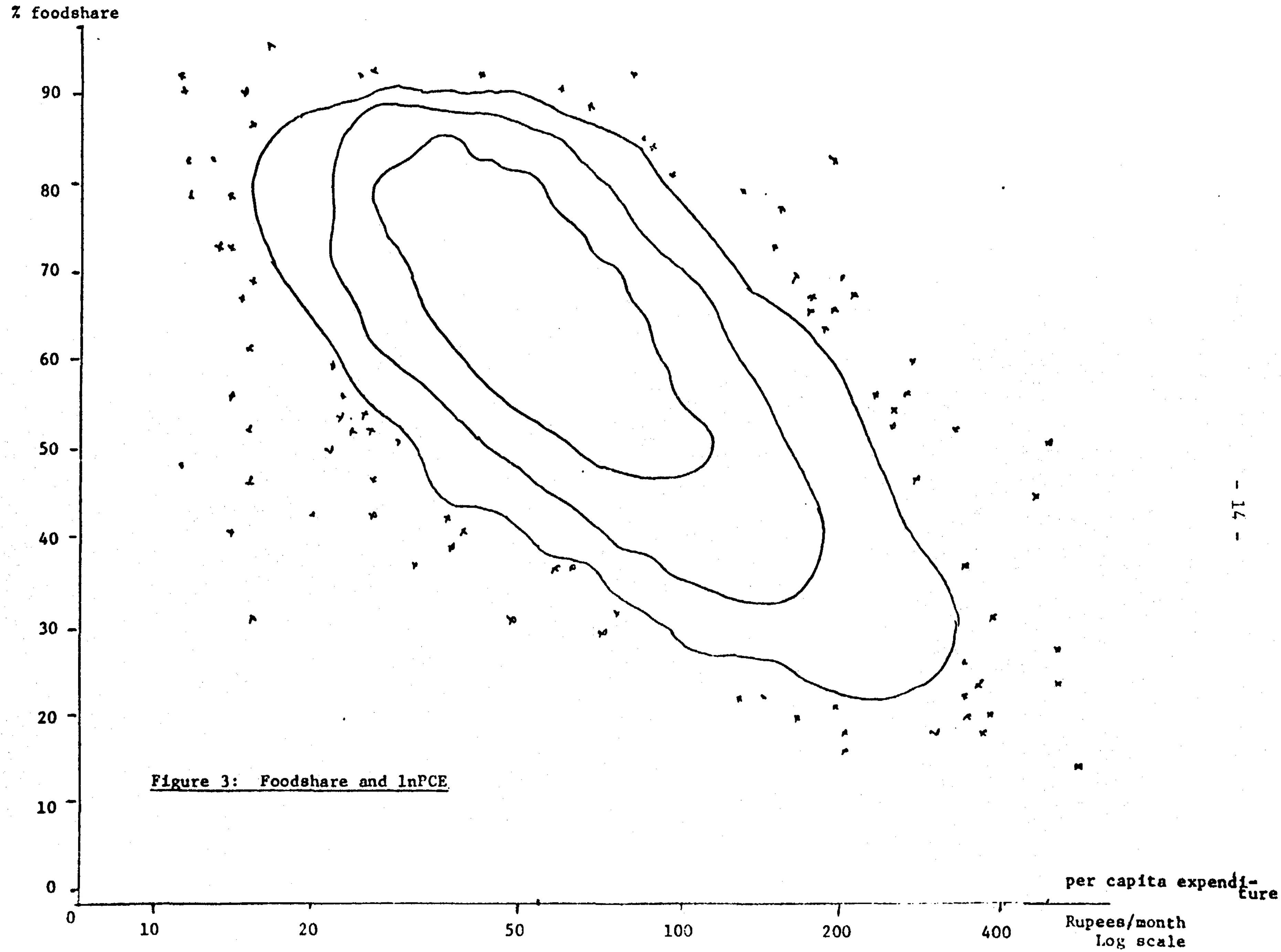


Figure 3: Foodshare and lnPCE

of consumers. For more details and further discussion of the model see Deaton and Muellbauer (1980, Chapters 1, 3 and 5).

Equation (2) may be estimated by ordinary least squares to yield:[1/]

<u>Urban</u>

$$\hat{w} = \underset{(127.1)}{1.3905} - \underset{(-71.6)}{0.18482} \ln PCE \qquad R^2 = 0.561 \quad \hat{\sigma} = .0917 \tag{2a}$$

<u>Rural</u>

$$\hat{w} = \underset{(97.1)}{1.2606} - \underset{(-46.5)}{0.15310} \ln PCE \qquad R^2 = 0.372 \quad \hat{\sigma} = .0933 \tag{2b}$$

<u>Estates</u>

$$\hat{w} = \underset{(58.2)}{1.2322} - \underset{(-26.4)}{0.14270} \ln PCE \qquad R^2 = 0.260 \quad \hat{\sigma} = .0987 \tag{2c}$$

where $\hat{w}$ is the predicted value of w by the regression line, $\hat{\sigma}$ is the equation standard error (compare with the mean values of w in Table 2) and the numbers in brackets are t-values. The equation standard error is (significantly) larger on the Estates even though (see Table 2) the variability of the food share there is less; the combination of these two effects produces the low R^2. Clearly, there are other important factors influencing the food share in that sector. By contrast, in the urban sector, the high R^2 (compared, say, with the rural sector) is largely a consequence of the greater inequality there and the consequent high variance of the share.

1/ Note that for the predicted food-share to lie between 0 and 1, ln PCE and PCE must lie within a limited range. It may easily be checked that the range given by the estimated equations is very wide and that no household in the sample lies outside it.

These results can be used to make first estimates of the total expenditure elasticity of food. From (2), the derivative of lnw with respect to the logarithm of total expenditure is β_1/w which, in turn, is the elasticity e less unity. Hence, e is calculated from

$$e = \beta_1/w + 1 \tag{3}$$

if we use a value of 0.70 for w as (roughly) model for all three sectors, the elasticity estimates are 0.736, 0.781, 0.796 for urban, rural and estates respectively. It should be noted that formula (3) implies that increasing total outlay will cause the elasticity to fall (since $\beta_1 < 0$ for a necessity which also means that w falls with PCE). However, in the present calculations, the same value (0.7) was used for w in (3) for all three sectors. Hence the results suggest that β_1 itself may be a declining function of PCE. The next section will confirm this interpretation.

D. DEMOGRAPHIC EFFECTS AND MULTIVARIATE ANALYSIS

1. Family Size and Composition

In Section B per capita expenditure was used as the main indicator of welfare and the variable determining the food share. In this section, rather more sophisticated constructions are considered. Clearly, PCE is likely to be more satisfactory than total household expenditure (THE) in that some allowance is made for household size. However, a crude head count is likely to overstate family needs. In particular, children may have lesser needs than adults, and social customs may indicate unequal allocations among adults, for example, as between men and women. At the same time, there are likely to be economies of scale in the costs of maintaining a household, if only because of the presence of overhead costs independent of family size and which provide facilities shared by all members.

Moreover, the presence of several adults and of older children in the household may generate services many of which are not monetized or officially imputed and so do not show up in either THE or PCE. Housemaking services, childminding, and so on are the obvious examples. To the extent that these consequences of larger families generate welfare unrepresented in THE or PCE, the food share may be lower than predicted from PCE alone. Stated another way, the apparent costs associated with larger families are lower than their head counts would suggest. Considering these factors together presents a rather complex picture of the relationship between household size and needs; it is undoubtedly too simple to regard household needs simply as a sum of individual needs, males counting as unity, females as something rather less, and children a fraction of unity depending on their age.

One relatively straightforward possibility is to relate the food share to total household expenditure for different family types separately, in the hope that, for a given family type, compositional effects will be held constant. Figure 4 illustrates three (independently estimated) regression lines for three different household types in the urban sector (2 adults and 2 children, 4 adults and 1 child, and 2 adults and 0 children). Ideally, children should be disaggregated by age, but this would reduce the numbers of households in each class to unacceptably low figures. Even so, the results shown are rather interesting. All three regressions have very similar slopes, so that although the smaller family has a smaller food share at all levels of THE, most of this can be ascribed to the difference in intercept. The regression lines for the two large family types are very close, suggesting that, for these two groups, PCE is a good enough determinant of the food share, in spite of their differences in composition. Further, if each of those family types has THE increased by a factor of 2.5, so

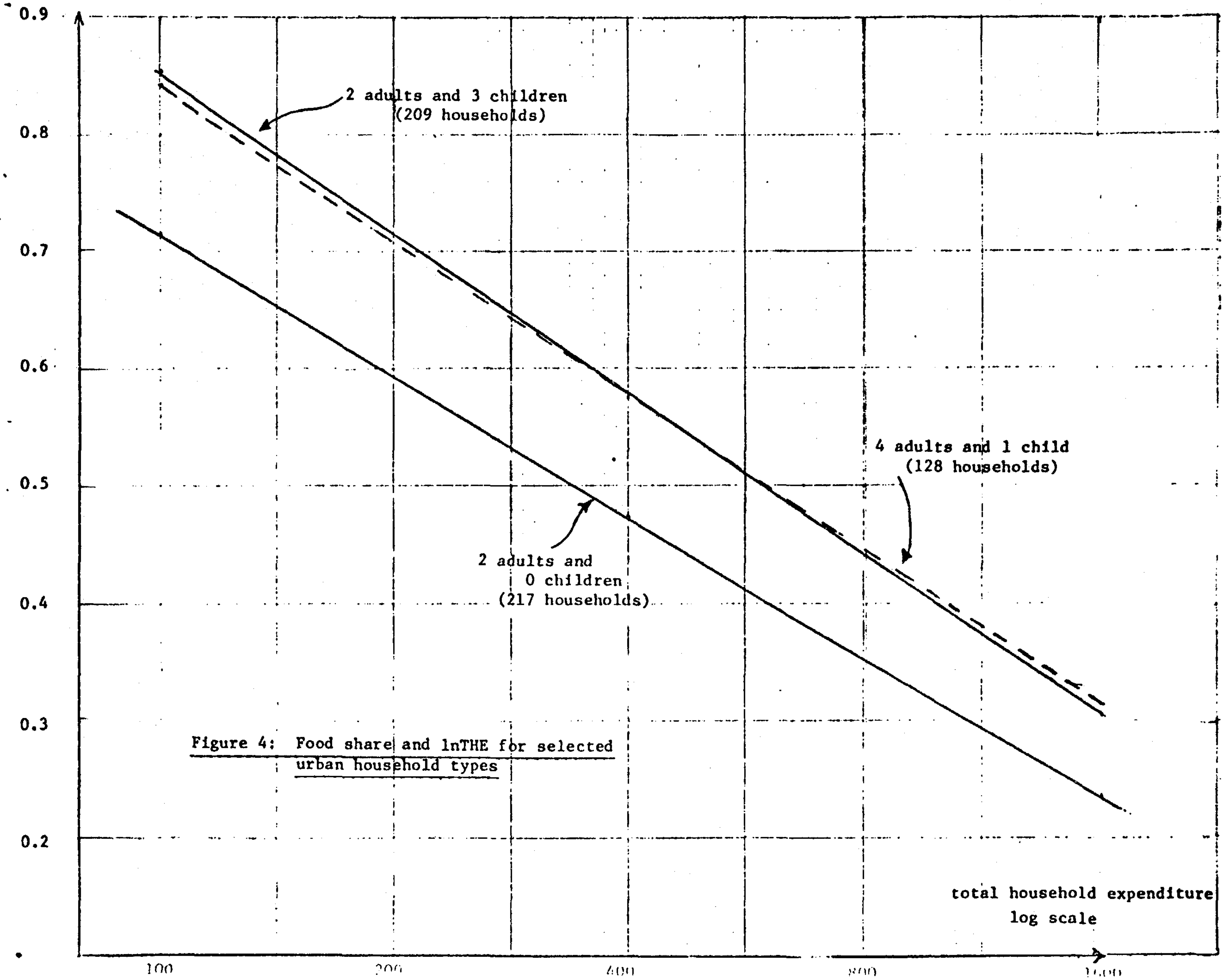


Figure 4: Food share and lnTHE for selected urban household types

that they have the same PCE as the smaller family, the two higher lines are moved down by 0.12 throughout, which brings all three lines very close together. This rather limited evidence, then, suggests that PCE works surprisingly well, at least as a first approximation.

However suggestive, it is not really practical to analyze each family type separately, at least in Sri Lanka. In many developed countries, a small number of family types would cover a large proportion of households, but this is not true of LDC's. Table 4 gives the figures for the three sectors of Sri Lanka. Although two adult families are the most common, the dispersion over types is very wide so that in order to fit separate Engel curves for the most important types, a large number of different cases would have to be considered. And even this would take no account of sex or of children's ages. Clearly, a more economical approach must be adopted.

Table 4: Adult† and child† composition of Sri Lankan families by sector (%)

1. **Urban**

% of households (4022)

No. of adults \ No. of children	0	1	2	3	4	≥5	Σ
1	*	0.3	0.2	0.2	0.1	0.2	1.0
2	5.4	4.6	4.9	5.2	3.7	4.8	28.6
3	4.7	3.5	3.3	2.7	2.6	3.0	19.8
4	3.8	3.2	2.9	2.6	1.9	2.7	17.1
5	3.0	2.5	2.1	2.4	1.6	1.9	13.5
≥6	4.3	4.1	3.3	3.1	2.2	2.8	19.8
Σ	21.2	18.2	16.7	16.2	12.1	15.4	100

% of persons (25358)

No. of adults \ No. of children	0	1	2	3	4	≥5	Σ
1	*	0.1	0.1	0.2	0.1	0.3	0.7
2	1.7	2.2	3.2	4.2	3.6	5.9	20.8
3	2.3	2.2	2.6	2.6	2.9	4.3	16.9
4	2.5	2.5	2.8	2.9	2.5	4.2	17.3
5	2.4	2.4	2.3	3.1	2.3	3.3	15.8
≥6	4.8	5.2	4.0	5.0	4.0	5.6	28.6
Σ	13.7	14.6	15.0	17.9	15.3	23.6	100

2. **Rural**

% of households (3652)

No. of adults \ No. of children	0	1	2	3	4	≥5	Σ
1	*	0.4	0.4	0.1	0.3	0.2	1.4
2	6.4	5.7	6.7	5.8	4.7	4.7	34.0
3	5.9	3.7	3.3	2.6	2.8	3.6	21.9
4	3.5	3.5	3.1	2.2	2.5	2.9	17.7
5	2.7	2.4	2.3	1.7	1.2	2.1	12.4
≥6	2.7	2.6	2.7	2.2	1.1	1.3	12.6
Σ	21.2	18.3	18.5	14.6	12.6	14.8	100

% of persons (21464)

No. of adults \ No. of children	0	1	2	3	4	≥5	Σ
1	*	0.1	0.2	0.1	0.2	0.2	0.8
2	2.2	2.9	4.6	5.0	4.8	6.1	25.5
3	3.0	2.5	2.8	2.7	3.3	5.2	19.6
4	2.4	2.9	3.2	2.6	3.4	4.9	19.4
5	2.3	2.5	2.8	2.4	1.8	3.6	15.4
≥6	3.1	3.4	3.8	3.5	2.3	3.2	19.4
Σ	13.0	14.4	17.4	16.2	15.7	23.2	100

3. **Estates**

% of households (1989)

No. of adults \ No. of children	0	1	2	3	4	≥5	Σ
1		[illegible].8	0.6	0.3	0.3	0.0	2.0
2	11.0	8.1	8.3	6.4	4.0	3.7	41.5
3	6.6	3.8	3.5	2.8	1.9	1.2	19.8
4	4.8	3.0	3.2	2.0	1.4	1.5	15.9
5	2.6	2.8	2.1	1.7	1.2	1.0	11.4
≥6	1.5	1.8	1.7	2.0	1.0	1.4	9.4
Σ	26.5	20.3	19.4	15.2	9.8	8.8	100

% of persons (10347)

No. of adults \ No. of children	0	1	2	3	4	≥5	Σ
1	*	0.3	0.3	0.2	0.3	0.0	1.2
2	4.3	4.7	6.5	6.0	4.6	5.3	31.4
3	3.8	2.9	3.4	3.2	2.5	2.1	18.0
4	3.7	2.9	3.7	2.6	2.1	2.9	18.0
5	2.5	3.2	2.8	2.6	2.0	2.1	15.4
≥6	1.8	2.6	3.1	4.0	2.2	2.5	16.2
Σ	16.1	16.7	19.7	18.6	13.8	15.0	100

* Single person households were excluded from the sample.

