MARKETING OF FOOD CROPS AND PURCHASE OF INDUSTRIAL COMMODITIES BY NYANZA SMALLHOLDERS, 1970-1971.

By

M. David

WORKING PAPER NO. 247

INSTITUTE FOR DEVELOPMENT STUDIES
UNIVERSITY OF NAIROBI
P.O. Box 30197
Nairobi, Kenya.

September, 1975.

Contents

I am most grateful to the Central Bureau of Statistics, and particularly to Messrs. Parmeet Singh, G.H. Olum, and James Otto, for their assistance in helping me with this study. Hopefully this is the first in a series of analyses of rural smallholders in Kenya. The basic descriptive findings relative to these data will be published by the Government of Kenya as a report on the Rural Household Budget Survey: Nyanza Province 1970/1971.

MARKETING OF FOOD CROPS AND PURCHASE OF INDUSTRIAL COMMODITIES BY NYANZA SMALLHOLDERS, 1970-71

by

Martin David

ABSTRACT

We consider the allocation problem of a rural household whose income is available in the form of $c+r$ inventories of farm produce and an amount of cash income from non-farm activity. The inventories available may be sold to earn cash at a price of τ. The inventories may be supplemented by purchase at market prices ω. The budgeting problem of the household is the problem of optimizing both sales of inventories and the consumption of commodities. The level of purchases are structurally related to inventories through the level of sales and cash income.

The optimum sales - consumption allocation depends heavily on the demographic composition of the household. An household with few members needs to reserve less of the inventory for subsistence than a household with many members. The analysis emphasizes the effect of changing household composition on the allocative choices made.

The model set forth is a combination of a system of purchase equations used to explain allocation of consumption and a multivariate logit model used to explain the proportion of inventories sold. The novel features of the model include:

(a) Conditioning the behavior of the household on its initial stocks;

(b) The use of transactions cost to define discrete differences between sellers and buyers of foodstuffs;

(c) Testing for interrelationships between the structure of income sources and the willingness to sell foodstuffs; and

(d) Structuring the model to make the consequences of demographic change clearly evident, even when the existing population is restructured into more or fewer households.

I. Notation

There are I households indexed i = 1.... I. Each household completes the agricultural production cycle with inventories of q food crops and r industrial crops. Industrial crops must be sold for processing. Food crops can be consumed directly. In addition p consumer goods other than food can be imported from the industrial sector; and z investments can be made in productive capital. The table below defines the variables to which we refer in our model:

Item	Number of components	Consumption vector	Production Disposed	Sales	Price of goods sold	Purchases	Price of purchases
Food crops	q	C(1)	Q(1)	S(1)	τ(1)	P(1)	ω(1)
Industrial crops	r	0	Q(2)	S(2)	τ(2)	0	0
Consumer goods	p	C(3)	0	0	0	P(3)	ω(3)
Investments	z	C(4)	Q(4)	S(4)	τ(4)	P(4)	ω(4)
Total		C	Q	S	τ	P	ω

In each row the symbols depict a vector of commodity flows. Zero is the null vector for those transactions that are definitionally ruled out. Industrial crops can not be directly consumed and would not be purchased as a final consumer product. Consumer goods are not produced in the rural household and are not sold to others. Q(4) represents capital construction by the rural household; and C(4) represents gross capital investments in productive facilities.

An accounting identity links the flows shown

(1) $C = Q - S + P$

<u>Income relationships</u>. Given that all of the inventories might be sold the full income available is

(2) $\tau'Q + E + R = Y(1)$

where $\underline{E}$ represents earned non-farm income and $\underline{R}$ represents remittances, rents, and other transfers. The amount of cash realized, Y(2), is less than Y(1) because some goods are directly consumed so that

(3) $\tau'S + E + R = Y(2)$

<u>Other variables.</u> F = // F(j) // is a (kx1) vector that describes the demographic structure of the family.

j =	Definition F(j)
1	Number of infants < 1 year of age
2	Number of children aged 1 - 4
3	Number of children aged 5 - 9
4	Number of children aged 10 - 12
5	Number of girls aged 13 - 17
6	Number of boys aged 13 - 17
7	Number of females aged 18 or over
8	Number of males aged 18 or over

A = // A(j) // is a vector that describes the access of the family to various amenities and the market

j =	Definition A(j)
1	Distance to nearest market
2	Distance to nearest public transport
3	Distance to water (dry season)
4	Distance to nearest primary school
5	Distance to nearest govt. secondary school
6	Distance to nearest district headquarters

Production characteristics. The Province of Nyanza comprises a large number of ecological zones, each with different climate and rainfall characteristics. For purposes of this analysis the differences are dichotomized according to the number of harvests generally expected in the sublocation where data were collected. Thus H = 1 where only one harvest can be obtained; and H = 2 in areas that regularly crop their land twice a year.

Differences in production can also be expected to lead to differences in purchasing behaviour. Households in areas that produce only one crop will be more likely to import foodstuffs, to seek urban employment and to sell a large proportion of their food crops (owing to deterioration during storage) than households in two-crop areas.

K denotes an index of household durable stocks. G denotes storage available for food crops.

$$M = // m_{jk} // \text{ where } m_{jj} = 0 \qquad j = 1 \ldots q$$

$$m_{jk} = q_k - s_k \qquad j \neq k$$

$$j, k = 1, \ldots, q$$

In the discussion below $\hat{}$ takes the indicated vector into a conformable diagonal matrix.

2. The Model: Theoretical Development in The Zero Saving Case

The problem facing the household can be set forth as a simple optimization problem. Given a utility function U(C) the problem is to find P(1), S(1), P(3) ≥ 0 that maximize U subject to the q + 1 constraints:

(1) $$C(1) - S(1) + P(1) - Q(1) = 0 \qquad \text{(inventories)}$$

$$E + R - \omega' P + \tau(1)' S(1) = 0 \qquad \text{(cash budget)}$$

To simplify presentation we consider E as exogenous; later we will relax that assumption. The Lagrangian to be maximized is

(2) $$L = U(C) + \lambda[E + R + \tau(1)' S(1) - \omega' P] + \mu' [C(1) - S(1) + P(1) - Q(1)]$$

The Kuhn-Tucker conditions for a maximum are

(3.1) $$\frac{\partial U}{\partial C(1)} - \mu \leq 0 \qquad C(1) \geq 0$$

(3.2) $$\frac{\partial U}{\partial C(3)} - \lambda\omega(3) \leq 0 \qquad C(3) \geq 0$$

(3.3) $$\lambda\tau(1) - \mu \leq 0 \qquad S(1) \geq 0$$

(3.4) $$-\lambda\omega(1) + \mu \leq 0 \qquad P(1) \geq 0$$

(3.5) $$C(1)' \frac{\partial U}{\partial C(1)} - \mu + C(3)' \frac{\partial U}{\partial C(3)} - \lambda\omega(3) + S(1)' (\lambda\tau(1) - \mu) + P(1)' (-\lambda\omega(1) + \mu) = 0$$

plus the constraints (1). Conditions (3.2) reflect the usual condition that marginal utility of additional purchases divided by price must equal λ, the shadow price or marginal utility of cash. (The condition is derived by using the identity C(3) = P(3).) Conditions (3.3) - (3.5) can be combined to give insight into the management of the commodity inventories. If for the j^{th} commodity, $S_j(1) = P_j(1) = 0$, then

(4) $$\lambda\tau_j(1) \leq \mu \leq \lambda\omega_j(1)$$

If $C_j(1) > 0$ then (3.5) implies $\frac{\partial U}{\partial C_j(1)} = \mu_j$ and therefore

(5) $$\lambda\tau_j(1) \leq \frac{\partial U}{\partial C_j(1)} \leq \lambda\omega_j(1)$$

In other words the marginal utility of the commodity divided by the shadow price of cash lies between the buying price and the selling price of the commodity.

It is not possible for $S_j(1)$, $P_j(1) > 0$. If $S_j(1)$, $P_j(1) > 0$, condition (3.5) implies that

$$\lambda \tau_j(1) = \mu \text{ and } \lambda \omega(1) = \mu.$$

Since $\tau_j(1) < \omega_j(1)$, μ cannot have two different values simultaneously. If follows that:

(6) $S_j(1) > 0$ implies $\lambda \tau_j(1) = \mu$ and $P_j(1) = 0$

$P_j(1) > 0$ implies $\lambda \omega_j(1) = \mu$ and $S_j(1) = 0$

Assuming that $C_j(1) > 0$, we then have

(7)
$$\frac{\partial U}{\partial C_j(1)} = \lambda\omega_j(1) \qquad \text{If } P_j(1) > 0,\ S_j(1) = 0$$

$$\lambda\omega_j(1) \leq \frac{\partial U}{\partial C_j(1)} \leq \lambda \tau_j(1) \qquad \text{If } P_j(1) = S_j(1) = 0$$

$$\frac{\partial U}{\partial C_j(1)} = \lambda \tau_j(1) \qquad \text{If } S_j(1) > 0,\ P_j(1) = 0$$

As the marginal utility of commodities is assumed to decline with increasing consumption, the conditions (8) imply that a larger quantity of the commodity is consumed when surplus is sold than when deficits must be purchased.

Figure 1 illustrates the implications for the relationship between inventory of a crop and consumption. If the initial inventory is less than Q_o the deficiency is purchased at price ω as the marginal utility of consumption divided by price would otherwise exceed the shadow price of cash. If the initial inventory exceeds Q_1 the excess will be sold as the marginal utility of consumption divided by price falls short of the shadow price of cash. Between Q_o and Q_1 no action will be taken to buy or sell.

Now imagine that we consider the impact of changing the proportion of total income Y(1) that is in a commodity inventory. Define $\tau' Q/Y(1) = \rho$. Consider a two commodity world with one C(1) and one consumer goods C(3). C(3). Then Figure 2 shows the expected relationship of C(1)/C(3) to ρ.

The foregoing argument motivates the treatment of cash income and inventories in our model: Surpluses will be sold on the basis of selling prices $\tau(1)$ and Y(1). Inventories will be retained for consumption on the basis of buying prices. Purchases will supplement inventories to an extent determined by buying prices $\omega(1)$ and cash income.

