76-27,631

CHAE, Hi Kyung, 1939-
TWO ESSAYS ON THE RESIDENTS OF NAIROBI, KENYA: I. INTRAURBAN MIGRATION. II. SCHOOLING, EARNINGS, AND EXPERIENCE.

University of Oregon, Ph.D., 1976
Economics, general

Xerox University Microfilms, Ann Arbor, Michigan 48106

TWO ESSAYS ON THE RESIDENTS OF NAIROBI, KENYA:

I. INTRAURBAN MIGRATION

II. SCHOOLING, EARNINGS, AND EXPERIENCE

by

HI KYUNG CHAE

A DISSERTATION

Presented to the Department of Economics
and the Graduate School of the University of Oregon
in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

June 1976

APPROVED: W. Ed Whitelaw

(W. Edward Whitelaw)

VITA

NAME OF AUTHOR: Hi Kyung Chae

PLACE OF BIRTH: Seoul, Korea

DATE OF BIRTH: December 9, 1939

UNDERGRADUATE AND GRADUATE SCHOOLS ATTENDED:

Seoul National University
University of Oregon

DEGREES AWARDED:

Bachelor of Arts, 1963, Seoul National University
Master of Arts, 1974, University of Oregon
Master of Science, 1975, University of Oregon

AREAS OF SPECIAL INTEREST:

Urban and Regional Economics
Economics of Human Resources
International Economics
Public Finance

PROFESSIONAL EXPERIENCE:

Graduate Teaching Fellow, Department of Economics, University of Oregon, Eugene, 1972-1974

ACKNOWLEDGEMENTS

I am deeply indebted to Professor W. Edward Whitelaw and Professor Chulsoon Khang. Dr. Whitelaw, the chairman of my dissertation committee and friend, helped me to pursue these topics with severe criticism and enlightening suggestions. Dr. Khang is the one who opened the door to the science of economics for me when I was plagued with uncertainty as to what I should study in the United States. I owe much of my lean knowledge of economics to him.

I am also grateful to Professor James N. Tattersall and Professor Donald Watson for their valuable suggestions and timely responses.

I also wish to express my appreciation to my friends, Ron and Dona Lehr, Lenny Muraya, Jim Scherzinger, Suk Young Yoon, and especially Mrs. Joyce Sweeney who helped my study in one way and another.

My final thanks go to my wife who had to master the art of patience.

PREFACE

The dissertation is composed of two separate essays: I. Intra-urban Migration; and II. Schooling, Earnings, and Experience.

The idea of researching two different topics together came up with the convenient accessability to an invaluable set of data that contains abundant information on urban households.

The sample used in both essays is a part of the data of the Nairobi Household Survey. The survey was conducted in the Spring of 1971 by the Institute for Development Studies under the direction of W. Ed Whitelaw. The original data consists of information on 2148 individuals comprising 1365 middle and low-income households in Nairobi, Kenya, a city of 509,000 residents in 1970.

The general assumption underlying both essays is that the models developed in this study may apply to the urban areas of developing countries where many small employment centers are scattered and the educational level of the residents is low compared with the urban areas of developed countries.

TABLE OF CONTENTS

LIST OF TABLES

I. INTRAURBAN MIGRATION

II. SCHOOLING, EARNINGS, AND EXPERIENCE

LIST OF FIGURES

I. INTRAURBAN MIGRATION

I. INTRAURBAN MIGRATION

A. INTRODUCTION

It is widely recognized that intraurban migration--changing residential location within an urban area--plays a significant role in altering spatial pattern and social structure of cities. If we want to change the pattern and the structure of cities to improve the efficiency of urban transactions, it is essential to understand how the altering occurs. I believe that analysis on the moving behavior of households would be valuable in formulating government policy. As Mills puts it, "The desirability of almost every public policy depends on qualitative and quantitative effects which can only be predicted on the basis of considerable understanding of the way the system works."[1]

Intraurban migration has long been a favorite research field for sociologists, dating back to the beginning of this century.[2] Recently urban geographers have shown interest and have reported some

---

[1] Edwin S. Mills, Urban Economics, Glenview: Scott, Foresman and Co., 1972, p. 2.

[2] For a guide to the old literature on intraurban migration, see Theodore Caplow, "Incidence and direction residential mobility in a Minneapolis sample," Social Forces, May, 1949.

interesting research findings (1,6). One can scarcely find any literature on intraurban migration in economics, however, even though it offers tools as good as, if not better than those offered by other social sciences.

Each social science has different views regarding the decision to move. Some human ecologists see the motivation for changing residence as an element in a large pattern of the process of growth and succession (16), whereas "stress" seems to be the key word for sociologists (63). Most economists, of course, view the move as a maximization of a utility function. At any rate the decision to move involves many complex factors. The economic factors, however, that seem to be major factors influencing the decision to move were virtually ignored by most of the researchers in their studies. The failure to take this into account is the significant drawback of other social science studies on the subject. Examining the effect of various economic factors on the decision to move is the main purpose of this paper.

Following the introduction this study is divided into four sections. Section B describes the characteristics of movers by comparing the mean value of socio-economic factors such as income, rent, age, and various distances from the residential location. Section C develops a theoretical model based on the consumer theory and offers a probability model of a binary choice that employs most of the characteristics presented in section B. The empirical results estimated by ordinary least squares, weighted least squares, and discriminant analysis also are reported in this section. Section D explores the question of whether there are biases in direction and distance when people move. For directional bias, the hypothesis of mental map suggested by Adams is revised and tested. For distance bias, a model of the distance of move is formulated and estimated. Finally, section E summarizes the major findings and contributions of this study and suggests areas for further research.

B. THE CHARACTERISTICS OF MOVERS

People are mobile. Almost everyone experiences more than one change of residential location during his life time. Mobility, however, varies among individuals. Residents of an urban core area, for example, are more mobile than residents of a suburban area. Among the urban core residents, some people appear to change their residential locations more frequently than others. Many questions, however, remain unanswered. Are the movers younger and wealthier than non-movers? Do they have many children and live far from their jobs? Or is there no characteristic difference between movers and non-movers? These are some of the questions to be addressed in this section.

The behavior of intraurban migration stems from the change in quantity of housing demanded. As many previous studies have reported (11, 35, 62), changes in the quantity of housing demanded are related to income, price of housing, and tastes that include such characteristics as age and family size. Also amenity and access to central business district (CBD) are reported to be important variables in housing demand. (24, 31).

Among the large number of variables contained in the data, I selected as many variables that appear to be conceptually relevant to the behavior of intraurban migration as possible. Each variable can be classified in one of the following three categories: economic, taste, or distance characteristics. Average monthly income, monthly rent, and the income-rent ratio reflect the economic characteristics while age, relationship of household head to his own parents (oldest son or not) and number of children are taste variables. Distance variables (distance from job location to residential location; distance from CBD to residential location; distance from school to residential location) represent access variables to various places. These variables may be regarded not only as the proxies for economic attributes based on daily or weekly travelling cost but also the proxy for locational preference of the household.

In my sample of 755 observations, 219 household heads changed their residential location during the year of 1970 in which all the information used in this paper was available. The following table presents the mean value of various characteristics of household heads in each case. The table reveals that movers are, in general, poorer and, thus, pay lower rent but have a higher income-rent ratio than non-movers. The high income-rent ratio of movers shows that their low housing expenditure takes an even smaller share of their average income than that of non-movers. The movers are relatively younger and have fewer children. Their residential locations are closer to downtown but farther from their job locations.

Table B.1: Mean Values of Various Factors

Mean (Standard Deviation)

	MOVE	NOMOVE	COMMON
AVY	385.31 (291.80)	461.12 (382.77)	439.13 (360.41)
RENT	50.73 (52.73)	62.87 (62.16)	59.35 (59.71)
YRAT	10.34 (8.40)	9.84 7.18)	9.98 (7.56)
KIDH	0.70 (1.32)	1.39 (1.92)	1.20 (1.70)
AGE	29.84 (8.23)	37.13 (10.58)	35.02 (10.49)
DJH	18.48 (19.82)	13.26 (12.56)	14.78 (14.92)
DCH	13.54 (7.35)	19.03 (20.36)	17.78 (17.78)
TDSH	9.09 (11.49)	13.05 (11.80)	11.91 (11.85)

Where,

MOVE = Mover
NOMOVE = Non-mover
AVY = Average monthly income
RENT = Monthly rent
YRAT = Income-rent ratio
KIDH = Number of children with household head
AGE = Age of household head
DJH = Distance from job location to residential location
DCH = Distance from CBD to residential location
TDSH = Distance from school to residential location multiplied by the number of children in school

The numbers of oldest son and the household heads who changed jobs during the year of 1970 in each group are presented in the following table.

Table B.2: Distribution of JOB and OSON

	MOVE (219)	NOMOVE (536)	TOTAL (755)
OSON	74 (33.8%)	226 (42.2%)	300 (39.8%)
JOB	60 (27.4%)	59 (11%)	119 (15.8%)

Where

OSON = Household head who is the oldest son in his family.
JOB = Household head who changed job location.

There are 119 (15.8%) household heads who changed their job locations in the total sample of 755 (100%) observations and slightly more than half of them also changed their residential location. As a proportion to each group, they compose 27.4 percent in MOVE group and 11 percent in NOMOVE group respectively. For oldest son, the proportion is 33.8 percent in MOVE group and 42.2 percent in NOMOVE group.

C. THE DETERMINANTS OF INTRAURBAN MIGRATION

To contribute to intelligent formulation of public policy toward housing, it is important to find out what are the major factors that influence the decision to move. Knowledge of the motivations for changing residence should contribute to effective housing policy. If one knows the determinants of the decision to move, one should be able to estimate the volume of intraurban migration in a certain urban area in a certain period of time. In turn, this information will be essential for estimating the demand and supply of housing.

This section is divided into four subsections: Subsection C.1 describes the theoretical basis for intraurban migration, and a probability model of binary choice is developed in subsection C.2. Subsection C.3 and C.4 present the empirical work employing various techniques, and subsection C.5 examines the performance of each technique employed.

C.1. Theoretical Basis

As discussed in the previous section, intraurban migration depends on the change in quantity of housing demanded, the supply of housing, and some other factors like job change and access to the job. Thus the functional relationship between the decision to move as an independent variable and changes of income, price of housing, and job location as independent variables may be intuitively obvious to most economists in the framework of consumer utility maximization. Muth's work especially provides an excellent theoretical basis for the analysis of the decision to move. (44).