† Adults are persons aged 15 or over, children those aged 0-14.

2. Multivariate Analysis

Given the results of Figure 4, it seems sensible to go back to the PCE formulation and modify it by entering the numbers of household members of each type as separate variables. Hence, the basic equation (2) might become:

$$w = \beta_o + \beta_1 \ln PCE + \sum_{2}^{K} \beta_k n_k + u \tag{4}$$

where n_k is the number of persons of type k in each household, i.e. number of male and female children of different age groups, number of old people and so on. Note that (4) is not supposed to represent any specific model of how needs are generated in the household; rather it is in the spirit of a "flexible functional form" where the important variables are allowed at least one unrestricted parameter each and which can be thought of as a suitable linearization of whatever is the true (complex) process linking family composition, welfare, and the food share.

In practice, (4) was modified still further prior to estimation. First, in order to allow for the remaining curvature in the relationship between the food share and lnPCE, the square of lnPCE was also introduced. Second, it must be noted that the survey was conducted in four rounds, the first in April, May and June of 1969, the third in October, November and December 1969 and the fourth and last in January, February and March 1970. To the extent that there are seasonal patterns in food consumption, the results of different rounds may be different; there are also seasonal fluctuations in prices and availability of some foodstuffs and these too may induce seasonal fluctuations in the food share. Hence, the regression specification used in this section is:

$$w = \beta_o + \beta_1 \ln PCE + \sum_{2}^{K} \beta_k n_k + \gamma(\ln PCE)^2 + \sum_{2}^{4} \xi_k d_k + u \tag{5}$$

where d_k, $k = 2,3,4,$ is a dummy corresponding to rounds 2, 3, and 4.

The maximum disaggregation of household members was into eleven classes as follows:

(i) males aged 15-59

(ii) females aged 15-59

(iii) old persons aged over 59

(iv) male children aged 0

(v) female children aged 0

(vi) male children aged 1-4

(vii) female children aged 1-4

(viii) male children aged 5-9

(ix) female children aged 5-9

(x) male children aged 10-14

(xi) female children aged 10-14

A priori, it is uncertain which or how many of these variables need to be separately distinguished; an ideal procedure would be to begin with a very general procedure including all possibilities. Such a regression, however, would be extremely large given that various interaction terms may also be important. Inevitably, then, the model search procedure tends to result in looking at different issues one at a time; whether it is necessary to distinguish male from female adults, and then what is the necessary disaggregation of children, for example. Regression is, of course, a multivariate technique, so that alterations in one variable in the regression will inevitably have consequences for the parameter estimates elsewhere, and this can make sequential inference dangerous. In the present example, however, the majority of the variables are not highly interrelated and, although coefficients change, the <u>significance</u> of the various determinants very rarely does. If this had not been so, it would have been necessary to go back and recheck previous results whenever new variables or new combinations of variables were introduced. Occasionally this was done, but regressions with several thousand observations tend to be expensive to compute so that unnecessary runs are to be avoided.

Table 5 presents the first set of regression results which investigate the effects of the age and sex composition of adults in the household on the food share. As usual, results are presented separately for each of the three sectors, urban, rural, and estates. Starting near the foot of the table, note that there is a limited amount of seasonal variation in the food share and that this varies as between the sectors. In the urban sector, the third quarter tends to be low vis-a-vis the second, while in the countryside, a dip occurs in the fourth quarter. On the estates, however, the second quarter (in these regressions the excluded dummy) is lower than the other three. Looking next at the effects of PCE, the effects noted in the previous section are confirmed here in that the quadratic term in lnPCE is significantly different from zero in all sectors. The curvature is as anticipated, with the relationship between the food share and lnPCE dipping more sharply for better-off households. If this is correct, the elasticity of food demand is not only less than unity, but also declines with PCE. This tendency is most marked in the estates (as would be expected if the imputation argument in Section B were correct) but it should be noted that the <u>slopes</u> of the share to lnPCE relationships differ less than the table might suggest. For example, when lnPCE is 4 (PCE $\approx$ 55 rupees), the slopes for urban, rural and estates are -0.19, -0.17 and -0.15, respectively.

The effects of family composition are given in lines 4-10 of the table.

Table 5

Regression of Food Share and Family Composition
(t-values[1] in parentheses)

Variable	Urban			Rural			Estates		
Constant	1.237 (20.2)	1.236 (20.2)	1.237 (20.2)	0.738 (9.9)	0.738 (9.9)	0.743 (9.9)	0.578 (4.8)	0.578 (4.8)	0.579 (4.8)
$\ell nPCE$	-0.0740 (-2.65)	-0.0732 (-2.62)	-0.0738 (-2.64)	0.1468 (3.95)	0.1467 (3.95)	0.1457 (3.92)	0.1870 (3.14)	0.1869 (3.14)	0.1870 (3.14)
$(\ell nPCE)^2$	-0.0145 (-4.55)	-0.0146 (-4.58)	-0.0146 (-4.56)	-0.0393 (-8.49)	-0.0393 (-8.48)	-0.0392 (-8.46)	-0.0418 (-5.69)	-0.0418 (-5.69)	-0.0418 (-5.70)
# Male Adults[2]	-0.0049 (-3.96)	-	-	-0.0053 (-3.58)	-	-	-0.0014 (-0.60)	-	-
# Female Adults[2]	-0.0075 (-5.68)	-	-	-0.0061 (-3.77)	-	-	0.0026 (1.01)	-	-
# Old Persons	-0.0035 (-1.57)	-0.0037 (-1.63)	-	0.0006 (0.25)	0.0006 (0.23)	-	0.0023 (0.50)	0.0028 (0.60)	-
# Adults[2]	-	-0.0061 (-7.93)	-	-	-0.0057 (-5.97)	-	-	0.0004 (0.28)	-
# Adults + Old	-	-	-0.0059 (-7.98)	-	-	-0.0051 (5.51)	-	-	0.0006 (0.42)
# Small[3] Children	-0.0098 (-5.95)	-0.0098 (-5.98)	-0.0100 (-6.12)	-0.0060 (-3.20)	-0.0060 (-3.21)	-0.0066 (-3.51)	-0.0075 (-2.61)	-0.0074 (-2.58)	-0.0075 (-2.61)
# Large[3] Children	-0.0090 (-9.32)	-0.0090 (-9.32)	-0.0092 (-9.70)	-0.0077 (-7.11)	-0.0077 (-7.11)	-0.0083 (-7.82)	-0.0015 (-0.81)	-0.0015 (-0.82)	-0.0017 (-0.92)
d_3 (July, Aug., Sep.)	-0.0169 (-4.07)	-0.0171 (-4.10)	-0.0170 (-4.07)	-0.0033 (-0.72)	-0.0033 (-0.71)	-0.0034 (-0.75)	0.0247 (3.74)	0.0246 (3.74)	0.0246 (3.73)
d_4 (Oct.,Nov.,Dec.)	-0.0046 (-1.15)	-0.0047 (-1.18)	-0.0047 (-1.18)	-0.0128 (-2.97)	-0.0128 (-2.97)	-0.0132 (-3.08)	0.0186 (3.02)	0.0188 (3.06)	0.0188 (3.06)
d_1 (Jan.,Feb.,Mar.)	-0.0024 (-0.59)	-0.0026 (-0.64)	-0.0026 (-0.64)	-0.0055 (-1.27)	-0.0054 (-1.26)	-0.0056 (-1.30)	0.0162 (2.57)	0.0164 (2.60)	0.0164 (2.60)
R^2	0.5863	0.5861	0.5860	0.4053	0.4053	0.4043	0.2830	0.2830	0.2825
$\hat{\sigma}$	0.0891	0.0891	0.0809	0.0909	0.0908	0.0909	0.0974	0.0974	0.0974

1/ See Appendix 1 for interpretation and comments on t-ratios.

2/ Adults are aged 15-59.

3/ Small = 0-4; Large = 5-14.

Note that all significant (non-zero) effects are negative implying that, for all categories considered here, the crude head count embodied in PCE overstates the costs of maintaining the family. Even for male adults, except on the estates, increases in numbers are accompanied by some economies of scale or other welfare enhancing effects. In both urban and rural areas, the coefficient on women is absolutely larger than that on men suggesting either that women have lower needs or that they make a contribution to family welfare not included in PCE. (However, as will be seen, this difference is not a significant one). The coefficient on old people is insignificant in all the regressions; apparently, their contribution to the head count is an adequate representation of their costs. Perhaps most interesting is the insignificance of any adult effects on the estates. This may well reflect the limited opportunity for informal economic activity in those areas as well as the relatively equal treatment of women in the labor market. In all sectors, the children coefficients are absolutely larger than those for adults, presumably because the head count overstates children's needs by even more than it overstates those of adults. Surprisingly, perhaps, the coefficients on small and large children are typically very close; perhaps the extra needs of larger children are matched by their extra ability to undertake welfare generating activities.

The second column in each of the sectors shows the effects of combining males and females into a single count; the third, that of combining males, females and old people into a single count. Inspection of the changes in R^2 and the standard errors suggests that these combinations are not rejected by the evidence, and formal calculation of F-ratios confirms this. For column two over column one, the three F-ratios are 1.70, 0.09 and 1.10 for urban, rural and estates, respectively. According to the null hypothesis of equal coefficients for males and females these are distributed as $F_{1,4011}$, $F_{1,3641}$, and $F_{1,1978}$, respectively, all with 1% critical values of 6.63; clearly, the null can be accepted. For column three over column one, the ratios are 1.40($F_{2,4011}$), 3.00 ($F_{2,3641}$), and 0.67 ($F_{2,1978}$), once again indicating that the number of old people can be absorbed into the adult head count (critical value of 1% is 4.6). However, it is also acceptable to set the coefficient on old people to zero (counting them in the deflator of PCE alone) and this alternative is adopted in the subsequent regressions.

The second set of regressions (which are not reported in detail) are concerned with possible disaggregation of the children effects. The regressors were lnPCE, $(\ln PCE)^2$ and the seasonal dummies as before, with the number of adults (excluding old people) representing the three adult variables originally considered in Table 5. This time, however, for the first regression the last eight categories given above were entered as separate variables so that children are disaggregated by age and sex. At the next stage, all sex effects were suppressed so that the child regressors were numbers in the age groups only, i.e. 0, 1-4, 5-9, and 10-14, irrespective of sex. Finally, both age and sex effects were removed, leaving a single regressor, i.e. number of children. Table 6 lists the F-ratios testing each of these specifications against the unrestricted alternatives, together with their respective degrees of freedom. None of these suggest that the null hypothesis should be rejected. Hence, for all three regions in Sri Lanka, demographic effects on the food share are adequately captured by including, in addition to PCE, the numbers of adults (excluding the old) and the numbers of children. This is presumably a feature of this particular data set and there is no reason to suppose that the result will hold elsewhere. It should also be noted that the equality of the coefficients within adults and within children is not the only hypothesis that could be investigated and might be accepted by the data. However, it is a very convenient one and it is not rejected by the data.

Table 6

F-tests for Age and Sex Effects of Children on the Food Share

(degrees of freedom in brackets)

	Urban	Rural	Estate	1% Critical Value
Exclusion of Sex	0.60 (4,4007)	1.52 (4,3637)	2.66 (4,1974)	3.32 (4, ∞)
Exclusion of Age	0.50 (7,4007)	0.99 (7,3637)	2.26 (7,1974)	2.64 (7, ∞)

The final set of regressions examine the possibility of interactions between the demographic variables and PCE. This is potentially important since, if the costs of children as a fraction of PCE vary systematically with PCE, (which is likely if, for example, some of the costs of children are fixed costs), inequality comparisons between households with different incomes and family compositions will be biased if such variation is not allowed for. A simple way of checking for such effects is to estimate a general quadratic form relating the food share to lnPCE, n_a, the number of adults and n_c, the number of children, i.e. the two surviving demographic effects. Hence, ignoring seasonals, the estimated equation is

$$w = \beta_o + \beta_1 \ln PCE + \beta_2 (\ln PCE)^2 + \beta_3 n_a + \beta_4 n_c + \beta_5 n_a n_c + \beta_6 n_a^2 + \beta_7 n_c^2$$

$$+ \beta_8 n_a \ln PCE + \beta_9 n_c \ln PCE . \qquad (6)$$

In none of the separate sector regressions were the coefficients on n_a^2 or n_c^2 significant; other interaction effects were so, in at least some of the sectors. Hence n_a^2 and n_c^2 were dropped as regressors with the results shown in the left hand columns of Table 7. The new dummy variables z_2, z_3 and z_4 represent the different regional zones of the island; z_1 the omitted "base" zone is Colombo, Kalütara, Galle and Matara, z_2 is Hambantota, Moneragala, Amparai, Polonnaruwa, Anaradhapura and Puttalam, z_3 is Jaffna, Mannar, Vavuniya, Trincomalee, and Balticalea, and z_4 is Kandy, Matale, Nuwara Eliya, Budulla, Ratnapura, Kegalle and Kurunegala. Zone 4 has a typically lower food share in both urban and rural areas while Zone 2 has a lower share in the rural sector. On the interaction terms, that betweeen the number of children and lnPCE is significantly negative everywhere suggesting that the higher PCE is, the greater is the understatement of welfare produced by deflating by total numbers in the family. This issue will be pursued further in the next section.

