(831)

R.D.R. 96
J. J. Oloya
10th February,
1970.

SOME SPECIFICATION PROBLEMS IN DEMAND ANALYSIS: AN EMPIRICAL STUDY OF TWO MARKETS IN KAMPALA 1967/1968.

Note: Rural Development Research Papers are written as a basis for discussion in the Makerere Rural Development Research Seminars. They are not publications and are subject to revision.

SOME SPECIFICATION PROBLEMS IN DEMAND ANALYSIS: AN EMPIRICAL STUDY OF TWO MARKETS IN KAMPALA 1967/1968*

by

J. J. Oloya

Introduction:

The measurement of demand is a complicated subject. It is often necessary for example to have a thorough knowledge of the economic factors that affect the commodity and to obtain adequate data on which to base the analysis. It is also necessary to reinforce this with an adequate comprehension of economic theory and to employ the appropriate analytical tool for the derivation of demand functions.

In this paper some attempt will be made to discuss some of the specification problems frequently encountered in demand analysis in urban areas. The discussion will be confined to demand analysis of consumption expenditures and not time series although the data is based on statistics collected over a period of one year in Kampala markets during 1967/68 and therefore could lend itself readily to time series analysis.

Identification:

The first specification problem is that of identification. This was first established by Working[1] in his paper "What do Statistical 'Demand Curves' show"? In this paper Working pointed out that when a research worker begins a demand study, he is confronted with a set of dots. For instance Chart I gives a scatter diagram of the quantity/price relationship for Sweet Potatoes in Mulago market, during the period of our study.

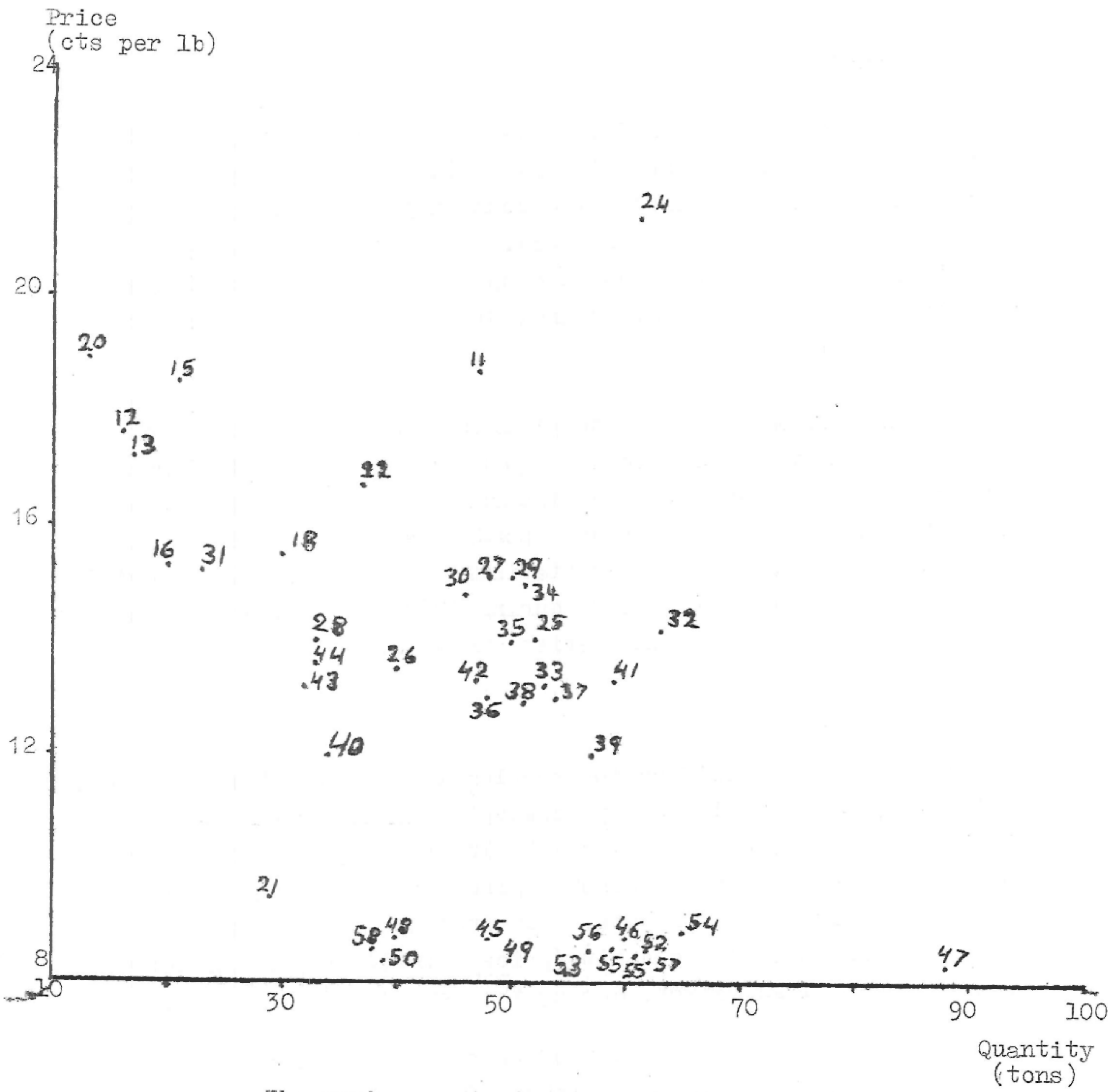
Each dot can be thought of as the intersection of a demand and supply curve, but without further information neither curve can be determined from the data.