I summarize Muth's analysis on the equilibrium of the household in urban space, and then I relax some of his assumptions. Muth assumes the prices of all commodities other than housing and transport are the same everywhere in the city. Muth divides all commodities into two mutually exclusive and exhaustive groups, housing and all other commodities except transportation. In Muth's model all employment is concentrated in the CBD. The household is assumed to act in such a way as to maximize an ordinal utility function subject to a budget constraint as is typical in economic analysis of consumer behavior. The utility function can be written as $U = U(x, q)$, where q is consumption of housing and x is dollars of expenditure on all commodities except housing and transportation but including leisure. The budget constraint can be written as $g = x + p(k) + T(k,y) - y = 0$, where P is the price per unit of housing, a function of distance, k_1 from the CBD; T is the cost per trip, a function of k and income, y, multiplied by a given number of CBD trips for the household. By equating the first partial derivatives of the Lagrangian function $L = U - \lambda g$ to zero, the first order conditions for household equilibrium are found. They are:

$$\frac{\delta L}{\delta x} = U_x - \lambda = 0 \qquad (1)$$

$$\frac{\delta L}{\delta q} = U_q - \lambda = 0 \quad (2)$$

$$\frac{\delta L}{\delta k} = -\lambda(qp_k + T_k) = 0 \quad (3)$$

$$\frac{\delta L}{\delta \lambda} = y - \{x + p(k)q + T(k,y)\} = 0 \quad (4)$$

The first two equations together imply the condition

$$U_x = \frac{Uq}{P(k)}$$

which says that the household consumes housing and other commodities in such a way that the marginal utility per dollar spent is the same for all commodities. Equation (3) implies:

$$-qp_k = T_k,$$

which states that in the equilibrium location, the household's net savings on the purchase of a given quantity of housing and transport costs that would result from a very short move would be equal to zero. Equation (4) expresses the condition that the household's expenditure on everything must exhaust his income. The relationships can be seen easily in the following figure.

Figure C.1.1: Equilibrium of the Household

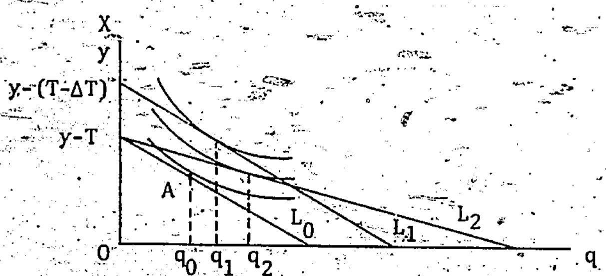


Figure C.1.1 shows the usual indifference curves, the coordinate axes being consumption of housing, q, and X. For a household located at k, the budget line, L_o, intersects X-axis at {y - T(k)} and utility is maximized at point A where the slope of the budget line is equal to the slope of the indifference curve. Thus he consumes (q_o, x_o).

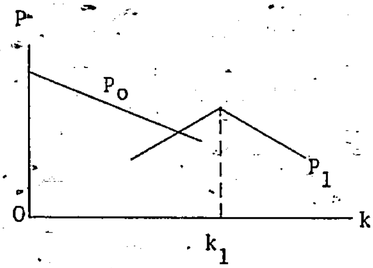
Suppose the household's disposable income after transportation cost increased either by increased income or decreased transportation cost. The budget line will shift up from L_o to L_1 in the figure C.1.1. To maximize his utility, he has to consume more housing like q_1. A similar explanation can be applied to a situation in which income less transportation cost decreased.

If the price of housing decreased due to a deterioration in the local environment like an increase in pollution, or the crime rate, the budget line becomes flatter like L_2 in the figure. The household must consume housing like q_2 in this framework, holding other factors constant and ignoring the moving expense.

A single concentration of employment, CBD, has been assumed so far. Suppose, now, that a secondary concentration of employment (I will call this secondary business district (SBD)) exists along a certain radial from the CBD at a distance k_1 from it, as shown in figure 2.

In figure 2, P_o shows the declining housing prices with distance from the CBD along the radial because transport costs increase. For the same reason, housing prices paid by workers employed at SBD decline from k_1, as shown by P_1.

Figure C.1.2: Effect of SBD

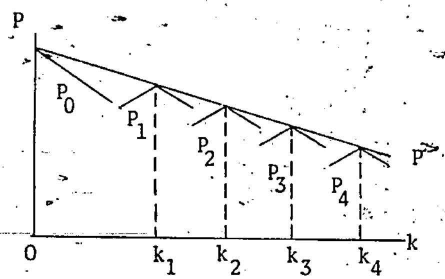


Suppose further that there exist n SBD's, that is, many concentrations of employment. Since I will be using data collected from lower income households in Nairobi where business districts and residential districts are less clearly divided than the cities of developed countries and lower income people work usually outside the CBD, the assumption may not be as unrealistic as the assumption of one concentration of employment.[1] If there were almost continuous employment centers along the radial from the CBD at k_1, k_2, k_3, k_4, assuming k_1 is bigger employment center than k_2 and so on, in the following figure 3, then the general trend of the housing price would be declining with the distance from the CBD as shown by the bold line P, which is somewhat similar to the average cost curve for an industry in the theory of the firm. In short, the housing price depends on the distance from the other employment centers as well as on the distance from the CBD.

---

[1]In my data, only 124 people out of 755 worked in CBD.

Figure C.1.3: Effect of SBD's on Housing Price



Since I am assuming many employment centers, the distance from the job location, j, is the major factor in determining the transportation cost while the distance from the CBD, k, is a minor factor.

The equalibrium conditions under these new assumptions are essentially the same as Muth's. But the differences are that k's are replaced by j's in equations (3) and (4). Thus

$$\frac{\delta L}{\delta j} = -(qP_j + T_j) = 0$$

$$\frac{\delta L}{\delta \lambda} = y - \{x + P(j)q + T(j,y)\} = 0.$$

The last equation is again a mere expression of the condition that the household's income must be exhausted. This condition can be written as $q = \frac{1}{p(j)} \{y - x - t(j,y)\}$.

One can see that any change of one variable holding the others constant in right hand side of the condition will most likely affect the demand of the quantity of housing. This in turn will affect the household's decision to move. Suppose income changed. If the change of income were positive, the household must either improve its present housing or move to better housing in this framework. If the change was negative, the household must either rent a portion of his housing or move to a cheaper housing. Similar things can be said for other variables in the condition.

C.2. The Model

Based on the theory developed in the previous section, it is hypothesized that the decision to move is a function of changes in income, price of housing, and transportation cost. There are some more factors that are intuitively relevant to the model, however. They are age, size of family, transportation costs for children in school, and the family relationship of the household head to his parents.

Age represents the stage in one's life cycle, during the latter part of which the family is more likely to be settled into its final equilibrium location. An increase in family size would tend to reduce family mobility on the one hand, but increase the demand for housing and thus mobility on the other. Transportation cost for children in school is obvious for its relevancy in the model. The household head's relationship to his parents -- whether he is the oldest son or not -- may be a particular variable which can be applied to a model for most underdeveloped countries, where it is a common tradition that the oldest son has the duty of supporting his parents, usually in the same household, who have great influence in the head's decision making.

In doing empirical analysis, I cannot avoid a serious problem with the data. The problem is that my data do not have what I need such as changes in income, price of housing, and transportation cost. Hence I will have to use proxy variables for some independent variables. My initial hypothesis is that the probability of intra-urban migration (1 = household move was made, 0 = household move was not made) is a function of eight explanatory variables: average monthly income, rent for housing, the number of children which shows family size, age of the household head, the relationship of the household head to his parents, the distance from job location to residential location, the sum of the distances from the school locations to residential location and the distance from CBD to residential location. In algebraic form, the model may be written as

P(MOVE) = f(AVY, RENT, KIDH, AGE, RSON, DJH, TDSH, DCH)

where

P(MOVE) = Probability of intraurban migration (MOVE = 1, NOMOVE = 0).

RSON = Household head's relationship to his parents (oldest son = 0, if not = 1) and other variables are as previously defined.

There are four methods -- regression, discrimination, probit and logit analysis -- that can be used for the estimation of models of binary choice. I will employ the first two techniques for which canned computer programs are conveniently available at the University of Oregon Computing Center.

C.3. Regression Analysis

Linear regression analysis is undoubtedly the most widely used technique of all econometric methods. I do not attempt to explain it in any detail in this study.

I estimate the probability function of the model by ordinary least squares (OLS) regression analysis. There are two theoretical specifications that could be tested. The first specification is a linear relationship between the probability of MOVE and the independent variables. The second specification is a multiplicative relationship. But the second specification is not appropriate for this study since some of my independent variables have the value of zero.

In the model, the dependent variable takes the value of either one or zero representing the certainty of MOVE or NOMOVE and, thus, the conditional expectation of the dependent variable may be interpreted as the conditional probability of MOVE. But there are two major drawbacks in using regression analysis to estimate binary choice. First, the predicted value of the dependent variable can be greater than one or less than zero. Watson pointed out this problem clearly: "In cases where predictions form an integral part of the analysis, the potential inconsistency of the predictions with the probability interpretation of the dependent variable is a serious obstacle to the use of regression analysis techniques."[1] This problem can be

---

[1]Peter L. Watson, "Choice of estimation procedure for models of binary choice," Regional and Urban Economics, 4, 1974, p. 189.

minimized, however, by specifying the model correctly and using a sufficiently large sample.

Second, the model violates the assumption of homoscedasticity of the regression model, i.e., the assumption that all the disturbance terms of each vector of observations have the same variance. This problem of different variances, namely heteroscedasticity, comes from the limitations of the dependent variable which in turn lead to restrictions on the values that may be taken by the disturbance terms. There are two solutions that give the same result to this problem of heteroscedasticity: a weighted least squares solution (WLS) and a solution by transformation. Both of the techniques are sesigned to satisfy the assumption of homoscedasticity. The details of these techniques are discussed in appendix I. I prefer to use the weighted least squares solution since, in the computer program (RAPE) I use, it takes less generation of new variables than the solution by transformation.

The model is estimated by OLS and WLS, and the results are compared. I employ eight explanatory variables first and some variables are replaced and added to examine the structural change of the model. Table G.3.1 presents the first equation for the probability of intraurban migration estimated by each technique.

I am interested in the closeness of the fit of the model and the t statistic of each coefficient as are most researchers who use regression analysis. The goodness of the fit represented by the coefficient of determination (R^2) is a precise measure of the strength of the model, while the t statistic shows the appropriate level of a

Table C.3.1: Model of Intraurban Migration (Equation 1)

Probability of Migration

	OLS		WLS	
	Coefficient	t-ratio	Coefficient	t-ratio
AVY	-0.00001	-0.256	0.00002	0.355
RENT	-0.0009	-2.916	-0.0011	-3.663
KIDH	-0.0217	-2.486	-0.0267	-3.778
AGE	-0.0129	-8.807	-0.0150	-11.552
RSON	0.0721	2.354	0.0580	2.172
DJH	0.0044	41460	0.0043	4.734
TDSH	-0.0045	-3.564	-0.0054	-4.612
DCH	-0.0040	-4.746	-0.0049	-7.875
Intercept	0.8443	12.796	0.9441	14.402
Degrees of freedom	746		690	
R^2	0.1910		0.2695	

variable in the model.

The R^2s by both techniques are reasonably good. OLS explains 19 percent of the total variation of the model while WLS explains approximately 27 percent while every parameter is significant at the 5 percent level except AVY. Between the two techniques every parameter except RSON obtained by WLS shows higher t ratios than by OLS. The magnitudes of parameters except RSON are larger in WLS model than in OLS. Based on the theoretical reasoning of homoscedasticity, the better R^2 and the t statistics of WLS than those of OLS, the magnitudes of parameters obtained by WLS are more reliable.

Finally, notice the decrease of the degrees of freedom in WLS model in comparison with OLS model. This change comes from the problem of the zero-to-one-probability range. Since the variance, which is used as the weight in WLS, has the form of $\sqrt{1/p(1-p)}$ where P is the predicted value, it becomes an imaginary number if the predicted value is negative or greater than one. The computer simply gives zero values to imaginary numbers. Hence these weights of zero reduce the sample size and the degrees of freedom.

