CER Discussion Paper
No. 12.
July, 1965

"The Use of Sample Household Surveys in Economic Planning with Some Empirical Results for East Africa"

I. INTRODUCTION

The basic point of this paper is that sample surveys, particularly household expenditure surveys, provide an efficient, low-cost method of predicting aggregate final consumption demands during periods of rapid structural change in a developing economy. In fact, it is likely that the analysis of sample household surveys is the only basis on which reasonably accurate predictions of final consumer demand can be made. At present in East Africa, no use is being made of this approach, and the few surveys that are being carried out are intended primarily to provide the weights for price indices or to give indications of the general level of living of various wage-earning groups. Thus they are not appropriately designed to provide all the information desirable for predicting changes in consumer demand as incomes rise.

The expenditure side of the national accounts is very incomplete in East Africa. Collecting additional detail will be expensive, coverage will not be uniform over different sections of the country and different groups, and, even if such a program is started soon, it will be several years before reliable aggregate time series on expenditures are available. There are more serious objections to using aggregate time series data as the basis for demand predictions which will be discussed below.

What is needed in East Africa is a well planned coordinated series of household surveys covering the major consumer groups (racial, tribal, geographical) which, on obvious a priori grounds, can be expected to exhibit different patterns of expenditure relative to income. These surveys could be quickly and relatively inexpensively carried out because of their very specific objectives, and would provide the basis for extrapolating future aggregate demands for consumer goods by methods which take fully into account changes in income distribution among households and major sub-groups of the population. The methods of analysis which might be used are described in the second section of this paper.

It might be hoped that the results of demand studies carried out in other underdeveloped areas could be applied to East Africa, or carried over from one East African country to another. The available evidence indicates that this cannot be done. Houthakker, in his very interesting article "An International Comparison of Household Expenditure Patterns....." (1), found wide ranges of income elasticities for the

.....sixty-four

sixty-four countries, regions and time periods covered: total food .344 to .731; clothing .918 to 1.784; housing (including fuel and light but not furniture) .346 to 1.114; miscellaneous 1.081 to 1.879. With reference to Kenya, Marion Forrester (2) has found strong tribal differences in expenditure patterns. Table 1 presents some comparisons derived by the present author from income elasticities of demand and percentages of income spent for Nairobi, Kenya, Kampala, Uganda, and three semi-urban towns in Uganda (combined). These again exhibit substantial differences in income elasticities. While many of these differences might be expected a priori, they serve to emphasize the dangers inherent in applying the results of one country or region to another.

II. The Use of Sample Expenditure Surveys

It is unnecessary to comment upon the general principles involved in sampling techniques. Sampling (as opposed to complete enumeration) permits cost reductions and frequently increased accuracy because of the attention which can be given to obtaining accurate answers from respondents. In studying expenditure patterns, it is generally agreed (see (3)) that the household is the relevant decision-making unit and therefore is the unit from which data should be gathered. To assure observation of a sufficient range of variation of the characteristics of the households, the sample is usually stratified by income and, preferably, by family size. On the basis of the data so gathered, it is then possible to make point or interval estimates of population parameters of interest such as the mean household income, mean levels of expenditure on different commodities, etc.

These point estimates are not, however, what is needed for planning purposes. While it is of interest to know the characteristics of the population of households today, what is needed is a system of relationships which will permit prediction of the changes in the economic characteristics of the households as economic development takes place. Of particular importance will be predictions of new levels of expenditure and saving as household incomes rise.

The ideal approach to making these predictions would be to estimate statistically from the survey data household demand functions of the general form

$$(1) \qquad q_{ij} = f\left(\frac{x_j}{p_o}, \frac{p_i}{p_o}, n_j\right)$$

.....where

TABLE I

Comparison of African Household Expenditure Patterns[1]

Expenditure Type	Percentage of Cash Income Spent			(Cash) Income Elasticity at Mean Income		
	Nairobi	Kampala	Mbale, Fort Portal Gulu	Nairobi	Kampala[2]	Mbale, Fort Portal Gulu
total expenditure	97.1	87.8	98.2	.75	1.03	.93
total food	42.3	55.4	55.3	.17	1.01	.83
basic starches	11.4	23.5	19.7	-.18	.88	1.01
sugar (confectionary)	2.7	3.4	*	.10	1.44	*
meat and fish	13.1	11.4	19.0	.33	1.84	1.16
alcoholic beverages and tobacco	7.1	9.3	12.6	1.01	1.36	*
clothing (including footwear)	8.3	7.3	12.5	1.00	1.05	1.30
transport service and equipment	5.4	.3	3.3	1.80	6.11	2.03
household goods and operation	2.6	*	6.1	1.35	*	*
school fees	2.6	1.5	6.6	1.76	3.35	.78

[1]For data sources and forms of the Engel curves from which the clasticities were computed, see the first three table of Appendix A.

[2]For Kampala, it was not possible to correct for household size.