Table 7 also reports the regressions without any quadratic terms. These results are clearly inferior to those in the first columns but give an indication of the overall marginal effects of each of the variables in the food share. They will also be used in the next section as a contrast with the quadratic model. Note finally the "All-Island" results in the final two columns. In Appendix 1, pooling tests are carried out which suggest that it is not possible to accept the hypothesis that the coefficients in the quadratic regression are identical across sectors. Even so, provided this is borne in mind, the All-Island regression provides a useful overall summary of the results.

Table 7

Quadratic and Linear Regressions
(t-values in parentheses)

Variable	Urban		Rural		Estate		All Island[1/]	
Constant	0.9104 (11.05)	1.5117 (111.0)	0.3493 (3.57)	1.3953 (85.8)	0.1170 (0.68)	1.2634 (42.8)	0.5853 (10.9)	1.4430 (148.3)
ℓnPCE	0.0565 (1.62)	-0.2012 (-72.6)	0.3136 (6.96)	-0.1740 (-48.0)	0.3890 (5.04)	-0.1512 (-24.0)	0.1938 (8.30)	-0.1869 (-90.9)
$(\ell n\text{PCE})^2$	-0.0262 (-7.08)	-	-0.0552 (-10.5)	-	-0.0631 (-7.18)	-	-0.0407 (-15.7)	-
n_a	0.0054 (0.80)	-0.0059 (-7.71)	0.0204 (2.28)	-0.0054 (-5.72)	0.0054 (0.31)	0.0006 (0.39)	0.0145 (2.97)	-0.0050 (-8.92)
n_c	0.0222 (3.38)	-0.0093 (-11.8)	0.0355 (4.57)	-0.0078 (-9.05)	0.0644 (4.15)	-0.0047 (-3.21)	0.0319 (6.97)	-0.0087 (-16.0)
$n_a n_c$	0.0014 (3.60)	-	0.0009 (1.65)	-	-0.0005 (-0.54)	-	0.0009 (3.09)	-
$n_a \ell n$PCE	-0.0035 (-2.36)	-	-0.0071 (-3.28)	-	-0.0008 (-0.18)	-	-0.0053 (-4.70)	-
$n_c \ell n$PCE	-0.0089 (-5.79)	-	-0.0119 (-6.04)	-	-0.0174 (-4.46)	-	-0.0109 (-9.72)	-
d_3	-0.0188 (-4.57)	-0.0174 (-4.19)	-0.0027 (-0.59)	-0.0015 (-0.32)	0.0249 (3.79)	0.0237 (3.57)	-0.0041 (-1.47)	-0.0032 (-1.14)
d_4	-0.0060 (-1.53)	-0.0053 (-1.32)	-0.0130 (-3.06)	-0.0115 (-2.66)	0.0178 (2.89)	0.0191 (3.07)	-0.0039 (-1.48)	-0.0032 (-1.21)
d_1	-0.0037 (-0.93)	-0.0033 (-0.82)	-0.0056 (-1.32)	-0.0033 (-0.76)	0.0164 (2.61)	0.0156 (2.46)	-0.0005 (-0.18)	0.0002 (0.08)
Z_2	-0.0107 (-1.87)	-0.0094 (-1.62)	-0.0174 (-3.80)	-0.0169 (-3.64)	-0.0012 (-0.07)	-0.0004 (-0.02)	-0.0150 (4.39)	-0.0131 (-3.77)
Z_3	0.0050 (1.26)	0.0077 (1.93)	-0.0035 (-0.71)	-0.0027 (-0.54)	-0.0295 (-0.67)	-0.0259 (-0.58)	0.0005 (0.16)	0.0038 (1.19)
Z_4	-0.0304 (-6.93)	-0.0311 (-7.01)	-0.0268 (-7.43)	-0.0275 (-7.51)	-0.0073 (-0.97)	-0.0064 (-0.83)	-0.0181 (-8.35)	-0.0182 (-8.24)
R_2	0.5989	0.5899	0.4240	0.4036	0.2890	0.2701	0.5022	0.4864
$\hat{\sigma}$	0.0877	0.0887	0.0895	0.0910	0.0971	0.0983	0.0910	0.0924

1/ Simple unweighted regression; see Appendix 1 for more sophisticated treatment.

E . EQUIVALENCE SCALES

As discussed in the introduction, the identification of the food share as an indicator of welfare allows the construction of measures of the costs of children. These measures, known as equivalence scales, give the fraction by which total expenditure must be increased in order to "compensate" a household for its size relative to some base, or reference household. If, for example, the reference household contains two adults, and it costs twice as much to maintain a family of two adults and three children at the same welfare level, then the equivalence scale for such a family would have a value of 2.

In the present context, the procedure can be illustrated using the version of equation (6) excluding quadratic terms. Hence, a household h with n^h persons, $n_a{}^h$ adults aged 15-59, $n_c{}^h$ children aged 0-14 and <u>total</u> expenditure x^h will have a food share w^h given by

$$w^h = \beta_o + \beta_1 \ln\left(\frac{x^h}{n^h}\right) + \beta_3 n_a{}^h + \beta_4 n_c{}^h \ . \tag{7}$$

If the reference household (h=0) has two adults, no children and a total expenditure of x^o,

$$w^0 = \beta_o + \beta_1 \ln\left(\frac{x^o}{2}\right) + 2\beta_3 \ . \tag{8}$$

The essential hypothesis is that if household h is given sufficient x^h to equalize w^h and w^o, they will have the same welfare level, so that x^h/x^o will measure the relative costs of reaching the same welfare level and hence, the equivalence scale, s^h, for example. Setting (7) equal to (8), the ratio of h^x to x^o can be solved to be

$$s^h = \frac{n^h}{2} \div \exp \ \{\theta_a (n_a{}^h - 2) + \theta_c n_c{}^h\}, \tag{9}$$

where

$$\theta_a = \beta_3/\beta_1 \text{ and } \theta_c = \beta_4/\beta_1 \tag{10}$$

This formula is readily interpretable. Since the reference household has two members, $n^h/2$ is the crude measure of costs obtained by taking ratios of head counts. This is then modified by the second term. Since β_3, β_4 and β_1 are all negative in the regressions (extra adults and children cost less than the head count would suggest), θ_a and θ_c are both positive, so that for households with at least two adults, the modifying term is greater than unity and the scale is less than the head count ratio. Table 8 gives the values of the scales for the figures given in the right hand columns of Table 7. Note first that the scales are quite close to half the total number of persons, reflecting the dominant role of PCE in explaining the food share. This means that they tend to be considerably larger, particularly with regard to children, than those obtained using weighting factors based on nutritional or calorific needs. Second, there is relatively little difference across sectors, although they are higher in the estates; this last reflects the negligible (or even positive) coefficient on adults in the food share equation for the estates which results in each adult effectively contributing an extra 0.5 to the scale.

Formula (9) and Table 8, although simple, do not take account of the interactions or the quadratic lnPCE term, which turned out to be important in Table 7. Using the more general model complicates the formulae but does not affect the principle. The more complex forms of (7) and (8) can be equated to one another and a solution for the scale derived by solving a simple quadratic equation. To do this, replace (7) by its quadratic form, i.e.

$$w^h = \beta_o + \beta_1 \ln(\frac{x^h}{n^h}) + \beta_2(\ln \frac{x^h}{n^h}) + \beta_3 n_a{}^h + \beta_4 n_c{}^h + \beta_5 n_a{}^h n_c{}^h , \tag{7a}$$

Table 8

Adult Equivalence Scales for Combinations of #Adults and #Children: I (reference: 2 adults)

# Adults		1	2	3	4	5
# Children 0 Sector:	U	0.52	1.00	1.46	1.89	2.29
	R	0.52	1.00	1.45	1.89	2.28
	E	0.50	1.00	1.51	2.02	2.53
1 Sector:	U	0.98	1.43	1.85	2.25	2.62
	R	0.99	1.43	1.85	2.25	2.61
	E	0.97	1.45	1.95	2.44	2.94
2 Sector:	U	1.41	1.82	2.21	2.58	2.92
	R	1.41	1.83	2.22	2.58	2.92
	E	1.40	1.88	2.36	2.84	3.32
3 Sector:	U	1.79	2.18	2.54	2.87	3.19
	R	1.80	2.18	2.54	2.87	3.19
	E	1.82	2.27	2.74	3.21	3.67

Table 9

Adult Equivalence Scales for Combinations of #Adults and #Children: II (reference: 2 adults)

# Adults		1	2	3	4	5
# Children 0 THE:	50	0.52	1.00	1.45	1.88	2.29
	100	0.53	1.00	1.43	1.83	2.19
	400	0.53	1.00	1.42	1.80	2.13
1 THE:	50	1.01	1.48	1.94	2.38	2.82
	100	0.98	1.41	1.82	2.20	2.55
	400	0.96	1.37	1.74	2.07	2.37
2 THE:	50	1.47	1.95	2.42	2.89	3.35
	100	1.37	1.78	2.17	2.54	2.90
	400	1.31	1.67	2.00	2.30	2.58
3 THE:	50	1.93	2.42	2.91	3.40	3.89
	100	1.73	2.12	2.51	2.88	3.24
	400	1.59	1.92	2.22	2.50	2.76

ignoring the zonal and seasonal effects. Hence, the food share of the reference household is predicted as:

$$w^o = \beta_o + \beta_1 \ln(\frac{x^o}{2}) + \beta_2(\ln \frac{x^o}{2})^2 + 2\beta_3 \qquad (8b)$$

For any given level of x^o, equation (8b) yields a w^o and setting the right hand side of (7a) equal to this value yields two solutions for x^h corresponding to any given n^h, $n_a{}^h$ and $n_c{}^h$. The larger of these two solutions corresponds to the relevant part of the Engel curve, and dividing by x^o yields the scale. The major difference as compared with (9) is that s^h now depends not only on $n_a{}^h$, $n_c{}^h$ and n^h, but also on the reference expenditure level x^o. This is a natural consequence of the quadratic specification with its implication that the costs imposed by family size are not proportionally the same at all levels of welfare. Table 9 replicates Table 8 but for the quadratic model, and instead of giving separate scales for each sector, uses the All-Island results at three different levels of x^o, 50, 100 and 400 rupees (total household expenditure). When there are few or no children in the household, the effects of changes in THE is rather limited since the interaction between PCE and the number of adults is small. However, the coefficient on $n_c$lnPCE in Table 7 is much larger than that on $n_a$lnPCE so that the decline in the scales with THE is much more marked for families with many children. Clearly, if such figures are accepted, the measurement of inequality in terms of expenditure per equivalent adult is likely to give quite different results than inequality in either per capita or total household expenditure. Such comparisons, however, must be left for another paper.

APPENDIX 1: ECONOMETRIC CONSIDERATIONS

1. Heteroscedasticity

The statistical inference in Section C above was implicitly based on the assumption of homoscedasticity in the residual variance of the regression equation. Formally, if the regression is written

$$y_i = x_i' \beta + u_i \tag{10}$$

for food share y_i, independent variables x_i, parameter vector β and residual u_i, all for observation i, then the standard assumptions for u_i are

$$E(u_i) = 0 \qquad E(u_i^2) = \sigma^2 \tag{11}$$

where σ^2 is not indexed on i. In cross-section regressions of the type discussed in this paper, the homoscedastic assumption is rather unlikely to hold. As is well known, heteroscedasticity does not affect the consistency of the least squares parameter estimates although it does reduce their efficiency. In relatively large samples such as the present, the loss of efficiency may not be thought to be very serious. However, heteroscedasticity also implies that the variance-covariance matrix of the parameter estimates, as usually estimated, viz.,

$$V = \hat{\sigma}^2 (X'X)^{-1} \tag{12}$$

is inconsistent for the true variance-covariance matrix. This may have very serious consequences, especially in work such as the present where model search is being made using t-ratios and F-statistics to simplify the structure of the regression. In particular, there is the disturbing potential analogy with serial correlation in time-series regression, where the inconsistent estimates of t's and F's are frequently ten to twenty-fold out giving rise to the classic nonesense (or spurious) regression phenomenon (see, e.g., Yule (1926) and Granger and Newbold (1973)). Without checking against it, heteroscedasticity may produce very similar effects in cross-sections.

The first necessity is to be able to detect heteroscedasticity by means of a suitable test statistic. One such, based on the efficient score statistic, has recently been proposed by Breusch and Pagan (1979). This is a parametric test, which amends (11) to read

$$u_i \sim N(0, \sigma_i^2) \qquad \sigma_i^2 = h(\alpha_o + \alpha' z_i) \tag{13}$$

where $h(\ )$ is some unspecified function, α_o is a constant, α is a parameter vector and z_i is a vector of heteroscedasticity generating variables, which may overlap with (or be identical to) the x variables. Breusch and Pagan show that the efficient score statistic may be computed by computing two regressions. The first is the OLS original regression under the null hypothesis of homoscedasticity. This generates OLS residuals - for example - e_i and an estimated equation standard error $\hat{\sigma}^2$. The second regression regresses $(e_i^2/\hat{\sigma}^2 - 1)$ - with a conditional expectation of zero under homoscedasticity - on the z_1 variables. The explained sum of squares from this second regression - which should be small under homoscedasticity and large under heteroscedasticity - when divided by two is asymptotically distributed as χ^2 with degrees of freedom equal to the number of elements in Z.

The Breusch/Pagan procedure was carried out routinely in running the regressions reported earlier in the paper. The simplest specification was adopted in which the z_i variables were taken to be the x_i's. Given this, heteroscedasticity always appeared to be present. For example, in the final quadratic regressions reported in Table 7, the score tests took values of 291.8 (All-Island), 85.5 (urban), 132.6 (rural) and 101.1 (estates), each of which should be distributed as χ^2_{13} under the null hypothesis. Interestingly, the seasonal dummies were always extremely significant in the second regression suggesting a much stronger seasonal effect on the regression variance than on its mean. This may well result from seasonal variation in some of the regression parameters. The zonal dummies were occasionally significant, as was lnPCE and the number of adults.

On the basis of such results, it would seem of some moment to "correct" the heteroscedasticity. But this is by no means straightforward. The Breusch/Pagan procedure, although suggestive of variables, tells nothing about the form of the h( ) function in (13) and this would be needed to construct a weighted regression. And even if linearity were assumed, fully efficient estimation would require an iterative (and rather expensive) procedure. Fortunately, this turns out to be unnecessary since White (1980) has provided an easily computed consistent estimator of the variance-covariance matrix. Specifically, if D is a diagonal matrix with e_i^2 on the diagonal (i.e., an $n \times n$ matrix with the squares of the OLS residuals on the diagonal), then

$$V^* = (X'X)^{-1} X'DX(X'X)^{-1} \tag{14}$$

is a consistent estimate of the variance-covariance matrix. (Note that given the use of X for Z in the Breusch/Pagan procedure, V^* and the score test are intimately related). Using (14), it is possible to check whether the usual standard errors are seriously misleading and to make correct inferences.

