INVESTIGATIONS INTO SEVERAL ECONOMIC AND DEMOGRAPHIC DECISIONS IN RURAL KENYAN HOUSEHOLDS: A RESEARCH PROPOSAL

By

Caroline Swartz

WORKING PAPER NO. 301

INSTITUTE FOR DEVELOPMENT STUDIES
UNIVERSITY OF NAIROBI
P.O. BOX 30197
NAIROBI, KENYA.

March 1977

IDS/WP 301

INVESTIGATION INTO SEVERAL ECONOMIC AND DEMOGRAPHIC DECISIONS IN RURAL KENYAN HOUSEHOLDS: A RESEARCH PROPOSAL

by

Caroline Swartz

ABSTRACT

In view of the extreme importance of understanding the relationships between economic and demographic behaviour in rural Kenya, this paper outlines a proposed econometric research undertaken, utilizing the data from the Kenya Government's integrated rural survey. The proposed analysis focuses on the income, saving and demographic decisions made at the household level and the interactions between these decisions and the roles children play in the household. Parents are assumed to allocate their children's time between education and work in the household enterprise, with each of the alternatives having different implications for household behavior. Additional research into the determinants of each child's main activity, the determinants of woman's fertility experience, and the expenditure behavior of the household is described. Finally, a non-technical description of the empirical methodology to be employed is presented.

1.0 INTRODUCTION [1]

The role of rural areas in economic development is receiving increased attention. A large share of the population reside and work in the rural areas. Migration into urban centers is in part a consequence of conditions in the rural areas. Overall high rates of population growth are maintained or aggravated by high fertility among rural residents. Increased pressure on rural resources and declining levels of human welfare are already presenting severe problems in some areas. The rural and agricultural sector nevertheless has considerable potential to accelerate the development process. In providing productive employment, expanding food and raw material supplies and in generating foreign exchange, for example, the rural sector has vital interactions with other sectors in the process of successful growth. In Kenya, the importance of the rural sector has been recognized as evidenced by emphasis on rural and agricultural development. Considerable resources are being expended on agricultural extension and related work, transportation planning and a series of government projects designed to increase the infrastructure and productivity in the rural areas.

This research examines the rural areas on a microeconomic level, using the household and the individual as the unit of analysis to investigate the nature of the economic and demographic decision making process. The main issue of the study is an examination of farm production relationships and particularly the role of children and its impact on household production, saving and composition. Since education is voluntary, the household may a allocate its children's time between work in the household enterprise, work off ~~the holding~~, education and leisure. The relative importance of these roles has implications for the household's behavior. Essentially, children who work directly augment household income, thus decreasing the cost of children and perhaps increasing their numbers in the household. On the other hand, children in school are essentially non-productive in the current period, making them relatively expensive and therefore, decreasing the number of children

---

1. The author is grateful to Dr. Peter Hopcraft for his comments on an earlier draft of this paper. Financial support for this research is provided by a traineeship in demography at Duke University from the National Institute of Health in the Federal Government of the United States and the Ford Foundation.

in the household.

The investigation of the impact of children is carried on in several ways. First, a simultaneous equation model of the household is formulated to highlight the interactive nature of household decisions concerning income, saving and household size. This approach highlights both the determination of children's roles and their consequences. The model, however, has the disadvantage of being unable to examine decisions concerning the individual since it employs the household as the unit of analysis. Two sets of decisions which employ the individual as the unit of analysis are examined outside the model. First, the determination of the child's main activity is investigated. The child is assumed to either attend school or work on the holding, leading to an investigation of the various individual - and household-specific - attributes which influence this decision. The second individual specific investigation is, in a sense, consideration of the consequences of the role of children. The fertility experience of each women is in part a consequence of the roles her children play and the costs and benefits associated with these roles. The final subject of analysis proposed is an investigation of household expenditures on various categories such as food and clothing, in relation to total household expenditures and the demographic characteristics of the household. This will enable the measurement of the income elasticity of demand in several consumption categories and also the investigation of economies of scale in household consumption as a function of household size and composition. To the extent that scale economies exist, the cost of child rearing are decreased and therefore, the household is expected to have relatively more children.

The remainder of this paper describes the proposed research in detail, motivating and specifying the equations to be estimated, indicating expected results and identifying potential policy implications. The final section is a non-technical description of the estimation techniques to be employed.

2.0 A MODEL OF THREE SETS OF HOUSEHOLD DECISIONS

The research described here considers the household decisions relating to income, saving and household size. The household determines the level of income generated on the holding by combining its endowment of physical assets

such as land and tools, with labor. Thus, the labor input, the technology employed and the level of output and income are all simultaneously determined. Since the role of children in farm production is a major determinant of household decisions in the area of family size and education, the composition of the family labor input receives particular attention.

Three aspects of household saving behaviour are considered: the level of total saving and the household's investment in physical and human capital. Saving in the form of financial assets is assumed to be negligible and therefore is not investigated. The level of savings is related to the level of household income, the stock of physical assets and the size and composition of the household. The question of the effect of additional children on household savings is considered in detail. If children are a net drain on savings, continued high rates of population growth may mean a decline in the rates of rural saving and investment. Alternatively, children may cause a redistribution of consumption within the household, making their impact on saving negligible. The distribution of household savings between physical and human capital depends on the expected rates of return, the timing of the returns and the household's rate of time preference. The actual distribution question is approached indirectly via the investigation of the level of investment in physical and human capital as separate variables.

The demographic decisions investigated deal with several components of household size. Children are divided by age and sex into five categories in order to highlight their age- and sex-specific impacts in other sections of the model. Adults are divided into two age groups, 15-19 and 20 or older, to allow for possible school attendance by the younger group. The number of individual in each category is related to the household's demand for labor as represented by its stock of physical assets. The supply of older children is included in equations for younger children on the hypothesis that it decreases the demand for younger children.

The model presupposes that the decisions made by the household are interrelated. Income depends on household size and composition, but it is also a determinant of household size. Savings is a function of income and household size while the allocation of savings between human and physical capital influences the children's role in the household. Children's roles

in turn have an impact on the levels of income and saving.

Each of the three sets of decisions are discussed in detail below. The equations are specified and the expected signs for the coefficients are indicated and discussed. Finally, some of the implications which can result from the analysis of the model are identified.

2.1 Income

This section of the model examines the level of household income, focusing on income earned on the holding as the major component of income. This emphasis on production in the household enterprise entails an analysis of the technology and input mix employed. One aspect of the technology decision, the supply of labor from the household members to the household enterprise, is examined in detail to determine the nature of the labor supply decision and its effect on the roles of adults and children in the household.