---

*The raw data used in this analysis was collected in the course of a research project conducted by T. T. Poleman and the author with financial assistance provided by the Rockefeller Foundation. Grateful acknowledgement is due Dr. S. Rao of the Institute of Statistics and Applied Economics, Makerere University College.

A Scatter diagram of Quantity - Price Relationship of Sweet Potatoes in Mulago Market: 1967/1968

Chart I:



The number against dots refer to weeks

However, if the demand curve has shifted over time but the supply curve has remained relatively stable the dots trace out a supply curve. Conversely if the supply curve has shifted, but the demand curve has remained stable the dots trace out a demand curve (see Chart II).

In many analyses of the demand for agricultural products, factors that cause the demand to shift over time - income, price of other commodities and so on - are included as separate variables in a multiple regression equation. In effect we are then able to derive from our estimating equation an average demand curve.

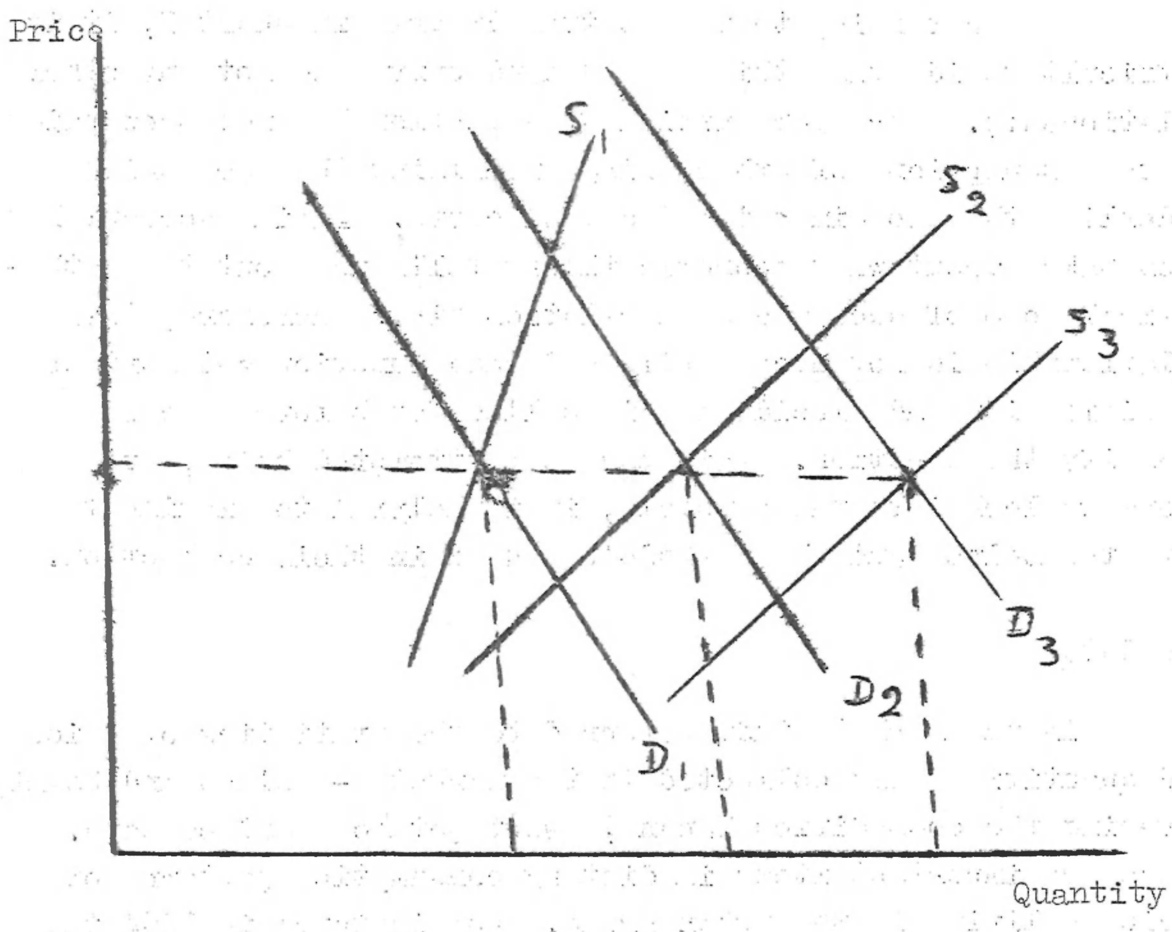
Thus returning to the scatter diagram in Chart I, it is difficult to identify this as a demand curve and not any other relationship. Moreover sufficient attention has not been paid to the assumptions of other things remaining the same which underlies the two dimensional demand curve. Again, account has not been taken about what economic theory tells us about the mathematical form of such a demand relationship. Therefore, this relationship is not necessarily a demand function and further qualification and elucidation is called for before we can identify the function. This has been attempted below, but before we look into these models, it is helpful to examine the data themselves and the methodology used in their collection.

<u>The data</u>:

In the market study referred to the statistics of prices and quantities were collected in two markets - Mulago and Nsambya- covering the commodities banana, sweet potatoes and cassava. These are important items in dietary consumption patterns of Africans living in the urban areas. For instance in 1964 the total expenditure on these items comprised nearly 50% of the total household expenditure on food[(2)]. The primary data were collected on a weekly basis, over a period of one year thus providing a simple timeseries data. The explanatory endogenous variables are therefore quantity and price.