Discussion

Assume the household wishes to maximize its utility U(C(1), C(3)) subject to the constraints :

(1')
$$C(1) = P(1) - S(1) + Q(1)$$
$$E + R = \omega'P + \tau(1)' S(1)$$

We can eliminate the inventory constraints and write the Lagrangian to be maximized as

(2')
$$L = U[P(1) - S(1) + Q(1), P(3)] + \lambda [E + R - \omega'P + \tau(1)' S(1)]$$

The Kuhn-Tucker conditions for an optimum are:

(3')
$$\frac{\partial L}{\partial P(1)} = \frac{\partial U}{\partial P(1)} - \lambda\omega(1) \leq 0 \qquad P(1) \geq 0$$
$$\frac{\partial L}{\partial S(1)} = \frac{\partial U}{S(1)} + \lambda\tau(1) \leq 0 \qquad S(1) \geq 0$$
$$\frac{\partial L}{\partial P(3)} = \frac{\partial U}{\partial P(3)} - \lambda\omega(3) \leq 0 \qquad P(3) \geq 0$$
$$\frac{\partial L}{\partial \lambda} = E + R - \omega'P + \tau(1)'S(1) \leq 0 \quad \lambda \geq 0$$

and
$$\frac{\partial L}{\partial P(1)} P(1) - \frac{\partial L}{\partial S(1)} S(1) + \frac{\partial L}{\partial P(3)} P(3) = 0$$

From our earlier argument it is clear that no more than q of the first 2q inequalities may hold as equalities as $P(1)'S(1) = 0$. Then Q can be partitioned:

$$Q' = (\bar{Q} \ \vdots \ \underset{\sim}{Q} \ \vdots \ Q_o)$$

where $\bar{Q}$ correspond to j for which $S_j(1) > 0$;

$\underset{\sim}{Q}$ corresponds to j for which $P_j(1) > 0$; and Q_o corresponds to j for which $C_j(1) = Q_j(1)$. Partition P and S conformably. Then

(7')
$$\frac{\partial U}{\partial S(1)} = \lambda\tau(1)$$
$$\frac{\partial U}{\partial P(1)} = \lambda\omega(1)$$

From these relations we determine that λ is the weighted average of marginal utilities afforded to those commodities that are purchased less the weighted marginal utility of those commodities that are sold per unit of exogenous cash income:

(8) $$\left[\left(\frac{\partial U}{\partial \underline{P(1)}}\right)' P(1) - \left(\frac{\partial U}{\partial \overline{S(1)}}\right)' \overline{S(1)} \right] / \left[E + R \right] = \lambda$$

It is now possible to visualize the indirect utility function that corresponds to U:

(9) $$\begin{aligned} \underline{P(1)} &= f^{(1)} (Q, E+R, \omega, \tau(1)) \\ \overline{S(1)} &= f^{(2)} (Q, E+R, \omega, \tau(1)) \\ 0 &= f^{(3)} (Q, E+R, \omega, \tau(1)) \\ P(3) &= f^{(4)} (Q, E+R, \omega, \tau(1)) \end{aligned}$$

For $Q = 0$ we have the classical consumer allocation problem. No sales are possible, the set $\{j \;/\; \overline{S(1)} > 0\}$ is empty. $\tau(1)$ is not relevant.

Write demand for consumption as

(10) $$\begin{aligned} C(1) &= g^{(1)} (E+R, \omega) \\ C(3) &= g^{(3)} (E+R, \omega) \end{aligned}$$

Two polar cases of (9) are easy to relate to (10):

Case A. No surplus inventories: $0 \leq Q(1) < C_A(1)$

Case B. All inventories in surplus: $C_B(1) \leq Q(1)$

<u>Case A.</u>

The amount consumed will be determined by valuing inventories at their purchase prices, and substituting in (10)

(11) $$\begin{aligned} C_A(1) &= g^{(1)} (E+R + \omega(1)'Q(1), \omega) \\ C_A(3) &= g^{(3)} (E+R + \omega(1)'Q(1), \omega) \end{aligned}$$

so that

(12) $$\begin{aligned} \underline{P(1)} &= f^{(1)} (Q, E+R, \omega, \tau) \\ &= g^{(1)} (E+R + \omega(1)'Q(1), \omega) - Q(1) \end{aligned}$$

Case B.

The amount consumed will be determined by valuing inventories at their sale prices and substituting in (10)

$$(13) \qquad C_B(1) = g^{(1)}\left(E+R + \tau(1)'Q(1),\ (\tau(1)',\ \omega(3)')'\right)$$
$$C_B(3) = g^{(3)}\left(E+R + \tau(1)'Q(1),\ (\tau(1)',\ \omega(3)')'\right)$$

so that

$$(14) \qquad \overline{S(1)} = f^{(2)}\left(Q,\ E+R,\ \omega,\ \tau(1)\right)$$
$$= Q(1) - g^{(3)}\left(E+R + \tau(1)'Q(1),\ (\tau(1)',\ \omega(3)')'\right)$$

Comparison of (12) and (14) reveals a number of difficulties where $0 < Q_j(1) < C_{A,j}(1)$ for some j and $C_{B,k} < Q_k(1)$ for some k :

(a) Clearly the inventories of some commodities should be valued at purchase prices, while others should be valued at sales prices.

(b) The relative prices of goods shift as some become surplus so that the prices necessary to ascertain a proper allocation include some elements of $\tau(1)$ as well as $\omega(1)$.

To proceed further we must ascertain to what extent the programming feature of this model, namely that $P'(1)\ S(1) = 0$ is valid. This requires examination of actual behavior in the sample of Nyanza households.

3. Prevalence of production, purchase, and sales

To ascertain whether the programming conditions derived in (3) are relevant for the expenditure data of hand , we examine the prevalence of production, sales, and purchases of major food crops in Nyanza. Figure 3 indicates substantial variation in the prevalence of the various food crops among households, and considerable variation in the proportion (by value) of the crop sold. Production of vegetables, cereals, pulses, and meat is almost universal among these households. Roots, fruit, and dairy products are almost as common while the production of sugar, fish and nuts is restricted to a small minority of households.

Among the crops vegetables, cereals, and roots are produced almost exclusively for home consumption. A substantial portion of pulses, fruit, and sugar is sold; but only dairy, meat, fish, and nuts are commercial products where more than 40 percent of the value produced is sold.

The programming model described above suggests that the household would not engage in both purchases and sales, because the margin between buying and selling prices ensures that real income falls whenever a single commodity is sold and repurchased. Table 1 indicates that none of the crops conform strictly to the programming model. A significant number of households are engaged in both purchase and sale of the same food crop.

It is not possible to ascertain the cause of this seemingly inconsistent behavior, but three contributing factors may be cited. Each of the crop categories is an aggregate, so that sale of one and purchase of another is to be expected.(E.g. sale of green maize and purchase of millet would both be included under cereals). Secondly ,some perishable commodities may be sold during a local harvest while they are repurchased at other times (e.g fruit, vegetables). Thirdly, it is possible, in areas with two crops per year, that one crop was in surplus, while the second was deficient to meet family needs.

What stands out in Table 1 is that sugar, pulses, fish, and nuts reflect more consistent behavior than other crops. In combination with the relatively high proportion of these crops that are sold, consistency appears to be an indication of the commercialism associated with the crop.

For dairy and meat production the combination of high proportion sold, and high proportions of producers in both buying and selling markets, appears to relate to the discontinuities in production and the indivisibilities in output, associated with ownership of relatively small herds.

The principal implication of Table 1 for further analysis is that it is not essential to restrict estimation in the manner suggested by $P(1)'S(1) = 0$. For an appreciable fraction of the population both sales and purchases of the same commodity group will occur.

As we have not developed a theoretical argument to explain simultaneous purchases and sales, our approach is necessarily pragmatic. The approach of estimating an extended system of demand equations such as is suggested by (11)and (13) is one possibility. That approach has three shortcomings. (a) Information on purchases and sales does not enter the relationships. (b) The inventories of food crops produced will necessarily be highly correlated to consumption of food because of the identity (1). Thirdly, it is infact not possible to ascertain the correct value of income that should be included in the demand functions. In the general case,

$$C = g(Y_o, \omega)$$

where Y_o is a measure of income that is in the interval

$$E{+}R + \tau(1)'\,Q(1) \leq Y_o \leq E{+}R + \omega(1)'\,Q(1) = \bar{Y}(1)$$

The value of Y_o will depend on which crops are in surplus and which are deficient to meet desired consumption.

As an alternative, the original problem in (2) can be modified to assert that

(15) $$P'(1)\,S(1) \geq h\,(H, G) \geq 0$$

The logic for this constraint is that continuity in rainfall and multiple harvests increase the periods when the holding is self-sufficient. Therefore $\frac{\partial h}{\partial H} < 0$. Similarly the capacity to store inventories should increase self sufficiency, implying $\frac{\partial h}{\partial G} < 0$.

With this new constraint simultaneous purchases and sales are determined by supply considerations. As modification of (3) gives few insights, we develop further implications of (15) heuristically.

A casual asymmetry suggests a natural order for the decisions S(1) and P(1). The harvest must mature before sales can be consumated. Sales imply a reduction of inventories that may be allocated to consumption in the present or to either sales or consumption in the future. Purchases may augment the inventories remaining at present to provide for present consumption.

This temporal sequence appears more rational than alternatives that are logically possible. Because of spoilage it would not appear advantageous to purchase in advance of the time of consumption. It would not appear logical to purchase when available inventories exceed current consumption requirements. And the household should know its remaining inventories at the time purchases are undertaken.

Given this reasoning, sales would appear to be conditioned on the entire inventory of food commodities in relation to spoilage and the future demands of the household. Purchases, on the other hand, depend only on the level of retained inventory after sales. It follows that we should modify (12) and (14) to

$$P(1) = f^{(1)}\ (Q-S\ ,\ E+R\ ,\ \omega) \quad (12')$$

$$S(1) = f^{(2)}\ (Q,\ E+R\ ,\ \omega,\ \tau) \quad (14')$$

4. Family Composition

Up to this point we have not explicitly introduced family composition into the model. One of two approaches can be taken:

(a) The demography of the family can be taken as an explicit determinant of the utility function (i.e. a measure of taste).

(b) The effects of family structure may be viewed as the effect of aggregating like preferences of individuals into a single household demand.

We discuss the second approach first.