[3]The mean income per adult equivalent (see Appendix A for scale)in the Nairobi sample was EAS 125.51 per month or EAS 370.25 per household. For Kampala, mean household cash income per month was EAS 72.87. For the Mbale-Fort Portal-Gulu, the sample mean cash income per adult equivalent was EAS 26.81 or EAS 75.07 per household.

where q_{ij} is the quantity of the ith commodity demanded by the jth household whose money income is x_j, with n_j members of the household. The variable p_0 is an index of consumer prices and p_i is the price of the particular commodity. A set of such demand functions for the important consumer commodities would permit predictions to be made of the quantities demanded by households of different characteristics under various assumptions about incomes (x_j) and prices (p_i, p_0) in accordance with changes anticipated during the development planning period. There exist procedures (to be discussed below) for then aggregating the household demands to get total quantities demanded.

This ideal approach usually has to be compromised for lack of data on quantities and prices. Future surveys can probably provide these data, but it is usually easier to keep records of expenditure (as opposed to quantities), and prices are frequently difficult to define where markets are quite imperfect - often characterized by bargaining. The compromise usually takes the form of fitting relationships between expenditures on a commodity and household income and household size. Such a relation is called an Engel Curve (or function) after the German Ernst Engel (1821 - 1896) who published (1857) what is probably the first econometric study of the relation between household income and expenditure. The general form is

$$(2) \qquad y_{ij} = g(x_j, n_j)$$

where y_{ij} is the expenditure of the jth household on the ith commodity. Since there are usually several possible ways of defining household income (before or after tax, with or without the value of certain employment perquisites, with or without the value of home grown goods concerned), total expenditure is frequently substituted for income. When incomes are low and savings nil, the differences are negligible.

If one had a set of relations like (2), how is it possible to get predictions of total (economy-wide) expenditures? This is one form of the oft-discussed aggregation problem. First let it be clear that there will be population sub-groups whose expenditure behaviour can be expected to be quite different on a priori grounds. In East Africa some of these groups would be:

.....urban

1. urban European households
2. urban Asian households
3. urban African households
4. non-urban Coastal tribes
5. non-urban Central Region tribes
6. non-urban Rift Valley tribes
7. non-urban Western and Nyanza tribes.

Other groups, e.g. Masai, North Eastern Region, could be added as they come to participate to a greater degree in the money economy. It will clearly be necessary to have a set of relations like (2) for each of these sub-groups. The procedures that follow will then apply to each sub-group.

Having our Engel curve for some commodity, it would be delightful <u>if</u> we could assume that

$$(3) \qquad \mu_y = g(\mu_x, \mu_n)$$

where μ's indicate the population (of household) means. If (3) were invariably true, a simple survey of households would provide us with estimates of μ_x and μ_n, call them $\bar{x}$ and $\bar{n}$, and we could compute

$$(4) \qquad \bar{y} = g(\bar{x}, \bar{n}).$$

Then knowing the total number of households, N, would permit us to compute an estimate of total expenditure

$$(5) \qquad \hat{Y} = N\bar{y} = N \cdot g(\bar{x}, \bar{n})$$

Unfortunately, (3) isn't true (it holds for a few special cases) so we can't avail ourselves of this straightforward procedure. To see that (3) doesn't hold in general, consider first a simple case of only two households for which expenditure depends only on household income, i.e.

$$(6) \qquad y = f(x)$$

If y represents, say, total food expenditure, (6) would have a graph as shown in figure 1.

Clearly in this case, $\bar{y} \neq f(\bar{x})$ and the true $\bar{y}$ depends upon the "spred" of x_1 and x_2 about $\bar{x}$, i.e. upon the distribution of the x's about $\bar{x}$.