The one definition and four behavioral equations of this section, represented in functional form are*:

(1) $Y = Y_{HH} + Y_O$

(2) $Y_{HH} = f_2$ ($LAND_F^+$, $LAND_O^+$, $LABOR_F^+$, $LABOR_H^+$, $TOOLS^+$, $BUILD^+$, $LSTOCK^+$, $OTHER^+$, AGE^+, $EDUC^+$)

(3) $LABOR_F = f_3$ ($ADULTS_W^+$, $CHILD_W^+$, $KIDS0\text{-}4^?$, $LAND_F^+$, $LAND_O^+$, $TOOLS^+$, $BUILD^+$, $LSTOCK^+$, $OTHER^+$, $WAGE^+$)

(4) $ADULTS_W = f_4$ ($ADULTS\ 15\text{-}19^+$, $ADULTS20^+$, $KIDS0\text{-}4^?$, $LAND_P^+$ $LAND_O^+$, $TOOLS^+$, $BUILD^+$, $LSTOCK^+$ $OTHER^+$, $WAGE^+$, $CHILD_W^?$)

(5) $CHILD_W = f_5$ ($BOYS5\text{-}9^+$, $BOYS10\text{-}14^+$, $GIRLS5\text{-}9^+$, $GIRLS10\text{-}14^+$, $KIDS0\text{-}4^?$, $LAND_P^+$, $LAND_O^+$, $TOOLS^+$, $BUILD^+$, $LSTOCK^+$, $OTHER^+$, $WAGE^+$, $ADULTS_W^?$, $FEES^-$ $DIST_{FS}^+$)

* The superscript signs indicate the sign expectations for the estimated coefficient of the variables ("?" indicates an uncertain sign expectation). Endogenous variables on the right hand side of each equation are underlined.

where

Y = total household income;

Y_{HH} = total value of production in the household enterprise;

Y_O = income from all other sources;

$LAND_P$ = area of land on the holding planted in perennial crops;

$LAND_O$ = area of land on the holding available for other (non-perennial) crops;

$LABOR_F$ = man-day equivalents of labor supplied to the household enterprise by family members;

$LABOR_H$ = man-day equivalents of labor supplied to the household enterprise by hired workers;

TOOLS = value of all tools and equipment owned by the household;

BUILD = value of all buildings on the holding;

ISTOCK = value of livestock owned by the household;

OTHER = cost of all purchased inputs for the household enterprise other than labor plus an imputed value for miscellaneous inputs supplied by the household;

AGE = age of the head of household;

EDUC = formal education of the head of household;

$ADULTS_W$ = number of adult family members employed in the household enterprise;

$CHILD_W$ = number of children family members employed in the household enterprise;

WAGE = prevailing wage for adult agricultural workers in the area;

KIS0-4 = number of children under age 5 in the household;

ADULTS15-19 = number of adults ages 15 - 19 in the household;

ADULTS20 = number of adults age 20 or older in the household;

BOYS 5-9 = number of boys ages 5 - 9 in the household;

BOYS10-14 = number of girls ages 10 - 14 in the household;

GIRLS5-9 = number of girls ages 5 - 9 in the household;

GIRLS10-14 = number of girls ages 10 - 14 in the household;

FEES = household expenditures on children's school fees;

$DIST_{PS}$ = distance from the holding to the nearest government primary school.

Total household income is used as the primary income measure because of its comparative ease of construction from cross-sectional data and because of its likely correlation with other income or welfare measures. Only income from the household enterprise is considered as endogenously determined within the model; other income, Y_0, is taken as exogenous. Endogenous consideration of the components of Y_0, such as income earned outside the household enterprise and the value of remittances received, would enrich the model, but it would also enlarge it and make estimation and interpretation more difficult.

A Cobb-Douglad production function describes the total revenue or aggregate value of output for the household enterpise. Any increment in one of the inputs, ceteris paribus, is expected to increase output, thus leading to positive coefficients on the input variables. Land planted in perennial crops is considered separately from other land since perennial crops represent a substantial investment in the land. Aggregation of the two variables would be possible if prices which accurately reflected the value of the land and the perennial crops planted on it were available. Such prices are not readily available so that disaggregation in this manner is necessary to avoid problems of misspecification.

The use of two labor variables enables further analysis of the household's labor-supply decisions. The man-day equivalents measure in both cases is derived by assuming children are half as productive as adults. While the relative productivity of adults and children would fruifully be subjected to empirical investigation, the data are unfortunately only available in man-day equivalents.

The valuation of tools, buildings and livestock poses a difficult problem since it is the flow of services from these assets which enters the production process but it is the value of the stock which is known. Ideally the rental values for the services would be employed in the production function, but reliable rental prices are not available. Consequently, these three separate asset categories are included in the production function to allow the ratio between the value of the stock and the flow of services to vary between categories. Valuations of buildings is particularly problematic since buildings represent a mixture of producer and consumer durables. The total value of all buildings on the holding is arbitrarily selected for use in the model.

The final physical factor, the cost of other inputs, represents the value of those inputs not otherwise included in the production function. It includes the value of inputs purchased by the household other than labor plus an imputed value for miscellaneous inputs owned by the household such as seed saved from previous crops. If the relative importance of different inputs varies markedly between households, this could be a source of misspecification error.

Two human capital variables, the age and education of the head of the household, are included in the production function. These variables help ensure the comparability of results across households by controlling for characteristics which may influence the productivity of labor or the efficiency of the household enterprise.

Thirteen binary variables based on zonal divisions formulated by the Central Bureau of Statistics as part of their survey operations are also included in the production function. These variables are used to account for differences in production as a result of ecological or cultural location-specific characteristics. The assumption of factor homogeneity across Kenya is thus avoided.

The three equations describing the family labor input which complete this section of the model will be discussed together because of their similarity. In the first equation, the total labor supplied by family members is expected to be directly related to the number of family adults and children working on the holding. In the second labor equation, the number of family adults working on the holding is expected to be directly related to the number of adults in the household. The adults are divided into two age groups, 15 - 19 and 20 and older, for reasons discussed in the savings section below. For the third equation, potential child workers (those aged 5 - 14) are divided into four age-sex groups to enable analysis of the relationship between these characteristics and the children's role in production. The number of family children working on the holding is expected to be directly related to the size of each of the four age-sex groups.