Assume that the indirect utility function $C^* = g(Y^*, \omega)$ applies to an adult male. For two adult males in the same household, each with Y^* income to spend consumption would be $2C^*$. However, the household income constraint is $Y = 2Y^*$ so that we can write household demand C as

$$C = 2C^* = 2g\left(\frac{Y}{2}, \omega\right)$$

or

(16) $$C^* = \frac{C}{2} = g\left(\frac{Y}{2}, \omega\right)$$

If all members of the household were identical and each were able to spend a pro-rata share of the household income, estimating a generalization of (16) would clearly be the appropriate way to account for family structure in an Engel curve. Thus

(17) $$C^* = \frac{C}{N} = g\left(\frac{Y}{N}, \omega\right)$$

where N is the number of family members. However, it is clear that members of the family differ in their consumption needs and demands. A child of 6 years of age needs less food and clothing than an adult. A crude approach to this problem is to assume that there are d types of consumption demand for each of d different age-sex characteristics and that consumption demand of the j^{th} type is a proportion γ_j of the adult male level.

Then family structure is summarized by the vector F and aggregate demand in the household is

$$(18) \qquad C = \gamma F\, g\left(\frac{Y}{\gamma' F}, \omega\right)$$

where $\gamma' = (\gamma_1 \ldots \gamma_d)$

The use of $\frac{Y}{\gamma' F}$ as the measure of income density appropriate to each individual's choices is debatable. It follows only if we are willing to assume an equalitarian distribution of resources within the family.

The notion that demand for each individual commodity bears the same relationship to the adult male demand is also debatable. For that reason Prais and Houthakker generalized (18) to

$$\hat{\omega} C = \Gamma F\, g\left(\frac{Y}{\gamma' F}, \omega\right)$$

where Γ is a full matrix subject to the constraint

$$\iota' \Gamma = \gamma'$$

To summarize this approach, (A) aggregation of members in a household implies that each is subject to spending constraints that are a fraction of the household constraint. (B) The differences in needs of different household members are parameterized as multiples of the adult male demand. Nothing in this theory demands continuity in the nature of demand by household members that are similar in age. Nothing in the theory allows for economies of scale that arise in housing, durable equipment, and clothing.

As the alternative to this approach one may include the family composition directly in the utility function as an argument, i.e $U(Y,\omega,F)$. The only theories that constrain the form of U are the so-called extended linear expenditure systems that suggest that family structure operates through the determination of a minimum consumption bundle, but does not affect the marginal propensity to consume out of supernumerary income (Betancourt, Leuthold). This approach is not only suspect on a priori grounds, it is also conceptually difficult as empirical work in the past has often established subsistence allowances related to family size that imply the poorest individuals in the sample have undefined preferences!

In our initial investigations we adopt a modification of both approaches:

A. Calorie requirements for the family γ_o are used to estimate the γ in (18)

B. The family structure is also incorporated as an argument of the per-equivalent-adult (PEA) expenditure function g. We hypothesize that per capita expenditures are a declining linear function of the calorie requirement (to account for economies of scale), and an increasing function of the diversity of family members, Δ.

Under these assumptions (18) generalizes to

$$(19) \qquad \hat{\omega} C = \gamma_o{}'F \left[g\left(\frac{y}{\gamma_o{}'F}, \omega\right) + \delta_1(\gamma_o{}'F) + \delta_2\Delta \right]$$

where Δ is the mean square of the difference between individuals' requirements and the average requirement of individuals in the family

$$(20) \qquad \Delta = \left[\sum_{i=1}^{k} (\gamma_{oi}F(i) - \gamma_o{}'F/N)^2 \right] / \left[\gamma_o{}'F \right].$$

Two arguments can be made for relating expenditures to the Δ defined in (17). The inclusion of a measure of the heterogeneity of family members could be expected to account for the more diverse requirements of a household that includes infants, adolescents, and adults; conversely, the uniformity of a household that includes only adults would be reflected in a lower level of expenditures to meet a given level of living.

Implicitly, the measure provides an index of effective child-spacing, after the deaths of young children and the decision to extend the family by including relatives has been taken into account._/

_/ The choice of Δ rather than (a) $\Delta \gamma_0' F$ or (b) $(\Delta \gamma_0' F)^{\frac{1}{2}}/(\gamma_0')$ is arbitrary. The correlation between Δ and $\gamma_0' F$ is -.25 and highly significant; it could potentially be reduced by a measure such as (b) which has dimensions similar to a coefficient of variation.

The logic for (19) can be seen by studying the relationship between γ_o'F and various expenditures. Most items of expenditure increase to a significant extent with increases in the estimated calorie requirements of the family. A few bear no relationship to increased calorie requirements:

Significant Positive Relationship	R^2	Null relationship	R^2
Clothing	(.108)	Drink and tobacco	(.011)
Soap	(.105)	Meals bought	(.011)
Cereals	(.079)	Roots	(.010)
Sugar	(.054)	Household equipment	(.009)
Fuel	(.052)	Vegetables	(.008)
Dairy	(.045)	Pulses	(.001)
Meat	(.042)		
Fish	(.040)		
Other food	(.027)		
Legal/other	(.014)		
Fruit	(.009)		

Analysis of variance for 10 groups based on the calorie index shows that all the positive relationships are significant at the .005 level, except fruit and legal/other which are significantly positive at the .05 level. The number in parentheses indicates the proportion of variance in expenditures that can be explained by a simple linear regression on the mean value of the calorie index in each of the ten groups. Mean values of C/γ'F for the ten groups based on γ_o'F are shown in Table 2.

The absence of a relationship between total expenditures on some items and the calorie index can be easily explained by the fact that there is a strong positive correlation between the calorie index and household income. 0.163 of the variance in total income can be explained by a positive relationship to the calorie index._/

_/ The relationship between Y/γ_o'F and γ_o'F is much smaller, accounting for .093 of the variance in Y/γ_o'F .

Some indication of the source of the relationships between total income and the calorie index is given by the correlation between the calorie index and income components:

Source of income	Symbol	Zero-order Correlation to CALORIE
Farm sales (food crops)	S(1)	.381
Farm sales (industrial crops)	S(2)	.244
Own produce consumed	Q(1)-S(1)	.296
Earnings	E	.202
Remittances	R1	.290
Rents	R2	.090
Farms costs	-	.290

Because one would anticipate a positive relationship between income and expenditure, *ceteris paribus*, the null relationship observed in simple classification of families by the calorie index is the result of offsetting income and family size effects. To study those effects requires multiple regresssion or similar techniques.

The simple relationships between calorie and per capita expenditures are shown in Table 3.

5. Derivation of ω, τ

The data used in this study were collected from May, 1970, through May, 1971. Over the period there were substantial variations in prices, in fact in April, 1971, near famine conditions resulted in no goods being marketed in Siaya District. In addition to inter-temporal variation, there were also differences in prices at the various markets in the four districts of Nyanza.

For purposes of this study most of this variation has been ignored. Produce consumed directly by the household was valued at the mean price prevailing in the Province over the entire period. This has the effect of undervaluing inventories in areas where there is relative scarcity, and overvaluing inventories in areas of relative surplus. Nonetheless there appeared to be some merit to exploring household behaviour using this approximation, rather than attempting a complex revaluation of inventories at the outset.

In fact, the most pertinent way to view ω, τ is as a two-part tariff. There is a fixed cost to the household for accessing the market. We assume that cost increases linearly with distance from the market. Once access to the market is attained the household may enter the market directly, either as a buyer or seller so that a single price vector $\bar{\omega}$ governs activity in the market.

Under these asssumptions the vector A (See page 3) must bear a negative relation to both sales and purchases. While it can not be said that A is a proxy for the difference between ω and τ in any given market, it is clear that for each day of activity in the market, total costs of purchases must be augmented by the cost of access. Alternatively, income must be reduced by cost of access. If the household thinks in terms of the value of a typical day at the market, cost of access defines a differential between buying and selling prices.

A simplified example is shown in Figure 4. Sales of surplus commodities are shown on the horizontal axis; purchases of the other commodities are shown on the vertical axis. The cost of accessing the market is $O\alpha$. The expected volume of sales $\hat{S}$ will then finance $\hat{P} = \hat{S} - O$ of purchases. The slope $\emptyset = O\hat{P} / O\hat{S}$ becomes an index of the expected differential between buying and selling prices. Clearly as $O\alpha$ increases, $\emptyset$ falls and the household must set a greater value on the purchases to embark on market activity.

6. Proportion of produce sold

The argument in section 3 suggests that sales must be a function of other income (E + R), the market price τ, and the quantities in inventory Q. The ultimate value of inventories when sold is $\tau'Q$, and a simple approach would be to use full income Y(1) as a conditioning variable.

The functional form of the sales relationship is constrained by the fact that expected sales should not exceed production. Moreover expected sales should be non-negative. In the present analysis these conditions are achieved by estimating a logit function describing proportion of crop sold for producers only. Expected sales are then estimated as the product of expected proportion sold and the value of production. (This product will be zero for all non-producers.)

In order to limit the number of sales functions that were to be estimated food crops were aggregated into two categories:

(a) Cereals, roots, and vegetables

(b) Meat and dairy products

The logic for this aggregation is revealed by Figure 3 Group (a) are nearly universal items of crop production, and are largely consumed on the farm. Group (b) are somewhat less prevalent, but are marketed to a substantial extent, undoubtedly because meat and dairy products are perishable. Together (a) and (b) account for 1405/ of 1703/ of farm production of food and animal products. (See table 3.)Relationships for these two categories cover four-fifths of the food available for sale, on average. At the same time differences in perishability clearly influence the level of sales.

The logits were estimated iteratively by the method of Duncan and Walker (1967) which has also been shown to converge to maximum likelyhood (ML) estimators. A useful starting value for the first iteration are Ordinary Least Squares (OLS) estimators. The existence of any sub-group in the data for whom conditioning variables are identical and for whom proportion sold is zero or unity causes the estimation procedure to breakdown. Thus a premium is placed on limiting the number of arguments in the logit.

The functional form for the logit was selected on the basis of the correlation matrix in Table 4. Total income ($Y/\overline{1}$/) was generally not related to the proportion sold. This finding may be due to the fact that a minimum threshold must be reached before income affects sales positively. That hypothesis has not het been tested.

The nutritional requirements of the family, as measured by catorie ($Y_o'F$) do not affect proportion sold, though we had hypothesized that greater domestic needs would inhibit sales. The presence of older members in the family or several adults was deemed to make it easier to allocate one person to marketing crops; hence a positive correlation to sales was hypothesized. This relationship only appeared for pulses.