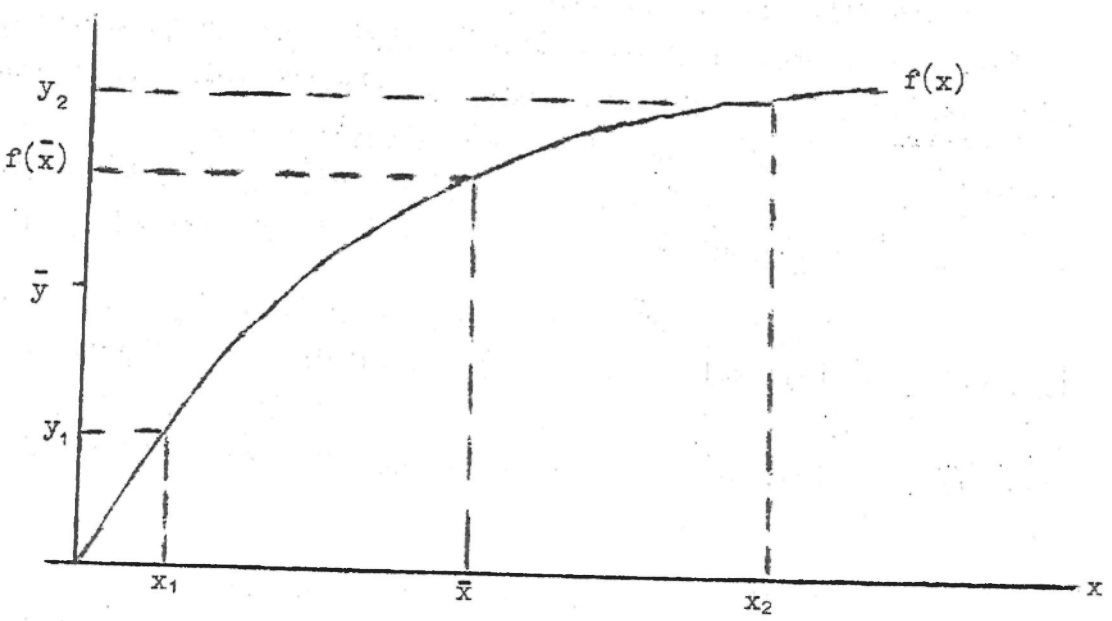

Figure 1

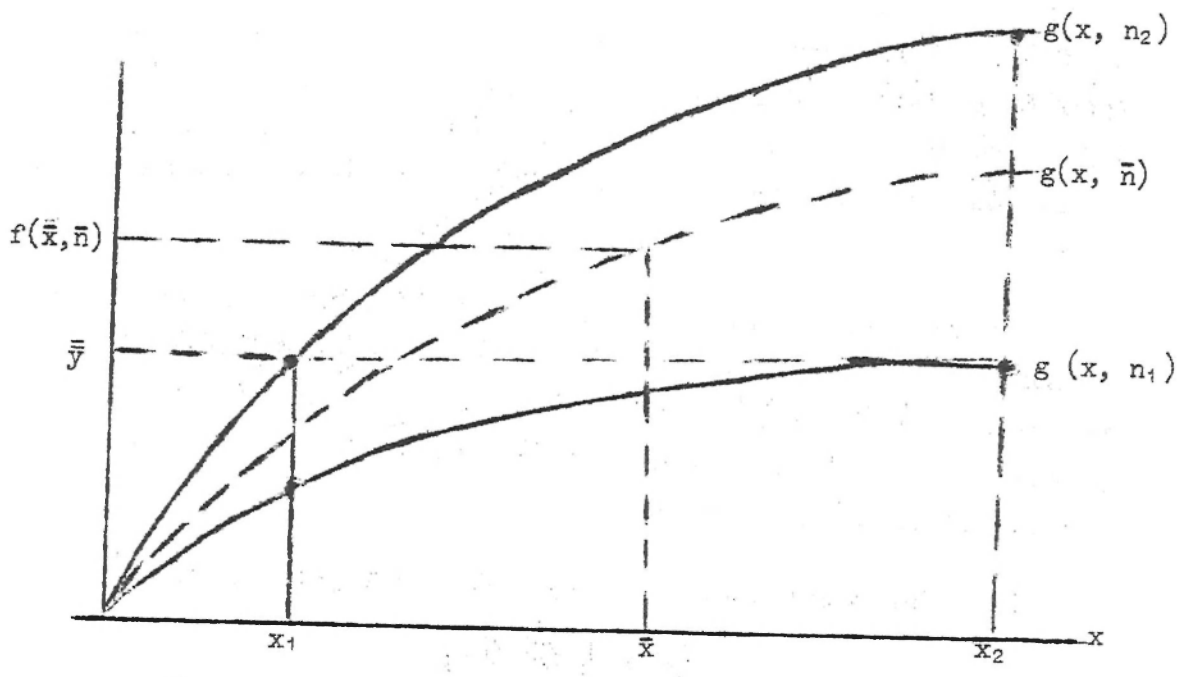

Figure 2

A slightly more complicated but more realistic example will similarly show that when expenditure depends upon both income and household size, the true mean expenditure will depend not only on how the x's are distributed about the mean income level, but also on how household incomes are "paired off" with the corresponding household sizes, i.e. on the joint distribution of household income and size. (See Fig. 2.)

If our "population" of two households had the paired attributes (x_1, n_1) and (x_2, n_2) it can be seen that the true $\bar{y}$ would be quite near $f(\bar{x}, \bar{n})$. However, if the population were characterized by (x_1, n_2) and (x_2, n_1) the true mean would be $\bar{\bar{y}}$, a value far different from $f(\bar{x}, \bar{n})$.

The point of these simple demonstrations is simply that what is needed to utilize the Engel curve approach to predicting aggregate consumer demand is a knowledge of the joint distribution of household incomes and sizes, i.e. a two-way table showing the percentages of the household population falling within the various income and household size classes. The finer the gradation of the income and size classes, the smaller will be the approximation error. For each income - household size cell, the mean expenditure can be approximated by $f(\bar{x}_{ij}, \bar{n}_{ij})$ and the cell total expenditure approximated by $N_{ij} \cdot f(\bar{x}_{ij}, \bar{n}_{ij})$. While some error still is involved within each cell, this error tends to zero as the income and size gradations become finer. Summing over the cell totals then yields the estimate of aggregate expenditure.