The exogenous variables were assumed away since seasonal fluctuations due to weather and other climatic factors were taken into account by extending the survey period over 12 months

Chart II: Identification Problem: Hypothetical Demand(D) and Supply(S) Curves



The model:

On the basis of this simple model, the demand function was therefore attempted namely,

$$q_t^d = \alpha + \text{Bpt} + \text{ut}$$

q_t^d = quantity demanded at time t

α = constant

pt = price at time t

ut = error at time t (the disturbance or stochastic element)

It is to be noted that in such a simple model no attempt has been made at specification of the stochastic term. Instead it is assumed that the demand for primary agricultural products such as banana or sweet potatoes is stable over time. By contrast, supply being dependent largely on the vagaries of weather, is subject to a great deal of fluctuations. Supply as a function of price, or even other conventional economic variables, is therefore a highly variable function from season to season depending on complex exogenous variables. It follows that the disturbance variance in demand equation will be small relative to that of supply. It should be noted, however, that although this assumption has been made for simplicity, specification of the nature of the random disturbance is often nevertheless important in demand analyses since it may itself be a method of achieving identification.

Our model can, therefore, be simply written as:

$$q_t^d = \alpha + \text{Bpt}$$

This is the model attempted at first, although, as will be seen later, an alternative specification was also tried.

The data was processed using T.S.P. computer program first developed at Berkeley University for use in time series analysis and now being used also at Stanford University.

RESULTS:

Table I gives the results of the computation using ordinary least square regression.