The number of children under age 5 in the household is included in each of the three labor equations but expectations with regard to the signs of the coefficients are uncertain. Young children may lead to a decrease in

the labor input and the number of workers if their care requirements necessitate the withdrawal of one or more workers. Alternatively, child care and employment on the holding may be compatible activities so that young children may have no effect on the quantity of household labor supplied. Finally, the family members may increase their labor input in order to pay for the consumption of the young and perhaps to help increase saving for future large expenditures on the young.

The non-labor inputs to the production process, land planted in perennial crops, other land and the value of tools, buildings, livestock and other inputs, are included in the labor equations. If these factors are complements to labor in the production process, they can be expected to have positive coefficients.

Finally, in all three equations, the local wage for adult agricultural labor is included to represent relative labor scarcity and the price of substitute labor. If wages are relatively low, the household can be expected to hire a larger share of its labor and decrease the family labor input. The wage inevitably contains an element of the attractiveness of off-holding employment, but the majority of small holders are likely to have few better alternatives to working their own holdings. They are likely to seek outside employment only when the returns to additional labor on the family holding compare unfavorably with alternative earning opportunities.

The equation for the number of adult (child) family workers includes the number of family children (adults) working on the holding. In general, negative coefficients are expected since adult and child labor are to some extent substitutes. Reservations, however, must be noted. Some tasks are regarded as the domain of a particular age group, thus decreasing the substitutability of adults and children, even if either could adequately perform the task. If children require instruction or supervision in their work, adult and child labor become more complementary and less substitutable.

The final two variables in the child labor supply equation, expenditures on school fees and the distance to the nearest government primary school, represent the attractiveness of education, the principal alternative use of children's time. The more the household spends on school fees, the more children

it can be expected to have in school and the fewer children working on the holding. The distance to the nearest primary school measures both the household's accessability to education and the likelihood that some children will be enrolled in Standards I - IV for which there are no fees.

2.2 Saving

The level of household savings and the related expenditures on physical capital accumulation for the household enterprise and on children's school fees are the subjects of analysis in this section of the model. Investigation of the level of household saving seeks to identify its major determinants, highlighting the effect of children in the household. The allocation of saving between physical and human capital formation has a major influence on the role of children in the household.

The three equations of this section can be represented as:

(6) $S = f_6(Y^+, AGE^{2+}, AGE^{2-}, \underline{KIDS0\text{-}4}^?, \underline{BOYS5\text{-}9}^?, \underline{BOYS10\text{-}14}^?, \underline{GIRLS5\text{-}9}^?, \underline{GIRLS10\text{-}14}^?, \underline{ADULTS15\text{-}19}^?, \underline{ADULTS20}^?, LAND_P^+, LAND_O^+, TOOLS^+, BUILD^+, LSTOCK^+)$,

(7) $I_P = f_7(Y^+, LAND_P^+, LAND_O^+, TOOLS^+, BUILD^+, LSTOCK^+, \underline{LABOR}_E^\pm)$

(8) $FEES = f_8(Y^+, \underline{KIDS0\text{-}4}^-, \underline{BOYS5\text{-}9}^+, \underline{BOYS10\text{-}14}^+, \underline{GIRLS5\text{-}9}^+, \underline{GIRLS10\text{-}14}^+, \underline{ADULTS15\text{-}19}^+, EDUC^+, \underline{CHILD}_W^-, DIST_{PS}^+)$

where

S^2 = total household savings;

AGE^2 = age of the head of household squared;

I_P = ivestment in physical capital for the holding;

other symbols as defined above.

Detailed asset and liability data for the household are used to formulate a measure of saving and investment in physical capital. It is a comprehensive attempt to evaluate the difference between total production and consumption, but the details of the measurement will not be discussed here. The household's expenditures on school fees are estimated using information on government school fees, the number of children in each household attending school, their ages and the actual age-grade distribution of the children in

school in the province. Several sources of error or bias in this measure can be readily identified. The measure is employed in spite of its weaknesses because an analysis of the household's investment in human capital is one of the central themes of this research.

Income is generally the principal determinant of the level of saving. The income coefficient in the equation represents the marginal propensity to save; i.e., the fraction of a unit increment in income which would be saved, and so is expected to be between zero and unity.

The two age variables for the head of household are included to represent the life cycle pattern of saving, if it exists. Theoretically, the head is expected to save during his most productive years and to dissave at relatively young and old ages. While there is evidence of a life cycle pattern of saving in some developed countries, its occurrence in developing countries is questionable. It is often argued that because of differences in time horizons, time preference and social structures between developed and developing countries, the life cycle hypothesis finds support only in the former. The coefficients of these age variables will provide further information on this question.

The five variables for children are included to capture the age - and sex-specific impact of children on saving. Additional children may be net consumers thus decreasing household saving. However, additional children may cause the household to reallocate its existing consumption bundle, reducing average consumption but maintaining the household's previous saving level. Finally, children may cause the household to save more to either invest in their education or to augment the holding's physical capital to enable the children to work productively on the holding. This is considered to be the least likely possibility, especially where farming is already quite intensive so that the marginal product of labor is low and the household is living near subsistence.

The number of adults in the household is divided into two age groups, 15 - 19 and 20 or older to allow for possible school attendance by the younger group, thus making their impact on savings more positive (or less negative) than the older group's. At a given level of income, an additional adult may increase the household's consumption demands, thereby

decreasing household savings. Alternatively, an additional adult may cause a redistribution of the existing consumption with no effect on saving. The expected signs for these two adult variables are therefore uncertain.

The household's physical assets land planted in perennial crops, other land, and the value of tools, buildings and livestock, are included as measures of household wealth which is directly related to the household's ability to save and its savings.

The household allocates its saving between two main uses: investment in physical capital and investment in children's school fees. The level of physical capital investment has an effect on output and the demand for labor in the future and therefore indirectly influences the household's ability to save and the roles its members play.

Since income is the main determinant of household saving, it is expected to have a strong, direct relationship to the level of investment in physical capital. The coefficient on income represents the household's marginal propensity to invest in physical capital, i.e., the fraction of a unit increment in income which will be used for the purchase of capital.

The physical assets of the household are included in this equation, with positive coefficients expected. As measures of household wealth, they indicate the household's ability to save and its ability to invest. It is possible that a household with a relatively large capital stock may elect to concentrate its expenditures on human capital investments; but in general, a direct relationship is expected to prevail.