In the present context, although the standard errors given by (14) are, in general, larger than those from (12), the differences are not very large. For example, for the urban quadratic regression in Table 7, the heteroscedastic-consistent t-values, with the originals in brackets, are 9.19 (11.05), 1.32 (1.63), -5.63 (-7.08), 0.83 (0.80), 3.06 (3.38), 3.78 (3.60), -2.39 (-2.36), -5.12 (-5.79), -4.36 (-4.57), -1.47 (-1.53), -0.91 (-0.93), -1.84 (-1.87), 1.20 (1.26), -6.78 (-6.93). These results are typical of those obtained from the other regressions.

Although a mechanical approach to inference, rigidly adhering to a specific significance level, could use these two sets of numbers to give different results and different model search procedures, the differences are not such as to render the conventional values badly misleading *in this particular case*. Hence, in the text the standard formulae were used in order to keep the methodology as familiar as possible. However, in similar studies, the effects of heteroscedasticity on inference should be investigated and are ignored at some peril to the results.

2. Pooling

Although nearly all of the results in the paper have been quoted for the three sectors separately, it is of considerable interest to test whether the final results can be pooled, giving a single regression for the whole island. This is not to deny that there are major differences in many respects between the sectors; such differences are documented earlier in the paper. The question is much more limited: Does there exist a single regression summarizing the behavior of the food share? If so, that representation is not only convenient but it also implies that the known differences across sectors in the food share can be adequately captured by the variables included in the regression.

The standard pooling test for several regressions is to compare the residual sum of squares for the pooled regression with those for the separate regression by means of an F-ratio. However, such a procedure assumes that all three regressions have a common error variance, something which is far from obviously true in the results in Table 7. It is thus necessary to test the common variance assumption first. This may be done using a likelihood ratio test, see e.g. Mood, Graybill and Boes (1974, p. 438); the test statistic is

$$t = \sum_{j=1}^{3} n_j \log \hat{\sigma}_j^2 - n \log \tilde{\sigma}^2 \qquad (15)$$

where n_j is the number of observations in sector j $(j=1,2,3)$, n is the total number of observations and $\tilde{\sigma}^2$ is the pooled variance estimate given by

$$\tilde{\sigma}^2 = \sum_{j=1}^{3} n_j \hat{\sigma}_j^2 / \sum_{j=1}^{3} n_j . \tag{16}$$

Asymptotically, under the null of common variances, t has a χ_2^2 distribution. Although, in the reference given, the test is derived for three normal populations, it may easily be checked that it is also correct for regression models with normal errors. Inserting the values of $\hat{\sigma}_j$ from Table 7 into (15) and (16) gives a test statistic of 22.9 which is too large to be consistent with the null. It it therefore certainly <u>not</u> true that the error variance in the regression is the same across sectors; in particular it is larger in the estates, even though the total variation in the food share is much less in that sector.

Even so, inequality in the variances does not imply that the regression coefficients themselves may not be equal. This can be tested by means of a weighted pooled regression in which both independent and dependent variables are divided by the standard error corresponding to the sector from which they come. If the residual sum of squares from this pooled regression is RSS_p, the conventional F-test would be

$$F = \frac{(RSS_p - \sum_{j=1}^{3} RSS_j)/2k}{\sum_{j=1}^{3} RSS_j/(n - 3k)} \tag{17}$$

where RSS_j are the individual sector residual sum of squares and k is the number of variables in the regression. However, in the individual weighted regressions corresponding to the pooled weighted regressions, the residual sum of squares is simply $(n - k)$, by the definition of the estimated equation standard error. Hence, (17) simplifies to

$$F = (RSS_p - n - 3k)/2k \tag{18}$$

which, under the null, is distributed as F with 2k and (n-3k) degrees of freedom. Note that this distribution is not exact since estimates of the individual equation standard errors have been used; even so, the asymptotic theory should work quite well with the current sample sizes.

The F-ratio corresponding to (18) has a value of 12.62 compared with a 1% critical value of 1.76, so that comparably to the situation with the variances, it is possible to reject the hypothesis that the regression coefficients are identical across sectors. However, it should be noted that (17) and (18) are not likelihood ratio tests in this context since the unrestricted estimates of the individual equation standard errors were used for weighting. A more favorable, but asymptotically equivalent test, is computed by using the pooled (common) parameter values to calculate separate equation standard errors for each sector. These can be used to reweight, and new parameter estimates can be calculated which, in turn, give new standard errors and so on. Following this iterative scheme made no perceptible difference to the test statistics, suggesting that with nearly ten thousand observations, asymptotic approximations are quite good. Note finally that the tests reported here are all computed under the (false) assumption of heteroscedasticity within each sector. There seems to be no reason, however, to suppose that the tests would be any more misleading than those for the individual equations.

APPENDIX 2: EQUIVALENCE SCALE THEORY

The basic theory of equivalence scales is laid out in Chapter 8 of Deaton and Muellbauer (1980). The model used here corresponds to the Engel model discussed there. Formally, the cost function for a household with characteristics a is written $c(u, p, a)$, i.e. the minimum cost of reaching utility u at prices p and characteristics a. For a cost-minimizing (or utility-maximizing) consumer, this is equal to total expenditure x. Models of household behavior impose some structure on this function. In particular, the Engel model assumes that

$$c(u, p, a) = m(u, a)\,\gamma(u, p) \tag{19}$$

for two functions m, of u and a alone, and γ of u and p alone. For a reference household, $m(u, a)$ is unity so that $\gamma(u,p)$ is the cost function of the reference family. If households face the same prices, the equivalence scale $m(u,a)$ can be calculated by finding a household with the same u as the reference, so that $c(u, p, a^h) = x^h$, $\gamma(u, p) = x^o$ and $m(u,a) = x^h/x^o$.

The food share (or any other share) is a welfare indicator in this model. This follows from (19) since the share of any commodity is given by differentiating $\ln c(u, p, a)$ with respect to the $\ln p$ of that good. Hence

$$w = \frac{\partial \ln c(u,p,a)}{\partial \ln p_f} = \frac{\partial \ln \gamma(u,p)}{\partial \ln p_f} \tag{20}$$

for price of food p_f. Clearly, if households face the same prices, w varies directly with u and is thus a direct welfare indicator. Regarding functional form, w will depend on x and a through its dependence on u where u is solved from

$$x = m(u,a)\,\gamma(u,p). \tag{21}$$

For a single commodity, and given certain monotonicity conditions, the functions $m(u,a)$ and $\gamma(u,p)$ can always be chosen to give <u>any</u> functional form for w so that it is convenient to reverse the process and start from a convenient functional form for w, as in the text.

APPENDIX 3: DATA REQUIREMENTS

These are mostly obvious from the text but it may be convenient to provide a separate list in one place:

(i) food expenditure, including imputed values of rations or food received in lieu of payment or domestically produced, all at market prices. Imputations should be recorded separately in case market prices are not appropriate shadow prices and some correction needs to be made. Ideally, prices of principal foodstuffs should be recorded or the survey supplemented by regional/district pricing of a suitable food bundle.

(ii) total expenditure, inclusive of the above, also including imputations and again with imputations listed separately. Lumpy purchases such as durables should ideally be priced at user cost, or more crudely spread over the life of the good. Similar techniques may also be appropriate for health expenditures. Failing this, it is probably better to exclude such lumpy purchases than to include them.

(iii) household composition, the numbers, ages (or age groups) and sex of members of the household. The ages of children are more important than those of adults even though, in this study, the variable was not influential on the food share.

(iv) non-household variables, regional, district and seasonal variables may all be important and should be recorded for each household. Particularly when there is no separate information on prices, regional and seasonal dummies may be good proxies.

II. INEQUALITY AND NEEDS: SOME EXPERIMENTAL RESULTS FOR SRI LANKA

A. INTRODUCTION

It is generally recognized that, in view of the variable needs and demographic composition of households, neither per capita expenditure nor total household expenditure provide an adequate measure of household or individual welfare. Although the move from total to per capita expenditure makes some allowance for variations in household size, the crudeness of a simple head count excludes considerations of differential needs by sex or by age, as well as any allowance for household "economies of scale" to the extent that such exist. A more sophisticated approach can be based on nutritional needs, usually calorie counts which are specific by sex and by age; accordingly, children count as fractions of an adult male, as do women in most schemes, so that family size can be expressed in terms of "equivalent adult units" given the demographic composition of the household. However, there are considerable difficulties in objectively establishing what calorie intake is indeed necessary, since much depends on environment, climate, health, work habits and so forth. Furthermore, such measures of need concentrate entirely on food, ignoring other elements of consumption, not all of which can be classified as dispensible luxuries. Nor can the calorific measures take account of social dimensions of consumption, so that very real but non-dietary needs may be ignored however important they may be within the social structure inhabited by the household or individual.

Since it is difficult to envisage obtaining reliable or useful information from direct questions about needs, an alternative indirect approach is worth exploring. This is provided by basing the calculation of equivalent adult scales on the analysis of household behavior using the cross-sectional data provided by a household survey. In this essay I discuss the basic theory underlying such attempts and report on some first experiments for Sri Lanka using the socioeconomic survey of 1969-70 and the econometric results reported in the first essay of this working paper. Section A explores the link between observable behavior, welfare and equivalence scales. The basic behavioral theory is given and alternative approaches discussed. I argue that the simplest method, due originally to Engel (1895), is likely to be the most appropriate for developing countries, at least given the current state of knowledge. Section C gives the results of applying the method to the empirical evidence revealed by the Sri Lankan survey. The procedure for calculating the scales is discussed in some detail and it is shown how a "corrected" per capita expenditure figure may be calculated for each household. The behavior of the correction factors is analyzed as a function of demographic composition and of per capita expenditure itself. Finally, in Section D the corrected per capita expenditure figures are used to calculate inequality measures to be contrasted with those resulting from the conventional, uncorrected figures. As might be expected, the overall measures of inequality (with a few exceptions) are not very sensitive to the correction, although the ordering of households in the distribution is quite different by the two criteria. The results are very similar to those obtained by Visaria (1979) who experimented with different equivalence scales based on calorific requirements. Hence it would appear that the crucial issue to be studied using equivalence scales is <u>not</u> the measurement of inequality but rather the identification of the poor.

B. BEHAVIORAL MODELS OF EQUIVALENCE SCALES

Behavioral equivalence scales are based on two quite distinct elements: The first is the empirically observable relationship between consumption patterns and demographic composition; the second is an essentially untestable assumption linking behavior to welfare. Any given model of the welfare effects of household composition will have implications for observable behavior. These can be tested against the data so that, in principle, any scheme for measuring adult equivalences can be rejected if its implications are not borne out. However, once a scheme is found which does not contradict the evidence, it is easy to show (see below) that there is an infinite number of such schemes with identical empirical implications. Hence, in one sense, it is not possible to identify the welfare implications of family composition from observable behavior and it is perfectly intellectually respectable (although not very helpful!) to refuse to try to do so. However, given that we recognize that an untestable assumption is involved, it is still possible to try to reach agreement on 'reasonable' versus 'unreasonable' schemes, or if there are many of the former, to explore their consequences.

As in all demand analysis, we can begin either from the demand functions or from a specification of preferences, deriving the implied demands as a second stage. If the first course is adopted, it is useful to restrict attention to functions which are consistent with the existence of preferences so that a basis for discussing welfare does, in fact, exist. In the second case, it is extremely important that the functional form for preferences be general and flexible enough to allow the demand functions to model the main empirical regularities in the data. For this reason, in the earlier empirical work underlying the results in Sections C and D of the first essay in this paper, the main focus was on the specification of the demands, with relatively little attention paid to the preferences underlying the model. As we shall see, this is essentially all that is necessary for the methodology to be proposed, but for explication of the theory, it is convenient to begin from preferences.

The essential assumption behind all interpersonal comparisons of welfare is that all households (for the moment the reference unit) have identical tastes once various conditioning factors have been allowed for. Hence, if preferences are represented by a utility function with value u, we write

$$u = \upsilon(q, a) \tag{1}$$

where $\upsilon(., .)$ is the common utility function, q is the vector of commodities purchased, and a is a list of factors or characterisitics conditioning tastes. For the current discussion, it is natural to focus on a as a vector of demographic variables (e.g., number of males, females aged 15-59, number of small male children, etc.), although in practice other important influences (region, race, religion, seasons) would be included. The household maximizes (1) subject to a price vector for goods, p, and a total expenditure level, x, with the vector a <u>taken as fixed</u>. This assumption means that children are treated as gifts (or curses) from the gods, over which the family has no control, and this inevitably limits applicability of the theory. Nonetheless, it can be argued (see, for example, Deaton (1980) or Deaton and Muellbauer (1980, Chapter 8)), that such a model can still adequately represent the real economic costs of maintaining children once they exist and that this is the measure required in assessing poverty and inequality in LDC's.

Corresponding to the utility function (1), and any price vector p, there can be defined a <u>cost function</u> $c(u, p, a)$ which gives the minimum cost for a household with characteristics a of reaching welfare level u given a price vector p. For a cost-efficient household, this will equal total expenditure x, so that we can write

$$c(u, p, a) = x. \tag{2}$$

This function is then the basis for comparisons of welfare between households with different compositions; For if we take household zero, with characteristics a^o, as the reference, then the ratio

$$m(a^h, a^o; u, p) = c(u, p, a^h)/c(u, p, a^o) \tag{3}$$

measures the relative costs of reaching the same welfare level at the same prices for a household with characteristics a^h vis-à-vis the reference household. For example, if a^o is simply one adult, a^h is two adults, and if there are no economies of scale, $m(a^h, a^o; u, p)$ might take the value two for all u and p. Hence, (3) can be thought of as defining the equivalence scale corresponding to the utility function (1).

How does this relate to observable behavior and how can we get from the fitted demands back to preferences and scales? The link comes through the cost function, the derivatives of which, with respect to the prices, are the quantities demanded. Alternatively, in logarithmic form (see, for example, Deaton and Muellbauer (1980, pp. 40-3)), we can write

$$w_i = \partial \ln c(u, p, a)/\partial \ln p_i \tag{4}$$

where w_i is the share of the budget devoted to good i. Now, given a cost function, equation (4) gives the budget shares as a function of prices, characteristics and utility. Since the latter is not observable, equation (2) has to be inverted to get utility in terms of prices, total expenditure and characteristics, i.e.

$$u = \psi(x, p, a) \tag{5}$$

so that substitution of (5) into the right-hand side of (4) gives the budget shares in terms of total expenditure x, prices p and characteristics a, all of which are observable. Write this as

$$w_i = f_i(x, p, a); \tag{6}$$

it is these functions which can be fitted to the data.

The relationship between the data, essentially equation (6), and the equivalence scale, equation (3), should now be clear. If preferences are taken as a starting point, the original specification is of the cost function in (2); this leads to the scale (3) and to the demands (6) and the latter must conform to the data, hence guiding the original specification of the cost function. Alternatively, it is possible to start from (6), and subject to certain integrability conditions, recover the cost function and hence the scale (3). Unfortunately, this can never be done uniquely, so that many scales are consistent with any given demand functions. To see this, consider the two cost functions

$$x = c(u, p, a) \tag{7.1}$$

$$x = c(\phi(u, a), p, a). \tag{7.2}$$

These are essentially distinct since the level of welfare associated with any given x, p and a will be different under (7.1) than under (7.2). Similarly, the scale given by (3) will be different as between the two specifications.