For certain special cases there are esthetically more pleasing and more accurate ways of proceeding from the Engel curve to predictions of aggregate demand. One of these special cases which may approximate actual conditions for some commodities quite well is that in which it is assumed that:

(1) the Engel curve in (2) has the log-log form:

$$y = \beta_0 x^{\beta_1} n^{\beta_2};$$

(2) household income and household size have a joint log-normal frequency distribution with parameters

$$\sigma_{\log x}, \quad \sigma_{\log n}, \quad \text{and} \quad \rho_{\log x,\ \log n}.$$

Once the Engel curve has been estimated from survey data (say by weighted least-squares techniques) so that estimates of β_0, β_1, and β_2

.....are available

are available, the following relationship permits one to estimate the population mean (household) expenditure from estimates of mean household income and size (see Appendix B for the derivation):

$$(7) \quad \log \bar{y} = \log \beta_0 + \beta_1 \log \bar{x} + \beta_2 \log \bar{n} + \tfrac{1}{2}(\beta_1^2 - \beta_1)\sigma^2_{\log x} + \tfrac{1}{2}(\beta_2^2 - \beta_2)\sigma^2_{\log n} + \beta_1\beta_2\rho\sigma_{\log x}\sigma_{\log n}.$$

Knowing $\bar{y}$ and the number of households in the population then permits computation of the aggregate expenditure, $Y = N\bar{y}$.

III. Shortcomings of the Engel Curve Approach

The primary weakness of this approach is the omission of relative prices from the analysis. While it may be difficult to observe price differences in a cross-section survey, more attention should be paid to obtaining price data in future surveys.

It may well be that longer-term income elasticities differ from those indicated by cross-section surveys. Duesenberry's relative income hypothesis indicates that rising general levels of income and expenditure will induce higher levels of household expenditure than would occur if only one household in an unchanged environment experienced higher income. Tastes may also change in the longer term.

Some empirical evidence for East Africa is available on this point. Engel curves were fitted to African household survey data for Nairobi and to the combined observations from surveys in Mbale, Fort Portal, and Gulu, Uganda. The latter areas are in many ways similar and all have much lower household incomes than Nairobi. The expenditure elasticities derived from Engel curves fitted to the Uganda data were then evaluated at the mean Nairobi income level. The results are presented in Table 2.

It seems that there are definite biases involved in trying to extrapolate over incomes far in excess of those used in the fitting. Table 2 indicates a definite overstatement of elasticities for those commodities which might be considered necessities (food, starches, and possibly meat and fish) and understatement of elasticities for "luxuries" (clothing, transport, school fees). The practical meaning of this is that a given fitted function well approximates the actual expenditure - income relationship only over a limited income range. New surveys

.....will have

TABLE 2

A Comparison of Income Elasticities of Demand Evaluated at Nairobi Income Levels

Expenditure Category	Nairobi Elasticity	Mbale-Fort-Portal-Gulu Elasticity
total expenditure	.750	.966
total food	.175	.370
basic starches	-.177	.402
meat and fish	.329	.426
clothing	1.000	.457
transport	1.802	.531
school fees	1.756	.402

Note: Average household income in Nairobi survey was EAS 370.25 per month. The average for the Mbale-FortPortal-Gulu surveys was EAS 75/07 per month cash or EAS 109/73 per month including value of home produce consumed.

will have to be taken more frequently in areas of rapid income change (say, every four to five years if household incomes grow at six per cent) than in slower growing areas (about every eight years at a three per cent rate).

A shortcoming shared by all approaches relying on expenditure data is that expenditure is the product of quantity and price. If markets are competitive, price can be taken as a measure quality. Thus as we observe increasing expenditures with rising incomes, the increase is compounded of an increased demand for quantity and a demand for improved quality. The latter effect can be substantial (3, Chapter 8), (5), (6), especially for meat and fish and other commodities which are typically available in the market in a great variety of qualities. Separate studies of quality elasticities using methods such as those referred to in the above references can provide factors of correction to convert the Engel curves to quantity-valued functions. Such studies may not prove feasible, however, because of the increased demands they place on an accurate data collection. The lack of proportionality between quantity and expenditure can be reduced by using more narrowly defined commodity categories.

IV. Empirical Results for East Africa

The results presented here refer to African urban households in Nairobi and Kampala and to "semi-urban" households in Mbale, Fort Portal and Gulu. By "semi-urban" is meant a setting where heads of household work for wages in a town but live on a shamba which provides a significant proportion of the family's food needs. The present results are primarily of interest for comparing behaviour in different parts of East Africa and for testing certain hypotheses regarding African consumer bahaviour. Planning purposes will require the analysis of a large number of new surveys as explained in Part II of this paper.

The basic date used were taken from published tabulations of survey results (7) - (13). Households generally had been classified according to household income level, with average household size being indicated for each income group. Thus the basic observations were group averages. The family size heterogeneity within income groups may have been quite large, although the use of the indicated average family size for each income class significantly improved the statistical explanation. The main problem is using tabulated data of this sort was the small number of observations provided. (The underlying household data are presumably available on worksheets at the relevant statistical offices.) The statistical technique used was weighted least-squares regression, the weights being the numbers of households in each income class[1].

Ideally, the functional form for the Engel curve should be like (2). Because of the summary form in which the data were published and the resultant small number of degrees of freedom available, it was decided to incorporate the household size variable by deflating both income and expenditure variables by the average number of adult equivalent persons in each income group. That is, instead of (2), we have actually used

$$(8). \qquad \left(\frac{y_{ij}}{a_j}\right) = f\left(\frac{x_j}{a_j}\right)$$

where a_j is the average number of adult equivalents per household in jth income group. Plotting the original data and the data deflated in this way indicated that deflation resulted in a much more regular pattern of expenditure versus income. The adult equivalent scales used are indicated in the footnotes of Tables A1 and A2.