Distance to transport and distance to district headquarters are inversely correlated to sales, confirming the hypothesized effect of access. All crops do not show a significant correlation, but only one, pulses, contradicts the hypothesis with a non-negative (and significant) correlation.

Number of harvests is strongly correlated with proportion sold for five out of the seven crops.

Number of building on the holding was though to be a proxy for the quality of storage. In that case, ceteris paribus, more buildings should inhibit sales. In fact both buildings and the index of durable wealth have a positive correlation with sales of cereals and pulses, suggesting that these variables may be proxies for wealth of the household.

The expected level of purchases of goods other than food was correlated with sales to determine whether demand for products not available on the farm enhanced sales. This variable was significantly correlated only with sales of cereals, roots, and vegetables.

Variables significantly correlated to proportion sold were included in the logit formulations. While this procedure suffers from pre-test bias, it was the only one available in view of the limited computing capabilities available. The procedure can be validly criticized for failing to discover significant partial correlations that may be obscured in zero-order correlations. For example the calorie needs and income of the household are sufficiently correlated that neither may significantly affect sales unless the effect of the other has been controlled.

The logits estimated are shown in Table 5. Only a small proportion of the total variance is explained by the logit. The dominant effects are the inhibiting of sales by greater distance from transport and the greater marketing achieved by householders in 2-crop areas of Nyanza. The possibility of an additional effect of wealth, as measured by the durables index needs further investigation. (When durables were included in the logit for grains, roots, and vegetables no significant effect could be estimated.)

Sales of grain, roots, and vegetables appear to be positively related to school fees and household purchases as measured by EPUR. The significance of this effect is that it suggests a natural dynamic in marketing food crops: Increased family demands (say for school fees) lead to increased sales; and increased sales lead to increased expenditures on food and other purchases.

7. Purchases and Consumption by the rural smallholder

The estimates of the purchase relationships fitted to the Nyanza data are presented in Table 6. A few conceptual points should be made at the outset.

1) The purchase relationships for food items are not expenditure functions; they can be converted into relationships describing expenditure per equivalent adult (PEA) by adding the inventory per equivalent adult to both sides of the equation in question.

2) The five purchase relationships fitted are part of a larger system that includes purchases of farm inputs and purchase and sales of capital assets (both physical and financial). Estimates of the farm input equation will be presented in Table 10. Asset transaction equations were not fitted for several reasons: a) Adequate information on the value of a number of capital items was not available; b) The reliability of reporting of financial asset transactions is uncertain; c) Cash and inventory positions of the household at the beginning of the year were not measured and a stock adjustment model would appear most relevant. Logically the sum of certain coefficients in the purchase equations is constrained. The marginal propensity to allocate cash income must be unity. The effect of all other variables on purchases, ceteris paribus, is to reallocate among purchase flows and asset transactions. Hence there is some presumption that the sum of all coefficients should be zero. These constraints imply values for the sum of the coefficients in equations that were not estimated. These implications will be discussed below.

3) Because of the manner in which sales of crops enter the equation systems, farm costs are not netted out of gross income items, with the consequence that marginal propensity to spend on household consumption must be less than unity.

Table 6 provides the basis for several generalizations about the purchase behavior of households. Results pertaining to endogenous variables will be discussed first; results pertaining to demographic factors will be presented next; this section concludes with a discussion of the remaining influences on purchases.

The equations pertaining to food purchases were estimated in two forms, the structural model (A) that conforms to the earlier theory developed in (12'), (14') and a modification (B) that tests the validity of excluding the inventory of the food crop for which purchases are estimated from the purchase relationship. An alternative model of behavior that links full income (C) with consumption was fitted to test the conceptual model presented here. (Results for that model are presented in Table 7.) The latter model constitutes the maintained hypothesis that relates to more conventional theory of household allocation.

Endogenous Variables

Cash inflow per equivalent adult is highly significant in all purchase equations. The relatively low values of the marginal propensity to spend cash on household and farm inputs can best be assessed by adding the coefficients of INCCAL3; this has been done in Table 8. The resulting sum implies that .7 of every shilling of cash is expended on the omitted capital asset transactions. This value may appear more reasonable when one recalls that the 879/- per household of cash income compares to a retained inventory of food for consumption of 1331/-. The marginal propensity to spend the latter is unity so that the weighted propensity to spend (on non-capital items) out of gross income becomes .724.

It is of interest to contrast these results to those of Massell (1972), the only other study of rural expenditures in Kenya. Massell fitted a log-linear model in total expenditures, family size, and proportion of income consumed in kind to expenditures by rural households in Central province in the year 1963-1964. The elasticities derived from the propensity to purchase out of cash income are much smaller than the elasticities estimated by Massell. However the elasticities relating increases in consumption of food out of total income (including income in kind) are larger than Massell's. (See Table 9.) For the household goods the elasticity remains smaller. It is apparent from the wide range of elasticities that can be derived from the structural model that there is more information to be derived from that approach.

To test the value of the structural model an alternative model was estimated in which income without regard to source is aggregated and used to explain consumption rather than purchases. Results are presented in Table 7. The information included in the relationship when sales of food crops are included provides a highly significant increase in explained variance (Table 11). The knowledge of sales of food crops makes a marginal improvement in the explanation of purchased farm inputs.

Slightly less than half of the variance explained can be attributed to large differences in the propensity to spend for farm inputs out of different components of income (See also Table 10.) The findings bear out a strong orientation to purchase inputs on the part of those who have industrial crops (coffee, tea, sisal, cotton, hides, firewood). However, in the structural model (A) the relationship or purchased farm inputs to sales of food crops is

---

_/ The full income equations are estimated by OLS. The reasoning behind that procedure is that we may treat cash income other than earnings and food sales as predetermined. Furthermore it is assumed that earnings are not simultaneously determined by purchases or sales. This assumption will be tested in future work; it assures that the system is recursive so that no instrument for earnings is required, and OLS on cash income is unbiased. As food production is also pre-determined total income can also be teated as independent of purchases.

also substantial, with the propensity to purchase inputs at 0.2 of each shilling received. Another item of income that generates purchase of inputs is "other" which includes both land rent and equipment hire. The contrast between these coefficients and the 2.6 percent of earnings that is spent for farm inputs is striking.

The foregoing findings pertain to items of cash income. The role of retained inventories in the structural model can now be discussed. It was hypothesized that the availability of a large inventory of one type of food might act as a substitute (or complement) to inhibit (augment) purchases of another food crop. The structural model allows for six such effects and none was significant.

The theory developed in section 2 posits a relationship between cash and purchases and predicts no relationship between the inventory of a commodity retained for consumption and purchases of that commodity. To test the latter ressume the modified structural model was estimated. The results partially support the hypothesis. Significantly negative effects appeared in the grain and in the other food equations. While significant effects were not expected, they would appear to be a natural result of deficiencies in the instrument used for sales. If errors in the instrument are positively correlated with the volume of production, actual inventories retained will be greater and purchases will be less than predicted on the basis of the instrument. This would generate a negative coefficient on the retained inventory. As a negative coefficient implies that a part of retained inventories are not actually consumed, the modified structural model must be rejected as inadmissable.

Further confirmation that the structural model (A) is appropriate comes in the equation relating to purchase of household goods other than food and school fees. On theoretical grounds it was hypothesized that cash income would increase such purchases while inventories retained for consumption would not. This hypothesis is borne out in the equation (B) shown in Table 7. The coefficient attaching to retained inventories (after sales of food crops have been deducted) is essentially zero.

Demographic Variables

As indicated in (19) the purchase equations were estimated as linear relationships between purchases per equivalent adult (PEA), incomes and inventories per equivalent adult, and other variables. The inclusion of the equivalent adult measure (CALORIE) in the equation allows the elasticity of responses to income PEA to increase or decrease. Alternatively the term may be thought of as providing a quadratic term in CALORIE in relation to total expenditures fitted to a homogeneous equation. That is,

$$\frac{P}{N} = a_o + a_1 \frac{Y}{N} + a_2 N$$

becomes

$$P = a_o N + a_1 Y + a_2 N^2$$

Thus a positive constant term in the equations of Table 6 assures a positive effect of CALORIE, $\gamma_o{}'F$, on total purchases. A negative coefficient on the CALORIE term assures diminishing marginal effect of CALORIE on total purchases.

The effects of CALORIE were negative for purchases of household goods and food, as hypothesized. Together with large and significant constant terms, that finding implies reductions of purchased farm inputs and savings with increasing CALORIE; the size of this effect declines as CALORIE increases.

The measure of dispersion in family requirements Δ, also affects purchases in the expected direction. The greater the diversity of calorie requirements of family members as measured by Δ, the greater the volume of purchases PEA. Over the life cycle, there is an inverse relationship between CALORIE and Δ as can be seen in the correlation of -.25 between the two variables. A hypothetical family development over the life cycle is portrayed in Table 12 to give a clearer picture of the complex relationship between CALORIE and Δ. It is clear that the combined effect of variables is to attenuate purchases below the level that would be forecast on the basis of family size alone. Furthermore an asymmetry is introduced between young and older families; younger families of a given size will undertake more purchases ceteris paribus.

The significantly negative effect of age of head on purchases of meat and dairy products, and purchases of household goods has two interpretations. It may be a life cycle effect in addition to those already described. The implication of that interpretation is that aging in a family produces even greater asymmetries in the purchases of young and old families than shown in Table 12. The alternative, and more probable, interpretation is that age of head connotes a cohort effect. Younger persons have had greater exposure to market relationships and will continue a greater rate of market transactions than the last generation. The correct interpretation can only be identified by investigating household behavior at another point in time.

Completion of primary education appears to be associated with greater diversity in the purchases of food products. More animal products and more fruit, pulses, fish, and other foods are purchased than by those with less formal education. It would be of interest to investigate this behavior more closely to ascertain whether better dietary balance is achieved by the more educated group. (Education was not included in the purchase equation for grains on the, perhaps erroneous, assumption that this factor is irrelevant to the level of carbohydrate intake that dominates the grains equation.)

Completion of primary education also enhances purchases of household goods. The arguments at the beginning of this section imply that it is not possible for all education effects to be positive, as increased expenditure in one area should correspond to reduced expenditure elsewhere. In fact, a counter-argument can be raised and should be investigated in further work — Better reporting of purchase and consumption behavior by persons with some education could well imply that all items of expenditure rise with attainment and that the education effect connotes no real behavioral differences, only differences in the reliability of reporting.