All of the parameter estimates except AVY have signs that can be reasonably explained. The older the household head, the less likely the family is to move, an expected result since greater age corresponds to later stages in the life cycle during which the family is likely to be settled down. A similar explanation applies to the rent variable. An increase in family size would tend to reduce family mobility on the one hand, but increase the demand for housing and thus mobility on the other. The coefficient on KIDH indicates that the first effect dominates.

The positive sign of DCH parameter suggests that the farther a household is from the CBD, the less likely it will change location, which again can be explained by stages in the life cycle. As incomes rise with the life cycle, the household's demand for housing may include a location farther from the CBD, and the greater this distance the more likely the household has arrived at its final equilibrium position. The negative effect of TDSH indicates that those people who have children going to school are less likely to move. If the household moves farther away from the school the children either have to make longer trips to school and, thus, increase the transportation cost, or have to take the inconvenience of being transferred to another school.

The parameter of RSON has the expected positive sign, since the oldest son has the duty of supporting his parents, who have great influence in household head's decision making as is a common tradition in most underdeveloped countries. The effect of the distance from job location is positive as expected. This indicates that the farther a household is located from employment location, the higher is the probability of moving his residential location. This seems to be a reasonable effect in a city where most of the residents either walk or take mass transit to their job.

The parameter of AVY is not only very small in comparison to that of rent but also it has an insignificant t-ratio. Also note that the sign of it is negative in OLS but positive in WLS. Thus I cannot reject the null hypothesis, the coefficient of AVY is zero, and conclude that there is not sufficient evidence to indicate that AVY is linearly related to the probability of move. It should be pointed out, moreover, that income is highly correlated with rent as reported in most studies of housing demand. The simple regression of RENT on AVY also reveals a correlation coefficient of 0.563. There results lead me to drop the variable AVY from the model in the next regressions (Equation 2) presented below.

Table C.3.2: Model of Intraurban Migration (Equation 2)

	Probability of Migration			
	OLS		WLS	
	Coefficient	t-ratio	Coefficient	t-ratio
RENT	-0.0009	-3.640	-0.0010	-4.263
KIDH	-0.0221	-2.551	-0.0264	-3.899
AGE	-0.0129	-8.809	-0.0149	-11.69
RSON	0.0719	2.349	0.0597	2.262
DJH	0.0044	4.474	0.0043	4.765
TDSH	-0.0046	-3.622	-0.0049	-4.832
DCH	-0.0040	-4.772	-0.0049	-8.022
Intercept	0.8417	12.91	0.9392	14.67
Degrees of freedom	747		693	
R^2	0.1909		0.2755	

The results are quite encouraging. While there are insignificant changes of R^2 and t-ratios in OLS model from equation 1, every t-ratio increases and R^2 improves in spite of using one less variable in WLS model. The magnitudes of parameters are almost the same as in the previous regressions using both techniques. It appears that the models of equation 2 are better than those of equation 1 in explaining the probability.

Whitelaw hypothesized that the change in job location is a major explanatory variable in his probability model (61). Using a dummy variable, he found the role of this variable to be significant. I have added the same dummy variable, JOB, representing change of job location to the first equation (yes = 1, no = 0). The second equation presented in Table C.3.3 shows the effects of adding JOB to equation 2.

Table C.3.3; Model of Intraurban Migration (Equation 3)

	Probability of Migration			
	OLS		WLS	
	Coefficient	t-ratio	Coefficient	t-ratio
RENT	-0.0008	-3.212	-0.0007	-2.818
KIDH	-0.0212	-2.462	-0.0306	-4.423
AGE	-0.0125	-8.477	-0.0087	-7.113
RSON	0.0698	2.287	0.0871	3.194
DJH	0.0034	3.257	0.0005	0.5212
TDSH	-0.0044	-3.543	-0.0044	7.457
DCH	-0.0039	-4.656	-0.0037	-6.716
JOB	0.1154	2.580	0.3203	7.457
Intercept	0.8126	12.32	0.6813	10.44
Degrees of freedom	746		696	
R^2	0.1918		0.2584	

The addition of the new variable, JOB, does not make substantial changes in the structure of OLS model but it does in WLS model. The following results are noticeable in WLS model. First the parameter of DJH becomes insignificant, thus, suggesting multicollinearity between JOB and DJH while other parameters are all significant at a 2 percent level. Second, R^2 drops .02. The highly significant t ratio of JOB, however, strongly supports the important role of the variable, JOB, in the model. Also notice the magnitude of the parameter, 0.3203. This implies that the decision of more than 32 percent of the households who moved in the sample was based on the job location change alone. There is no doubt, now, about the importance of the variable, JOB, in the model. But the question is whether the mere fact that a household head changes his job location should influence his decision to move or not. Suppose a household head found a new job not far from his old job, say, one block away. Would he still move? If other variables are held constant it seems to be reasonable that he is not likely to move.

In my next equation, I introduce a new variable: absolute value of the change in DJH due to job change, ΔDJH[1], which reflects the change in transportation cost. For those household heads who did not change their jobs or changed jobs but in the same grid square, the value of ΔDJH is zero. I replace ΔDJH with the dummy variable JOB in the following regressions.

The results of this regression are as good as those of equation 3. The newly added ΔDJH also appears to be a significant variable at 5 percent level and the R^2 (0.2679) by WLS appears to be much more

---

[1] $\Delta DJH = |(DJH)_{new} - (DJH)_{old}|$

accurate than that of other studies of binary choice problem.[1]

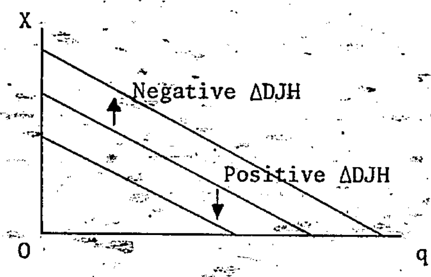
Table C.3.4: Model of Intraurban Migration (Equation 4)

	Probability of Migration			
	OLS		WLS	
	Coefficient	t-ratio	Coefficient	t-ratio
RENT	-0.0008	-3.283	-0.0009	-4.267
KIDH	-0.0207	-2.404	-0.0265	-3.807
AGE	-0.0126	-8.630	-0.013	-10.71
RSQN	0.0711	2.336	0.0566	2.183
DJH	0.0018	1.449	0.0016	1.535
TDSH	-0.0044	-3.526	-0.0048	-4.805
DCH	-0.0038	-4.581	-0.0046	-7.201
ΔDJH	0.0039	3.365	0.0024	2.265
Intercept	0.8380	12.94	0.9092	14.17
Degrees of freedom	746		692	
R^2	0.2030		0.2679	

[1]The statistical fit obtained by Stanley Warner in his study was R^2 of .167. See Stochastic Choice of Mode in Urban Travel: A Study in Binary Choice (1962, Northwestern Press), p. 60. For more examples of statistical fits of studies on binary choice problems see (34).

Since, among the variables I employ, ΔDJH is the only variable based on the theory discussed in C.1, I expected better results than what I obtained in equation 3. One thing I neglected when this variable was introduced is the direction of ΔDJH, whether one's new job is closer or farther than previous job from one's home. All the changes in job locations are treated as positive values (or zero if the changes were made in same grid) even when one found a job closer to his residential location than his old job was. But the effects of ΔDJH, which is a proxy for the change in transportation cost, on the budget line is opposite depending on the direction of ΔDJH, positive or negative value as shown in the following figure.

Figure C.3.1: Effect of ΔDJH on Budget Line



Positive ΔDJH (increased transportation cost) will shift down the budget line while negative ΔDJH (decreased transportation cost) will shift up the budget line. Both effects should increase the probability of move according to the previously discussed theory. But do people react the same way to both decreased budget and increased budget?

I now suggest a hypothesis that people adjust more rapidly to equilibrium when their disposable income increases than when their income decreases in Nairobi. In other words people move presumably to better housing when income increases due to a decreased transportation cost but they are reluctant to move when income decreases due to an increased transportation cost. In my next equation ΔDJH is replaced by ΔDJH^*[1] in which the direction is implied, i.e., if one's new job is closer or farther than one's previous job, then the value of ΔDJH is negative or positive.

The results of the regression are presented in the following table. By replacing this variable, I get better results. First notice the improvement in the fit of each model. R^2 increased 22 percent in OLS and 57 percent in WLS. Second all the parameters are significant at 1 percent level except RSON which is significant at 5 percent level. DJH which was insignificant in WLS of equation 3 and 4 is also highly significant. Now the model (WLS) is able to explain more than 42 percent of the variation.

The strict interpretation of the coefficient of ΔDJH^* would be that when people change to a closer job location than the previous one, the probability of changing residential location increases but, when the new job location is farther than the previous location, the probability decreases. The latter part of the interpretation, i.e., the case of further job location, sounds paradoxical. It would seem a realistic view that a person would be likely to move due to an

---

[1] $\Delta DJH^* = (DJH)_{new} - (DJH)_{old}$.

Table C.3.5: Model of Intraurban Migration (Equation 5)

Probability of Migration

	OLS		WLS	
	Coefficient	t-ratio	Coefficient	t-ratio
RENT	-0.0009	-3.659	-0.0008	-3.654
KIDH	-0.0210	-2.407	-0.0232	-3.418
AGE	-0.0121	-8.485	-0.0139	-10.84
RSON	0.0739	2.498	0.0561	2.205
DJH	0.0099	8.207	0.0104	10.78
TDSH	-0.0049	-4.009	-0.0049	-4.848
DCH	-0.0038	-4.009	-0.0049	07.794
ΔDJH*	-0.0079	-7.444	-0.0083	-7.948
Intercept	0.7369	11.42	0.8146	11.94
Degrees of freedom	746		679	
R^2 ($\bar{R}^2$)	0.2469 (0.2388)		0.4204 (0.4136)	

increased transportation cost when his new job location is drastically farther than the previous job location.

It was already discussed that both directions of ΔDJH would increase the probability of intraurban migration. The sign of the parameter of ΔDJH* seems to indicate that a negative ΔDJH* increases the probability of a move more than an equal positive ΔDJH*, however.

The following table presents additional information on ΔDJH*. It shows how low income level Nairobians reacted to their new job locations in 1970.

Table C.3.6: Distribution of JOB and ΔDJH*

	MOVE	NOMOVE	TOTAL
JOB	60 (50.4%)	59 (49.6%)	199 (100%)
ΔDJH* (+)	26 (32.1%)	55 (67.9%)	81 (100%)
ΔDJH* (0)	0	4 (100%)	4 (100%)
ΔDJH* (-)	34 (100%)	0	34 (100%)

The first line of the table shows that about half of the people who changed job locations changed the location of their residence. Among them, only about 32 percent of the people whose new job locations were further than previous ones moved while those who changed job locations but in the same grid square ΔDJH* (0) did not move at all. But every household whose new job locations are closer changed their residential locations. It could be coincidence, yet it obviously suggests the general trend of moving affected by ΔDJH*. In fact ΔDJH* along is able to explain 23 percent of the total variation of the probability of MOVE.[1]

In summary, ΔDJH* appears to be the most important variable in explaining intraurban migration in this model. I will examine the variables used as independent variables again in using discriminant analysis.

---

[1]The simple-WLS regression of ΔDJH* on P(MOVE) is:

$$P(\text{MOVE}) = \underset{(16.88)}{0.2958} - \underset{(-15.07)}{0.0031}\ (\Delta \text{DJH*})$$

$$R^2 = 0.2318, \quad \text{D.F.} = 732$$

The numbers in parentheses are t ratios.