---

[1] This procedure makes the weights of each observation (group mean) inversely proportional to the variance of the observation.

This compromise procedure is objecticnable on two grounds: (1) any adult equivalent scale is arbitrary and there should be different scales for different commodities (e.g. milk and whisky); (2) deflation, and even by a carefully designed measure of household size, represents a very special case among the ways in which household size may affect expenditures; i.e. it represents an a priori restriction on the relation which is not called for by consumer theory.

With regard to the first objection, if data are sufficient, there are statistical techniques for estimating scales of adult equivalents or unit consumer scales from the basic expenditure - income - household composition data. Existing techniques are quite complex (see (3), Chapter 9) and the data were insufficient to support such an analysis. Since deflation had been decided upon to conserve degrees of freedom, it still seemed more appropriate to use even an imperfect measure of adult equivalents than unadjusted family size.

That deflation of the data represents a special case of (2) can easily be seen. For functions like (2), it is true that

$$(9) \qquad \frac{y_{ij}}{n_j} = f\left(\frac{x_j}{n_j}, \; 1\right)$$

if and only if (i.e. by definition) the function f is homogeneous of degre one. Clearly, not all Engel functions meet this condition. If we take the log-log form of (2) which is frequently found applicable to expenditure on durable goods,

$$(10) \qquad y_{ij} = \beta_0 \, x_j^{\beta_{i1}} n_j^{\beta_{i2}} \; ,$$

empirical results (see (1), pp. 541-542) clearly show that $(\beta_1 + \beta_2) = 1$ is not the typical case,[1] i.e. the functions are not usually homogenous of degree one. If we assume that deflation is legitimate for the log-log function, we are assuming that (10) has the special form

$$(11) \; y_{ij} = \beta_0 \, x_j^{\beta_1} n_j^{(1-\beta_1)} \; .$$

Insofar as this assumption does not agree with the facts, biases will be introduced into the estimation of the Engel ourve parameters.

The income concept to be used in studies of the sort is not at all a settled matter. For urban households in Nairobi and Kampala where shamba produce for home consumption is negligible, the definition used in my analyses was:

> "cash income less taxes paid, plus value of housing provided or cash housing allowance, less gifts and remittances received."

[1] There are a priori (theoretical) constraints on the β_{ij} implied by the household budget constraint, but none which concern the homogeneity of individual functions.

The omission of gifts and remittances received was prompted by the absence of any regular pattern relating to income levels. These items or receipt seemed to be irregular and probably would not, therefore, enter into the making of ordinary expenditure decisions.

The "semi-urban" areas by definition display two distinct components of real income; cash income from employment and sales (beer, produce, cotton, etc.) and consumption of home produced food. In the surveys, local market prices were used to place a value figure on the latter. Various hypotheses can be formulated about the way in which the value of shamba produce consumed (VSPC) affects expenditure on market goods. The two hypotheses which were tested with the Mbale-Fort Portal-Gulu data were:

(1) that VSPC is indistinguisible from cash income in its impact on expenditures;
(2) that VSPC must be treated as a distinct type of income in explaining expenditures.

The first hypothesis is essentially a naive alternative to the second. There are many a priori reasons why hypothesis two should be true. Since consumption of shamba produce takes the form of food and related items, there is the specific displacement effect against market f. At the same time, there is a real income effect which should be especially strong in favor of non-food commodities.

It was found that the use of the variable (cash income + VSPC) resulted in large parameter changes when different subsets of the data were used. This was not the case when VSPC was used as a separate variable. Multiple correlations were significantly higher when VSPC was used separately. For example, total cash expenditure versus combined income yielded an $R^2 = .54$ while $R^2 = .97$ when cash income (less taxes and gifts and remittances received) and VSPC were introduced separately.

The results in Table A2 indicate fairly clearly that VSPC should be treated as a separate variable in the analysis of expenditures. Its impact on cash expenditure is significant in several cases (in spite of the very few degrees of freedom) and is different in magnitude and sometimes direction in its impact. The four cases in which the coefficient of VSPC is negative must be interpreted with care (alcoholic beverages and tobacco, clothing, household goods and operation, and school fees). Home produced goods could be a substitute for purchased alcohol and tobacco so that the displacement effect would dominate the income effect. This is clearly not the case for the other three

commodity groups. The income effect should predominate and the VSPC coefficients should be positive. It is likely that VSPC is positively correlated with distance from the market center, a factor making more difficult the purchase of beverages and clothing and the sending of children to school. In the case of school fees it is also likely that the larger shambas make greater use of child labor. Thus, the VSPC variable is acting as a surrogate for these factors which dominate the positive income effect.

The results that were presented earlier in Tables 1 and 2 indicate the dangers inherent in trying to extrapolate consumption behaviour over regions and wide ranges of income. Those commodities which would generally be considered luxuries (alcoholic beverages and tobacco, clothing, transport, household goods and school fees) exhibit elasticities in excess of 1 but still cover wide ranges. The elasticities for commodities commonly called "necessities", such as total food category and basic starches, seem to depend upon income levels. For starches, the elasticity ranges from -.18 in Nairobi to 1.01 in Mbale-Fort Portal-Gulu. Much more surprising is the wide range of elasticities for the category meat and fish. These differences in elasticities for high cost proteins must be related not only to income levels but to the protein content of the basic diet in each region. Nairobi with its basic diet of posho supplemented by beans has a much richer protein base than do the matoke based communities in Uganda.