One other test of demographic effects should be mentioned. The inclusion of the instrument for sales ESALES, without deflation by γ_0'F, allows for an interaction between sales of food surplus and family structure. This effect proved significant only for purchases of household goods. It may be interpreted to indicate that available surpluses increase expenditures on industrial consumer goods beyond what would be predicted by income per equivalent adult.

Asset Stocks

Durables and buildings are the two variables reflecting the family asset stocks that were included in the purchase functions. DURABLES, K, is crude index based on the frequency with which a variety of household and farm assets were reported. The procedure adopted is very crude and relies indirectly on the careful study of rural household wealth undertaken by Hunt(1975). Eight items in the inventory of household assets reported by the households were selected from a list of ten. (Motor cycles and automobiles were excluded as they were reported by only 5 - 7 households in all.) Ownership of ploughs and harrows was added to the list of household assets. The durables index was then computed by weighting each item proportionately to the reciprocal of its occurrence in the sample.

Item	Probability of owning 1 or more	Weight
Chairs	.972	1.029
Beds	.914	1.094
Tables	.847	1.181
Ploughs & Harrows	.335	2.985
Bicycles	.286	3.497
Watches	.181	5.525
Radios	.122	8.197
Pressure lamps	.119	8.403
Sewing machines	.035	28.571

(Hunt found an identical rank order for these items except sewing machines which were not included in her study.)

If each household acquires assets in the same order, then it can be argued that purchase is precipitated by acquisition of some minimal level of wealth, and that therefore wealth should be monotonically increasing function of the frequency of asset ownership. No defense of the precision of the this assertion can be made; and it would clearly be preferable to weight asset stocks by their depreciated value.

The durable index has a uniformly positive and highly significant effect on all household and farm purchases. By implication the rate of capital acquisition must be inhibited by increasing durable stocks, ceteris paribus.

The other measure of wealth, G, is a simple count of the number of buildings on the holding, with no weighting for construction details. A thatched mud hut is counted equally to a steel-roofed, stone building. This index of structures was thought to be a proxy for adequacy or storage among other things. For that reason G was thought to be associated with reduced purchases of grain. That effect was substantiated. Negative effects on other food purchases were not anticipated, and we have no explanation to offer.

Access:

Distance to various facilties plays a significant in both the purchase equations and the equations describing proportion of crop marketed./ The two effects should be clearly distinguished. Reduction in sales due to distance from market reduces cash income and thus automatically assures a reduced level of purchases. The effect of distance from transport in the purchase equations must be interpreted as an allocation effect. That is, the reduction in purchases of food and household goods effected by an additional mile of distance from transport must be offset by committing a larger proportion of cash income to capital asset acquisition.

While it is logically possible that purchased farm inputs are also affected positively by distance to transport, that eventuality has not yet been investigated and appears unlikely on a priori grounds. (One also cannot rule out some underreporting bias by persons less involved in market transactions, but the majority of that effect is probably captured in the age and education effects.)

The sum of the effect estimated for distance from transport in the purchase equations is -.464 shs/mile.

/ The effect was estimated on a categorical variable coded <1, 12,3 - 4, 5 - 8, 8+.

Harvests:

The role of harvests in the purchase equations must be given the same allocative interpretation as distance. (The impact of harvest on cash income is captured in the sales relationships.) A significantly positive effect amounting to 10.4 shs. for all purchases of household goods was not anticipated and remains to be adequately explained. It was hypothesized that the greater availability of foodstuffs in 2-crop areas would reduce purchases of meat and dairy products and other foods.

Variants

The structural model was retested including Distance from the local market, $\lambda(1)$, on the hypothesis that the market might be a more critical location for purchase of foods. In no case did distance to market appear significant when distance to transport was included, and the explained variance attributable to the former was less than distance to transport.

8. Earnings

At the outset, it was suggested that earnings may be simultaneosuly determined with purchases. That question has not yet been investigated. However, we felt it would be useful to investigate a reduced form for earnings in which access and demographic variables appear. Earnings include both cash and in kind components. .115 of total earnings was received in kind. The relationship estimated appears in Table 13.

Distance from district Hq. was an important influence on earned income, with every 5-mile increment in mileage reducing earnings by 51 shs. per year.

The impact of demographic structure on earnings was somewhat unexpected. It was hypothesized that the number of persons in the family over age 12 would significantly increase earnings, while the presence of young children under five years of age would inhibit earnings.

In fact the number of persons over 17 had no significant impact on earnings reported. A substantial increase was associated with the number of youths aged 13 - 17 as hypothesized. The effect of young children was to significantly enhance earnings contrary to hypothesis. The latter effect is not in some way associated with age of the head, which did not prove significant when included with other demographic variables. [1]

As expected, persons with primary education are employed to a greater extent than others. Durable ownership is slightly associated with increased earnings and the buildings index again shows a puzzling negative influence on earnings. All told, not much of the variance is explained with $R^2 = .104$.

9. Conclusions.

Table 14 summarizes the effect of access on proportion of crop sold. The effects are not large for sales of grain, but are relatively substantial for sales of meat and dairy products where both distance from a major center and distance from transport inhibit the proportion sold. If we consider the producer with a crop exactly equal to the mean, location at mile 10 from transport as opposed to mile 0.5 is associated with 70. shs less sales per year. This is about three per cent of total income.

A distance of 40 miles from district Hq. is associated, for this hypothetical farmer, with 34 shs. less sales of meat and dairy products than the farmer at 5 miles from Hq. In addition the earnings function predicts nearly 360 shs. less earnings for the farmer who is distant from Hq. Clearly the predicted effects of this latter measure of access have a major impact on income.

---

[1] E, γ_o'F, A(2), and distance to the nearest post office were included in the alternative specification and did not appear significantly. F(1) - F(4) were included only indirectly through γ_o'F.

The evidence in the purchase equations suggests a strong functional difference in the allocation decisions between cash income and income in the form of inventories retained for subsistence consumption. By implication, a large proportion of cash income is reserved for capital acquisition. Inventories do not influence purchases of market goods; their effect on purchases farm inputs is only a quarter that of the effect of food sales, despite the fact that both should be associated with the same derived demands for inputs.

Lastly, the model implies extensive and complex life cycle effects. Direct effects on allocation of purchases are coupled with powerful influences on the proportion of crops sold and the level of earnings.

This is clearly a preliminary analysis. The simultaneity between earnings and purchases needs to be explored. Functional form for both income and demographic effects should be tested. Disaggregation of purchase behaviour within the calendar year needs to be studied. Errors in variable models or income effects should be tested. Capital asset transactions should be investigated, and suitable methods for estimating the entire equation system, rather than single-equation techniques, should be applied. Finally an effort needs to be made to understand births, inclusion of adults, and separation of family members from the household as important endogenous responses of the unit to the economic and environmental forces discussed here.

Bibliographic Notes

Page 1. Interest in the problem of household allocation has resulted in several approaches to the stimation of the household utility function when allocation occurs according to a single budget constraint. See Brown & Heien (1972); Christensen, Jorgensen, and Lau (1975), and Phlips (1974). The hope in this study is to generalize to a larger number of constraints and to recognise the cost of converting inventory to cash. While work with the indirect utility function is undoubtedly appropriate here. the present investigation focusses more closely on income effects than substitution effects and is clearly only a beginning to the appropriate identification of the household preference function.

Page 6. This discussion was stimulated by Massell (1972).

Page 11. The descriptive data on this survey are reported in Kenya (forthcoming). An extremely useful background on smallholder agriculture appears in Heyer(forthcoming). In addition significant studies of interactions between producer and consumer allocations are reported in Kenya (1968) and Kenya (1969).

Specific analyses of Kisii farmers and their selling patterns were undertaken by Uchendu and Anthony (1975).

Nakajima (1969) attempts analysis of the allocation of small-holder labor that complements the theory developed here. Krishna's comments on that model insipred this more explicit modelling of the margin between subsistence and market production.

Page 15. The interpretation in (17) is provided by Brown and Deaton (1972) who have an excellent review of the demographic effects exploited by various students of consumer allocation. Despite the critique by Phlipps (1974, p) that Prais and Houthatker (1955) have an approach that is not readily integrated with economic theory, the extended linear expenditure systems of Lluch (1973) and Betercount (1973) can be seen as an extremely specific hypothesis that has not been adequately tested against more general roles for family composition in the utility function.

Page 17. Generality in demographic effects was achieved by Benus, Kemnta, and Shapiro (1975), but at the expense of deductive theory. The approach taken here imposes a certain subsistence theory on the role of demographic structure, and provides a logic for second order effects in estimation. The value of introducing more free parameters, as in Benus et al., should be tested here.

A more extensive setting for understanding demographic effects is provided by Mueller (1975).

Page 26. The simple linear approach taken in Section 6 needs to be revised to more comprehensive approximations as in Jorgenson (1975).

References

Amemiya , T. (1973) " The Equivalence of the Non-linear Weighted Least Squares Method and the Method of Scoring in Binary Regression Models " (Stanford University: mimeo)

Benus, J; Kmenta, J; and Shapiro H. (1975). "The Dynamics of Household Budget Allocations to Food Expenditures (mimeo: University of Michigan).

Betancourt, R. (1973) "Household Behavior in a Less Developed Country: an Economteric Analysis of Cross-Section Data" (mimeo: University of Maryland).

Brown, A and Deaton A. (1972) "Surveys in Applied Economics: Models of Consumer Behavior" Economic Journal 82 (1145 - 1236)

Brown, M. and Heien, D. (1972) "The S-branch Utility Tree: A Generalization of the Linear Expenditure System" Econometrica 40 (#4)

Christensen, L., Jorgeson, D. and Lau, L. (1975) "Transcendental Logarithmic Utility Functions" American Economic Review 65 (367-383); also Jorgenson below.

David, Martin (1961) Family Composition and Consumption (Amsterdam: N-Holland Publ.)

Gallant, A.R. (May, 1975) "Non-Linear Regression" American Statistician 29 (73-81)

Heyer, J. (1974) " A Survey of Agricultural Employment in the Small Farm Areas of Kenya since the 1920s" Working Paper 194 (Institute for Development Studies ,Nairobi University).

Hunt, D.M.(1975a) An Examination of the Distribution of Economic Status and Opportunity in Mbere, Eastern Kenya. Occasional Paper #11 (Institute for Development Studies, University of Nairobi) More detail on the asset index is included in the next item.