The final variable in this equation is the family labor input. The sign expectations are uncertain. Households with a relatively large labor input may wish to invest in labor-saving capital equipment in order to reduce the labor input and maintain or increase output. Alternatively, if the current availability of capital is a constraint in the productive utilization of household labor, those households with a relatively low level of labor input may be motivated to invest large amounts in physical capital. The coefficient on family labor will be positive if the high labor-input households

dominate, negative if the low labor-input households dominate and insignificant if the two are of approximately equal weight. The sign of the coefficient will indicate whether capital and labor are substitutes or complements.

In the third equation, income is expected to be directly related to the level of household expenditures on school fees since it again represents the household's ability to save. The coefficient can be interpreted as the marginal propensity to spend on school fees.

The five variables for children are included in the equation: the number of children under age 5, the number of boys in the 5 - 9 and 10-14 age groups and the number of girls in the same age groups. The impact of young children (under age 5) in the household is negative if their child care demands are sufficinet to necessitate the withdrawl of one or more children from school. If their demands for care are met by the household members working on the holding, the young children will have no effect on the level of expenditures on school fees. The number of older children in the household is expected to be directly related to the number of children in school and therefore directly related to the level of expenditures on school fees.

Adults 15-19 may attend secondary schools. To the extent that the young adults in the sample are enrolled in school, they will have a positive effect on spending for school fees. However, this sign expectation is tentative because the highly selective nature of the Kenyan education system and the costs of a secondary education mean that not many individual in the rural area will be enrolled in secondary schools.

The education of the head of household serves as both a taste and information parameter, with a positive sign expected. It is often argued that education alters the head's tastes so that he has a stronger preference for educated children. If this is the case, the education variable will have a positive coefficient. In Kenya, education is generally perceived as the key to obtaining formal sector wage employment. An educated head may attribute his failure to obtain such employment to his insufficient education. If this is the case, he is likely to educate his children even more to help ensure their success in the formal sector labor market.

The number of children working on the holding is inversely related to the number of children attending school and therefore inversely related to expenditures for school fees, assuming work on the holding and school attendance to be incompatible activities.

The final variable in this equation is the distance to the nearest government primary school. The closer the school, the easier access to it and the more children who can be expected to attend.

2.3 Demographic Decisions

The demogtaphic section of the model investigates the seven components of household size utilized: the number of children ages 0 - 4, the number of boys ages 5 - 9 and 10 - 14, the number of girls ages 5 - 9 and ages 10 - 14, the number of adults ages 15 - 19 and ages 20 or over. Endogenous consideration of these variables serves to close the model. The seven equations can be represented as follows, using the notation defined above:

(9) KIDS0-4 $= f_9 (\underline{Y}^?, AGE^+, AGE^{2-}, EDUC^?, FEES^-, LAND_P^+, LAND_O^+, TOOLS^+,$
$ISTOCK^+, \underline{BOYS5\text{-}9}^-, \underline{BOYS10\text{-}14}^-, \underline{GIRLS5\text{-}9}^-, \underline{GIRLS10\text{-}14}^-)$

(10) BOYS5-9 $= f_{10} (\underline{Y}^?, AGE^+, AGE^{2-}, EDUC^?, FESS^-, LAND_P^+, LAND_O^+,$
$\underline{TOOLS}^+, \underline{BUILD}^+, \underline{ISTOCK}^+, \underline{BOYS10\text{-}14}^-, \underline{GIRLS10\text{-}14}^-)$

(11) BOYS10-14 $= f_{11} (\underline{Y}^?, AGE^+, AGE^{2-}, EDUC^?, \underline{FEES}^-, LAND_P^+, LAND_O^+,$
$TOOLS^+, BUILD^+, LSTOCK^+)$

(12) GIRLS5-9 $= f_{12} (\underline{Y}^?, AGE^+, AGE^{2-}, EDUC^?, \underline{FEES}^-, LAND_P^+, LAND_O^+,$
$TOOLS^+, BUILD^+, LSTOCK^+, \underline{BOYS10\text{-}14}^-, \underline{GIRLS10\text{-}14}^-)$

(13) GIRLS10-14 $= f_{13} (\underline{Y}^?, AGE^+, AGE^{2-}, EDUC^?, \underline{FEES}^-, LAND_P^+, LAND_O^+,$
$TOOLS^+, BUILD^+, LSTOCK^+)$

(14) ADULTS15-19 $= f_{14} (\underline{Y}^+, AGE^-, AGE^{2+}, EDUC^+, LAND_P^+, LAND_O^+$
$TOOLS^+, BUILD^+, LSTOCK^+)$

$$(15) \qquad ADULTS20 = f_{15}\,(\underline{Y}^{+},\ AGE^{-},\ AGE^{2+},\ EDUC^{+},\ LAND_{I}^{+},\ LAND_{R}^{+},\ TOOLS^{+},\ BUILD^{+},\ LSTOCK^{+})$$

These equations, unfortunately, are not indicators of fertility decisions. Identification of the determinants of fertility is essential if Kenya is to take the necessary steps to slow its population growth. the model, however, utilizes the household as the unit of analysis while fertility needs to be examined at the level of the individual or conjugal couple. Its inclusion in the model is therefore impossible. Analysis of fertility histories of adult females is discussed as a separate topic in section 3.2 of this paper. The components of household size are examined . . . since they represent one of the possible sets of demographic variables which can be considered within the framework of the model.

The relationship between household income and the number of children in a given age-sex group is uncertain a priori. As income increases, the household is expected to consume more of all normal goods, including children. Rising incomes, on the other hand, are usually accompanied by increases in the cost of the household members' time and therefore by increases in the cost of child care. This increased cost, via the substitution effect will lead the household to have fewer children. Which of these two effects dominates is an empirical question.

The age of the household head and its square depict the life cycle of child bearing and rearing. The relationship is expected to be parabolic, concave downward. The maximum for each successively older group of children will be attained at successively older ages of the head as children grow into and out of the various groups.

The effect of the head's education on the number of children in his household is uncertain. Education may decrease the number of children the head has by exposing him to non-traditional values and life styles and even by informing him of contraceptive methods. Alternatively, the head may repay the extended family's investment in his education by housing members of the family and thus increasing the number of children in his household. The dominant effect of education can only be determined empirically.

School fees are a cost of raising children and so are expected to be negatively related to the number of children in a given group.