C.4. Discriminant Analysis

Discriminant analysis is a relatively less frequently used technique in the literature of economics compared to other disciplines like business administration, especially marketing analysis. The objective of a discriminant analysis is to define a set of equations that efficiently discriminate among two or more mutually exclusive populations or groups. Thus this technique is one of the most suitable techniques for the problem of binary choice.

In this model, one discriminant equation is constructed for the MOVE group and the NOMOVE group respectively as a linear function of discriminating variables. The general form of the discriminant equation associated with group i is

$$Z_i = \lambda_{io} + \lambda_{i1}X_1 + \ldots + \lambda_{ip}X_p$$

where Z_i is a discriminant score for group i, and the constant i_j is a coefficient of jth variable, X_j in group i.

I employ the same eight explanatory variables used in the regression model (equation 5) as discriminating variables. They are RENT, KIDH, AGE, RSON, DJH, DSH, DCH, AND ΔDJH*. Although I use the same variables in both models, there is a basic difference betweeen a discriminant model and a regression model. A discriminant analysis does not have a measureable counterpart to the dependent variable in regression analysis. It simply employs a multiple classification analysis of the data, and the Z_i values are simply group indicators. No attempt will be made, however, to explain the basic theory underlying discriminant analysis in this section for there are excellent descriptions on discriminant analysis available elsewhere (43, 46, 47).

I apply the data to the BMD07M[1] Stepwise Discriminant Analysis program. The results are presented in the following three tables.

Table C.4.1 gives the values of F to enter[2] which show the ability of each discriminating variable individually to classify the data.

Table C.4.1: F to Enter

RENT	6.41	KIDH	25.41	AGE	85.66	RSON	4.57
DJH	19.27	TDSH	17.80	DCH	14.82	ΔDJH*	10.52

Degrees of freedom 1,753

Since the critical F with 1 and 753 degrees of freedom for an allowable $\alpha = 0.05$ is about 3.85, all the variables are said to discriminate adequately between the groups in the analysis at the 5 percent level of significance.

Table C.4.2 presents the discriminant equations and U statistic.[3]

---

[1] W.J. Dixon, ed., BMD: Biomedical Computer Programs, Berkeley: University of California Press, 1973.

[2] F to enter values are statistics for testing the hypothesis
H_o: The association discriminating variable does not adequately classify the original data set, versus
H_a: It is a significant classification variable when taken alone. The largest F to Enter most strongly rejects the hypothesis and identifies the first variable to enter the model.

[3] The U statisitc tests the hypothesis

H_o: The variables which have been entered do not as a group adequately discriminate among the classes, versus H_a: At least one variable discriminates adequately among the classes.

Table C.4.2: Discriminant Functions

Variable	NOMOVE	MOVE
RENT	0.0299	0.0241
KIDH	0.0706	-0.0667
AGE	0.4134	0.3345
RSON	2.5511	2.9902
DJH	0.0763	0.1404
DSH	0.0771	0.0453
DCH	0.0736	0.0491
ΔDJH*	-0.0334	-0.0846
Constant	-11.417	-9.7129

Total sample size (N) = 755: n_1 = 536, n_2 = 219

U Statistic = 0.7498 Degrees of Freedom 8,1,753
Approximate F = 31.1* Degrees of Freedom 8,746
*significant at 1 percent level.

The interpretation of the discriminant coefficients is more or less similar to the means of two groups discussed in section B. The general tendencies of the MOVE group are that they are younger, pay lower rent, have fewer children, not likely to be oldest sons, live farther from CBD and their job locations, spend less time or money for journey to school, and find closer jobs than the NOMOVE group.

A complete table for the U statistic is not available but the U statistic can be transformed to a statistic which has an approximate F distribution. Since the Approximate F, 31.1, exceeds the critical F with 8 and 746 degrees of freedom at the 1 percent level, 2.53, I can say the variables as a group adequately discriminate between the groups.

Table C.4.3 presents the confusion matrix which summarizes the number of correct and incorrect classification of the original data set by the discriminant equations.

Table C.4.3: Confusion Matrix

Actual Group		Predicted Group	
		NOMOVE	MOVE
NOMOVE	536	506	30
MOVE	219	128	91

The confusion matrix shows that the discriminant equations perform well for NOMOVE group but poorly for MOVE group. They correctly classify 94.4 percent in NOMOVE group and 41.6 percent in MOVE group. Overall 597 cases or 79.1 percent of the samples are correctly classified. With this information I can now use a chi-square test by computing Q statistics[1] to determine the overall effectiveness of the discriminant model to classify the data set as opposed to chance classification of this set. The obtained statistic,

---

[1]Under the following hypotheses:

H_o: Discriminant model classification no better than randon classification, versus

H_a: Discriminant model performed better than chance, the statistic

$$Q = \frac{(N - mt)^2}{N(t-1)}$$

is distributed as a chi-square random variable with one degree of freedom, where

N = Total sample size
t = Group size
m = Total number of correct classification.

255, also far exceeds the significant level ($\chi^2_{0.005}$, 1 d.f. = 7.88) confirming that the discriminant model classifies the data set better than could be expected by chance.

The classification of the data will be discussed again in the next section along with the discussion of the classification by regression models.

C.5. Performance Tests of the Methods

Three methods of estimation, OLS, WLS, and discriminant analysis, were applied to the model developed in C.1. All of them appear to give significant estimates. It is difficult, however, to suggest the best method for empirical work among them without certain criteria. To evaluate the three methods of analysis, the performances of each method should be tested and compared with each other.

For each method two different tests -- the classification test and the prediction test -- are carried out. In the classification test, the probability of MOVE for each household is predicted by each technique using all the observations. Then the household heads were assigned to either MOVE or NOMOVE on the basis of P = 0.5 cut-off point. That is, a household head is classified as MOVE if the predicted probability is greater than 0.5 and as NOMOVE if the value is less than 0.5.

The prediction test is essentially the same as the classification test except that half of the data is used to generate the equation, and the probability for each household in the other half is predicted by this equation. The results of the classification test are presented in table C.5.1.

Table C.5.1: Classification Matrix

	SAMPLE	Percent	OLS	Percent	WLS	Percent	Disc.	Percent
NOMOVE	536	100	492	91.7	480	89.5	506	94.4
MOVE	219	100	103	47.0	117	53.4	91	41.6
Total	755	100	595	78.8	597	79.1	597	79.1

OLS classifies 492 cases (91.7%) out of 536 original observations correctly for NOMOVE group and 103 cases (47%) out of 219 original observations correctly for MOVE group and in total, 595 cases (78.8%) are correctly classified. Both WLS and discriminant analysis correctly classify 597 cases (79.1%), 2 cases more than OLS does in total while WLS classifies better than discriminant analysis does, 53.4 percent and 41.6 percent respectively, for MOVE group. Since my interest is in the MOVE group, WLS appears to be the best technique for classification among the three methods.

Given that WLS is the best technique among them, I run another classification test which considers prior probability. Since I know that there are 536 households who did not move, I order all the predicted values from the smallest one and set the 536th predicted value (0.3980) as the cut-off point.

I also employ a prediction test in which half of the sample is used to generate the equation. The sample is divided into two halves randomly by the built-in function of the computer. The coefficients estimated using the first half sample are then applied to the other half of the data set to produce estimates of the probability of each observation. The equations[1] generated from the first half of the data set are more or less the same as the equations from the entire data set.

The results of the tests, that predict the second half data set which contained 101 cases of MOVE and 276 cases of NOMOVE in the discriminant model, are presented in Table C.5.2.

Table C.5.2: Prediction Matrix

	OLS	Percent	WLS	Percent	Disc.	Percent
NOMOVE	269	97.4	260	94.2	253	94.8
MOVE	44	43.6	49	48.5	53	48.1
Total	313	83.0	309	82.0	306	81.2

The performances of prediction are slightly better than those of classifications in total percentage. But predicting MOVE is substantially better than classifying MOVE by all models. Among the three models, discriminant analysis gives less accurate results in overall performances. Thus it appears that regression models are better techniques to be applied than discriminant model in this study. I also prefer to use WLS to OLS not only because it gives better prediction and classification for MOVE group, but also because the theoretical reasoning is more sound, since heteroscedasticity is

---

[1]See Appendix II.

corrected for, although OLS performs almost as well as WLS in classification and even better in total prediction.

D. DIRECTION AND DISTANCE OF INTRAURBAN MIGRATION

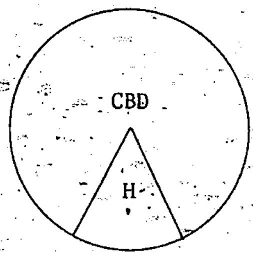
In the previous section it was shown that the probability of intraurban migration is a measurable behavior of urban households. Suppose a household is predicted to have a high probability of MOVE. Then where would it move to? Is intraurban migration a chaotic phenomenon or is there a systematic spatial pattern? If the latter is true, identifying relevant variables that determine distance and direction in intraurban migration would provide a useful guide for urban planning. The object of this section is to analyze the geographic aspects of the moving behavior of African households. I investigate the directional bias first and the discussion on distance bias follows next.

D.1. Direction of Intraurban Migration

Researches on directional bias in intraurban migration have been reported many times by researchers in different branches of social science (1, 6, 7). Among the studies, Adams' report on directional bias of intraurban migration in Minneapolis is particularly interesting, introducing the seemingly sound concept of mental map. Adams argues that a resident develops a mental map or perceptual image of his city based largely on his daily and weekly activities.

He argues further that the resident retains a relatively narrow mental map of his urban area like the pie slice shaped area shown in figure D.1.1 where H is the residential location.

Figure D.1.1: Mental Map

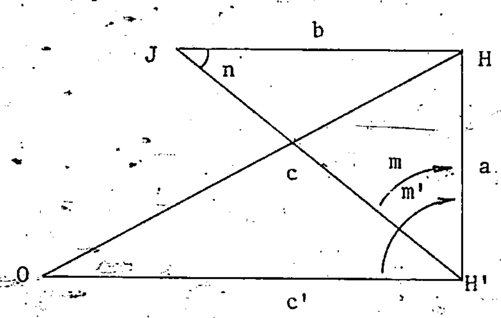


Based on his hypothesis, he suggested three attributes that describe an intraurban move. They are (i) the length of move, a, (ii) the distance of the origin, H', from the center, O, of the central business district, and (iii) the angle, m', of the move with respect to the location of the downtown center in the following figure.

According to Adams, the angle m', associated with intraurban moves would be a bimodal clustering around 0 degrees and 180 degrees. Using data from Minneapolis city directories, he confirmed the directional bias. Whitelaw, however, found nothing to support Adams' hypothesis in his study on the Nairobi households (61).

Although Adams' assumption of mental map sounds reasonable, his approach seems to have a drawback since he assumes that CBD is the

Figure D.1.2: Intraurban Move



H = Location of Present House

H' = Location of Previous House

J = Location of Job

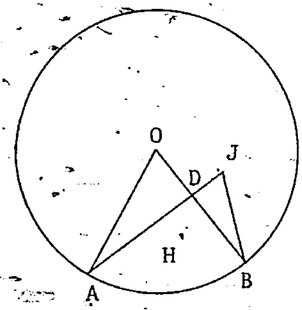
O = Center of Central Business District

center of daily activity for every household. A large city like Minneapolis, however, must have more than one business district.