__________ (1975b)"Methodological Issues and Selected Findings of an Analysis of Wealth and Income in Mbere Division, Eastern Kenya." Working Paper 212 (Institute for Development Studies, University of Nairobi).

Jorgenson, D.W. and Lau, L (1975) "The Structure of Consumer Preferences" Annuals of Economic and Social Measurement 4 (49-102).

Kenya. Statistics Division, Ministry of Economic Planning and Development, (1968) Economic Survey of Central Province - 1963/64

__________ (1969) A Report on Economic Studies of Farming in Nyanza Province, 1963 (Farm Economic Report // 26)

Kenya. Central Bureau of Statistics (forthcoming) Rural Household Budget Survey - Nyanza Province, 1970/1971

Kmietowicz, T. and Webley P. (1975) "Statistical Analysis of Income Distribution in the Central Province of Kenya" (mimeo: University of Leeds).

Lenthold J. (1964) (Unpublished Ph.D dissertation, University of Wisconsin - Madison)

Lluch C. and Powell A. (1973) "International Comparisons of Expenditure and Saving Patterns (Development Research Centre, IBRD)

Massell, B.F. (1972) "Rural Household Expenditure in Kenya: Consistent Estimation of Expenditure Elasticities

in S.P. Schatz

South of the Sahara (Philadelphia Temple University Press) pp. 168 -184.

Mueller Eva (1975) "The Economic Value of Children in Peasant Agriculture " Conference on Population Policy (Washington D.C: Resources for the Future).

Nakijima, C. (1969) "Subsistence and Commercial Family Farms: Some Theoretical Models of Subjective Equilibrium also "Comment" by Krishna, R. in

Clifford Wharton Subsistence Agriculture and Economic Development (Chicago: Aldine
165 - 189.

Nerlove, M. and Press J. (1973) Univariate and Multi-Variate Log-linear and Logistic Models
R-1306 EDA/NIH (Santa Monica: RAND 90406)

Phlips, L. Applied Consumption Analysis (N. Holland, 1974)

Prais S. and Houthakker (1955) Analysis of Family Budgets (Cambridge University Press)

Uchendu, V.C. and Authory (1975) Agricultural Change in Kisii District Kenya (Nairobi: East African Literature Bureau)

Walker, S. and Duncan D. (1967) "Estimation of the Probability of an Event as a Function of Several Variables" Biometrika 54 (167-79).

Table 1

Marketing behavior of producers of food crops.

Food crop	Producers reporting	
	Purchase and Sale	Subsistence only
Cereals	.840	.016
Pulses	.230	.206
Roots	.446	.134
Vegetables	.556	.026
Fruit	.417	.083
Fish	.215	.004
Dairy	.462	.066
Meat	.755	.000
Sugar	.287	.000
Nuts	.124	.012

TABLE 2

AVERAGE CONSUMPTION EXPENDITURE PER ADULT MALE EQUIVALENT

ITEMS	2 OR LESS	3	4	5	6	7	8	9	10 OR MORE
FOOD: VEGETABLE PRODUCTS									
CEREALS	174.2	142.33	113.48	100.49	93.46	96.16	76.88	79.39	51.37
PULSES	22.62	9.55	7.77	8.11	6.41	4.98	3.60	5.15	1.95
ROOTS & TUBERS	189.63	146.35	111.83	101.98	77.50	70.66	72.45	43.52	29.66
OTHER VEGETABLES	42.94	30.29	22.13	23.34	19.19	16.56	16.75	9.91	5.77
FRUIT	12.27	10.60	15.68	9.40	13.16	14.14	12.40	8.15	2.17
SUGAR	12.30	11.05	7.93	7.52	6.33	12.84	9.20	9.49	7.35
NUTS	0.23	0.60	0.23	0.26	0.20	1.27	0.14	0.53	0.05
MEALS BOUGHT	0.322	0.21	0.02	0.14	0.07	0.21	0.14	0.07	0.17
ANIMAL PRODUCTS									
FISH	43.46	20.21	14.93	12.25	16.05	10.90	10.72	16.11	8.33
DAIRY PRODUCTS	15.45	8.85	7.02	6.96	6.82	6.18	8.23	6.13	3.28
MEAT	75.87	47.50	41.40	29.98	30.93	37.75	32.93	30.67	21.03
OTHER FOOD	34.59	20.59	27.77	26.36	38.28	17.88	17.50	11.42	18.37
EXPENDITURE OTHER THAN FOOD									
DRINK/TOBACCO	15.02	7.9	5.55	6.47	5.33	4.78	4.82	3.32	3.68
SOAP	7.92	6.29	5.0	3.93	3.97	5.24	4.07	5.24	3.46
FUEL	9.28	5.73	4.28	3.54	3.16	3.46	2.89	3.63	2.73
HOUSEHOLD EQUIPMENT	8.78	5.72	6.11	4.64	4.12	3.67	4.03	3.79	2.25
PERSONAL CARE & HEALTH	2.94	5.52	2.65	3.52	2.19	3.21	3.31	4.88	2.83
TRANSPORT	5.08	5.58	4.11	3.92	3.13	6.35	4.32	7.71	6.92
SCHOOL FEES	3.08	6.52	6.55	9.40	9.47	9.79	12.82	19.16	11.57
CLOTHING	17.64	18.23	15.52	13.11	16.76	14.09	15.42	17.19	11.31
RECREATION	0.39	0.71	0.3	1.1	0.61	0.68	1.77	2.27	0.28
LEGAL & OTHER	7.69	3.42	1.97	1.65	1.91	2.65	1.12	1.69	2.96

Table 3

Average value of crop produced
per household and

Average value of crop produced
per producer

Nyanza Province 1970/1971

	Average value of crop		Share of total value
	Per Producer (Kshs per annum)	Per Household (Kshs per annum)	
Cereals	454	452	.265
Pulses	49	47	.028
Roots	585	533	.313
Vegetables	106	106	.062
Fruit	115	86	.050
Sugar	90	22	.013
Nuts	54	18	.011
Fish	81	20	.012
Dairy	50	44	.026
Meat	263	250	.147
Other	358	125	.073
Total	——	1703	1.000

Table 4

Simple correlation of selected variables to proportion of total food crop sold by type of crop - Nyanza Province 1970/71

Independent variable	Type of Crop: Cereals, roots, and Vegetables	Pulses	Fruit	Sugar	Nuts	Fish	Meat and Dairy products
Total income, Y(1)	-.044	+ .024	- .122 **	- .146	- .069	+ .067	- .053
Calorie, Y_o'F	-.028	+ .038	- .052	+ .010	- .066	+ .133	- .073
Δ	+.042	- .014	+ .010	- .035	+ .074	- .017	+ .051
Adults, F(7)+F(8)	+.012	+ .108 **	+ .021	- .046	- .032	+ .124	+ .058
Youths, F(5) + F(6)	+.038	+ .073	- .078	+ .011	- .036	- .022	- .026
Distance to: District Hq, A(6)	-.041	+ .079 *	- .040	- .053	- .034	- .221 **	- .135 ***
Market, A(1)	-.023	- .076 *	- .103 *	+ .028	- .037	+ .098	- .064
Transport, A(2)	-.103 **	- .068	- .060	- .029	- .096		
Buildings, G	+.103 **	+ .089 *	+ .087	- .098	- .068	- .092	+ .059
Durables, K	+.183 ***	+ .120 **	+ .024	- .145	- .167 *	+ .055	+ .003
Number of harvests, H	+.233 ***	+ .226 ***	+ .318 ***	+ .082	- .024	+ .204 **	+ .158 ***
Purchases, $P_1(3)+P_2(3)$	+.114 **	+ .006	- .003	- .070	+ .057	- .089	+ .021
Number of producers	704	686	525	172	228	177	692

*** Significantly different than zero at the .001 level.
** Significantly different than zero at the .01 level.
* Significantly different than zero at the .05 level.

Table 5

Estimated logit functions describing proportion of crop sold

Equation ($\bar{y}$, σy)	Endogenous variables		Exogenous Variables							R^2 (F)	N
	TOTINC1 Y(1)–	EPUR §§ $P_2(3)+P_2'(3)$	DIST A(6)	B27 A(2)	HARVEST H	ADULTS F(7)+F(8)	BUILD G	DURABLE K	CONSTANT		
Grain, roots, and vegetables (.103, .118)	—	+.0177 (3.62) E	—	-.0351 (2.27) E	+.513 (5.56) /	—	+.0867 (2.59) N	— *	-3.462	.076 (14.4)	704
Meat and dairy products (.556, .374)	—		-.0789 (2.87) E	-.0678 (3.48) E	+.425 (3.54) E	.0319 (.90) E	—	— *	.0858	.053 (9.92)	692
Pulses (.151, .250	—	—	+.091 (2.43) N	-.0953 (1.99) E	+.897 (5.56) /	*	*	*	-3.579	.054 (13.1)	686
Fish (.244, .394)	—	—	-.148 (1.79)	—	+.685 (1.76) /	—	—	—	-1.655	.059 (5.4)	177
Fruit (.310, .345)	-.00015 (3.75) N			B27 -.158 (4.14) E	+1.059 (7.30) /		+.127 (2.38) N		-1.930	.145 (22.1)	527

E signifies an expected effect; N, an effect contrary to hypothesis.

* Significant zero-order correlation, not included in regression

§§ The quantities were valued at market prices.