A bit can be learned about the financial behavior of households from the survey data, and again the finding is one of diversity. With regard to saving, positive saving appears to start at a household cash income of EAS 330.00 per month in Nairobi, EAS 56.00 in Mbale-Fort Portal-Gulu, and at very low incomes in Kampala. Gifts and remittances given were not included in the total expenditure, so the latter was probably somewhat understated and saving overstated. However it is not possible to tell what these remittances are used for by the recipients. A part is probably used for consumption purposes but some is intended for investment in improvements in the homestead or shamba. Net gifts and remittances given (those given less those received) show a regular pattern versus income. For instance, at low household income levels there is a net receipt of gifts and remittances. At a household income of about EAS 300/month, net gifts and remittances are zero and they rise asymtotically to about EAS 70/month at household incomes of about EAS 700/month.

Some information on debt and debt repayment is available in the Nairobi surveys. By debt (or debt repayment) we mean to include loans, hire-purchase credit, and shop credit. The data on gross increases in debt and gross decreases in debt, deflated by the number of adult equivalents,

showed very regular patterns versus income per adult equivalent. The gross debt increase was well represented by

$$(12) \quad GDI = \underset{}{-1.156} + \underset{(5.990)}{.390x} - \underset{(.00032)}{.00102 x^2}, \quad R^2 = .807$$

and the gross debt repayment was well represented by

$$(13) \quad GDR = \underset{(1.790)}{1.970} + \underset{(.013)}{.152 x}, \quad R^2 = .97$$

These results imply that at income levels below EAS 220/month per adult equivalent, households are typically going further into debt, while above that amount they are net repayers of debt.

V. Conclusions

From the sketchy empirical evidence presented here, there appears to be great diversity in consumption behaviour among different groups of East African consumers. This precludes the possibility of applying the results of analysis of one area to the prediction of consumption demands in another. Nonetheless household surveys and related aggregation techniques can be useful and relatively inexpensive tools for the prediction of aggregate consumer demands in rapidly changing economies.

Charles W. Howe
University College, Nairobi
JULY, 1965.

N a i r o b i

(Data from 1957/58 and 1963 Surveys)

Functional Forms for Engel Curves (logs to the base 10 used in Computation. e indicates elasticity)

a. linear: $y = a + bx$, $e = b\frac{x}{y}$

b. inverse: $y = a + b(\frac{1}{x})$, $e = -\frac{b}{xy}$

c. semi-log: $y = a + b \log x$, $e = .4343\ \frac{b}{y}$

d. log-log: $\log y = a + b \log x$, $e = b$

e. log inverse: $\log y = a + b(\frac{1}{x})$, $e = -2.3026\ \frac{b}{x}$

f. log-log inverse: $\log y = a + b(\frac{1}{x}) + c \log x$, $e = c - 2.3026\ \frac{b}{x}$

	Expenditure Category	Functional Form	Parameter Values and Standard Errors a	b	c	R^{2*}	Weighted** Mean Exp. p.a.e.	Thresh-hold Income	Elasticity at Means	% Income Spent at Mean
1.	Total Expenditure	linear	30.44 (5.613)	.728 (.0404)	-	.973	121.84	---	.7501	97.1
2.	Total Food	log - inverse	1.818 (.0184)	-9.51 (1.619)	-	.882	53.04	---	.1745	42.3
3.	Cereals & breads	log-log-inverse	2.309 (.7668)	-16.60 (13.23)	-.4813 (.3103)	.429	14.36	---	- .1768	11.4
4.	Milk & eggs	semi-log	-7.644 (.9563)	7.029 (.4646)	-	.962	6.75	12.23	.4522	5.4
5.	Confectionary	log-log-inverse	1.993 (0.400)	-26.44 (6.911)	-.5808 (.1622)	.386	3.45	---	.0960	2.7
6.	Meat, fish & Meals out	semi-log	-9.061 (5.676)	12.478 (2.758)	-	.695	16.48	5.32	.3289	13.1
7.	Vegetables & fruits	semi-log	-4.973 (1.669)	5.133 (.8110)	-	.817	5.53	9.31	.4030	4.4
8.	Oils & fats	semi-log	-4.459 (1.511)	3.598 (.7341)	-	.728	2.91	17.35	.5370	2.3