However, if we follow through the calculations (4) through (6), the quantity $\phi(u, a)$ in (7.2) simply replaces u in (7.1) so that the final demands (6) are unaffected. Hence, the two specifications (7.1) and (7.2), which <u>are quite different from a welfare point of view, are observationally equivalent on the data and cannot be empirically distinguished.</u> This is where the untestable assumption is required. My own preference, argued at greater length in Deaton (1980), is to assert that households which behave identically, have practical purposes, and identical welfare levels. According to this, welfare is u in (7.1) and $\phi(u, a)$, <u>not</u> u, in (7.2). Given this, any given set of demand functions leads back to unique equivalence scales.

The traditional models of equivalence scales can be viewed in terms of this analysis. For example, the model proposed by Barten (1964) can be thought of as specifying the cost function $c(u, p, a)$ as:

$$c(u, p, a) = c^o(u, p^*) \tag{8.1}$$

$$p_i^* = p_i m_i(a) \tag{8.2}$$

where $c^o(u, p)$ is the cost function for the reference household and the $m_i(a)$ are commodity specific measures of needs depending on the household composition a. In this model, demographic needs act as if to alter the price structure, so that, for example, soft drinks become "expensive" relative to liquor for a family with a large number of children. Specification of particular functions for $c^o(u, p)$ and $m_i(a)$ leads to commodity demands in terms of total expenditure and characteristics and these can be fitted to the data. An alternative model is that of Prais and Houthakker (1955). This begins from the demands which are written

$$q_i/m_i(a) = f_i(x/m_o(a)) \tag{9}$$

for commodity specific needs $m_i(a)$ and a general needs deflator $m_o(a)$. Hence, quantities relative to needs are made a function of resources relative to overall needs. Under highly restrictive conditions (see Muellbauer (1980) or Deaton and Muellbauer (1980, Chapter 8)), the Prais-Houthakker model is consistent with the preference formulation (8).

However, neither the Barten model nor the Prais-Houthakker model can yield equivalence scales when estimated on a single cross-section of households. The relationship between quantities, total expenditure and characteristics - assuming prices to be the same for all individuals - does not contain sufficient information to identify the needs functions $m_i(a)$ (see Muellbauer (1980) for details). This identification problem is quite distinct from that discussed above in passing from behavior to welfare, and is related rather to the specific formulations of the two models. Hence, the work on Indian data by Singh (1972) and by Singh and Nagar (1973), which uses the Prais-Houthakker model, is not well-founded; the empirical results they present are implicitly determined by the prior restrictions built into their estimation scheme. Since there is little reason to accept their prior restrictions over the many others which are possible, their results are arbitrary. Coondoo (1972) has suggested an ingenious method for identifying the model but the additional restriction he suggests cannot in fact hold good so that the problem remains. The Barten model, by contrast, can be identified on data with price variation, for example on several cross-sections (see the work on British data by Muellbauer (1977)). Such estimation, however, is complex, almost inevitably involving nonlinear techniques, and it requires data not commonly available in LDC's.

A much simpler and clearly feasible technique is that due to Engel (1895) who suggested that the budget share of food be used as an (inverse) indicator of welfare. Once again, this can be justified in terms of the analysis given above if the cost function is specified as

$$c(u, p, a) = m(u, a)\, c^{o}(u, p) = x \tag{10}$$

where $c^{o}(u, p)$ is the cost function of the reference household and $m(u, a)$ is a multiplicative factor measuring the extent to which costs are raised for a family with welfare level u and composition a. Clearly, from (3), $m(u, a)$ is the scale itself. Note that, in contrast to the Barten and Prais-Houthakker models, the effects of composition are not commodity specific; this is unrealistic and is clearly a defect of the model. As will be seen, however, the specification overcomes the very considerable difficulties associated with the estimation of the other two models in the context of LDC's. If equations (4) and (5) are applied, the budget shares are

$$w = \frac{\partial \ln c^{o}(u, p)}{\partial \ln p_i} = f_i\left(\frac{x}{m(u, a)}, p\right). \tag{11}$$

Hence, <u>if it can be assumed that prices are the same for all households</u>, the Engel model ascribes all variation in the budget allocation to variation in the single ratio x/m. Hence, if two households have the same w_i, they must have the same x/m. In particular, if one of the pair is the reference household, so that x^* is the expenditure level at which the reference household has the same budget pattern as the household under consideration, say household h, the scale is given by

$$m(u, a^h) = x^h/x^*. \tag{12}$$

Note that use of this formula requires only the fitted Engel curves, bypassing the need to specify preferences. If the model (10) is correct, then by (11), all commodity budget shares should be a function of the single quantity x/m. In Muellbauer's work on Britain, this was found not to be the case. In poor countries, however, the budget is dominated by the behavior of the food share (e.g., around 70% in Sri Lanka as opposed to 20% in Britain), so that it seems sensible to adopt Engel's original suggestion in the LDC context, ignoring other elements in the budget. This has the additional advantage of restricting the econometric methodology to no more than the use of ordinary least squares.

Note finally that the natural way of using the scales to measure welfare is to divide total expenditure by the number of equivalent adults given by the scale to obtain expenditure per equivalent adult. By (3), since total expenditure x^h equals $c(u^h, p, a^h)$, expenditure per equivalent adult is simply $c(u^h, p, a^o)$. But this is the concept referred to as "money metric utility"; see Samuelson (1974) for the original term, and Deaton (1980) for a general plea for its use as a welfare measure. By the definition of the cost function, it is the minimum total expenditure required by the reference household to take it to the welfare level of household h and is to be thought of as the welfare indicator for the latter. In terms of behavior, it is the total expenditure required by the reference household to give it the same behavior as household h. This, of course, is a quite general definition holding for all the behavioral models. For the Engel model specifically, and the rest of this paper will be concerned only with this, money metric utility is the x^* in equation (12), that is the total expenditure level at which the reference household has the same food share as the household under consideration. The next section turns to the calculation of such quantities for Sri Lanka.

C. EQUIVALENCE SCALES FOR SRI LANKA

The starting point for calculating equivalence scales is the estimation of a set of Engel curves relating the food share in the budget to household total expenditure and household composition. Such regressions for Sri Lanka, based on the Sri Lankan 1969-70 socioeconomic survey, are reported in detail in Essay No. I. Apart from seasonal and regional effects, it was found that a quadratic specification in terms of per capita expenditure and the numbers of adults and numbers of children gave an adequate explanation for each of the three sectors of the economy. Hence, if x is total household expenditure, n the total number of persons (including children and the old), n_a the number of adults aged 15-59, and n_c the number of children aged 0-14, the food share w is given by:

$$\begin{aligned} w = \ & \beta_0 + \beta_1 \ln(x/n) + \beta_2\{\ln(x/n)\}^2 + \beta_3 n_a + \beta_4 n_c \\ & + \beta_5 n_a n_c + \beta_6 n_a \ln(x/n) + \beta_7 n_c \ln(x/n) \\ & + \text{seasonal} + \text{regional effects.} \end{aligned} \tag{13}$$

The parameter estimates for the three sectors, urban, rural and estate, are reproduced below in Table 1, together with their t- values, equation R^2 statistics and equation standard errors. Tests showed that these three regressions could <u>not</u> legitimately be pooled into a single all-island regression. However, in all three sectors, further disaggregation of household composition, by age and sex of children, by sex of adult, and by allowing for old persons separately, produced very little additional explanatory power. This is in contrast to results for British data (see in particular Muellbauer (1977)), where age effects among children exerted a powerful influence. This difference may be due to the greater unrecorded productive input to family welfare by older children in a country such as Sri Lanka.

Table 1

Parameter Estimates for the Food-Share Equation*

	URBAN	RURAL	ESTATE
β_0	0.9104 (11.05)	0.3493 (3.57)	0.1170 (0.68)
β_1	0.0565 (1.62)	0.3136 (6.96)	0.3890 (5.04)
β_2	- 0.0262 (-7.08)	- 0.0552 (-10.5)	- 0.0631 (-7.18)
β_3	0.0054 (0.80)	0.0204 (2.28)	0.0054 (0.31)
β_4	0.0222 (3.38)	0.0355 (4.57)	0.0644 (4.15)
β_5	0.0014 (3.60)	0.0009 (1.65)	- 0.0005 (-0.54)
β_6	- 0.0035 (-2.36)	- 0.0071 (-3.28)	- 0.0008 (-0.18)
β_7	- 0.0089 (-5.79)	- 0.0119 (-6.04)	- 0.0174 (-4.46)
R^2	0.5989	0.4240	0.2890
$\hat{\sigma}$	0.0877	0.0895	0.0971

* See Essay No. I, Table 7, p. 29.

Figure 1

Schematic Representation of Calculation of Equivalence Scales

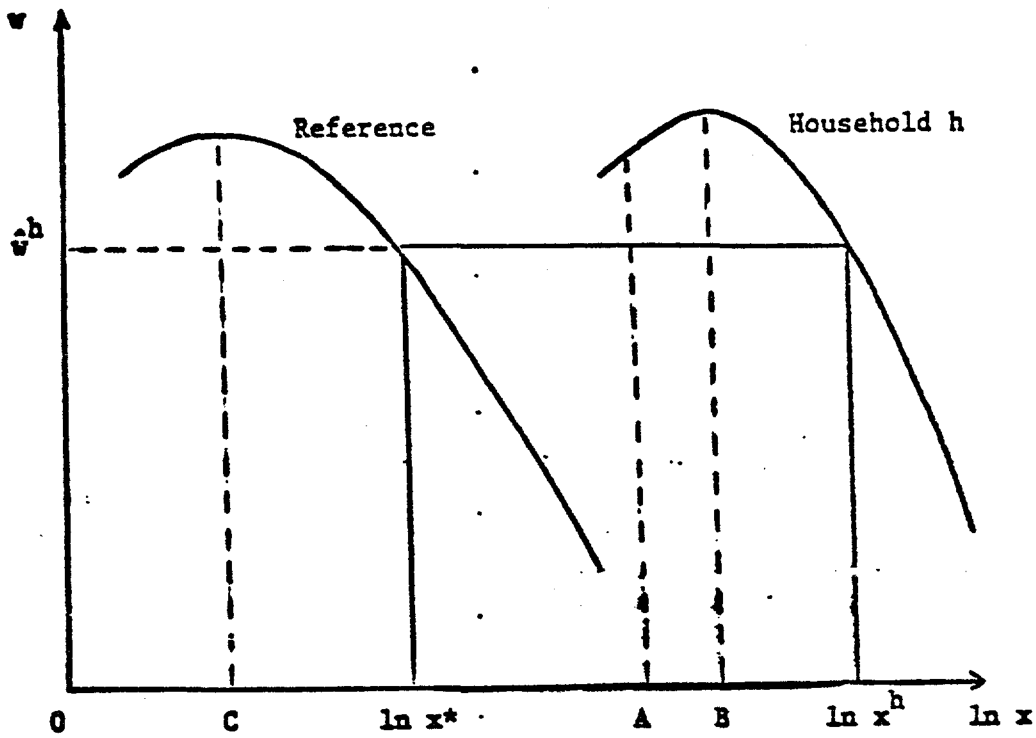


Figure 1 gives a graphical exposition of the procedure for calculating the equivalence scale. The left-hand curve shows the relationship between the food share and total expenditure for the reference household; that on the right is for a (larger) household h for which an equivalence scale is to be calculated. Starting from household total expenditure x^h, we move up to the predicted food share $\hat{w}^h$ on the right-hand curve. Moving then to the left-hand curve, x^* is read off as that level of total expenditure which would generate a predicted $\hat{w}^h$ by the reference household. The equivalence scale is then x^h/x^*, and the corrected welfare measure x^*.

In practice, a household of two adults and two children was chosen as reference; there are a sizable number of such households in each of the three sectors, although it is certainly not the dominant type (4.9% urban, 6.7% rural, and 8.3% estates). Hence, from (13), x^* is defined by the equation:

$$\begin{aligned}\beta_2\{\ln(x^*/4)\}^2 &+ (\beta_1 + 2\beta_6 + 2\beta_7)\ln(x^*/4) + 2\beta_3 + 2\beta_4 + 4\beta_5 - \beta_1\ln(x^h/n^h) \\ &- \beta_2\,(\ln x^h/n^h)^2 - \beta_3 n_a^h - \beta_4 n_c^h - \beta_5 n_a^h n_c^h - \beta_6 n_a^h \ln(x^h/n^h) \\ &- \beta_7 n_c^h \ln(x^h/n^h) = 0\,. \end{aligned} \tag{14}$$

For each set of values for x^h, n^h, n_a^h and n_c^h, this equation gives two roots for x^*, the larger of which is the relevant one, as is clear from Figure 1. Note in particular that the ratio of x^h/x^*, the equivalence scale, will not in general be independent of the value of x^h. This is quite natural: for better-off households, the cost of a larger family may well take up a smaller proportion of total resources, particularly if there are fixed costs associated with the presence of children.

At this point it is necessary to note two practical difficulties which arise in the implementation of the method discussed above. These both occur because of the use of the quadratic relationship between the food share and the logarithm of total expenditure. First, some very poor households may have a level of $\ln x^h$ lower than that which generates the maximum food share for a household of that demographic type (see, for example, point A in Figure 1). Since the bending back of the quadratic is essentially artificial and is not a real feature of the data (only 5 households out of 9663 are in this position), it is inappropriate to base scales on this portion of the curve. As an alternative, the ratio OB/OC, i.e., the scale at the peak food share, is used for these households. A more frequent difficulty occurs when the equation (14) has no solution for x^*. Graphically, this occurs when the predicted food share for household h is larger than the maximum possible food share for the reference household. In theory, if the Engel model is correct, this cannot happen; see equation (11) above which guarantees a unique maximum for the food share for all household types. In practice, a nonlinear restriction on the parameters of each of the three equations would be required, but this was thought to be too complicated for the current exercise. Instead, the ratio OB/OC was also used for such households.

One of the most notable features of the empirical results for Sri Lanka is the extent to which household per capita expenditure explains most of the variation in the food share with only a minor role played by the additional demographic variables. This implies that the behavioral scales are relatively close to the ratios of the head count in each household to that in the reference household, i.e. that $x^h/x^* \simeq n/4$ as a first approximation. Hence, it is convenient for presentation to stay as close as possible to the per capita expenditure concept. Hence, I use $x^*/4$, the "corrected" per capita expenditure, as the basic welfare measure; this is the per capita expenditure which, if given to the reference household, would cause it to display the same food share as household h. Correspondingly, correction factors c^h may be defined by

$$c^h = \frac{x^*/4}{x^h/n^h}. \qquad (15)$$

Hence, multiplying the actual per capita expenditure of household h by c^h leads to corrected per capita expenditure, this latter incorporating the behavioral equivalence scale.