Suppose Mr. K resides at point H' and has a job at the second business district like point J, instead of CBD, point O. Then he may have a mental map of AODJB shown in figure D.1.3. But the image of AJB should be stronger than that of AOB since he goes more often to his job location than to the downtown.

Given that the revised mental map and the hypothesis of bimodal clustering are correct, then the relevant attributes are distance c and angle m instead of distance c' and angle m' in Figure D.1.2. It is not surprising that Adams' assertion was supported by weak

Figure D.1.3: Revised Mental Map



empirical evidence in his report and no support at all from Whitelaw who analyzed angle m'. My investigation on 206 samples which made intraurban migration, however, does not support Adams' bimodal clustering in terms of angle m either.[1] The frequency of the samples is almost evenly distributed all over 10 degree intervals except the interval from 170 degrees to 180 degrees as presented in Table D.1.1.

---

[1]To calculate the angel, m, of the move in degrees, we first compute the cosine of m in radians by the following formula:

$$\cos m = \frac{a^2 + c^2 - b^2}{2ac}$$

and find m by taking the arccosine of cosine m, then multiplies the result by $180/\pi$ to transform the angle to degree.

Table D.1.1: Frequency Distribution and Proportion of Angle m

Angle	Freq.	Percent	Angle	Freq.	Percent	Angle	Freq.	Percent
0-10	9	4.4	67-70	9	4.4	121-130	10	4.9
11-20	20	9.7	71-80	11	5.3	131-140	10	4.9
21-30	14	6.7	81-90	9	4.4	141-150	7	3.4
31-40	11	5.3	81-90	12	5.8	151-160	12	5.8
41-50	8	3.9	101-110	13	6.3	161-170	24	11.6
51-60	16	7.7	111-120	11	5.3	171-180	0	0.0
0-60	78	37.9	61-120	65	31.6	121-180	0	30.6

Now the question is whether or not the angle m or m' is a relevant variable to investigate for verifying the validity of mental map hypothesis. In figure D.1.3, Mr. K could move any angle from 0 degrees to 360 degrees without violating Adams' hypothesis. Clearly the relevant angle is not the angle m but n in Figure D.1.2. The following table presents the distribution of angle n.

Table D.1.2: Frequency Distribution and Proportion of Angle n.

Angle	Freq.	Percent	Angle	Freq.	Percent	Angle	Freq.	Percent
0-10	84	40.8	61-70	3	1.5	121-130	2	1.0
11-20	32	15.5	71-80	0	.0	131-140	2	1.0
21-30	19	9.2	81-90	1	0.5	141-150	2	1.0
31-40	22	10.7	91-100	1	0.5	151-160	5	2.4
41-50	11	5.3	101-110	1	0.5	161-170	12	5.8
51-60	5	2.4	111-120	4	1.9	171-180	0	0
0-60	173	83.9	61-120	10	4.9	121-180	23	11.2

More than 40 percent of the intraurban migration occured within a 10 degree span, and almost 84 percent of the sample falls in 60 degrees. This skewed distribution appears to suggest that people do have a perceptual image of the urban area constructed along their journey-to-work line and the migrants' behavior of searching new housing is constrained within their mental maps.

D.2. Distance of Intraurban Migration

One's decision of how far to move may depend on the supply situation of housing in a particular geographic area, presumably within his mental map. Since the situation of housing supply in Nairobi is not known to me, I will simply assume that housing for rent is evenly available all over the city.

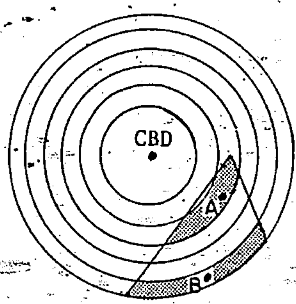
The primary reason for intraurban migration is, as discussed in C.1, the change in one's budget. But the change in income is probably gradual rather than drastic. Then the new housing one is looking for is more or less at a similar price level to that of present housing.[1] Since the price of housing depends upon the distance from CBD as previously discussed, I can now suggest a hypothesis that the distance of move depends on where he lives, i.e., how far he lives from CBD. According to this hypothesis, household A's (or B's) probable new residential location is in the shaded area where the price of housing is in the neighborhood of the present housing price

---

[1]The income elasticity of housing is reported to be around 1. See F. de Leeuw (5).

in the following figure, given that they have the same mental map.

Figure D.2.1: Probable Area of Move



Besides the distance from CBD (DCH), there may be some other factors that affect the distance to move. Rent is one of them. Since high rent housing is relatively more scarce than low rent housing, the distance of move of high rent payers will probably be further than that of low rent payers.

Finally the change in DJH (ΔDJH) seems to be an appropriate variable when one changes his job location because it will directly affect his taste.

The regression equation for the distance of migration, obtained by the method of ordinary least squares on 219 observations, i.e., those who changed residential location in 1970, is presented in Table D.2.1.

The parameter of RENT is significant only at the 10 percent level but others are significant at the 1 percent level. The fit of the model is reasonably good: $R^2 = 0.41$ and all the signs are as expected. Thus the model supports the previously discussed argument.

Table D.2.1: Determinants of Intraurban Migration Distance

	Coefficient	t-ratio
RENT	0.0191	1.892
DCH	0.8517	11.89
ΔDJH	0.0582	2.618
Intercept	-5.776	-4.580
Degrees of freedom	215	
R^2	.4138	

The parameter of DCH is the main concern of this subsection. The parameter (0.81) is striking in both its statistical significance and its relative weight in the equation. It suggests that the effect of DCH is most significant in affecting the distance of MOVE after accounting for the difference in RENT.

I summarize the results of this study and suggest areas for further research in the next section.

E. SUMMARY AND CONCLUSION

The previous sections discussed three major aspects of intraurban migration behavior -- who, why and where -- in the city of Nairobi, Kenya.

The characteristics of movers in section B show that they appear to be relatively younger and poorer, thus pay lower rent but have a higher income-rent ratio than non-movers. They have fewer children and live closer to downtown but further from their job locations. About 50 percent of the household heads who changed job locations also changed their residential locations, while only 21 percent of the household heads who are oldest sons to their parents did so.

A theoretical model was developed to investigate the determinants of intraurban migration in section C. Although the model, which assumes that intraurban migration is a function of the change in income, price of housing and transportation cost, is a simple extension of consumer theory, it clearly shows why people move.

In developing the empirical model, I have employed proxy variables for the change in income, price of housing and transportation cost and implemented the model with the variables that appear to be relevant from the analysis of section B. That is, the empirical model assumes that the probability of intraurban migration is a function of income, rent, number of children with household head, age of household head, distance from job location to residential location, distance from CBD to residential location, distance from school to residential location multiplied by the number of children in school and the household head relation to his own parents. The probability function of the model is estimated by ordinary least squares regression and weighted least squares regression under the specification of linear relationship between the probability, which takes the value of either one or zero, and explanatory variables.

The results of both regressions showed poor fits. Due to the apparent multicollinearity problem between income and rent, and insignificant t ratio of income, AVY was dropped from the model. This improved the fit of the model in spite of using one less variable. In succeeding regressions, three variables of the same nature: job change (JOB), absolute value of change in distance from job location to residential location (ΔDJH) and signed change in distance from job location to residential location (ΔDJH*) were added to the previous equation (equation 2). Addition of JOB or ΔDJH resulted in a significant change of the fit but the addition of ΔDJH* caused a much improved fit with all significant t ratios of the independent variables in WLS.

I also employed discriminant analysis using the same independent variables as wereused in regression analysis (equation 5). Each variable shows a significant F value at the 5 percent level, i.e., all the variables are significant classification variables. The U statistic and Q statisitc also confirm that the discriminant model classifies the data set better than could be expected by chance.

Performance tests, i.e., the classification test and the prediciton test were carried out for the three techniques employed, OLS, WLS and discriminant analysis. Each technique performed well compared to other studies of binary choice. The best performance was shown by WLS which correctly classified 79.7 percent and predicted 82.0 percent.

I prefer WLS to OLS on theoretical grounds, and prefer WLS to discriminant analysis on grounds of better classification and prediction.

The relatively low explanatory power of the model (equation 5, $R^2 = 0.42$) may be due to some combination of measurement error, omitted variables or poor specification of the relationships. But this fit is much better than that of previous studies of similar nature. Also a high R^2 may not be expected in a cross-section study as Theil argued in his textile example that "when we consider cross-section data and run a regression for textile expenditure by individual households, we should expect an $\bar{R}^2$ which is much smaller, say .5 or even less." (56, p. 181).

I have demonstrated that the probability of intraurban migration is not something unmeasurable. It is found that economic variables appear to be important factors in intraurban migration. My analysis supports Whitelaw's denial of Simmons' claim that "all studies reject job location as an important reason for moving." (52,60). I also question the report by Steines and Fisher that employment has little effect on residential location (53). On the other hand, my findings generally follow the implication of recent theoretical and empirical analysis of Muth's residential location in urban areas. One difference is that people seem to adjust their housing when disposable income increases but are reluctant when their income decreases as suggested in equation 5 of the regression model.

Directional and distance biases in intraurban migration were examined in section D. For directional bias, Adam's "mental map" hypothesis is criticized and revised. The empirical findings indicate that his basic hypothesis appears to be valid, although his approach is not quite convincing. Under the assumptions developed in section C, a model of moving distance is formulated. The model suggests that the moving distance depends on the distance from the CBD. Combining the directional and distance biases, Simmons' claim[1] that: "The best factor(s) to predict the location of a new residence is (are) the location of the former house (and his job location)" (52) may have some truth in it if the words in parentheses are added.

The most significant findings of this study are that the change in journey-to-work distance is the key determinant of the probability of intraurban migration while the distance from CBD to residential location is the key determinant of the distance of intraurban migration.

Including these two findings, I consider the verification of the

---

[1]This claim was rejected by Whitelaw (60, p. 22).

determinants in the two models, the probability model and the distance model, as the major contribution of this study. Perhaps I can now safely say that the decision of intraurban migration substantially depends on the change in economic factors, but that where to move depends on the taste of the household.

The research undertaken in this study has been based on the theoretical and empirical work of previous investigations of intraurban migration. The emphasis of the study was placed particularly on economic factors, and reasonably good results were obtained in empirical analysis. It is clear, however, that only a portion of the complex problem has been explained in this study. In further research, the following suggestions are offered:

First, other relevant factors besides economic factors such as social characteristics, perception of the urban people, and physical characteristics of the urban area should be included both in theoretical and empirical analysis. Second, more accurate data are needed rather than proxy variables. It is my feeling that I could have obtained better explanatory results in the probability model had I possessed data on the change in income. It would be also valuable if one could follow the movement of individual families within an urban area over time. Third, probit analysis or logit analysis is recommended in the estimation of the probability model since both of them are theoretically sound for binary choice problems.

II. SCHOOLING, EARNINGS, AND EXPERIENCE

A. INTRODUCTION

Education produces dual benefits. One can expect to receive extra earnings in later life and intellectual satisfaction as a result of having gone to school. The first benefit makes education an investment good, while the second makes it a consumption good.

That education is an investment good has been central to the extensive literature on human capital or economics of education in the past 15 years. Education has been found to be an important factor for the long run economic growth of the economy (12, 66). Further, it has been found to affect the structure of wages and thereby the structure of relative earnings (65). In developing a rational public policy toward investment in educational and manpower planning, it is argued that the rate-of-return approach can be crucial because it tests the worth of education in the market system (5, 45). The major portion of the research on the economics of human resources is found to follow this line of reasoning (2,8,10,20,21,26).