...../2

	Expenditure Category	Functional Form	Parameter Values and Standard Errors a	b	c
9.	non-alcoholic beverages	semi-log	-1.163 (.8331)	1.549 (.4047)	-
10.	Alconolic beverages & tobacco	linear	-0.1265 (.7493)	..0718 (.0054)	- -
11.	Clothing, including footwear	semi-log	-38.59 (4.876)	23.928 (2.369)	- -
12.	rent, rates, and water	semi-log	-44.522 (9.727)	30.226 (4.725)	- -
13.	furniture & furnishings	log-log	-2.207 (.3196)	1.385 (.1553)	- -
14.	transport and transport equipment	log-log	-3.040 (.4699)	1.802 (.2283)	- -
15.	Fuel and light	inverse	5.309 (.3195)	-110.50 (28.15)	- -
16.	Recreation and entertainment	semi-log	-10.176 (1.124)	5.923 (.5459)	- -
17.	Medical and personal health	log-log	-2.846 (.2884)	1.561 (.1401)	- -
18.	Household operation	log-log	-2.366 (.4001)	1.352 (.1944)	-
19.	Misc services including education	log-log	-3.317 (.8449)	1.756 (.4104)	-

Footnotes: *R^2 corrected for degree of freedom

**Weighted mean income is sh. 125/51 per adult equivalent (p.a.e.). years = 0.6, females 16 and over = 0.8, and males over 16 = 1.0.

***Positive saving appears to start at an income p.a.e. of sh. 112/- adult equivalents, this means a household income of sh. 330/- per

R^{2*}	Weighted** Mean Exp. p.a.e.	Thresh-hold Income	Elasticity at Mean	% Income Spent at Mean
.618	2.01	5.63	.3350	1.6
.952	8.89	1.76	1.014	7.1
.919	10.39	41.00	1.000	8.3
.820	17.35	29.71	.7570	13.8
.898	5,59	--	1.3850	4.5
.874	6.77	--	1.8020	5.4
.608	4.19	20.81	.2100	3.3
.929	1.95	52.25	1.3190	1.6
.932	2.96	--	1.561	2.4
.843	3.26	--	1.3520	2.6
.670	3.21	--	1.7560	2.6

The adult equivalent scale used is: children less than 16

per month. At the weighted mean household size of 2.95

month.

TABLE A2

Mbale - Fort-Portal - Gulu
(Data from Mbale Survey 1958, Fort Portal 1960, Gulu 1959)

Functional Form for Engel Curves: $y = a + b \log$ (cash income p.a.e.) $+ c \log$ (VSPC p.a.e.).

	Expenditure Category	Parameter Values and Standard Errors a	b	c	R^2(1)	Weighted[4] Mean Exp. p.a.e.[5]	Thresh-hold Income	Cash Elasticity[(2)] and VSPC Elasticity	% Cash Income Spent at Mean
1.	total cash expenditure	-56.704 (6.478)	56.043 (3.650)	3.844 (2.114)	.971	26.32	(3)	.9248 .6342	98.2
2.	total cash food	-38.301 (12.140)	28.343 (6.841)	12.636 (3.962)	.682	14.82	(3)	.8306 .3703	55.3
3.	basic starches	-17.329 (10.510)	12.332 (5.930)	5.031 (3.430)	.248	5.29	(3)	1.0120 .4130	19.7
4.	meat and fish	-18.244 (3.013)	13.612 (1.700)	4.000 (.985)	.899	5.09	10.45	1.1614 .3413	19.0
5.	Alcoholic beverages and tobacco	1.190 (6.482)	2.203 (3.651)	-.867 (2.115)	0	3.39	-	-	12.6
6.	clothing	- 7.832 (3.020)	9.982 (1.701)	-2.786 (.985)	.665	3.34	1.09	1.2980 -.3622	12.5
7.	household goods and operation	.142 (1.801)	1.188 (1.015)	-.181 (.588)	.028	1.64	-	-	6.1
8.	transport service and equipment	- 5.841 (1.692)	4.167 (.953)	.687 (.552)	.691	.89	16.64	2.0336 .3353	3.3
9.	school fees	4.204 (7.355)	3.150 (4.142)	-6.611 (2.399)	.577	5.34	9.09	.7774 -1.6314	6.6

(1) R^2 corrected for degrees of freedom

(2) Elasticities with respect to cash income p.a.e. and value of shamba produce consumed (VSPC) p.a.e.

(3) Threshold incomes less than Sh. 1.00

(4) Weighted mean cash income p.a.e. is sh.26.81 and VSPC p.a.e. is sh.12.38. Saving appears to start at a cash income of about sh. 20.00 p.a.e. when VSPC is 12.38. At the average household size of 2.8 adult equivalents, this would mean a household income of sh. 56.00 per month.

(5) The Oxford Scale of adult equivalents has been used in these studies. That is: head of household = 1.0, other adults and children 14 or over = 0.7, and children less than 14 = 0.5.

K A M P A L A

(Data from 1953 and 1957 Surveys)

Functional Forms for Engel Curves: with the exception of the total expenditure category, to which a linear function was fitted, all the following functions are semi-log. The data are on a household basis and have not been deflated for household size. No information was given in the surveys on household size or on the value of any shamba produce consumed, if any.