Table 2 lists the calculated correction factors by number of adults and number of children for three different household per capita expenditures, all for the urban sector. Figures for the rural and estates sectors show similar patterns. In each section of the table, the 2-children, 2-adult entry is unity since, in this case, household h and

Table 2

Correction Factors for PCE by Numbers of Adults and Children
Urban Sector: Sri Lanka

1. Household per capita expenditure = 50 rupees per month

Adults:		1	2	3	4	5
Children	0	0.848	0.892	0.938	0.985	1.035
	1	0.907	0.945	0.985	1.025	1.067
	2	0.968	1.000	1.033	1.066	1.100
	3	1.033	1.057	1.082	1.108	1.134
	4	1.100	1.117	1.134	1.151	1.168

2. Household per capita expenditure = 100 rupees per month

0	0.811	0.857	0.904	0.953	1.005
1	0.884	0.926	0.970	1.015	1.061
2	0.962	1.000	1.039	1.079	1.120
3	1.045	1.078	1.112	1.146	1.182
4	1.133	1.160	1.188	1.217	1.246

3. Household per capita expenditure = 200 rupees per month

0	0.787	0.833	0.881	0.931	0.984
1	0.869	0.914	0.960	1.007	1.057
2	0.958	1.000	1.043	1.088	1.135
3	1.054	1.093	1.133	1.174	1.217
4	1.157	1.192	1.229	1.266	1.303

the reference household are identical so that PCE and corrected PCE are also identical. For larger families, and especially for those with a high proportion of children, the correction factors are greater than unity. This is because of economies of scale on the one hand and because of the fact that children cost less than adults, so that the larger the family, the more does actual PCE understate corrected PCE. This correction factor also increases with the level of household PCE since for better-off families the proportional costs of children of large households are less. Conversely, for households smaller than the reference, the same reasons cause the correction factors to be less than one and to decrease with increases in PCE. Clearly, using corrected PCE as a welfare measure may give rather different results than using PCE itself. The next section pursues this in the context of measuring inequality.

D. INEQUALITY AND IDENTIFYING THE POOR

The simplest way of seeing the differences between welfare measures based on per capita expenditure and corrected per capita expenditure is to cross-tabulate the decile positions of households according to each criterion. Table 3 gives the full matrix for the urban sector. This shows just over 42% of the 9663 households being reclassified from one decile to another. Although a few households suffer major reclassification (one household is the bottom decile according to PCE and the top decile according to corrected PCE!), most reclassifications are between adjacent deciles. As might be expected, a larger proportion of households are reclassified in the central deciles with the large majority of households in the outer deciles not changing position. The matrix is also close to symmetric so that roughly equal numbers of households move from decile i by PCE to decile j by corrected PCE as more from j by PCE to i by corrected PCE. These general features are retained in the rural and estate sectors although the total percentage of reclassifications is different, 46% in the rural sector, and only 35% in the estates corresponding to the smaller dispersion of family size and per capita expenditure in that sector.

Table 3

Proportions of Households in each Decile by PCE and Corrected PCE: Urban Sector (4022 households)

Corrected PCE decile / PCE decile	1	2	3	4	5	6	7	8	9	10
1	0.89	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.09	0.71	0.19	0.01	0.00	0.01	0.00	0.00	0.00	0.00
3	0.01	0.16	0.51	0.30	0.02	0.00	0.00	0.00	0.00	0.00
4	0.01	0.03	0.21	0.44	0.29	0.02	0.00	0.00	0.00	0.00
5	0.00	0.00	0.08	0.21	0.40	0.27	0.04	0.00	0.00	0.00
6	0.00	0.01	0.00	0.03	0.23	0.39	0.29	0.05	0.00	0.00
7	0.00	0.00	0.01	0.01	0.04	0.26	0.42	0.24	0.01	0.00
8	0.00	0.00	0.00	0.00	0.00	0.04	0.22	0.49	0.22	0.02
9	0.00	0.00	0.00	0.01	0.01	0.00	0.02	0.21	0.63	0.12
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.86

Note: Rows and columns may not add to unity because of rounding errors.

Table 4 gives the percentages of households reclassified in moving from PCE decile to the decile by (i) corrected PCE (equivalent to predicted food share), (ii) actual food share, (iii) Indian weights per equivalent adult expenditure. The Indian and Taiwanese equivalent scales were taken from Visaria (1979) and are given for comparison between behavioral and nutritional scales. For the rural and urban sectors, the behavioral scales reclassify a fraction

Table 4

Percentage of Households Reclassified by Alternative Criteria

From PCE decile to decile by:	URBAN	RURAL	ESTATES
(i) predicted food share	42.6	46.3	35.7
(ii) actual food share	74.5	79.8	81.7
(iii) per equiv. capita Indian units	34.5	41.5	45.6
(iv) per equiv. capita Taiwan units	49.8	57.6	62.7

of households between the fraction reclassified by the Indian and Taiwanese scales. For the estates sector, many fewer are reclassified by the behavioral scale, perhaps because of its dependence on PCE which is comparatively equally distributed on the estates. The second row of the table, giving the largest number of reclassifications, uses the actual food share as a ranking device. If the equations for the food share had been predicted perfectly, rows 1 and 2 would be identical; as it is, ranking by actual food share is rather different from ranking by predicted food share. The former would be preferable if the actual food share could reasonably be taken as a measure of welfare. In a survey covering only two weeks' expenditure, however, the actual share is subject to a good deal of random fluctuation of no great significance. Presumably, then, if the food share could be predicted better, but still ignoring random fluctuation, the number of reclassifications would lie somewhere between the figures in the first and second rows.

Finally, per capita expenditure and corrected per capita expenditure can be used to calculate measures of inequality between households and individuals. Two measures are computed, the Gini, g, given by

$$g = 1 + \frac{1}{H} - \frac{2}{HY} \sum_{1}^{H} \rho(h) y^h \tag{16}$$

where H is the number of units (households or individuals), y^h is the quantity over which inequality is being measured, Y is its total, and $\rho(h)$ is the rank assigned to household h ranked by y. The second measure is Theil's entropy coefficient scaled to lie between 0 and 1, i.e.

$$t = 1 - \frac{Y}{H} \exp\{ - \Sigma \frac{y^h}{Y} \ln y^h \}. \tag{17}$$

These are computed for both per capita expenditure and corrected per capita expenditure over both households and individuals. By the former , PCE in one version or the other is taken as the welfare measure of the <u>household</u>, so that inequality is inequality of household welfare. By the latter, each individual in each household is ascribed the PCE for the household as a whole, so that inequality is now inequality of individual welfare. It is also interesting, for the reasons discussed in the previous paragraph, to compute inequality based on the <u>actual</u> food share. However, since the food share is nonlinearly related to PCE, this cannot be done directly to give a number comparable with the other measures. The simplest way to overcome this is to use the relationship between the food share and PCE for the reference household to transform actual food shares into hypothetical PCE's. This should be understood solely as a scaling device; it clearly has no behavioral significance.

The various measures are presented in Table 5. Qualitatively, the behavior of the entropy measures is similar to that of the Gini coefficients; the former is considerably easier to calculate and has the theoretical advantage over the Gini of being more sensitive to the welfare of the poorest, the greater inequality is (see Blackorby and Donaldson (1978)). By all measures, both PCE and corrected PCE are most equally distributed on the estates and least equally distributed in the urban areas. The move from PCE to corrected PCE always decreases inequality; this is presumably because larger families typically have lower PCE and the correction partially compensates for this. The effect on measured inequality of the correction is somewhat more than the effect of moving to Indian weighted equivalent units but somewhat less than that of moving to Taiwanese units (figures not shown in the table). Note finally that the inequality among the actual food shares is rather large and, remarkably, is largest on the estates. This partly reflects the fact that the food share regressions fit worst on the estates so that there is a marked contrast between inequality in PCE 1 and inequality in PCE 2. However, it is a genuine feature of the data that PCE is more equally distributed than is the food share and that the difference is most marked on the estates. Clearly, there is likely to be a considerable pay-off to further study of the determinants of the food share, particularly in this sector.

Table 5

Inequality Measures for Sri Lanka

ENTROPY MEASURES		URBAN	RURAL	ESTATE	ALL ISLAND
Households	PCE	.1681	.1150	.0966	.1335
	PCE 1	.1548	.1067	.0861	.1258
	PCE 2	.2162	.1935	.2273	.2110
Individuals					
	PCE	.1637	.1094	.0903	.1287
	PCE 1	.1565	.1087	.0826	.1281
	PCE 2	.2162	.1943	.2309	.2117
GINI'S					
Households	PCE	.3259	.2683	.2356	.2599
	PCE 1	.3124	.2592	.2222	.2535
	PCE 2	.3854	.3636	.3975	.3686
Individuals					
	PCE	.3170	.2582	.2254	.2224
	PCE 1	.3115	.2593	.2133	.2376
	PCE 2	.3831	.3619	.3962	.3600

Notes: PCE 1 is corrected per capita equivalent, PCE 2 is hypothetical PCE corresponding to the actual food share.

E. CONCLUSIONS

This essay has given a brief overview of the theory of equivalence scales together with an application of Engel's method to the Sri Lankan socioeconomic survey of 1969-70. A welfare measure was defined consisting of a correction to per capita expenditure depending on the number of adults and children in the household and on the level of per capita expenditure. This correction reflected economies of scale in consumption, the cheapness of children relative to adults, and the increasing importance of these effects with increasing per capita expenditure. It was shown that the ranking of households according to corrected per capita expenditure differed significantly from the ranking according to uncorrected per capita expenditure. Hence, many households classified as poor by one criterion would not be so by the other. Overall measured inequality was somewhat reduced by the correction but the difference was not so large as to suggest that allowing for adult equivalents is of major importance in inequality comparisons. Instead, further work is necessary to explore the link between corrected per capita expenditure and other aspects of poverty and deprivation. The current paper has shown only that the correction is feasible in both theory and practice. What still remains to be done is the demonstration that allowing for demographic effects in this way gives a more accurate and more useful definition of poverty.

III. ON A METHOD FOR MEASURING THE COSTS OF CHILDREN

A. INTRODUCTION

There are many different procedures for measuring the costs of children, even if we confine ourselves, as I shall do here, to schemes based on observing consumer behavior. Most "modern" academic work uses formulations derived from the work of Prais and Houthakker (1955), of Barten (1964) or of Gorman (1976): see in particular Muellbauer (1977, 1980), Lazear and Michael (1980a, 1980b) (who, although not saying so, use the Barten model) and Pollak and Wales (1981). The Prais-Houthakker, Barten and Gorman models, while relatively sophisticated, are not invulnerable to theoretical challenge (for example, none offers any explicit account of how goods are allocated within the household), and all are difficult to implement, impossibly so by the standards of most LDC's. All involve non-linear estimation, there are serious identification problems to be overcome, and data requirements are considerable, for example, time series of cross-sections, or, as in the Lazear and Michael study, a cross-section and an extraneously obtained (and by no means obviously compatible) set of parameter estimates for a complete system of demand equations. In consequence, these models cannot offer a robust and readily applied methodology for measurement and, indeed, even in developed economies, have not been used for practical measurement in a policy context, remaining instead experimental and academic.

By contrast, the oldest methodology, that of Engel (1895) retains wide currency and, for example, is still embodied in the official United States definition of the poverty line. By adopting the food share (or the share of some group of necessities) as an (inverse) indicator of welfare, families with different numbers of children but with the same welfare level can be identified so that the difference in their outlays gives an immediate behavioral estimate of the costs of the additional children corresponding to that particular welfare level. This procedure has the virtue of great simplicity and can be readily implemented with minimal data and econometric technology. The problem lies with the theory, since there is no very convincing reason why the food share should indeed correctly indicate welfare. Certainly, following Muellbauer (1977), the Engel methodology can be given a theoretical rationale (see Section C below), but the assumptions required only serve to reinforce the basic point that the hypothesis is not rooted in any plausible model of behavior.

The purpose of this essay is to discuss an alternative to the Engel methodology, equally straightforward in practice, but based on rather different and hopefully more plausible assumptions. The method appears to have been first suggested by Rothbarth (1943), has since been used in one form or another by Nicholson (1949), Henderson (1949-50a,b) and Garganas (1977) (who combines it, not entirely legitimately, with the Prais-Houthakker model) and is intimately related to much earlier attempts to measure the cost of children such as that of Dublin and Lotka (1947): see in particular the references in Espenshade (1976). Rothbarth divides commodities into children's goods (group A say) and goods not consumed by children (group B) and suggests that total outlay on group B (he uses luxuries including saving) may be taken as an indicator of welfare. Hence, instead of identifying households with identical food shares, households with identical B group expenditures are found and then, as in the Engel model, differences in total outlay give the measure of the cost of children. Comparison of the individual A group commodity expenditures between households also allows the total cost of children to be broken up into commodity components. It should be immediately clear that this procedure, although operationally similar to the Engel method, is conceptually quite different. It is based on different assumptions and in general gives different figures for the cost of children. However, it might be interesting to note that a common audience reaction to presentation of the Engel model is an intuitive justification in terms of children "needing" food. Such an assumption is quite at home within the Rothbarth method but, as we shall see, it is directly contrary to the central supposition of the Engel method, that the food share correctly indicates welfare. Even so, I suspect that the very frequency of the misapprehension suggests that the Rothbarth model is essentially more plausible than that of Engel, an issue to which I shall return below.

The plan of the paper is as follows. Section B discusses the theoretical basis of the Rothbarth procedure, and I shall be particularly concerned with issues of restrictiveness, plausibility, and testability. Section C compares the Rothbarth with the traditional Engel procedure and I show that the methods are essentially incompatible, with the Engel method always leading to higher estimated costs of children if the same empirical evidence is used by both. Section D contains some illustrative numbers based on my earlier work on Sri Lanka, Essay No. I. For example, two children are estimated as 0.82 "couple equivalents" for Sri Lankan urban households according to the Engel methodology, but only as 0.21 "couple equivalents" according to the Rothbarth procedure, a difference of some importance for evaluating welfare. Finally, Section E contains a summary comparative evaluation of the two methodologies. My personal preference for measuring the costs of children is the Rothbarth procedure although if it were possible to accept its assumptions, the Engel methodology would offer a potentially much more powerful and general tool for welfare measurement.

B. THE THEORETICAL BASIS OF THE ROTHBARTH PROCEDURE

The earliest work on measuring the costs of children simply compared the expenditures of households with different numbers of children. This is sound only if the households involved have identical living standards; otherwise, demographic and welfare effects are confounded. Rothbarth's suggestion is that expenditure on some group of goods, luxuries, or commodities such as alcohol, tobacco, or adult clothing, serve as an indication of welfare independently of the number of children. If this is accepted, the cost of children can be estimated by comparing expenditures or incomes of households with identical outlays on this group. What is required is that all commodities be partitioned into two groups, A and B, the former being explicitly associated with the needs of children. Households with the same total outlay but different numbers of children will not spend the same on group B but the differences will be due solely to the income effects of the group A needs of the children.