Like the studies mentioned above, this study is concerned with schooling as an investment in Nairobi, Kenya. For a country which was listed 64th among 75 countries ranked according to development of human resource,[1] the importance of effective educational planning is obvious. An appropriate evaluation of the rate of return to education is a prerequisite to such an effective planning.

The rates of return to education suggested by earlier studies vary so widely that one is tempted to question the theory, or its application, or both. Estimates of the return to secondary schools range from 11.4 percent (Japan) to 36.5 percent (Mexico) and widely varying figures also are reported in the United States.[2]

Although there is no reason to believe that estimates in different areas should be similar, the substantial disparity in these estimates can be attributed to the different specification of biased earnings functions. A number of well known studies have been carried out under the assumption that earnings are a function of years of schooling, age, and some socio-economic characteristics. Lately the role of mental ability differences in an earnings function has been considered in several investigations (19,54). It has been argued that a measure of the contribution of education to income that ignores the mental ability variable will be biased upward if education and ability are positively related. This argument seems theoretically sound although some empirical results using an intelligence quotient or some other kind of test score as a measure of mental ability suggest no

---

[1]F.H. Harbison and C.A. Myers, Education, Manpower and Economic Growth, New York, McGraw-Hill Book Company, 1964, p. 33.

[2]See Appendix III.

significant bias (19).

If the major determinants of income are indeed the years of schooling and IQ, why does a brilliant recent college graduate not draw a salary as high as an old executive? The answer is not that he is simply young, but that he has a lack of experience and seniority, even in a case where demand for college graduates is high.

It is surprising that this obvious relationship between income and experience was not considered in the analysis of earnings function until J. Mincer suggested it only recently (39). Even after his suggestion, there has been little systematic attempt to analyze the impact of experience in the estimates of the rate of return to education or earnings function. The role of experience is often ignored or considered minor in earlier literature. If experience is one of the major determinants of income, omitting such a variable would result in substantially biased estimates. The failure to recognize the importance of work experience to earnings may be a basic shortcoming of other similar studies. One purpose of this paper is to implement the existing models of earnings function by emphasizing the role of experience. The other purpose is to provide new evidence regarding the rate of return to education in Nairobi and a set of estimates which examines the income differentials, if any, among different ethnic groups.[1]

---

[1]The population of Kenya is composed of forty-three different tribes among which four tribes are dominant.

The organization for the rest of this paper is as follows. Section B discusses the theoretical framework and variables to be used. Section C presents the empirical results obtained by ordinary least squares (OLS) with various specifications. Finally, in section D, the findings and the contribution of this paper are summarized.

B. THE MODEL

It is assumed that an individual's earnings (Y) are a function only of his ability (A) to perform the assigned task well. Thus

$$Y = f(A).$$

But the ability to perform the task well depends on i) educational training (E), ii) experience (EXP), and iii) intelligence (I). Other sociological factors and institutional factors further influence the income determination. Thus

$$A = g\,(E, EXP, I)$$

$$Y = f \cdot g(E, EXP, I) + e$$

where e accounts for the effects of sociological and other institutional influences and assume $E(e) = 0$.

On the basis of the above, the general form of the earnings function to be estimated initially is then:

$$Y = \alpha + \beta E + \gamma EXP + \delta I + \Sigma\, \varepsilon_i s_i + e.$$

where s_i stands for a relevant institutional variable.

I will examine briefly the nature of two key independent variables, education and experience, along with other variables to be used.

Education. Education may be divided into formal education and informal education. The former refers to the mental, moral, or perhaps physical development obtained through a school and the latter refers to the development obtained outside of school (parental guidance for example). In an earnings function, formal education, which is best represented by years of schooling (S), plays two different roles. One is socially accepted credential of formal education and the other is its contribution to productivity. S is usually used as the best index for the quality of human capital in a job market. A number of jobs require a certain level of S regardless of whether one is capable of doing the job or not. The lack of educational credentials is a barrier to entry to well paying jobs. For wage earners, S appears to be one of the most important factors determining one's starting wage. Promotion, then, mainly depends upon one's productivity afterwards.

<u>Experience</u>. Work experience increases an individual's earning capacity by developing his job skills and through the acquisition of job related information. Experience also provides seniority which has a positive effect on earnings if one stays at the same job. These are some of the direct influences of the experience embodied in human capital on earnings.

There may be certain attributes of education that are not realized without several years of experience. Just as a diamond should be polished to show its beauty, the knowledge obtained in schooling should be also polished to show its value and be utilized in the world outside of school. Experience does the job of polishing or adjustment.

There appears to be another influence of experience on earnings. As schooling is used as a screening device (55), experience also is used as a screening device. In a job market where a certain level of schooling is required, one frequently finds that a certain length of experience is specified in the announcement of job openings. A number of public agencies in the United States substitute years of experience for years of schooling on a one-to-one basis in selecting candidates for vacant positions.

Age often is employed as an independent variable in an earning function estimate. Although it may be a reasonable proxy for experience, age itself does not have anything to do with earnings. Employers do not pay higher wages to older workers than to younger workers when other attributes are equal. Johnson's assertion, "Whatever the precise mechanism of causation, both age and education should be included as determinants of an individual's potential earnings,"[1] is too strong a statement, or misleading.

In a statistical analysis of a sample, if all independent variables are constant other than the experience variable, the years of experience should be positively correlated with earnings. But what if the sample is drawn randomly from among individuals with different schooling backgrounds? That is, should one year of experience as a janitor (low schooling) be treated the same as that of a lawyer (high schooling) in earnings? The experience variable needs to be adjusted to show its quality, and its contribution to earnings.

---

[1]See Johnson (27), p. 9.

Variables To Be Used

One major objection to the usual estimates of the contribution of schooling to earnings is that the true effect of schooling is likely to be overestimated because of its intercorrelation with the relevant but unincluded variables. Although my sample does not contain all the variables that I would like to have, it does provide most of the basic variables needed in this type of study. The following table compares my sample with that of other similar studies.

Table 1

Variables Used in Some Studies

Author	Year	Variables Used						
		Gross Income	Wage Rate	Experience	Other Training	IQ or Test Score	Family Bkgrd.	Other Socio-econ. status
Hu*	71	✓				✓	✓	✓
Griliches*	72	✓		✓	✓	✓	✓	✓
Johnson	73		✓	✓				✓
Mincer	73	✓		✓				
Welch	73	✓		✓				
This Study	76	✓	✓	✓	✓		✓	✓

*coauthored

Among the samples of other researchers, that of Griliches and Mason (19) is comparable with, or perhaps better than, mine in terms of the richness of the relevant variables. Since the contribution of ability, represented by Armed Force Qualification Test results, was found to be minute in their study, I do not feel that missing such a variable will affect my results significantly. My sample has at least two advantages over others. First, it has both concepts of income, gross income and hourly wage rate, so that the results can be compared with each other. Second, the information on experience in my sample may be more accurate than that of other studies. While experience is obtained usually by subtracting years of schooling and preschooling age from age in other data, it is obtained here by subtracting the year the respondent claimed to get his first job in Nairobi from the year my data were gathered (1970). The experience of migrants in rural areas is not included for obvious reasons.[1] All the independent variables are, strictly speaking, proxy variables that partially reflect the true variables. The major characteristics of my sample and the variables are summarized in table 2. The definition and measurement of most of the variables are standard.

<u>Dependent Variables</u>. Two variables are employed as dependent variables respectively. They are income (DY) and hourly wage rate (HW). Income is actual gross earnings in Shillings in December 1970. Hourly wage rate is computed by dividing the income by the hours worked during the month. The data also provide gross earnings in 1970 but the hours

---

[1]Ten years experience of farming, for example, is not appreciated by the employer of a shoe factory.

worked in that year are not reported. I experimented with both specifications of earnings, getting somewhat better results for the December income. This paper reports only those estimates for the December income and the hourly wage rate. Both ordinary and log of these two variables are used.

Independent Variables. Eleven independent variables are included in this study. They are years of schooling (YS); years of experience (EXP); other training (OT); school type (GS: government aided school or not); schooling-experience interaction (INT); father's schooling (DS); low-income area (LA); ethnic groups (KIK, KAM, LUH, LUO).

YS is an obvious proxy for formal education and GS is included to show the difference, if any, in school quality. INT, years of experience multiplied by years of schooling, attempts to measure one's productivity based on experience. DS is intended to account for informal education assuming that a more educated father provides a better informal education than a less educated father does. The variables of GT, and the vector of ethnic groups represent an individual's socioeconomic background.

YS, OT, DS are measured in years while LA, GS, and ethnic groups are dummy variables. INT is considered an index number.

The simple correlation coefficients between the major variables of my sample are listed in table 3. There are several correlations to be observed. First, schooling is highly correlated with the earnings variables. The correlation of log earnings is slightly higher than that of ordinary earnings. In fact all the independent variables other than age and experience have higher correlations with log earnings than with ordinary earnings. Second, experience is positively correlated with age and the interaction variable. The high correlation between experience and age suggests that age may be a reasonable proxy for experience, and also that they should not be used together in an earnings equation because of multicollinearity. The correlations of experience with other variables, however, are all negative. The negative correlation between experience and years of schooling is expected since at any given age one with more schooling is bound to get less work experience than one with less schooling.

Table 2

Mean and Standard Deviations of Major Characteristics

Variable	Mean or Percent in Sample	SD	Symbol
Years of Schooling	7.68	3.84	YS
Other Training	4.08	4.71	OT
Age	29.78	7.56	AGE
Work Experience	6.73	6.88	EXP
Father's Schooling	1.37	2.45	DS
Sex	95%		MALE
School Type	78%		GS
Low Income Area			
Pumani	3%		PUM
Mathare	13%		MAT
Ethnic Group			
Kikuyu	39%		KIK
Kamba	21%		KAM
Luhiya	18%		LUH
Luo	15%		LUO
Other tribes	7%		OTH
Hourly Wage Rate	2.82	2.35	HW
December Income	518	411	DY
Log of HW	0.72	0.81	LHW
Log of DY	5.97	0.76	LDY

Table 3

Correlations Between Selected Variables

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Years of Schooling	1.000	0.358	-0.481	-0.460	0.135	0.316	0.539	0.521	0.602	0.580
(2) Other training	---	1.000	-0.154	-0.148	0.130	0.136	0.293	0.306	0.332	0.354
(3) Age	---	---	1.000	0.665	0.392	-0.130	-0.073	-0.058	-0.069	-0.069
(4) Work Experience	---	---	---	1.000	0.660	-0.224	-0.050	-0.042	-0.021	-0.025
(5) Interaction	---	---	---	---	1.000	-0.085	0.299	0.310	0.336	0.336
(6) Father's Schooling	---	---	---	---	---	1.000	0.220	0.236	0.240	0.260
(7) Hourly Wage Rate	---	---	---	---	---	---	1.000	0.967	0.876	0.850
(8) December Income	---	---	---	---	---	---	---	1.000	0.848	0.875
(9) Log of HW	---	---	---	---	---	---	---	---	1.000	0.958
(10) Log of DY	---	---	---	---	---	---	---	---	---	1.000

C. EMPIRICAL RESULTS

Since this study emphasizes the role of experience in an earnings function, it may be appropriate to present the relations between earnings and experience first. As already shown in table 3, earnings is negatively correlated to experience in the data. To see the true relationship, the sample is divided into subgroups based on the years of schooling and simple regressions of experience on December income are performed for the subgroups that have reasonably large observations. Also the variable frequently used as experience in other studies (experience = age - years of schooling - preschooling age, see 19, 40) is regressed. The results are presented in table 4. The brackets show the results obtained by using the definition of experience in other studies.