The five asset measures, land in perennial crops, other land, and the value of tools, buildings and livestock are included as measures of wealth. Annual income for the household is in part a result of the use if these physical assets, but income for a given year is subject to a variety of fluctuations arising, for example, from temporary price changes or weather patterns. Annual income may therefore give an incorrect indication of the household's ability to support a given household size. The household size decision may rely more on the household's stock of physical assets as a measure of the household's long run average income and its ability to maintain a given household size.

Finally, the numbers of children older than the group under consideration are included on the assumption of a finite desired household size, regulated mainly by fertility adjustments. Households with relatively many older children are expected to be near their desired size and so will have fewer younger children, and conversely.

In the two equations for the number of adults in the household, traditional family ties account for the expected direct relationship between the number of adults in the household and household income. Relatively affluent members of the extended family are likely to support other members. As discussed above, physical assets may be a better measure of the household's ability to support a given household size and so the asset measures are also expected to have positive coefficients.

The age terms for the head of the household are included to test the hypothesis that the number of adults in the household follows a life cycle pattern with the minimum being reached in the head's middle productive years when he has sufficient experience and physical strength to manage the holding The head may receive assistance from the extended family in his youth when he is relatively inexperienced and again in his old age to compensate for his declining vigor. Finally, the education of the head is expected to increase the number of adults in the household. The head may have an obligation to house and support some of the family members arising from family financing of

his education.

2.4 Policy Implications

The model enables examination of several areas of importance to Kenya. The investigation of the level of agricultural output and its determinants suggests the viability of increasing production to feed the expanding population and to generate foreign exchange. Analysis of the production relationships also indicates the feasibility of increasing productive employment in the rural areas. If the rural areas can productively absorb the growing numbers of unemployed, economic growth and perhaps a large increase in total output could result. Examination of household saving levels provides estimates of the level of private investment in the rural areas. This knowledge can be employed to plan government programs and expenditures to complement or to further stimulate the development of the rural areas. The amount of investment in physical capital is particularly relevant to consideration of investment levels and output expansion since much of this capital is not readilly transferable to urban employment. Only a small portion of those educated in the rural area remain there since human capital is relatively portable and adaptable to urban employment. The level of human capital accumulation is therefore of little direct importance to the output decision in the rural areas. Since household saving is allocated mainly between physical and human capital the opportunity cost of expenditures on education is physical capital accumulation which could directly increase agricultural output. Furthermore, formal education increases the probability that an individual will migrate to the urban area and further aggravate the already severe problem of the educated unemployed. Finally consideration of the household size decisions enables some estimates of the demographic composition of rural areas in the future to be made. This too has implications for the levels of output and investment because of the strong traditional roles of individuals according to age and sex. Because the model is specified as interactive, both direct and indirect influences on a given variable can be identified, providing a wider range of possible policies.

Simultaneous equation systems assume that all changes in the endogenous variables are the result of changes in one or more of the exogenous variables. While the interactions of the endogenous variables are of interest, it is the exogenous variables which are the potential policy instruments. Derivation of the policy implications is relatively complex since each of the

exogenous variables in indirectly related to the dependent variables as a result of the model specification and the estimation procedure employed (see section 4.1). The total effect of a change in an exogenous variable on a dependent variable is usually indeterminate a priori although some components of the total can be identified. The discussion here is concerned with the general interactive nature of the model, some of the policy areas on which this research can provide evidence and on several of the key exogenous variables in the model. A more exact determination of the interrelationships among the variables will be carried out and analyzed after the model has been estimated.

A change in any exogenous variable causes a series of adjustments in the model which eventually lead the system to re-equilibrate. The exact nature of the process depends upon which exogenous variable has changed but in general, the interactions within the model are the same. Virtually every aspects of production on the holding is affected by a change in an exogenous variable. The adjustment in the quantity of family labor employed may influence the number of adult and child-family workers on the holding. In the savings section of the model, a change in an exogenous variable will cause a change in the level of saving and of investment in human and physical capital. It may also influence the relative shares of school fees and physical capital accumulation in total saving. In the demographic section, the size and composition of the household will respond to a change in an exogenous variable. Interactions within the demographic section occur as the number of older children in the household influence the number of younger children.

The sections of the model interact to cause further adjustments. A change in the level of income affects the level of saving and investment and also household size and composition. The adjustment in the family labor input for the household enterprise influence the number of children attending school and therefore the level of household spending on school fees. The change in expenditures on education implies a change in the average cost of children and therefore a change in their numbers. The level of investment in physical capital influences the technology decisions on the holding. Adjustments in the size and composition of the household affect the amount of family labor employed on the holding and the levels of saving and investment in school fees.

Several decisions are of particular interest. In the rural area, the levels of saving and investment and the growth of agricultural employment and output are all closely related. To the extent that agricultural employment is constrained by the availability of tools and equipment, high rates of investment in physical capital will increase the employment and output on the holding. The behavior of the rural household and, in particular, its continued investment in education indicate a disparity between the expected social and private returns to alternative investments. The social returns to formal education, when the result is migration to an already oversupplied urban market for educated people, are low. Largely as a result of distortions in the labor market, private returns are still assumed to be high. There is, however, evidence that additional formal education is not the best means of increasing agricultural productivity. This is especially relevant to the current situation in Kenya since the slow growth rate of formal-sector urban employment implies that school leavers must seek employment in agriculture. If the private returns to alternative investments as viewed by the household could be realigned with the social returns, the over all productivity of investments would increase. While this study does not examine the rates of return to alternative investments, it can indicate several ways to influence the distribution of household savings.

A second issue, somewhat related to the first concerns the labor input decision in the household which has implications for the technology employed and the labor intensity of rural production. The high level of unemployment and population growth and the slow expansion of formal sector employment in Kenya imply that a policy of promoting increasing labor intensity in the rural areas is a feasible means of more productively utilizing existing resources. The analysis of the labor input and composition decisions within the model provides evidence on viable means of encouraging increased labor intensity.

The final household decision to be discussed here concerns the debate over the impact of household income on the demand for children. Several fertility theories exist but there is little agreement about the income elasticity of demand for children, especially as development proceeds. Some evidence on the issue will be provided by the analysis of the effects of an exogenous change in income from sources other than the household enterprise, Y_0, on the numbers of children, and in particular on the number of children under age five in the household. The results will not be conclusive

since Y_0 is only one component of total household income but it represents better than any other variable the effect of an exogenous change in household income.

Since the exogenous variables are the potential policy instruments and since all changes in the system result from a change in an exogenous variable, analysis of adjustments in the system begins with an assumption about which variable has changed. Complete analysis of the model dictates that the implications of changes in each of the exogenous variables be examined. This will be carried out empirically once the model has been estimated. The nature of three sets of exogenous variables are discussed below.