[illegible] points emerge from an examination of the tabl

First, the small magnitude of correlation coefficient (coefficient of determination R^2) in table Ia - varying from zero to 0.144 would suggest that it is not so important to correlate prices in one market against those in the other. This is also evident from Chart III: if the corresponding prices were exactly the same, a straight line could be fitted which would pass through the origin (zero intercept). This general low correlation of prices must be accepted as evidence that allocation of commodities among the two markets is less than optinum even when allowances are made for error in the price observation as Jones and others found in Nigeria[10].

Secondly, in table Ib cassava and sweet potatoes quantities show high correlation coefficients with banana prices - nearly 0.8 and 0.9 respectively. This is partly explained by the possibility of substitution. Sweet potatoes and plantain, in particular, are close substitutes in the dietary habits of the Baganda who were the main tribal grouping in the area under review. It is probably also related to the possibility of cassava being a relatively inferior product and, hence, the high elasticity coefficient.

Alternate Specification of the Demand Function:

In the relationships we have discussed so far we have assumed more or less implicity that the points lie exactly on the demand or supply curve. In actual statistical analysis this is never true, since some variables that cause the curves to shift always are omitted and the precise shape of the curves to be fitted are not known. Thus we normally deal with stochastic rather than functional relations. A stochastic relation is one that includes a set of unexplained residuals or error terms whose direction and magnitude are usually not known exactly for any particular set of calculations, but whose behaviour on the average over repeated samples can be described or assumed, that is, they display normal distribution around a zero mean, with constant variance and the values of u are drawn independently of one another.

Secondly, the relationship in Chart I may not necessarily be linear.

It is thus necessary to explore alternative specifications of the demand functions.

/5

Table I: Estimated Regression Coefficients (n = 50)

I (a) Price:

Price (Mulago)	Price (Nsambya)	C (intercept)	R^2
Banana	.288 (.101)	7.829 (1.234)	0.144
Cassava	.286 (.113)	4.600 (7.440)	0.117
Sweet Potatoes	.002 (.133)	11.582 (1.716)	0.000

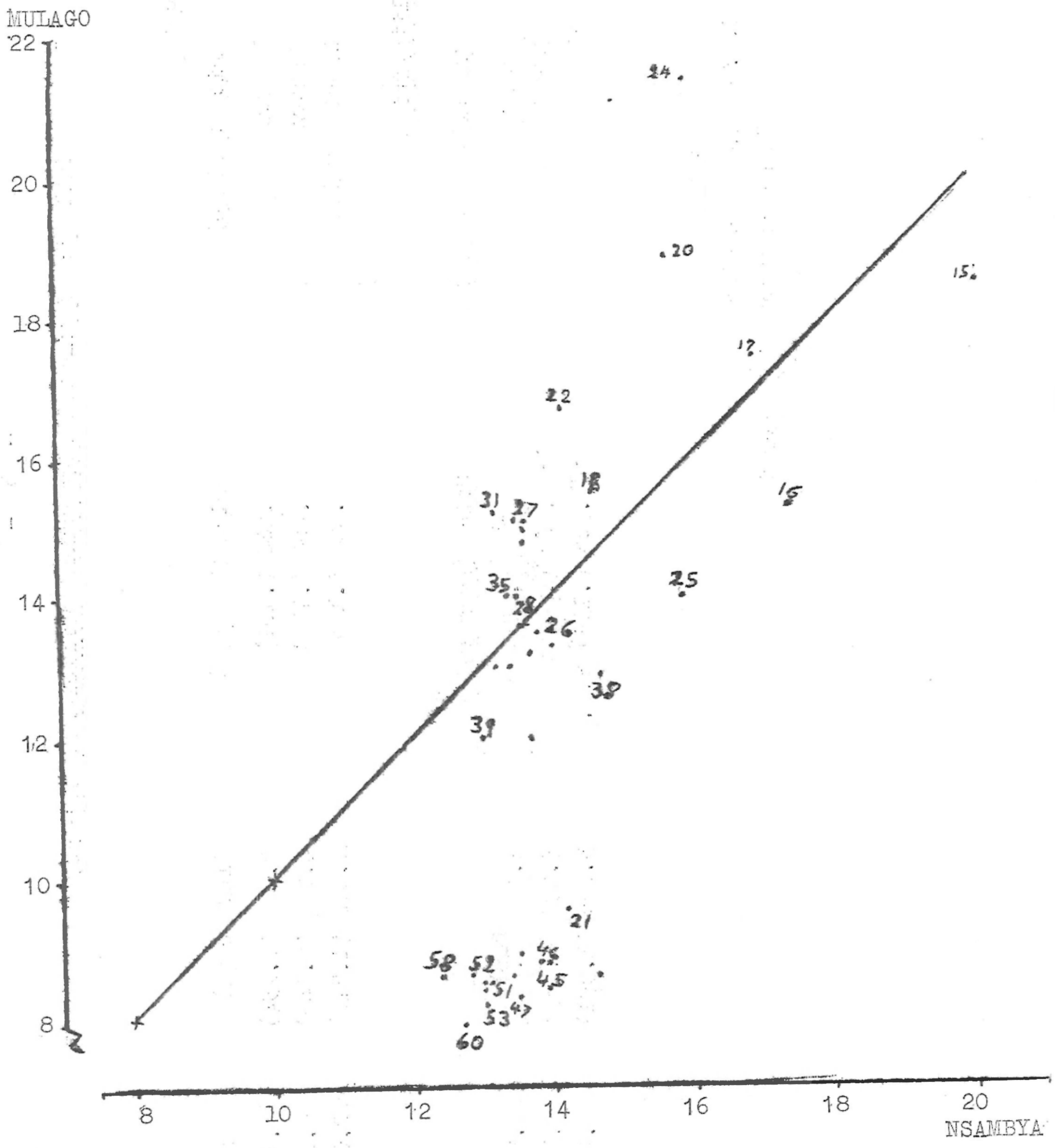
II (b) Quality/Price:

Quantity (Nsambya)	Price (Nsambya) B		
Cassava	.770 (.138)	1.419 (1.688)	.393
Sweet Potatoes	.867 (.265)	4.599 (3.233)	.183
Cassava	.586 (.154) S.P.*	3.025 (1.991)	.231

B = Banana
S.P* = Sweet Potatoes
Standard errors in bracket.

Pricesof Sweet Potatoes in Mulago and Nsambya Markets

CHART I II: (cts. per lb.)



The number against dots refer to weeks

An exponential (polynomial) form was therefore considered next. This is generally written

$$q_t^D = A\, p_t^{\alpha}$$

or $$\text{Log } q = \text{Log } A + \alpha \text{ Log } p$$

These two functions are, therefore, largely different and are similar only to the extent that the second model is linear in logs. The normal equations may be written:

$$n(\log A) + \alpha \sum (\log P) = \sum (\log q)$$

$$(\log A) \sum (\log p) + \alpha \sum (\log p)^2 = \sum (\log p)(\log q)$$

Thus it is possible in this way to solve for Log A and α where α is the coefficient of elasticity of demand

Log A is a constant

Log a is the dependent variable

Log p is the independent variable

No = No. of observations

At the time of writing the computer output for these regressions are being awaited, and it is not possible, therefore, to comment further on the results.

Price Elasticity of Supply:

Tables II and III provide a simple price elasticity of supply computations for sweet potatoes and cassava in Mulago market. This is simply calculated on the basis of proportionate quantity changes over proportionate changes in price. A time lag of one week in quantity was allowed for. It will be seen that in some cases the results are positive owing to the various factors which affect price - income, stock-holdings, consumer peculiarities, price of other commodities and so on. Price is therefore not only dependent on quantity, but on several factors.

To some extent, therefore, the figures are misleading and it is preferable to look at changes over the whole year as will be attempted in a later analysis.

Demand Functions which Include Income Variables:

In this study demand functions have been estimated which do not include income variables, whereas these are important determinants of demand. An investigation of consumption expenditure behaviour amongst urban consumers was planned at the design stage of the project but was unfortunately abandoned. The lack of studies in this area, therefore, leaves a vacuum in our knowledge of demand functions.** It is to be hoped that this gap will be filled in the near future by researchers interested in price analysis.

It follows that our models are not necessarily deterministic since so many variables that affect the demand function have not been taken into account. Progressively more complicated forms of demand functions are possible which include income as one of the explanatory variables as Messel and Hayer have done in Kenya(3).