Table 4. Simple Regressions of Experience on December Income*

Years of Schooling	Sample Size	Coefficient (t-ratio)		R^2	
0	20	2.47 (1.08)	[1.82] (0.73)	0.06	[0.03]
7 & 8**	64	28.94 (3.64)	[18.55] (2.98)	0.17	[0.12]
10	38	29.21 (3.40)	[11.39] (1.41)	0.24	[0.05]
12	63	100.5 (5.60)	[58.95] (4.19)	0.34	[0.22]

*Intercept terms are not reported in the table.

**The eight year primary school system was changed to the present seven year system in 1960. Primary school graduates under both systems are treated equally here.

There are three results to be observed from this table. First, the sign of the coefficients are all positive as expected. Second, not only the fit of the equation improves but also the contribution of experience increases as the years of schooling increase. These results suggest the existence of different qualities of experience or of an interaction between schooling and experience. Third, the experience defined in this study explains the variation of earnings better than that of the other study does.

Table 5 presents a number of regression results relating the earnings to selected variables. For equations 1 to 7, the dependent variable is December 1969 earnings in Kenyan Shillings. The semi-log relationship is given in equations 8 through 14.

Although log-linear equations provide better fits than linear equations do, the latter is used for the explanation simply because Shillings give a better feeling of the impact of the independent variables on earnings than logarithm of Shillings does.

Looking at the first three equations, note that the square of years of schooling is a better explanatory variable than YS in this model. The bases for this judgment are the R^2s and the t-ratios and the assumption that the earnings power of educated people may be quadratic in an education-poor country. Note also, in equation 2, that the contribution of YS to the fit is only 0.002 compared with equation 3.

I consider equation 3 as the basic estimate of the schooling which does not account for the effects of experience or interaction, other training, and father's schooling. By adding EXP to equation 3,

Table 5

Regression Equations with December Income as Dependent Variable*

Equation No.	YS	YS^2	EXP	INT	OT	DS	R^2
1	55,59 (9,94)						0,2705
2	-17,71 (-0,86)	5,38 (3,72)					0.3066
3		4,17 (10,82)					0.3046
4		5,11 (12,25)	16.09 (4,87)				0,3617
5		4.07 (11,21)		2,96 (5,94)			0.3862
6		3,81 (10,07)		2,81 (5,66)	10,28 (2.34)		0.3987
7		3,60 (9,06)		2,91 (5.83)	9,89 (2.25)	14.19 (1.66)	0.4050
8	0,1154 (11.65)						0.3368
9	0.0377 (1.02)	0.0057 (2,18)					0,3486
10		0,0082 (11.89)					0.3460
11		0.0103 (13,94)	0,0346 (5,94)				0.4277
12		0.0080 (12.52)		0,0059 (6,75)			0.4418
13		0,0074 (11.17)		0.0056 (6.42)	0,0251 (3.26)		0,4634
14		0.0069 (10,01)		0,0058 (6,65)	0,0243 (3.17)	0.031 (2,08)	0.4720

*Intercept terms are not reported.

the schooling coefficient increases about 18 percent.[1] This change indicates that other studies without an experience variable might have underestimated the schooling coefficient. There is one problem in the interpretation of the experience coefficient however. Equation 4 says that one year of experience will increase one's income about 16 Shillings whether one happens to be a lawyer or a janitor. This is unrealistic.

Since the qualities of experience are different among people due to their individual job status which in turn depends mainly upon education and experience, the experience should be weighted to provide plausible interpretation. The best candidate for the weighting at my disposal is, of course, the years of schooling. The weighted experience is also called interaction (INT) here as suggested by Mincer.

INT replaces EXP in equation 5.[2] The introduction of the INT variable leads to a drop of twenty percent in the schooling coefficient. This does not mean that the contribution of education decreased. The

---

[1] One minus the ratio of the schooling coefficient in subsequent equations to the corresponding schooling coefficient in equation 1 provides the proportionate bias in the schooling coefficient due to the ommission of a relevant factor.

$$1 - \frac{\text{coefficient (after)}}{\text{coefficient (before)}} = \text{Proportionate bias}$$

[2] Both EXP and INT were experimented in a regression. The results are insignificant t-ratio of EXP (0.64) and virtually no change in R^2 from equation 5 (0.3867).

contribution is actually bigger in equation 5 than in equation 3, except for people of no experience, or one year of experience and more than 14 years of schooling.[1] Notice also the role of experience. Since $\frac{\delta DY}{\delta EXP} = 2.96YS$, the contribution of experience income increases as education increases. The change of the fit is also noticeable. The introduction of INT increased R^2 by 0.0816 (relative to equation 3).

Accepting that equation 5 is the best so far, other relevant variables are added to this equation one by one. Other training is of course an important variable in earnings function in its own right. As expected, it causes a drop of 6 percent in the schooling coefficient but its own effect is substantial. The introduction of father's schooling to equation 6 shows how important one's family background is in Nairobi. Its effect is even bigger than other training in this equation. It may suggest that, in a city or a country where schooling opportunity is rare, family background is relatively more important than it is in developed countries.

Return to Schooling

The hourly wage rate of the individual is taken as the dependent

---

[1] From equation 2 and 5,

$$\frac{\delta DY}{\delta YS} = 8.34YS \qquad \frac{\delta DY}{\delta YS} = 8.14YS + 2.96\ EXP.$$

to be $8.34YS > 8.14YS + 2.96$ when $EXP = 1$, $YS > 14.8$.

Table 6

Regression Equations with Hourly Wage Rate as Dependent Variable*

Equation No.	Coefficient (t ratio) of YS	YS^2	EXP	INT	OT	DS	R^2
1	0.330 (10.48)						0.2915
2	-0.150 (-1.32)	0.035 (4.37)					0.3390
3		0.025 (11.59)					0.3347
4		0.030 (13.11)	0.094 (5.10)				0.3941
5		0.024 (12.00)		0.016 (5.80)			0.4034
6		0.023 (10.92)		0.015 (5.55)	0.047 (1.93)		0.4177
7		0.022 (10.01)		0.015 (5.64)	0.046 (1.87)	0.054 (1.12)	0.4205
8	0.1281 (12.35)						0.3635
9	0.0185 (0.48)	0.0080 (2.97)					0.3839
10		0.0093 (12.88)					0.3834
11		0.0116 (15.42)	0.0398 (6.68)				0.4721
12		0.0091 (13.65)		0.0063 (6.95)			0.4783
13		0.0085 (12.37)		0.0060 (6.65)	0.0210 (2.61)		0.4915
14		0.0082 (11.32)		0.0062 (6.78)	0.0204 (2.54)	0.021 (1.40)	0.4953

*Intercept terms are not reported.

variable in earnings functions in table 6. Conceptually hourly wage rate represents a person's stock of human capital better than income does. Also the slightly better R^2s in table 6 as opposed to the matching R^2s in table 5 indicate that the variation of hourly wage rate can be explained more than that of December income by the independent variables employed in this study.

Employing the logarithmic form of hourly wage rate in equations 7 to 14 provides two advantages in the interpretation of the coefficients. First, the partial coefficient of schooling can be interpreted as an estimate of the average rate of return to schooling. Second, the proportionate change in an individual's utility with respect to proportionate change in the real wage rate equals the elasticity of utility with respect to income as shown by Johnson.[1]

Equations 7 to 14 permit the estimation of different rates of returns at different levels of schooling. In equation 14, the marginal rate is

$$\frac{\delta(\ln HW)}{\delta YS} = 0.0164YS + 0.0062\ EXP.$$

When estimated at YS = 10, the marginal rates are 16.4 percent with no experience, 19.5 percent at 5 years experience, and 22.6 percent at 10 years experience. Again, the marginal rates, when estimated at EXP = 10, are 14.4 percent at 5 years schooling, 19.3 percent at 8 years of schooling, and 29.1 percent at 14 years of schooling.

---

[1]See Johnson (27), p. 14.

This result, the increasing marginal rate of schooling, is opposite of what was found in the United States. It suggests that the scarcity of educated people in this city, perhaps in Kenya too, gives highly educated people exponential earnings power.

Multicollinearity. Equation 14 (table 6) is the best equation obtained yet in terms of explanatory power ($R^2 = 0.495$). This kind of single equation least squares model, however, often violates one of the underlying assumptions of regression models: the assumption that the explanatory variables are independent of one another. If severe collinearity is present in equation 14, the contribution of each independent variable to the log of hourly wage rate may not be reliable.

I employ the Farrar and Glauber (15) technique for the diagnosis of the presence, severity, location, and pattern of interdependences among the explanatory variables. The detailed results are presented in Appredix IV. The Chi-square transformation for the matrix coefficient ($x^2(6) = 66.0$) shows the existence of multicollinearity but the overall severity is not extreme. Multiple correlations and associated F statistics show that INT is relatively stable, while OT and DS are moderately and YS^2 is most affected by multicollinearity.

The matrix of partial correltaion coefficients and associated t statistics show that YS^2 is collinear to some extent with OT and DS.

A regression using Mincer's specification,[1] in which two schooling variables are employed, is run on my sample to test for multicollinearity. The overall severity of multicollinearity is so extreme that the computer is not able to print out the Chi-square statistic and F statistics of schooling variables, implying that they are four digit numbers. If the earnings function is formulated for prediction purposes, the model may be acceptable. But structural questions cannot be answered for an obvious reason: the coefficients of independent variables are not reliable.

Income Differentials Among Tribes

The population of Kenya is composed of 43 different ethnic groups in which Kikuyu, Kamba, Luo, and Luhiya are the major tribes. Table 7 reports the mean values of the variables of interest.

The table shows that OTH (the members of minor tribes) earns the highest hourly wage, 4.3 Shillings, followed by LUH, KIK and KAM, while LUO earns only 2.1 Shillings per hour on the average. The differentials may be attributed to the difference in the level of schooling and experience. Notice that OTH has the highest level of schooling and LUH has the longest years of experience. LUO shows

---

[1]Mincer's specification (40, p. 92) is

$$LHW = \alpha + \beta_1 YS + \beta_2 YS^2 + \beta_3 EXP + \beta_4 EXP^2 + \beta_5 INT + e$$

using the variable names in this study.

Table 7

Mean (Standard Deviations) of Major Variables of Different Tribes

Tribes	Variable				
	HW	YS	AGE	EXP	INT
KIK	2.8 (2.3)	8.0 (3.7)	29.3 (6.9)	6.3 (6.3)	39.6 (38.1)
KAM	2.3 (1.8)	6.4 (3.5)	29.2 (6.5)	7.4 (6.6)	40.9 (43.2)
LUO	2.1 (1.7)	7.4 (3.6)	29.3 (9.9)	5.4 (6.1)	31.1 (32.1)
LUH	3.3 (2.5)	7.9 (4.2)	31.2 (6.8)	7.8 (7.5)	43.6 (40.2)
OTH	4.3 (3.3)	9.0 (3.2)	31.0 (8.8)	6.7 (8.4)	41.6 (47.5)
ALL*	2.8 (2.3)	7.6 (3.8)	28.7 (7.5)	6.7 (6.8)	39.5 (39.7)

*ALL = All tribes

distinct lack of experience and lowest level of INT in urban area. Most of the Luos, whose main territory is around the Lake Victoria, far from Nairobi, may have started the urban life in this city later than people of other tribes have.[1]

Separate regressions are fit for each ethnic group to detect any changes in the coefficients of independent variables, those of YS and INT in particular, employing the same independent variables used in equation 14, table 6. Table 8 reports the results.