First, the impacts of the household's physical assets are seen directly and indirectly throughout the model. Land planted in perennial cops, other land and the value of tools, buildings and livestock are considered together here because they appear together in the model. Their joint consideration does not imply that a unit change in one variable has the same consequences as a unit change in another variable. Separate investigations for each of these variables will be undertaken in the empirical analysis.

Physical assets have a dual role in the model and in the household. They represent factors of production and large share of the household's wealth. As factors of production, they increase output on the holding, draw forth family labor directly and indirectly as a result of the relationship between household size and labor input. As household wealth, the assets stimulate saving, influence the distribution of saving between physical and human capital and serve as an indicator of the household's long-run ability to maintain a given household size. The indirect effects of these assets are difficult to determine a priori and will be left for later analysis. Nonetheless, it can be seen that the household's stocks of physical assets have powerful and pervasive effect in the model.

The second exogenous variable is the education of the head of the household. The variable enters directly into the production function and the household size equations but its impact throughout the model, especially via household income, are expected to be significant. Analysis of the effects of education is of particular importance for two reasons. First, it is likely that in the forseeable future the educational attainment of the head of the

household will increase as more and more school leavers turn to the rural areas to seek employment. The impact of this can be analyzed within the framework of the model. Second, knowledge of the impact of increased education on household decisions provides evidence concerning the desirability of further expansion of the education system.

The final exogenous variables to be considered here are the age of the head of the household and its square. While age cannot be manipulated by government policy, it is investigated for two reasons. First, the high rate of population growth in the rural area and the resulting age distribution imply that the heads of households on average, will be younger in the future. The inclusion of the age variables facilitates the analysis of the implications of this change in the age distribution. Second, the inclusion of the age variables in the savings and household size decisions constitutes a test of the life cycle hypothesis of behavior as described above.

3.0 SUPPLEMENTARY RESEARCH

The investigation of three topics related to the model described above is proposed to supplement and enrich the conclusions reached by the use of the model. The first two topics, the determinants of the individual child's main activity and of the individual woman's fertility experience, are considered separately from the model since they employ the individual as the unit of analysis while the model utilizes the household. The third topic, the analysis of household expenditures, demands special estimation procedures which are incompatible with its inclusion in the model.

3.1. Allocation of Children's Time

Parents are assumed to allocate their children's time between education and work in the household enterprise, considering both individual characteristics such as age, sex, birth parity and parentage, and household characteristics such as income, household size and location, in their decision. The two activities, for the purposes of this research, are assumed to be mutually exclusive. The decision is analyzed by the use of the following equation:

(A) $$D = f_A(AGE_C^-, SEX^+, HEAD^+, Y^+, EDUC^+, DIST_{PS}^-, LAND_I^?, LAND_O^?, TOOLS^?, BUILD^?, ISTOCK^?, KIDS0\text{-}4^-, BOYS5\text{-}9^-, BOYS10\text{-}14^-, GIRLS5\text{-}9^-, GIRLS10\text{-}14^-)$$

where

D = binary variable = 1 if the child attends school,
= 0 otherwise;

AGE_C = age of the child;

SEX = binary variable = 1 if the child is a boy,
= 0 if the child is a girl;

HEAD = binary variable = 1 if the child is the son or daughter of the head of household,
= 0 otherwise;

other variables as defined above.

Productivity and education costs are directly related to the child's age so that both the opportunity cost and the direct cost of education increase with age, leading to a negative relationship between age and school attendance.

The expected positive coefficient on the sex variable is a reflection of the strong sex bias in favor of educating males which is exhibited in Kenya.

The variable describing the child's relationship to the head of the household is included to test the hypothesis that the head's children receive preferential treatment and are more likely to attend school than the other children in the household. However, the head may be bound by traditional family ties to treat all children in the household equally.

As described in an earlier part of this paper, household income and the education of the head of the household are directly related to school attendance while the distance from the household to the nearest school has a negative effect on the children's school attendance.

The household's physical assets are included in the equation, but no expectations concerning their signs are held. The assets are an indicator of household wealth and so may be directly related to the probability that any one child is enrolled. Furthermore, a household with a relatively large stock of assets may prefer to educate its children rather than invest in additional physical capital. Alternatively, if child labor and physical capital are complementary factors of production, the value of the household's assets and the probability that a child attends school are inversely related.

Finally, the five variables for the number of children in the household are included. The presence of young children in the household increases the likelihood of one or more children assisting in child care activities, thus decreasing the probability that he will attend school. The variables for the number of older children in the household indicate the extent of competition within the household for education. The more children in the household the stronger the competition and the lower the probability that each child attends school.

This analysis of children's activities when considered with the results of the model above fully describes the role of children in the household. The equation estimated here identifies the determinants of the children's main activity while the model takes the children's activities as given and investigates their impact on production, labor supply and saving in the household.

3.2 Fertility Behavior

The second topic investigated is the fertility experience of individual adult women. In countries where infant and child mortality are low, it is often sufficient to analyze the number of live births a woman has had to understand her fertility behavior. However, rural areas of developing countries have relatively high child mortality rates so that the number of children a woman has borne often differs from the number who are surviving. Both the number of surviving children and the number of children who have died will be investigated here, employing individual- and household-specific data. The equations to be estimated are:

$$(B) \qquad CS = f_B\,(Y^{?},\ EDUC_F^{-}, AGE_F^{+},\ L_F^{+}, AGED_F^{-},\ CHILD_D^{+},\ WIFE^{+},\ LAND_P^{+},\ LAND_O^{+},\ TOOLS^{+},\ BUILD^{+},\ LSTOCK^{+})$$

$$(C) \qquad CHILD_D = f_C\,(Y^{-},\ EDUC_F^{-},\ AGE_F^{+},\ AGE_F^{2-},\ CEB^{+},\ WIFE^{-})$$

where

CS = the number of children ever born to the woman who are still alive;

$EDUC_F$ = formal education of the woman;

AGE_F = age of the woman;

AGE_F^2 = age of the woman squared;

D_F = binary variable = 1 if AGE ≥ A* as defined below,
= 0 otherwise;

$AGED_F$ = product of AGE_F and D_F;

$CHILD_D$ = number of children of the woman who have died;

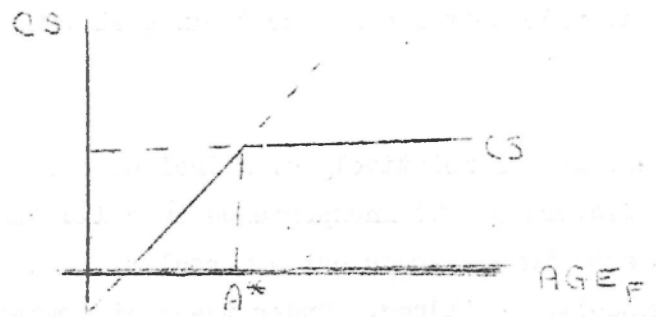
WIFE = binary variable = 1 if the woman is the wife of the head of household or she is the head of the household,
= 0 otherwise;

CEB = the number of children ever born to the woman;
other variables as defined above.