Let q be the vector of all commodities and let this be partitioned into two subvectors q_A and q_B of children's goods and other goods respectively. This partition is given a priori; I shall discuss how it might be determined below. The price vector p is conformably partitioned as (p_A, p_B). Total outlay is x with $x_A = p_A \cdot q_A$, expenditure on A goods and $x_B = p_B \cdot q_B$, expenditure on B goods. The vector (or scalar) a contains the relevant characteristics of children; at its simplest, a would be a scalar giving the number of children, but more generally, numbers will be broken down by age, sex or other characteristics. I shall also maintain the essential (but by no means practically innocuous assumption) that prices are the same for all households.

It is straightforward to show that x_B is an indicator of welfare independent of a if and only if the cost function can be written in the form

$$c(u, p_A, p_B, a) = \gamma(u, p_A, p_B, a) + \beta(u, p_A, p_B) \tag{1}$$

where u is utility, $c(u, p_A, p_B, a)$ is the minimum cost of reaching u at prices p_A and p_B for a household with children's characteristics a, $\gamma(u, p_A, p_B, a)$ is homogeneous of degree zero in p_B and homogenous of degree one in p_A and $\beta(u, p_A, p_B)$, as well as being independent of a, is monotone increasing in u, and linearly homogeneous in p_A and p_B jointly, i.e., linearly homogeneous in p. Given (1), expenditure on B-goods is given, for example, by

$$x_B = \sum_{i\varepsilon B} p_{Bi} \frac{\partial\beta(u, p_A, p_B)}{\partial p_{Bi}} = \theta(u, p_A, p_B). \tag{2}$$

Thus, given that p_A and p_B are the same for all households, x_B and u are monotonically related with the former correctly indicating the latter. Expenditure on A-goods is given by:

$$\begin{aligned} x_A &= \gamma(u, p_A, p_B, a) + \beta(u, p_A, p_B) - \theta(u, p_A, p_B) \\ &= c(u, p_A, p_B, a) - \theta(u, p_A, p_B). \end{aligned} \tag{3}$$

For any given u, p_A and p_B, costs of children are given by the differences in x_A corresponding to differences in a. If we normalize so that for the reference household, a^o, say $\gamma(u, p_A, p_B, a^o)$ is zero, then

$$CS = \gamma(u, p_A, p_B, a^h) \tag{4}$$

is the compensation required at p_A, p_B for household h on welfare level u to make up for the extra children embodied in a^h over those contained in the reference household. In this additive formulation, consumer surplus measures are naturally easier to work with than the ratios involved in equivalence scales, although the latter are easily enough obtained. Specifically, if θ_R is the scale based on the reference household,

$$\theta_R = \frac{c(u, p_A, p_B, a^h)}{c(u, p_A, p_B, a^o)} = \frac{\gamma(u, p_A, p_B, a^h) + \beta(u, p_A, p_B)}{\beta(u, p_A, p_B)} \tag{5}$$

In practice, the costs of children would be determined from the empirically estimated Engel curves for A- and B-goods. It is fairly clear that cost function (1) does not restrict the Engel curves except to ensure that group B as a whole is normal. Assume therefore that, on empirical grounds, an Engel curve

$$x_B^h = g(x^h, a^h) \tag{6}$$

is established. I have suppressed the prices p_A and p_B on the assumption that they are identical for all h and that only a single cross-section is available. The "money-metric" welfare measure for household h, x_*^h say, is the total outlay required by the reference household to generate an x_B equal to x_B^h. Hence x* is implicitly defined as a function of x^h, a^h and a^o by

$$g(x^h, a^h) = g(x^*, a^o), \tag{7}$$

the cost of children is $(x^h - x^*)$ and the eqivalence scale, θ_R is simply

$$\theta_R = x^h/x^* = \theta_R(x^h, a^h, a^o). \tag{8}$$

θ_R can be thought of as "adult" or "couple" equivalents, depending on the choice of a^o, but note the dependence on x^h; scales may rise or fall with income depending on whether costs of children rise more than or less than proportionately with income.

It is important in interpreting the Rothbarth model and its associated cost function to realize that it is expenditure on B-goods as a group which indicates welfare, not the expenditures on individual items. The preferences underlying (1) guarantee that changes in the number of children have only income effects on group B expenditures as a whole, but that does not prevent the structure of expenditures within the group from varying with children. For a household at a fixed welfare level, i.e., one which is fully compensated for the costs of any extra children it might have, expenditure on group B is independent of the number of children although the arrival of additional infants (or age changes of existing ones) may well cause rearrangements within the total. Although it is not difficult to think of cases where this might happen, it is perhaps natural to also examine the case when all goods in group B are independent of the number of children provided the household is fully compensated. This special case occurs when, in (1), the function $\gamma(u, p_A, p_B, a)$ is independent of p_B, not merely zero degree homogeneous (the two are, of course, the same when group B contains but a single good). The "individual commodities" version of the Rothbarth model thus requires the cost function

$$c(u, p_A, p_B, a) = \gamma^*(u, p_A, a) + \beta(u, p_A, p_B). \qquad (9)$$

Such a formulation gives a cost function which is additively separable in p_B and a so that, not only are the compensated demand functions of all B-goods independent of the number of children, but also the costs of children are independent of the prices of B-goods. This is not true of the "group version" (1); an increase in the price of cigars, even if compensated, causes substitutions amongst A-goods which, in general, changes (in one direction or the other) the cost of children. Even so, the restricted version (9) remains plausible and it may actually be easier in practice to identify a priori individual B-goods than to hazard a guess about a broad group.

How then are the A-and B-goods to be identified empirically, and is it possible to test whether there exists a group of goods for which children have no influence other than income effects? For the most general version, the cost function (1), <u>there are in fact no restrictions on behavior</u>. (This can easily be established by checking derivatives.) Hence, the partition of commodities into A-and B-groups must be given a priori and it is this prior knowledge which allows identification of the costs of children. Hopefully, the data will allow measurement of the effects of income and of children on the allocation between the two groups. Prior evidence that, for one group, the child effects operate only through income responses is exactly the information required to identify the effects of children on welfare and thence to measure the costs of children. No testing of this prior evidence is therefore possible, at least within the framework of the model. As Pollak and Wales (1976) have argued (see also Deaton (1980), p. 65)), the full effects of demographic variables on welfare cannot be inferred by only observing the effects of demographic variables on commodity demands. This theoretical underidentification can only be overcome by the use of prior information such as that embodied in the Rothbarth procedure. Such assumptions inevitably contain an arbitrary component; in the Rothbarth model different partitions into A and B goods will normally produce different estimates of costs and different scales. However, to the extent that the model relies on a fairly commonsensical proposition, that there are goods which are a priori identifiable as not required (directly or indirectly) by children, the underlying prior information may be accorded a relatively high degree of belief.

In practice there are likely to be several possible B-goods under consideration and each may be thought to be plausibly independent of the number of children. If so, the "individual commodities" version of the model, i.e. (9), can be applied and this does have testable implications. In particular, it is easily shown that, for all goods belonging to group B, i.e. for $i \in B$, the ratios

$$\frac{\partial q_i}{\partial a_j} \Big/ \frac{\partial q_i}{\partial x} \qquad \text{for all } j \tag{10}$$

are independent of i, where a_j is the jth component of a (if it has more than one, otherwise $j = 1$. Adding (or subtracting) goods from B which satisfy (10) should leave the estimated scales unchanged and conversely, (10) can be used as a criterion for defining the B group. There must, of course, be one "seed" commodity for which the independence from children cannot be tested, but given this, others can be sought empirically. Such testing can go some way to check prior beliefs about A- versus B-goods and also serve to estimate the robustness of the procedure. One example is not by itself convincing, but Garganas (1977) reports robust results for the method on British Family Expenditure Survey data. Unfortunately, two of the goods which Garganas uses as B-goods are drink and tobacco, both of which are notoriously under-reported in the FES. Such difficulties may well recur in other samples and it may be necessary to model the underreporting explicitly if the Rothbarth method is to be reliably implemented.

C. COMPARISON WITH THE ENGEL MODEL

An alternative procedure for estimating the costs of children, usually associated with the name of Engel 1/, also relies on partitioning commodities into two groups, usually food and non-food. The food share in total expenditure is then used as an inverse welfare indicator, or equivalently the non-food share as a direct welfare indicator. This would compare with the Rothbarth method's use of the non-food expenditure level, if food were taken as the prime children's good. The rationale for the Engel model appears to be entirely empirical; poorer families have larger food shares than do richer families while larger families with the same total outlay also have larger food shares than smaller families. Note, however, that such observations do not imply that the food share indicates welfare correctly and are themselves consistent with alternative models, in particular that food is more of a children's good than is non-food.

1/ Engel's method (see Deaton and Muellbauer (1980 Ch.8) for an exposition) assumes that a common deflator, a function of household size, can be applied to all commodities and total outlay. Such an assumption implies that the food share - and all other shares - are an indicator of welfare. Here I focus on the welfare indication aspect only.

The cost function for the Engel model is similar to that of the Rothbarth model but takes a multiplicative rather than an additive form. Again let A (usually food) and B be the two commodity groups (these may, of course, be quite different from the two groups used in the Rothbarth procedure above). The share in total outlay of either group will then correctly indicate welfare if and only if the cost function takes the form

$$c(u, p_A, p_B, a) = m(u, p_A, p_B, a)\,\phi(u, p_A, p_B) \tag{11}$$

where $m(u, p_A, p_B, a)$ - which is now the equivalence scale rather than the cost of children - is homogeneous of degree zero in <u>both</u> p_A and p_B, each taken separately, while $\phi(u, p_A, p_B)$ is linearly homogeneous in p_A and p_B together, i.e. in p. If w_A is the share of group A, clearly,

$$w_A = \sum_{i\varepsilon A} \frac{\partial \ln \phi(u, p_A, p_B)}{\partial \ln p_{Ai}} = w_A(u, p_A, p_B). \tag{12}$$

Identifiability and testability issues are technically identical to those arising for the Rothbarth model and discussed in Section B above. The cost function (11) has no implications which are refutable on demand data alone so that, once again, the prior assumption, in this case that the food share indicates welfare, is what identifies the scale and allows interpretation of the data. Restricted versions of (11), like restricted versions of (1), are testable and there are obvious possibilities; for example $m(u, p_A, p_B, a)$ can be made independent of either or both of p_A and p_B so that shares of A-goods, B-goods, or all goods indicate welfare. Choice between (11) and (1), the Engel and Rothbarth models, then rests entirely on the plausibility of the respective identifying assumptions. For my own part, the Rothbarth formulation is superior because I can understand and find plausible the distinction between children's goods and other goods. I cannot, on the other hand, think of any convincing reason why the food share should correctly indicate welfare and, in practice, I should have no basis for choosing between the food share, the cereals share, the durables share, the "basic needs" share, or any other share as the appropriate basis for measurement. For <u>all</u> to be simultaneously valid, the function $m(u, p_A, p_B, a)$ must be independent of both p_A and p_B. This denies any concept of specific needs for children and is entirely implausible. These difficulties arise much less severely with the Rothbarth procedure.

It should be clear that the two models give different results but this is worth demonstrating formally. For the purposes of argument, I assume that the A-goods and B-goods are the same for both procedures so that for the Rothbarth model, x_B identifies welfare, while for the Engel model, w_B, the share of B-goods plays the same role. Since neither cost function (1) nor (11) place any restriction on behavior, both are consistent with any estimated Engel curve, and so the same empirical evidence can be used as input for both methods. The argument below shows that the Engel procedure always gives the larger estimate of the cost of children, essentially because, since w_B is less elastic with respect to x than is x_B, compensation to equate x_B for extra children is less than the compensation required to restore w_B to its original level. Recalling equations (7) and (8), θ_R is defined as the ratio of x^h/x^* where x^* is the solution to

$$g(x^h, a^h) = g(x^*, a^0), \tag{13}$$

where, as before, $g(x,a)$ is the Engel curve for group B. Similarly, the Engel scale θ_E is defined by x^h/x^{**}, where x^{**} is the expenditure which equates $w_B(x^h,a^h)$ and $w_B(x^{**},a^0)$, i.e.

$$\frac{g(x^h,a^h)}{x^h} = \frac{g(x^{**},a^0)}{x^{**}} . \tag{14}$$

Hence, combining (13) and (14)

$$g(x^h,a^h) = g(x^*,a^0) = \theta_E\, g(x^{**},a^0). \tag{15}$$

Since, by assumption good B is normal, $\theta_E > 1$ must imply $x^{**} < x^*$, so that $\theta_R < \theta_E$. Hence, the Rothbarth model gives a lower estimate of costs than does the Engel model. As will be seen in the next section, the differences appear to be typically very large.

Finally, there is perhaps the most important of all the differences between the two procedures. Potentially at least, the Engel method can be applied quite generally with the food share acting as a welfare measure in many different situations, for example, to compare households in different regions, in different countries (and its dimensionlessness is a great convenience here), and with differing numbers of adults. By contrast, the Rothbarth procedure relates specifically to the behavior of households with children so that, for example, it would be absurd to suggest that two households with different numbers of adults but with identical outlays on B-goods had the same welfare level. The method cannot therefore be used to give equivalence scales which sumultaneously take account of the costs of children and economies of scale over both adults and children, as can the Engel model. Of course, this "advantage" of the Engel model is only an advantage if its assumptions are correct and the generality of the food share as a welfare indicator is due in no small measure to the lack of any convincing reason why it should be so in any circumstances. It may, for example, make sense to combine the models with an additive formulation for children and a multiplicative formulation for adults, i.e.

$$c(u, p_A, p_B, a) = \gamma(u, p_A, p_B, n_c) + m(u, n_a)\, \beta(u, p_A, p_B) \qquad (16)$$

for n_c = number of children and n_a = number of adults. Given (16), the Rothbarth procedure correctly indicates welfare for households with equal numbers of adults and can be used to measure the costs of children while the within-group shares of group B expenditures indicate welfare over households with different numbers of adults. I have not seriously thought about applying this model as yet, but it strikes me as being well worth thinking about.

D. SOME ILLUSTRATIVE CALCULATIONS

In earlier work, Essays Nos. I and II, I estimated food Engel curves for the three sectors of Sri Lanka based on the 1969-70 Socioeconomic Survey and calculated a number of equivalence scales using the Engel methodology. The simplest Engel curve considered was of the form

$$w_f = \beta_0 + \beta_1 \ln\left(\frac{x}{n}\right) + \beta_2 n_a + \beta_3 n_c \qquad (17)$$

for food share w_f, parameters β_0, β_1, β_2 and β_3, number of persons in the household n, number of adults (aged 15-59) n_a, and number of children (aged 0-14) n_c. The final preferred version of (17) contained also a squared term in ln (x/n) and interaction terms between n_a, n_c and ln (x/n). However, the extended form would complicate the illustrative calculations here without changing the results in any essential way. I am also going to identify food as the children's good although, if starting anew, experimentation with other definitions of group A would be desirable.