Table 6

Purchase Equations : Structural and Modified Structural Equations Compared
Nyanza Province: 1970/1971

Equation[a] ($\bar{y}i$, σ)	INCCAL3[a] E+R+S	INVD11[a] $Q_1(1)-S_1(1)$	INVD81[a] $Q_2(1)-S_2(1)$	INVOTHER[a] $Q_3(1)-S_3(1)$	ESALES τ(1)'S(1)	B27 A(2)	HARVEST H	CALORIE γ_o'F	VCAL Δ	C4 Age of head	C6 Primary Education of head=1	BUILD G	DURABLES K	Constant	R^2 (F)	Estimation method (N)
	Endogenous Variables					Exogenous Variables										
Purchases: grain, roots, vegetables, $P_1(1)$ A. Structural	+.0292 (6.00) E	-	+.0398 (1.71) /	-.00178 (.13) /	-	-.292 (2.19) E	4.247 (5.24) N	-.194 (3.02) E	+.117 (.77) E	-.0344 (1.13) E		-.743 (2.24) E	-	8.21	.159 (14.6)	2SLS (705)
B. Structural modified (10.37, 11.20)	+.0309 (6.46) E	-.0332 (5.31) N	+.0549 (2.40) /	+.00093 (.07) /	-	-.286 (2.18) E	3.501 (4.34) N	-.279 (4.28) E	.347 (2.24) E	-.0554 (1.84) E		-.803 (2.47) E	-	13.32	.191 (16.4)	2SLS (705)
Purchases: meat and dairy, $P_2(1)$ A. Structural	.0385 (8.55) E	.00004 (.00) /	-	.0171 (1.37) /	-	-.300 (2.39) E	.580 (.74) E	-.280 (4.41) E	1.008 (6.75) E	-.0610 (2.01) E	2.622 (2.39) E	-.444 (1.38) /	+.114 (4.29) E	9.46	.292 (28.0	2SLS (705)
B. Structural modified (10.19, 11.55)	.0383 (8.23) E	-.00009 (.00) /	.00385 (.17)	.0171 (1.37) /	-	-.304 (2.38) E	.588 (.75) E	-.280 (4.39) E	1.009 (6.74) E	-.0613 (2.01) E	2.620 (2.39) E	-.448 (1.38) /	+.114 (4.28) E	9.47	.292 (23.8	2SLS (705)
Purchases: other food $P_3(1)$ A. Structural	.0470 (9.82) E	-.0113 (1.80) /	-.00075 (.03) /	-	-	-.254 (1.92) E	+5.476 (6.83) N	-.369 (5.59) E	+.530 (3.42) E	-.0165 (.52) E	3.974 (3.50) E	-.165 (.49) /	.0968 (3.52) /	4.74	.309 (28.1	2SLS (705)
B. Structural modified (12.02, 12.12)	.0496 (10.34) E	-.0104 (1.67) /	-.0025 (.11) /	-.0474 (3.70) N	-	-.254 (1.94) E	+4.959 (6.15) N	-.372 (5.68) E	+.549 (3.57) E	-.202 (.64) E	3.772 (3.35) E	-.170 (.51) /	-	6.30	.322 (27.4	2SLS (705)
Purchases: household goods, P.(3) A. Structural		-	-	-												
B. Modified Structural (20.50, 20.79)	.0488 (5.56 E		+.00441 (.42) E		.00875 (3.12) /	+.381 1.62 N	.0988 (.02)	-.446 (3.58) E	+.995 (3.54) E	-.271 (4.85) E	+8.140 (3.96) E	- -	.125 (2.56) E	24.394	.228 (22.8)	2SLS (705)
Purchases: School fees, $P_2(3)$ (17.4, 47.9)	- N	-	-	-	-	-	-	-	-	-	C7[b] 11.67 (2.13)	-	.279 (2.27) E	11.011	.019 (6.55)	OLS (586)

[a] These are values per equivalent adult. The quantity symbols shown by each variable were in fact multiplied by appropriate prices and divided by γ_o'F.
[b] Education of wife. E signifies an expected effect; N, an effect contrary to hypothesis.

Table 7

Expenditure Equations : Full income and Modified Full income Model

Equation * ($\bar{y}i$, σy)	Endogenous Variables			Exogenous Variables									Cons-tant	R^2 (F)	Estima-tion method (N)
	INCCAL5* Y(1)	INCCAL2* E+R+S(2)		IN'EN* Q(1)	B27 A(2)	HARVEST H	CALORIE γ_δ'F	VCAL Δ	C4 Age of head	C6 Education	BUILD G	DURABLE K			
Expenditures:	.0512	--		-.0222	+.375	--	-.246	.887	-.263	8.712	-	.158	21.718	.212	OLS
household goods, C_1(3)	(5.70)			(1.86)	(1.58)		(2.22)	(3.22)	(4.70)	(4.23)		(3.25)		(23.4)	(7.05)
D. Modified Full income (20.50,20.79)															

* See note to Table 6

Table 8

Propensity to consume derived from structural purchase equations

Type of purchase	Propensity to consume out of: Gross cash receipts	Retained inventories	Gross income
Household items	.164	1.000	.667
Farm inputs	.142	.000	.056
Total	.305	1.000	.724

Table 9

Elasticities of expenditure

	This Study: Purchases out of cash per equivalent adult	This Study: Expenditures out of total income per equivalent adult	Masseli (1972) ≠ Expenditure out of total income
Cereals			.59
Roots	.219 (Cereals, Roots, Vegetables)	2.09 (Cereals, Roots, Vegetables)	.53
Vegetables			.79 a/
Meat	.294 (Meat, Dairy)	9.75 (Meat, Dairy)	1.20
Dairy			2.34
Sugar			1.06
Pulses	.304 (Sugar, Pulses, Fats - Oils)	6.446 (Sugar, Pulses, Fats - Oils)	.79
Fats - Oils			1.22
Housing			2.34
Tobacco - beverages			1.12
Misc. non-durables			1.09
Services	.185 (Housing to Fuel)	.483 (Housing to Fuel)	1.08
Durables			.99
Clothing			.89
Fuel			.69
Education	-		1.10 b/
Farm inputs	.981	-	-

≠ Elasticity in relation to total expenditure

a/ includes fruit also

b/ includes recreation also.

Table 10

Purchase of goods for farm inputs, including labour, $P_3(3)$

Equation	TOTINCl Y(1)	A14 τ(2)'S(2)	EINVXCAL τ(1)'(Q(1)-S))	ESALES τ(1)'S(1)	A9 E	A2 R2	A1 R1	CALORIE Y_0'F	DURABLE K	BUILD G	Constant	R^2 (F)	Estimation method
B. Modified Structural model	-	.241 (4.72)	.0387 (2.96)	.130 (3.35)	0224 (2.26)	.242 (3.12)	.150 (7.07)	2.727 (1.63)	1.860 (2.76)	-16.741 (2.07)	-48.166	.269 (28.4)	2SLS
C. Full income	.0326 (5.31)	-	-	-	-	-	-	8.252 (4.82)	2.770 (3.92)	-19.730 (2.30)	-54.696	.147 (30.1)	OLS
D. Full income modified	-	.242 (4.74)	.0567 (6.07)		.0224 (2.26)	.255 (3.29)	.164 (8.21)	2.740 (1.64)	2.033 (3.04)	-15.084 (1.87)	-56.566	.265 (31.3)	OLS
A. Structural model	-	.282 (5.66)	-	.203 (6.14)	0262 (2.67)	.236 (3.03)	.143 (7.05)	-	1.795 (2.66)	-10.897 (1.47)	.695	.255 (34.1)	2SLS

Table 11

Test of significance of explanation due to structure

Equation	Version of model: Naive	Modified
Household goods		
Explained sum of squares (d.f)		
Full income	–	64382 (8)
Structural	–	69447 (9)
Difference	–	5065 (1)
Error variance	–	337.7
F		15.0***
Farm inputs		
Explained sum of squares (d.f.)		
Full income	8636 x 10^3 (4)	15,625 x 10^3 (8)
Structural	–	15,860 x 10^3 (9)
Difference	–	235 x 10^3 (1)
Error variance	–	61,954
F	–	3.79*

* Significant at the P = .10 level

*** Significant at the P = .005 level.

Table 12

Effect of Life Cycle in a Hypothetical Family on CALORIE, Δ, and Purchases of Grain, Roots, and Vegetables PEA

Age of head	Life cycle Event	Size of family	CALORIE	Δ	Effect of CALORIE and Δ on Grain purchases PEA
18	Marriage	2	4.70	.10	-.90
20	First child	3	6.60	.14	-1.26
22	2nd child	4	7.70	.47	-1.44
24	3rd child	5	8.80	.46	-1.65
26	4th child	6	10.38	.32	-1.98
28	5th child	7	11.96	.23	-2.29
30		7	12.21	.25	-2.34
32		7	13.26	.18	-2.55
34		7	14.66	.10	-2.83
36		7	15.96	.10	-3.08
38	First child leaves home	6	14.38	.06	-2.78
40	Second child leaves home	5	12.22	.07	-2.36
42	Third child leaves home	4	10.08	.08	-1.95
44	Fourth child leaves home	3	7.58	.15	-1.45
46	Fifth child leaves home	2	4.70	.10	-.90
48		2	4.70	.10	-.90
50	Spouse dies	1	2.20	.00	-.48
52 etc		1	2.20	.00	-.48

Table 13

Determinants of Earnings (reduced form)

Variable	Regression Coefficient (t - ratio)
Distance to district Hq *, A(6)	- 51.0 (3.11)
Number of:	
Adults, F(7) + F(8)	+ 19.7 (.75)
Youths, F(5) + F(6)	+ 111.6 (2.39)
Older children, F(3) + F(4)	15.1 (.64)
Young children, F(1) + F(2)	163.8 (5.28)
Primary education of head (=1)	158.7 (2.78)
Durables	7.0 (2.78)
Buildings	- 69.8 (2.20)
Constant	196.1
R^2 (F)	.104 (10.1)
Estimation method	OLS
($\bar{y}$, σ)	(292, 982)
N	705.