	Expenditure Category	Parameter Values and Standard Errors a	b	R^2 (1)	Weighted Mean Expenditure (2)	Threshold Income	Elasticity at Mean	% Cash Income Spent at M e a n
1.	total expenditure	- 1.980 (12.29)	.905 (.163)	.856	63.96	---	1.0300	87.8
2.	total food	-133.127 (45.24)	93.827 (24.43)	.733	40.40	26.23	1.0086	55.4
3.	basic starches	- 47.426 (23.55)	34.919 (12.71)	.566	17.15	22.81	.8843	23.5
4.	sugar	- 12.795 (5.575)	8.266 (3.010)	.567	2.49	35.32	1.4416	3.4
5.	meat and fish	- 56.823 (17.90)	35.213 (9.663)	.711	8.30	41.08	1.8420	11.4
6.	ground nuts	- 7.019 (4.867)	5.719 (2.628)	.427	3.55	16.88	.7000	4.9
7.	vegetables and fruits	2.175 (5.796)	1.863 (3.129)	0	5.62	---	---	7.7
8.	alcoholic beverages and tobacco	- 32.222 (18.25)	21.077 (9.853)	.417	6.75	33.79	1.3561	9.3
9.	clothing	- 18.454 (8.970)	12.857 (4.843)	.547	5.32	27.25	1.0496	7.3
10.	transport	- 6.002 (1.750)	3.374 (.946)	.701	.24	60.10	6.1052	0.3
11.	school fees	14.580 (3.739)	8.484 (2.018)	.769	1.10	52.31	3.3496	1.5

(1) R^2 is corrected for degrees of freedom

(2) the weighted mean cash income is sh. 72.87.

Appendix B

The Aggregation of Household Expenditure: A Special Case.

Assume that the Engel curve has the form

$$(1) \quad y = \beta_0 x^{\beta_1} n^{\beta_2}$$

Taking logs and summing over all households in the population

$$(2) \quad \Sigma \log y_i = N \log \beta_0 + \beta_1 \Sigma \log x_i + \beta_2 \Sigma \log n_i$$

Taking anti-logs and Nth roots:

$$(3) \quad (\Pi y_i)^{1/N} = \beta_0 \left[(\Pi x_i)^{1/N} \right]^{\beta_1} \left[(\Pi n_i)^{1/N} \right]^{\beta_2}$$

This is a relationship between the geometric means of expenditure, income, and household size. If we let (*) variables represent geometric means, taking logs of (3) yields

$$(4) \quad \log y^* = \log \beta_0 + \beta_1 \log x^* + \beta_2 \log n^*.$$

For any log-normal variate (see [4]) like x, it is the case that

$$(5) \quad \log x^* = \log \bar{x} - \tfrac{1}{2} \sigma^2_{\log x}.$$

The simple rules of distribution theory also permit us to derive from the log form of (1) that

$$(6) \quad \sigma^2_{\log y} = \beta^2 \sigma^2_{\log x} + \beta_2^2 \sigma^2_{\log n} + 2\beta_1\beta_2 \rho \sigma_{\log x} \sigma_{\log n}.$$

Making the type of substitution (5) for $\log y^*$, $\log x^*$, and $\log n^*$ in (4) and then applying (6) yields the final result

$$(7) \quad \log \bar{y} = \log \beta_0 + \beta_1 \log \bar{x} + \beta_2 \log \bar{n} + \tfrac{1}{2}(\beta_1^2 - \beta_1) \sigma^2_{\log x} + \tfrac{1}{2}(\beta_2^2 - \beta_2) \sigma^2_{\log n} + \beta_1\beta_2 \rho \sigma_{\log x} \sigma_{\log n}.$$

References

(1) Houthakker, H. S.: "An International Comparison of Household Expenditure Patterns, Commemorating the Centenary of Engel's Law", Econometrica (1957), pp. 532 - 551.

(2) Forrester, Marion: Kenya Today: Social Prerequisites for Economic Development. International Institute for Social Studies: The Hague, 1962, p. 123.

(3) Houthakker, H.S. and S.J. Prais: The Analysis of Family Budgets. Cambridge (England): 1955. Chapters 2 and 3.

(4) Aitcheson and Brown, The Log-Normal Distribution. North Holland Press

(5) Bristol, R.: Unpublished Ph.D Thesis, Yale, 1955.

(6) Black, G.: "Variations in Prices Paid for Food by Income Level," Journal of Farm Economics, XXXIV (1952), p. 52.

(7) East African Statistics Department (Kenya Unit): "Patterns of Income, Expenditure, and Consumption of Africans in Nairobi, 1957/58 (1959)

(8) Kenya Government, Ministry of Finance and Economic Planning: "The Patterns of Income, Expenditure, and Consumption of African Middle Income Workers in Nairobi, July, 1963." (1964).

(9) East African Statistical Department: "The Patterns of Income, Expenditure, and Consumption of African Unskilled Labourers in Kampala, September, 1953". (1954).

(10) East African Statistical Department (Uganda Unit): "Patterns of Income, Expenditure, and Consumption of African Unskilled Workers in Kampala, February, 1957". (1957)

(11) East African Statistical Department (Uganda Unit): "The Patterns of Income, Expenditure, and Consumption of African Unskilled Workers in Mbale, February 1958". (1958)

(12) East African Statistical Department (Uganda Unit): "The Patterns of Income, Expenditure, and Consumption of African Unskilled Workers in Fort Portal, February, 1960". (1960)

(13) East African Statistical Department (Uganda Unit): "The Patterns of Income, Expenditure, and Consumption of African Unskilled Workers in Gulu, February, 1959". (1959).