The impact of income on the demand for children is uncertain, as discussed in the demographic section above. An increase in income leads to increased demand for all normal goods, but the increase in income may also lead to an increase in the cost of raising children, leading the household to substitute alternative expenditures for them.

The woman's education is expected to decrease her fertility either by altering her tastes or by informing her of contraceptive methods. The impact of the woman's education is likely to be attenuated by the traditionally passive role of women in the family size decision although educated women may be less passive. The woman's education may serve as proxy for her husband's education which may decrease his desired family size and therefore, her fertility.

The relationship between the age of the woman and her fertility is well established. For ease in estimation, the relationship will be approximated by two linear segments as shown in the figure below. This utilizes the binary variable, D_F , which when interacted with the age of the woman



produces a function of the desired shape. Below some minimum age, women are physically incapable of bearing children, but once they begin, children are

born at approximately regular intervals and child bearing is completed relatively rapidly. Beyond a certain age, A*, whether by choice or by physical incapacity, no more children are born and the CS function levels off. It is the A* point which defines the binary variable, D_F.

Under the assumption of a constant desired completed family size, children who die are expected to be replaced. This implies that the coefficient on the child deaths variables will be greater than zero. It is likely that the occurrence of a child death leads to an upward revision of the family's expectations of further child mortality and to over replacement of the dead child to ensure the attainment of the desired family size. (A zero coefficient would indicate exact replacement of the dead child; a negative coefficient indicates under replacement.)

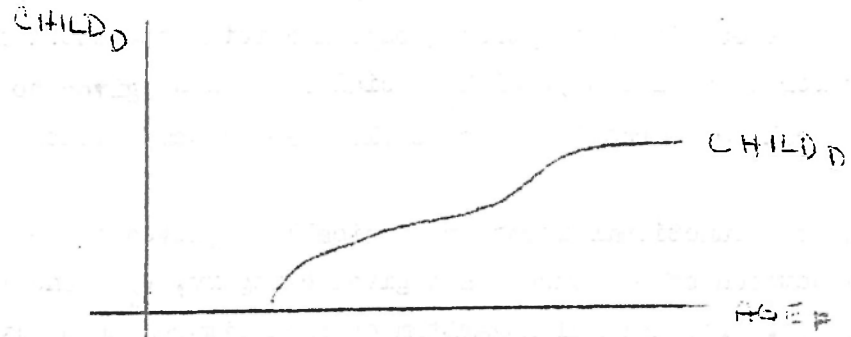
A binary variable for the wife of the head of household is included on the hypothesis that she is more economically secure and therefore has more children.

Finally, the physical assets variables for the household are included, with positive signs expected. Children may be regarded as sources of future labor. Therefore, under the assumption that the assets and child labor are complementary factors of production, the number of children born to women in the household and the stock of physical assets will be directly related.

Since the number of child deaths influences fertility as describbed above, investigation into the determinants of child mortality is closely related to the analysis of the number of surviving children. In the child deaths equation, income is expected to reflect a variety of factors such as nutrition and health care which would decrease child mortality. Educated women are thought to provide better care for their children, further decreasing child mortality.

Young mothers have a relatively high incidence of child mortality due to both physical factors and to inexperience in child care. Relatively old mothers may also experience higher child mortality if they are more likely to produce weak or unhealthy children. Under these circumstances child deaths would increase most rapidly in the mother's young and old ages, with

virtually no mortality experienced in the middle child bearing years. See the figure below for a graphic protrayal of the expected relationship.



The number of children born and therefore the number of children exposed to the risk of death is directly related to the number who die, leading to an expected positive coefficient on the children ever born variable.

Finally, the binary variable for the wife of the head of the household is included on the hypothesis that her children will receive preferential treatment in the household and are therefore less likely to die.

This investigation has several policy implications. The fertility of women in the household influences household size. The interaction of household size and other characteristics is investigated in the model described above. Furthermore, it is the individual fertility experiences, when taken together, which comprise the country's population growth rate. The importance of slowing this growth in Kenya for the economic well-being of the populace cannot be exaggerated. Several nutrition programs have been launched on the hypothesis that better nutrition reduces child mortality which in turn reduces fertility, again assuming a constant desired family size. The provision of nutrition, health and family planning information as a single service is also based on similar thinking. Although this investigation of child mortality does not assess the impact of the nutrition of the household on child mortality, it provides information to assess the crucial link between child mortality and fertility.

3.3. Household Expenditures

The level and composition of household consumption is related a variety of factors such as the level of household income and the demographic attributes of the household. Over the course of development, the composition of

household expenditures changes in a relatively predictable fashion the share of income spent on necessities such as food decreases and the share spent on luxuries increases. This research investigates the relationship between expenditures on specific consumption groups and total household expenditures (as a proxy for total income), with special attention given to the possible economies of scale as a result of household size or composition.

Several functional forms are typically employed to describe the relationship between expenditures on a given category, V_i , and total expenditures, V_0. The natural logarithm of expenditures on a given category V_i, is related to the logarithm of total expenditures, V_0 or to $1/V_0$. Alternatively, V_i is a function of V_0, $1/V_0$, or the log of V_0. The formulation selected for estimation here includes each of these possible relationships as special cases and allows for a variety of other relations. The equation to be estimated can be represented as:

$$(D) \quad \frac{V_i^a - 1}{a} = b + c \frac{V_0^d - 1}{d} + (\text{demographic and other factors})$$

where

V_i = expenditure on the i^{th} category;

V_0 = total expenditures;

a, b, c, d = parameters to be estimated.

Since a and d are parameters rather than known quantities, ordinary linear regression analysis cannot be utilized. Estimation requires that a likelihood function be specified and various combinations of values for a and d are tested to determine the pair which maximizes the likelihood function.