Islam(4) has summarized these various possible functions as follows (more complicated models have been omitted):

(o) $Y = A + Bx$
(i) $Y = A + B \log x$
(ii) $\log Y = A + B \log x$
(iii) $Y = A + B/x$
(iv) $\log Y = A + B/x$

where
Y = aggretate income
A = constant
X = aggregate expenditure
B = coefficient of elasticity

...../9

** There are exceptions to this, for example, Charles Frank's work on "The Sugar Industry in East Africa(9).

Table II: Price Elasticity of Supply for Sweet Potatoes:

Mulago Market 1967/1968

Week	Elasticity
1	1
4	4.67
5	.28
6	4.33
7	-5.28
8	2.61
9	-33.73
10	21.00
11	1.00
12	1.00
13	4.29
14	.54
16	-21.12
17	1.96
18	1.92
19	-2.53
21	0.31
22	-38.00
24	61.88
25	-4.20
26	2.92
28	1.00
29	8.10
30	-9.02
31	7.59
32	1.03
35	1.00
36	-0.56
37	1.57
38	2.49
39	-1.40
40	-0.57
41	30.66
42	-3.39
43	.24
44	.50
45	5.79
46	-0.33
[illegible]	[illegible]

Table III: Price Elasticity for Cassava in Mulago Market

1967/1968

Week	Elasticity
1	1.00
2	.29
4	.61
5	- .45
8	-2.55
9	-9.00
10	-0.98
11	3.27
12	1.00
13	4.25
14	5.48
19	-5.66
21	8.33
28	1.00
29	3.44
30	1.00
32	-0.57
35	1.00
36	9.07
37	0.25
38	-0.23
42	4.50
43	2.92
44	1.80
45	2.91
46	-1.34
47	-15.8
50	-14.0

In each case the selection of the most appropriate consumption function is made on the basis of R^2, the coefficient of determination.

The linear function (o) is the simplest form and assumes that the coefficient of elasticity tends towards unity as income increases indefinitely. The double logarithmic function (ii) implies constant elasticity and is satisfactory when the range of variation of income or aggregate expenditure is sufficiently narrow and when consumption is expressed in terms of quantities. The semilog function (i) yields an elasticity coefficient which is inversely proportional to the quantity or the value of consumption as the case may be. The log inverse function yields a saturation level for consumption which may appear plausible in the case of food, especially cereals(5).

Least square analysis in demand function:

In this study the demand functions are being estimated by the use of least squares. This is because, the use of simultaneous equations which are now much more frequent(6) would prove too complex and expensive. In any case since the variables included are so few such time consuming method is not warranted. This of course does not preclude their depolyment in the future as market - price research gathers momentum and more and more variables are incorporated in the models. Even then, least squares have still a useful role to play in demand analysis. They can be used in the estimation of elasticity of demand where data are available on purchases or consumption of individual consumers since prices that confront consumers are determined chiefly by factors other than those that affect their purchases. In this case consumption is taken as the dependent variable and retail prices of the various items, family income are taken as independent variables.

In addition least squares equations may yield satisfactory forecasts and at times better forecasts than any other method of fit depending on the particular circumstances that apply.

Citations:

(1) Working, Elmer 1927. What do Statistical "Demand Curves" show? Quart. Jour. Econ. 4: 212-235.

(2) Poleman, Thomas T, and Oloya, J. J. An Urban Food Consumption Survey in the Greater Kampala, R.D.R. 54, p. 4, table 3.

(3) Messel, Benton F., and Heyer Judith 1969. Household Expenditure in Nairobi: A statistical Analysis of Consumer Behaviour. Econ. Dev. and Cult.Change 17:2.

(4) Islam, Nurul 1966. Studies in Consumer Demand Vol. II. Bureau of Econ. Res. Dacca University.

(5) op. cit. p. 23

(6) Foote, Richard F. Analytical Tools for studying Demand and Price Structure, Agriculture Handbook, No. 146. U.S. Dept. of Agriculture, Washington D.C.

(7) Johnston J, 1963. Econometric Methods, McGrow Hill Book Co., London.

(8) Klein, Lawrence R, 1962. An introduction to Econometrics, Prentice Hall, Englewood Cliffs, N. J.

(9) Frank, Charles S. The Sugar Industry in East Africa, East African Studies No. 20.

(10) Jones, William O. 1969. Marketing of Staple Food Crops in Tropical Africa, Food Research Institute, Stanford University.