Following equation (14), the Engel scale θ_E is given by x/x** with x** given by

$$\beta_1 \ln (x^{**}/n^0) + \beta_2 n_a^0 + \beta_3 n_c^0 = \beta_1 \ln (x/n) + \beta_2 n_a + \beta_3 n_c \qquad (18)$$

where n^0, n_a^0, and n_c^0 are the number of persons, adults and children in the reference household. For further calculations I take $n_a^0 = n^0$ and $n_c^0 = 0$, i.e. the reference household contains n^0 working age adults, no old people and no children. Hence, rearranging (18), the Engel scale is

$$\theta_E = \left(\frac{n}{n^0}\right) e^{-\psi_a (n_a - n_a^0) - \psi_c n_c} \qquad (19)$$

where $\psi_a = \beta_2/\beta_1$ and $\psi_c = \beta_3/\beta_1$ both of which always turn out to be positive numbers. Hence, θ_E is always less than (n/n^0), the headcount ratio, but since ψ_a and ψ_b are generally small, (n/n^0) is a fairly good first approximation to the scale. In Sri Lanka, according to the Engel methodology, children are almost as costly as adults.

To evaluate the Rothbarth scale, we write x_B, expenditure on non-food, as

$$x_B = x(1 - \beta_0) - \beta_1 x \ln\left(\frac{x}{n}\right) - \beta_2 n_a x - \beta_3 n_c x \tag{20}$$

and $\theta_R = x^*/x$ with x^* defined, using the same reference household, as

$$x^*(1 - \beta_0) - \beta_1 x^* \ln\left(\frac{x^*}{n^0}\right) - \beta_2 x^* n^0 = x(1-\beta_0) - \beta_1\, x \ln\left(\frac{x}{n}\right) - \beta_2 n_a x - \beta_3 n_c x. \tag{21}$$

This equation does not yield an explicit solution for θ_R but on rearrangement gives

$$\ln \theta_R - \theta_R(\psi_0 - \ln \frac{x}{n} - \psi_a n_a - \psi_c n_c) + (\psi_0 - \ln \frac{x}{n^o} - \psi_a n_a^0) = 0 \tag{22}$$

where $\psi_0 = (1 - \beta_0)/\beta_1$, and which may readily be solved by elementary nonlinear solution techniques. An excellent first approximation is usually given by

$$\theta_R^{(0)} = (1-\psi_0 + \ln \frac{x}{n^0} + \psi_a n_a^0) \;/\; (1-\psi_0 + \ln \frac{x}{n} + \psi_a n_a + \psi_c n_c). \tag{23}$$

Updating can be carried out as necessary using the method of false position, i.e. according to

$$\theta_R^{(r+1)} = \theta_R^{(r)} - f(\theta_R^{(r)}) \;/\; \left\{(1/\theta_R^{(r)}) - \psi_0 + \ln \frac{x}{n} + \psi_a n_a + \psi_c n_c\right\} \tag{24}$$

where $f(\theta)$ is the left hand side of equation (22).

Note that, in these particular calculations, θ_R varies with x whereas θ_E does not. However, this is a feature only of this example and, in fact, both scales would vary with x if the full Engel curve with quadratic and interaction terms had been used instead of the simplified version (17).

In presenting the results, it must be remembered that the Rothbarth procedure is not designed to provide comparisons across households with different numbers of adults. Hence, it is convenient to take a series of reference households, each containing only adults, which are used as a basis of comparison. The scales then express the costs of maintaining a household with n_a adults and n_c children relative to the costs of maintaining a household with n_a adults alone, with separate scales for each value of n_a.

The parameters used are for the urban sector of Sri Lanka and are taken from Deaton (1980, Table 7). The values used are $\beta_0 = 1.5117$, $\beta_1 = -0.2012$, $\beta_2 = -0.0059$, $\beta_3 = 0.0093$. The Engel scales are given in Table 1 below; these are essentially those given in Table 8 of the earlier paper. Since equivalences are given relative to the unit consisting of adults only, the total number of equivalent adults is obtained by multiplying each number by the number of adults given in the first column. Note that, although in all cases, children are cheaper than adults, the difference is not great. By contrast, the Rothbarth scales are given in Table 2. These are calculated at a per capita expenditure level of 80 rupees per month (close to the mean); note that this implies that different households in Table 2 have different levels of total household expenditure. The scales are very much smaller than in Table 1 with children costing closer to 1/4 adults rather than whole adults as in Table 1. The costs of an extra child are estimated by the Rothbarth method as very much the same whatever the reference number of adults, while there appear to be considerable economies of scale to several children, a feature shared with the Engel

Table 1

Equivalence Scales for Sri Lanka: Urban Sector

Engel methodology

No of children		0	1	2	3
No of adults	1	1.00	1.88	2.71	3.44
	2	1.00	1.43	1.82	2.18
	3	1.00	1.27	1.51	1.74
	4	1.00	1.19	1.36	1.52

Table 2

Equivalence Scales for Sri Lanka: Urban Sector

Rothbarth Methodology

No of children		0	1	2	3
No of adults	1	1.00	1.23	1.36	1.44
	2	1.00	1.12	1.21	1.27
	3	1.00	1.08	1.14	1.19
	4	1.00	1.06	1.10	1.14

estimates. However, the Rothbarth estimates of children as costing rather less than a quarter of an adult are very much in line with the frequently heard suggestion that in many LDC's children are only a net burden to the family for the first few years of life. In any case, Table 2 makes more intuitive sense to me than does Table 1. Others may have different views.

I note finally that repetition of Table 2 for other levels of PCE shows a marked tendency for the scales to fall with increases in PCE. For example, the (1,1) figure in Table 2 of 1.23 (at 80 rupees) is 1.31 at 40 rupees and 1.18 at 160 rupees. But not too much weight should be attached to these numbers at this stage since the Engel curve without the quadratic term is not really flexible enough to support such calculations. In later work, I hope to recalculate the scales using the more flexible form.

E. CONCLUSIONS

In general, it is not possible to identify the costs of children from the effects of demographic composition on consumer demand unless a priori identifying assumptions are made. The Engel method, the oldest and most popular technique, assumes that the share of food (or some other necessity) in total outlay is a welfare indicator which is independent of demographic composition. This essay has discussed an alternative methodology, here accredited to Erwin Rothbarth. This identifies the costs of children by positing the existence of goods not used by children, the total expenditure on which is an indicator of welfare over otherwise identical households differing only in their numbers of children. For reasons given above, I find this particular identifying assumption more plausible and less arbitrary than that used by the Engel method. Technically, the methods are similar and both are easy to use. Experiments with Sri Lankan data yielded much lower, and to me more plausible estimates of child costs using the Rothbarth procedure than were given by the Engel formulation. Further work is required to examine the robustness of the method and to extend it to allow the general applicability of the Engel procedure as a device for measuring welfare to be combined with the plausibility of the Rothbarth procedure for identifying the costs of children. It seems to me that this work is well worth doing. The Engel method has been unchallenged in practice, not because of its plausibility, but because of its computational simplicity. The alternative discussed here is just as simple and, I believe, a good deal more plausible.

REFERENCES

Barten, A.P. (1964), "Family Composition, Prices and Expenditure Patterns," in P.E. Hart, G. Mills and J.K. Whitaker (eds.), Econometric Analysis for National Planning, London: Butterworth.

Blackorby, C., and D. Donaldson (1978), "Measures of Equality and their Meaning in Terms of Social Welfare," Journal of Economic Theory, Vol. 18, pp. 59-80.

Breusch, T.S. and A.R. Pagan (1979), "A Simple Test for Heteroscedasticity and Random Coefficient Variation," Econometrica, Vol. 47, pp. 1287-94.

Coondoo, D. (1972), "A Note on the Effect of Household Composition on Household Consumption Pattern," Calcutta: ISI mimeo.

Deaton, A.S. (1980), "The Measurement of Welfare: Theory and Practical Guidelines," Living Standards Measurement Study Working Paper No. 7, Washington, D.C.: Development Research Center, World Bank.

----------- and J. Muellbauer (1980), Economics and Consumer Behavior, New York: Cambridge University Press.

Dublin, L.I. and A.J. Lotka (1947), The Money Value of a Man, New York: Ronald Press Co.

Engel, E. (1895), "Die Lebenskosten Belgischer Arbeiter-Familien früher und jetzt," International Statistical Institute Bulletin, Vol. 9, pp. 1-74.

Espenshade, T.J. (1976), The Cost of Children in Urban United States, Westport, Connecticut: Greenwood Press.

Garganas, N.C. (1977), "Family Composition, Expenditure Patterns and Equivalence Scales for Children" and "Equivalence scales: a Methodological Review," Chapter 7 and Appendix VII in Fiegehan, G.C., P.S. Lansley and A.D. Smith, Poverty and Progress in Britain 1953-73, Cambridge: Cambridge University Press for the National Institute of Economic and Social Research.

Gorman, W.M. (1976), "Tricks with Utility Functions," in M. Artis and R. Nobay (eds.), Essays in Economic Analysis, Cambridge University Press.

Granger, C.W.J. and P. Newbold (1974), "Spurious Regressions in Econometrics," Journal of Econometrics, Vol. 2, pp. 111-20.

Henderson, A.M. (1949-50a), "The Costs of Children," Parts I-III, Population Studies, Vol. 3, pp. 130-150, Vol. 4, pp. 267-298.

-------------- (1949-50b), "The Cost of a Family," Review of Economic Studies, Vol. 17, pp. 127-148.

Lazear, E.P. and R.T. Michael (1980a), "Family Size and the Distribution of Real Per Capita Income," American Economic Review, Vol. 70, pp. 91-107.

------------------------------ (1980b), "Real Income Equivalence Among One Earner and Two Earner Families," American Economic Review, Vol. 70, (papers and proceedings), pp. 203-8.

Leser, C.E.V. (1963), "Forms of Engel Functions," Econometrica, Vol. 31, pp. 694-703.

Mood, A.M., F.A. Graybill and D.C. Boes (1974), Introduction to the Theory of Statistics, New York: McGraw Hill.

Muellbauer, J. (1977), "Testing the Barten Model of Household Composition Effects and the Cost of Children," Economic Journal, Vol. 87, pp. 460-87.

-------------- (1980), "The Estimation of the Prais-Houthakker Model of Equivalence Scales," Econometrica, Vol. 48, pp. 153-176.

Nicholson, J.L. (1949), "Variations in Working-Class Family Expenditure," Journal of the Royal Statistical Society (Series A), Vol. 112, pp.359-411.

Pollak, R.A. and T.J. Wales (1979), "Welfare Comparisons and Equivalence Scales," American Economic Review, Vol. 69, (papers and proceedings), pp. 216-21.

---------------------------- (1981), "Demographic Variables in Demand Analysis," Econometrica, forthcoming.

Prais, S.J. and H.S. Houthakker (1955), The Analysis of Family Budgets, Cambridge: Cambridge University Press.

Rothbarth, E. (1943), "Note on a Method of Determining Equivalent Income for Families of Different Composition," Appendix IV in C. Madge, War-time Pattern of Saving and Spending, Occasional Paper No. 4, National Institute of Economic and Social Research, London: Macmillan.

Samuelson, P.A. (1974), "Complementarity - an Essay on the 40th Anniversary of the Hicks-Allen Revolution in Demand Theory," Journal of Economic Literature, Vol. 15, pp. 24-55.

Singh, B. (1972), "On the Determination of Economies of Scale in Household Consumption," International Economic Review, Vol. 13, pp. 257-70.

-------- and A.L. Nagar (1973), "Determination of Consumer Unit Scales," Econometrica, Vol. 41, pp. 347-55.

Sri Lanka, Department of Census and Statistics (1973), Socio-economic Survey of Sri Lanka 1969-70, Statistical Tables, Vols. I and II, Sri Lanka: Department of Government Printing.

Visaria, P. (1979), "Demographic Factors and the Distribution of Income: Some Issues," in Economic and Demographic Change: Issues for the 1980s, pp. 289-320, Liege: IUSSP.

White, H. (1980), "A Heteroscedasticity-consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity," Econometrica, Vol. 48.

Working, H. (1943), "Statistical Laws of Family Expenditure," Journal of the American Statistical Association, Vol. 38, pp. 43-56.

Yule, G.U. (1926), "Why Do We Sometimes Get Nonesense Correlation Between Time Series?", Journal of the Royal Statistical Society, Vol. 89, pp. 1-64.

LSMS Working Papers (continued)

No. 28 *Analysis of Household Expenditures*

No. 29 *The Distribution of Welfare in Côte d'Ivoire in 1985*

No. 30 *Quality, Quantity, and Spatial Variation of Price: Estimating Price Elasticities from Cross-sectional Data*

No. 31 *Financing the Health Sector in Peru*

No. 32 *Informal Sector, Labor Markets, and Returns to Education in Peru*

No. 33 *Wage Determinants in Côte d'Ivoire*

No. 34 *Guidelines for Adapting the LSMS Living Standards Questionnaires to Local Conditions*

No. 35 *The Demand for Medical Care in Developing Countries: Quantity Rationing in Rural Côte d'Ivoire*

No. 36 *Labor Market Activity in Côte d'Ivoire and Peru*

No. 37 *Health Care Financing and the Demand for Medical Care*

No. 38 *Wage Determinants and School Attainment among Men in Peru*

No. 39 *The Allocation of Goods within the Household: Adults, Children, and Gender*

No. 40 *The Effects of Household and Community Characteristics on the Nutrition of Preschool Children: Evidence from Rural Côte d'Ivoire*

No. 41 *Public-Private Sector Wage Differentials in Peru, 1985–86*

No. 42 *The Distribution of Welfare in Peru in 1985–86*

No. 43 *Profits from Self-Employment: A Case Study of Côte d'Ivoire*

No. 44 *The Living Standards Survey and Price Policy Reform: A Study of Cocoa and Coffee Production in Côte d'Ivoire*

No. 45 *Measuring the Willingness to Pay for Social Services in Developing Countries*

No. 46 *Nonagricultural Family Enterprises in Côte d'Ivoire: A Descriptive Analysis*

No. 47 *The Poor during Adjustment: A Case Study of Côte d'Ivoire*

No. 48 *Confronting Poverty in Developing Countries: Definitions, Information, and Policies*

No. 49 *Sample Designs for the Living Standards Surveys in Ghana and Mauritania / Plans de sondage pour les enquêtes sur le niveau de vie au Ghana et en Mauritanie*

The World Bank

Headquarters
1818 H Street, N.W.
Washington, D.C. 20433, U.S.A.

Telephone: (202) 477-1234
Telex: WUI 64145 WORLDBANK
RCA 248423 WORLDBK
Cable Address: INTBAFRAD
WASHINGTONDC

European Office
66, avenue d'Iéna
75116 Paris, France

Telephone: (1) 723-54.21
Telex: 842-620628

Tokyo Office
Kokusai Building
1-1 Marunouchi 3-chome
Chiyoda-ku, Tokyo 100, Japan

Telephone: (03) 214-5001
Telex: 781-26838