The expenditure categories to be analyzed are household expenditures on all food, meats, and dairy products, fruits and vegetables, clothing, consumer durables (appliances and furnishings), fuels and transport.

Several demographic specifications will be tested to determine the impact of household size and composition on household structure. Adult equivalents, employing various weights for children are used to provide comparability between this research and other studies and to provide information on the usefulness of adult equivalent measures as opposed to alternative demographic specifications. One major weakness of the adult equivalent

measure is that it assumes no economies of scale in consumption. Three alternative demographic specifications will be employed to test for the presence of various scale economies. First, economies that result from household size are captured by the use of household size and its square. Second, the number of children, its square, the number of adults and its square are included to test for different economies for adults and children with constant economies within each group. Finally, the household members are divided into age groups as in the model described above. The number in each category and its square are included to test for differences in the economies of scale across different age groups.

This analysis of the household's pattern of expenditures enables the estimation of income and demographic elasticities for various expenditure categories. These elasticities provide information about the characteristics of current household demand and the changes in demand which can be expected in response to the economic and demographic changes which acompany developments.

4.0 ESTIMATION PROCEDURES

Several different econometric techniques are employed to estimate the equations described above. The simultaneous equation model requires two-stage least squares or similar techniques (indirect least squares, instrumental variables or k-class estimators). The analysis of the children's main activity employs the probit model for estimation with a qualitative dependent variable. Ordinary least squares will be used to estimate the two equations for the women's fertility experience. Finally, the analysis of household expenditures requires maximum likelihood procedures. Each of the techniques is discussed briefly here: references for the more technical aspects are provided.

4.1 Simultaneous Equation Estimation

The use of ordinary least squares analysis assumes that the variables appearing on the right-hand side of the equation are uncorrelated with the error term (2, p.162). In a system of simultaneous equations where endogenous variables appear on the right-hand side, it can be demonstrated that the endogenous variables are correlated with the error terms (2, p.305; 3, p. 445). If the system were estimated witn ordinary least squares procedures, the resulting coefficients would be inconsistent estimators while the two-stage least squares estimates are consistent.

The process of two-stage least squares is relatively simple. The reduced form equations for each of the endogenous variables is formed by writing each endogenous variable as a function of all the exogenous variables in the system. These reduced forms are estimated using ordinary least squares procedures to determine the coefficients for the exogenous variables. The coefficients are then employed to generate estimates for each of the endogenous variables. Since the estimates are constructed as a linear combination of the exogenous variables which are, by assumption, uncorrelated with the error terms, the estimates of the endogenous variables are likewise independent of the error terms.

In the second stage, the endogenous variables on the right-hand side are replaced by their estimates generated in the first stage of estimation. Since these estimates are not correlated with the error terms, ordinary least squares analysis is employed to estimate the coefficients.

If estimation is carried out using the procedure outlined above, the sum of squared errors and the t-statistics from the second stage are not the appropriate statistics. They are constructed on the assumption that the dependent variable is a function of the estimated values of the endogenous variables which appear on the right-hand side while the model assumes that the dependent variable is a function of the observed values of those endogenous variables. If the two-stage least squares routines of packages such as ESP, TTS or SOUPAC are used, the necessary adjustments are performed internally.

4.2 Estimation with a Binary Dependent Variable

The analysis of children's main activities, as described in section 3.1 above, requires estimation by some technique other than ordinary least squares. The use of ordinary least squares estimation results in a heteroskedastic error structure which violates one of the assumptions of the Gauss-Markov Theorem (2, p.249). Probit analysis transforms the binary dependent variable so that estimation is possible. Essentially, the dependent variable is re-specified as a function of a linear combination of the independent variables. It is then mapped into the cumulative normal distribution, the value of which depends upon the values of the coefficients for the independent variables. A likelihood function of these values

of the cumulative normal distribution if formulated and solved for the values of the coefficients which maximize the value of the likelihood function. (For a more technical description, see 2, pp. 249-251; 3, pp.630-631; 1, pp.4-5.)

4.3 Ordinary Least Squares Estimation

The assumptions of the Gauss-Markov Theorem which indicate when ordinary least squares estimation yields the best linear unbiased estimates of the coefficients are well known (2, p.164). These assumptions are approximately met for the equations describing women's fertility. The main problem is one of simultaneity between the fertility decisions and the variables such as income which are treated as endogenous in the simultaneous equation model described above. The solution to this is to employ the estimates of income and other endogenous variables from the first stage of estimation for the model under the plausible assumption that the exogenous variables of the model are not correlated with the error terms in the fertility equations. If this is done, the t-statistics and the sum of squared errors must be corrected in the same fashion as described for the second estimation stage in section 4.1.

The value for the D_F variable will be determined empirically by trial and error using the available a priori information. It is assumed that the second linear segment of the CS function should have approximately a zero slope. The definition of D_F which best satisfies this condition will be employed.
(For a detailed explanation of the theory of ordinary least squares estimation, see 2, pp. 156-181.)

4.4 Maximum Likelihood Estimation

The estimation procedures required for the analysis of household expenditures has already been discussed briefly. The equation is inherently non-linear, prohibiting ordinary least squares estimation. Instead, a likelihood function such as the one presented in logarithmic form below is specified. The partial derivatives of L, the likelihood function, with respect to the vector of coefficients and σ^2, the variance of the error term, are found and set equal to zero. The solution of these equations yield the maximum likelihood estimates for the coefficients and for σ^2.

$$L = -\frac{T}{2}\log(2\pi\sigma^2) - (2\sigma^2)^{-1}\left[\frac{Y_1^{\alpha}-1}{\alpha} - b - c\frac{Y_3^{\delta}-1}{\delta}\right]'\left[\frac{Y_1^{\alpha}-1}{\alpha} - b - c\frac{Y_3^{\delta}-1}{\delta}\right]$$

(For a discussion of maximum likelihood estimation procedures, see 2, pp. 179-180.)

5.0 BIBLIOGRAPHY

1. Falusi, B. "Multivariate Probit: Analysis of Selected Factors Influencing Fertilizer Adoption Among Farmers in Westen Nigeria." Nigerian Journal of Economic and Social Studies. XVI (1) March 1974, pp. 3 - 16.

2. Goldberger, A. Econometric Theory. New York , John Wiley and Sons, 1964.

3. Theil, H. Principles of Econometrics. New York., John Wiley and Sons, 1971.