Table 14

Expected rate of sales by distance from transport

Distance from transport	Grain, roots and vegetables	Meat and dairy products (DIST = 4.26)
10 miles	.080	.448
9	.082	.465
8	.085	.482
7	.088	.499
6	.090	.516
5	.093	.533
4	.096	.549
3	.099	.566
2	.103	.583
1	.106	.599
0.5	.108	.607
Mean value of crop/household	1091.	294.
Difference in sales rate mile 10 and mile .5	.028	.159

Distance from Hq.	(B27-3.322)
40	.510
35	.527
30	.544
25	.560
20	.577
15	.593
10	.610
5	.626
Mean value of crop	294.
Difference in sales rate mile 40 and mile 5	.116

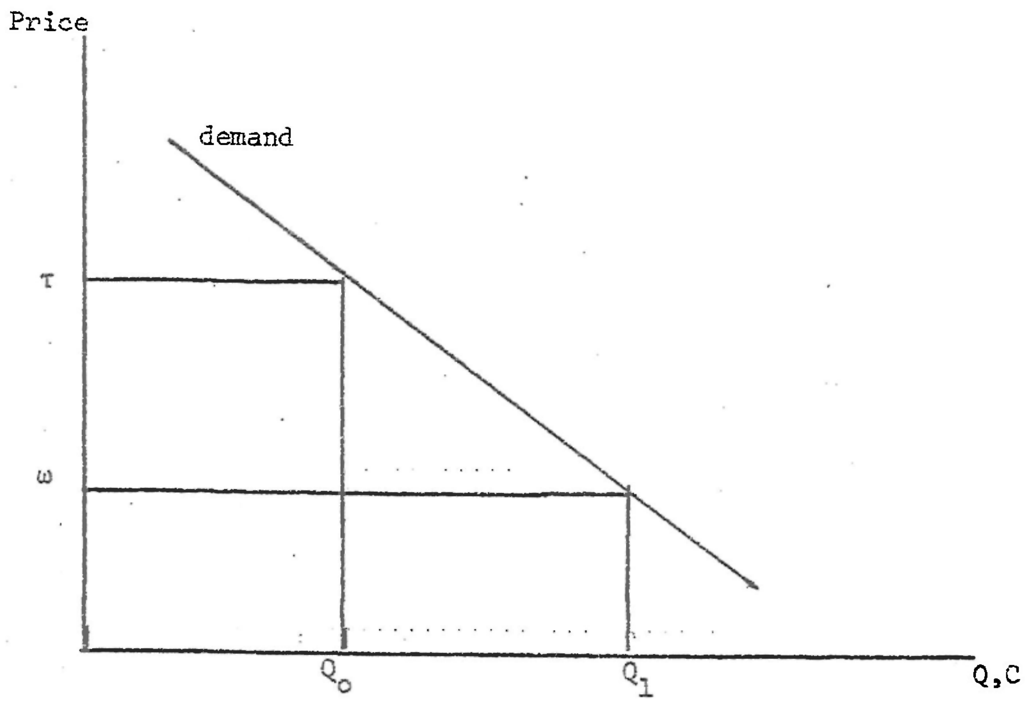

Figure 1

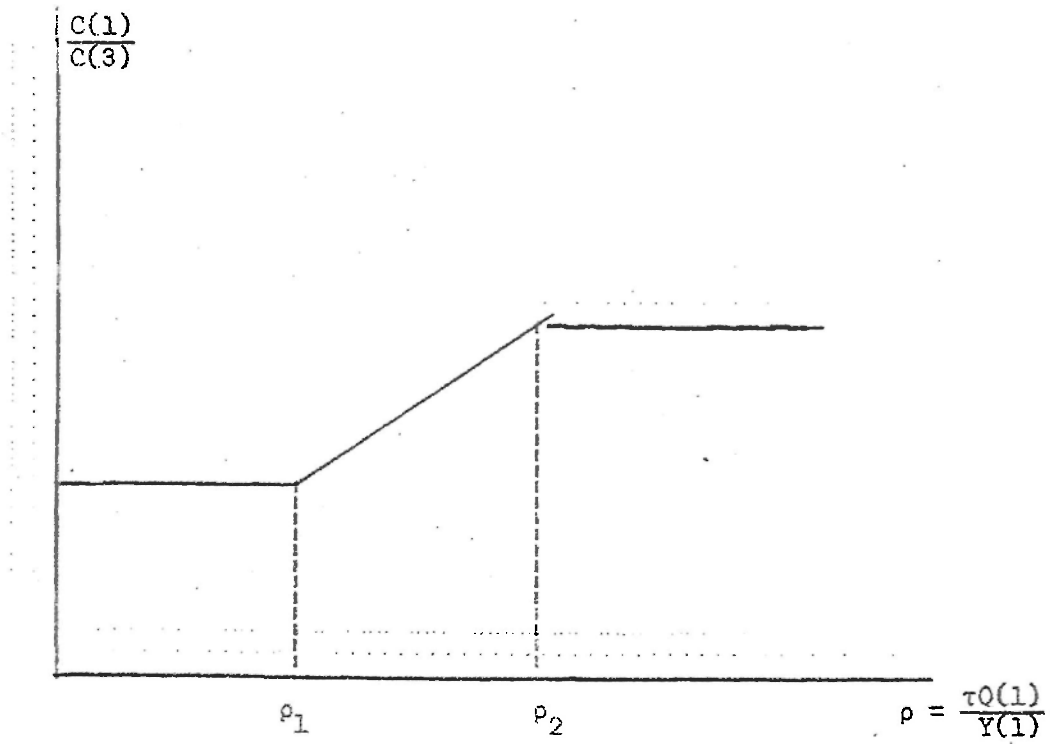

Figure 2

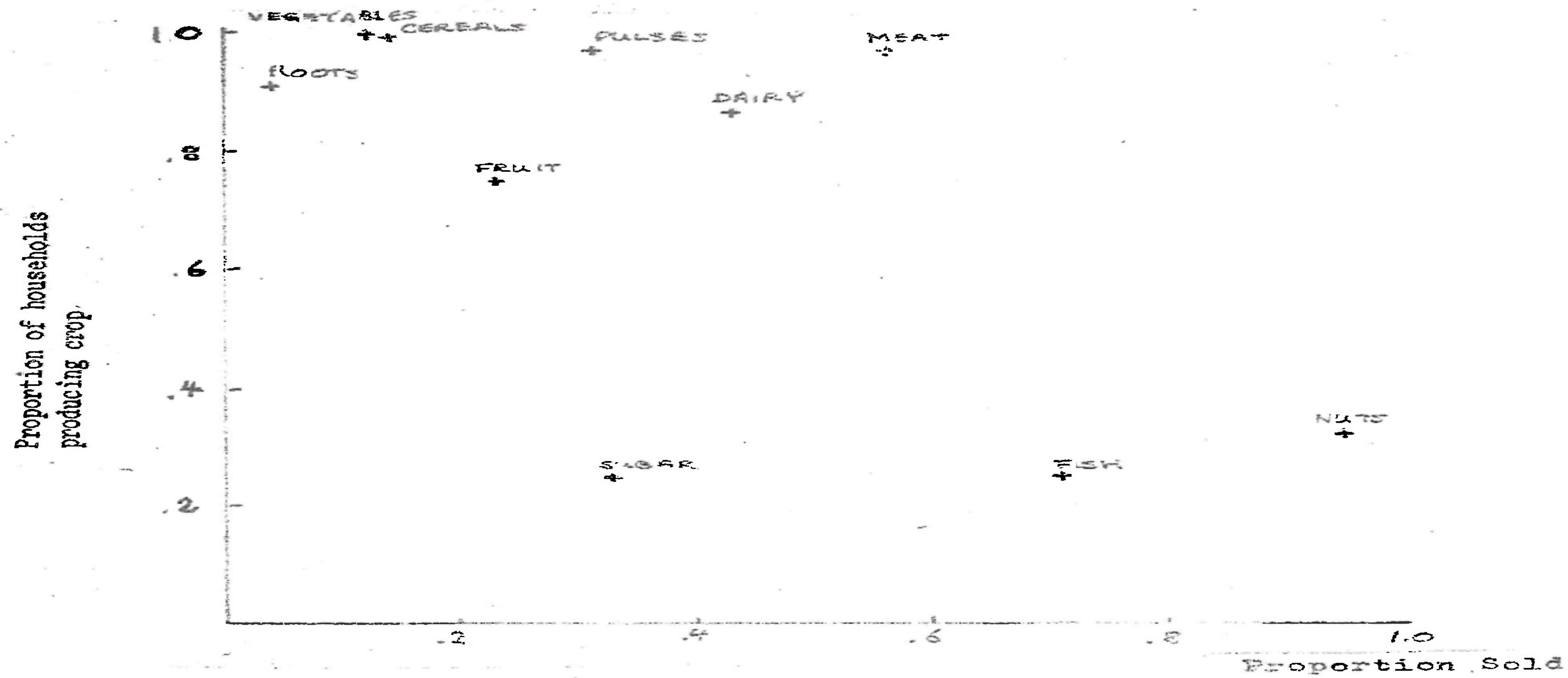

Figure 3. Prevalence of food crops and proportion of crop sold – Nyanza Province 1970/71

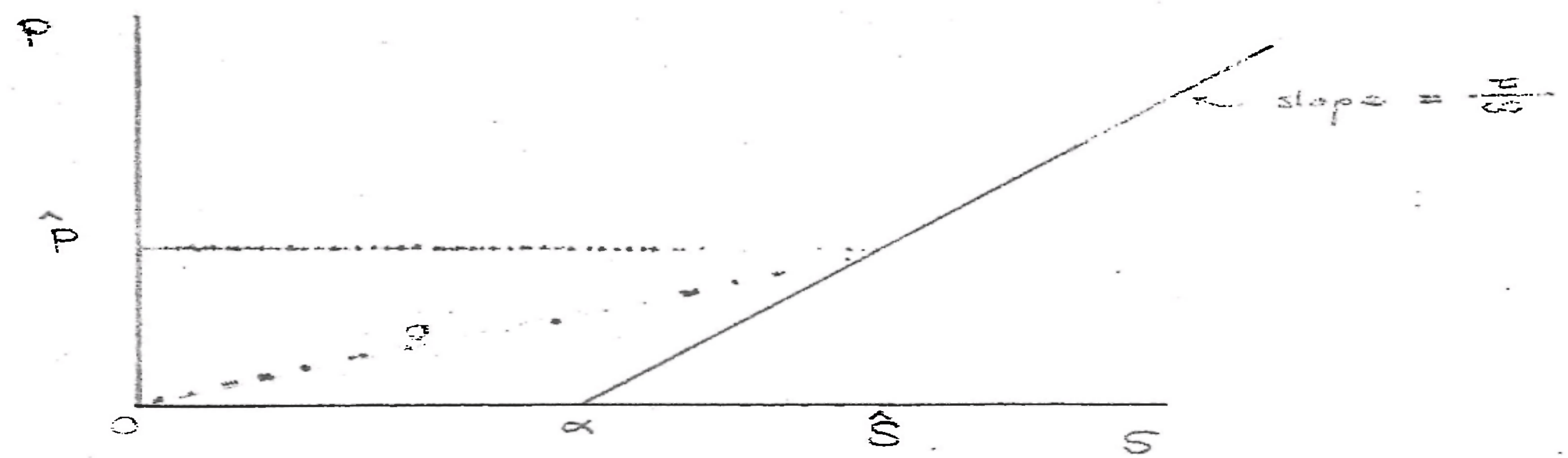

Figure 4