The different magnitudes of the coefficients of YS and INT indicate that the returns to schooling for each tribe are different.

---

[1] Most of the individuals in the original sample (97%) were born outside Nairobi. See Whitelaw (61) p. 9.

Table 8

Regression Equations with LHW as Dependent Variable for Each Tribe

	Sample Size	YS	INT	R^2
KIK	106	0.0085 (6.61)	0.0060 (3.94)	0.5912
KAM	57	0.0094 (4.91)	0.0091 (5.05)	0.6182
LUO	40	0.0078 (3.69)	0.0044 (1.70)	0.5784
LUH	49	0.0054 (3.41)	0.0015 (0.77)	0.5784
OTH	17	0.0078 (2.56)	0.0033 (1.14)	0.8276
ALL	269	0.0081 (10.26)	0.0062 (6.78)	0.4953

The Chow tests[1] performed for each tribe's equation with all tribes' equation rejects the null hypothesis that the coefficients are not significantly different in the two equations, except when the comparison is made with the Kikuyu's equation. The low returns to schooling for LUH are of interest. Recalling that their average earnings are highest among the major tribes, I conjecture that their occupations are probably self-employed, in which schooling plays a less significant role, rather than wage earners.

Discrimination

The income differentials among the major tribes can be attributed, at least partially, to the difference in the level of schooling and

---

[1]See Johnston (29), p. 207.

experience. But is a part of the differentials caused by discrimination? To estimate the statistical significance, dummy variables of ethnic origins (KIK, KAM, LUO, LUH) are added to the regression equations along with all the explanatory variables at my disposal.[1] Table 9 presents the results.

My interest here is the magnitude of the coefficients of ethnic dummy variables and their t-ratios. Since the regressor of other tribes variable (OTH) is included in the intercept term, the other coefficients of the ethnic group variables are interpreted as deviations from this regressor.

Equation 1 shows that the hourly wage rates of all four major tribes on the average are lower than that of other tribes as a whole. It is interesting that the people of two most dominant tribes in Kenya appear to be earning less than the people of other tribes after adjustments have been made for the influence of other variables on wages. Luos are earning 1.4 Shillings less and Kikuyus about 1.1 Shillings less than other tribes are. The t ratios of the coefficients are significant at 2 percent level. Kambas appear to be doing relatively better than Luos or Kikuyus but the small coefficient of LUH with low t-ratio suggests that Luhiyas, who are business oriented in general, are the tribe earning highest wage among the major four tribes.

---

[1] Other dummy variables, sex and Pumani are suppressed from the equation due to insignificant t-ratios.

Table 9

Regression Equations With All the Independent Variables

Independent Variable	Dependent Variable			
	HW	DEY	LHW	LDEY
YS^2	0.023 (9.60)	3.91 (8.83)	0.0081 (10.26)	0.0070 (9.22)
INT	0.014 (4.98)	2.62 (5.15)	0.0050 (5.52)	0.0047 (5.37)
OT	0.050 (2.04)	10.87 (2.47)	0.0246 (3.14)	0.0288 (3.82)
DS	0.057 (1.20)	14.37 (1.70)	0.0242 (1.60)	0.0325 (2.23)
GS	1.019 (3.49)	183.4 (3.51)	0.2291 (2.46)	0.2341 (2.60)
KIK	-1.08 (-2.32)	-146.5 (-1.75)	-0.3419 (-2.29)	-0.2508 (-1.74)
KAM	-0.92 (-1.87)	-115.4 (-1.30)	-0.2246 (-1.43)	-0.1137 (-0.74)
LUO	-1.40 (-2.71)	-185.3 (-2.11)	-0.3831 (-2.32)	-0.2689 (-1.63)
LUH	-0.60 (-1.21)	-75.37 (-0.84)	-0.0731 (-0.45)	-0.0064 (-0.04)
MAT	-0.69 (-2.06)	-106.0 (-1.767)	-0.4055 (-3.78)	-0.3534 (-3.42)
Constant	0.46 (-0.80)	56.29 (0.54)	-0.1402 (-0.76)	5.067 (28.47)
R^2	0.4758	0.4523	0.5588	0.5302

Table 10

Tribal Composition of Kenya, Nairobi and the Sample Used

	Kenya	Nairobi	Sample
KIK	27.2	48.5	39.4
KAM	11.0	15.0	21.1
LUH	13.3	15.9	18.2
LUO	13.9	15.5	14.8
OTH	13.0	5.1	6.5

The reasons for these results are open for speculation.[1] Some possible reasons could be: (i) the sample, which focuses on low and middle income people, systematically eliminates very successful Kikuyus and Luos who live in European neighborhoods and who, if they were included, would at least offset the negative coefficient. If the 1969 census figures were correct (see column 2 of table 10), the proportions of Kikuyus and Luos in the sample are less than true proportions, which might lend support to the biasedness argument in some way; (ii) especially for Kikuyus, they are systematically discriminated against by European and/or Asian employers who resent their role in the independence struggle. On the other hand, Luos are discriminated against by Kikuyus, who hold many important positions in the government and industry after independence.

[1] This portion relies on Johnson (27, pp. 22-24) and the discussion with Leonard Njuguna Muraya, a student from Nairobi, Kenya who participated in the survey and majors in economics at the University of Oregon.

D. SUMMARY AND CONCLUSION

The focus of the study was to derive and estimate the functional relation between the human capital of Nairobian workers and their earnings. In developing the earnings function, the use of one frequently ignored variable, experience, was discussed and included in the function. The following are the major findings of this study.

(1) Schooling is again found to be the most significant variable in explaining the inequality in the distribution of earnings. The years of schooling alone explains 29 percent of the variation of earnings in this study. Moreover the marginal rate of return to schooling is found to be increasing in Nairobi as opposed to the declining marginal rate of return to schooling in the United States. This indicates a severe gap of income between more educated and less educated workers and calls for effective educational planning.

(2) The role of experience in the earnings function is substantial as expected. It has the highest explanatory power among the independent variables when estimated at each schooling level. To see the true contribution of experience which is negatively correlated with earnings in the whole sample, years of experience is weighted by years of schooling. The coefficient of the weighted variable (INT) shows higher t-ratio than that of EXP. It also gives better R^2s (table 5 and 6).

(3) Other variables like other training and father's education are added in the function to test the sensitivity of the schooling coefficient. They did show significant impact on the schooling variable. That is, the coefficient of schooling is biased upward without these variables.

(4) The separate regressions for the different tribes indicate that there appear to be different rates of return to schooling among them.

(5) The log of earnings as a dependent variable gives a better fit. This agrees with previous reports. Between the log of income and the log of the hourly wage rate, although there is no firm basis to choose one as dependent variable, the log of the wage rate is preferred for the reason stated on page 68.

The analysis of this study was carried on with limited data of schooling: the highest schooling was only 14 years. The overall implication of the results is that education is immensely important in this city of an underdeveloped country. Assuming that the earnings function is correct (table 9), the marginal rate of return to schooling for college graduates without experience (16 years) would be 25.9 percent. I cannot think of any other investment that would match this rate.

In sum, my findings support the economic and statistical significance of schooling in the explanation of observed differences in earnings. They also point out that the contribution of experience is relatively high. The omission of expereince would result in an underestimated schooling coefficient holding other variables constant.

APPENDIX I

HETEROSCEDASTICITY AND VARIANCE

A. Heteroscedasticity

Assume that the Y_i has been generated by the simple regression model $Y_i = \alpha + \beta X_i + e_i$. The standard assumptions hold except that the errors are heteroscedastic; the i's are independent random variables, $e_i \sim N(0, \delta_i^2)$ and the variances are not all the same. Let us divide the ith equation by δ_i, then we obtain

$$\frac{Y_i}{\delta_i} = \alpha\left(\frac{1}{\delta_i}\right) + \beta\left(\frac{X_i}{\delta_i}\right) + \frac{e_i}{\delta_i}.$$

Let $v_i = \frac{e_i}{\delta_i}$, then $\mathrm{var}\,(v_i) = \frac{1}{\delta_i^2}\,\mathrm{var}\,(e_i) = \frac{\delta_i^2}{\delta_i^2} = 1.$

The v_i's are homoscedastic and, hence, all the standard assumptions for OLS are satisfied. Specifically what we do is, instead of minimizing the standard $\Sigma(Y_i - \alpha - \beta X_i)^2$, we minimize $\Sigma\left[\left(\frac{Y_i}{\delta_i}\right) - \alpha\left(\frac{1}{\delta_i}\right) - \beta\left(\frac{X_i}{\delta_i}\right)\right]^2 = \Sigma\, v_i^2.$

Running OLS on the transformed variables is one solution to take account of heteroscedasticity. But notice that the sum of v_i^2's can be written $\frac{(Y_i - d - \beta X_i)}{\delta_i^2}$. Here, each squared deviation,

e_i^2, is weighted by a factor $\frac{1}{\delta_i^2}$ before summing. This is weighted least squares (WLS) method. These two methods are equivalent.[1]

B. Variance

Consider the same equation

$$Y_i = \alpha + \beta X_i + e_i$$

where Y_i is either 1 or 0. Accordingly,

$$e_i = 1 - (\alpha + \beta X_i)$$

or

$$e_i = -(\alpha + \beta X_i).$$

Let $\alpha + \beta X_i = P$. Then

$$\begin{aligned} E(e) &= e_i \cdot f(e_i) \\ &= (1-P) P + (-P)(1-P) \\ &= 0 \end{aligned}$$

$$\begin{aligned} Var(e) &= (e_i - E(e))^2 f(e_i) \\ &= e_i^2 \cdot f(e_i) \\ &= (1-P)^2 P + (-P)^2 (1-P) \\ &= P(1-P). \end{aligned}$$

[1]For some unknown reason, the Regression Analysis Program for Economists (RAPE) does not give the same results. Since debugging the computer program is out of the scope of this study, the problem is not pursued.

APPENDIX II

MODELS FOR PREDICTION

A. Regression Models

Model of Intraurban Migration

Probability of Migration

	OLS		WLS	
	Coefficient	t-ratio	Coefficient	t-ratio
RENT	-0.0012	-3.588	-0.0012	-6.740
KIDH	-0.0186	-1.516	-0.0182	-2.307
AGE	-0.0112	-5.128	-0.0139	-6.685
RSON	0.0650	1.496	0.0814	2.238
DJH	0.0101	5.847	0.0111	6.812
TDSH	-0.0031	-1.745	-0.0024	-2.401
DCH	-0.0083	-3.422	-0.0050	-5.799
ΔDJH*	-0.0083	-5.541	-0.0099	-5.806
Intercept	0.7308	7.593	0.8044	8.481
Degrees of freedom	369			
R^2	0.2320			

B. Discriminant model

Discriminant Functions

	NOMOVE	MOVE
AGE	0.4040	0.3315
ΔDJH*	-0.0404	0.0879
DJH	0.0893	0.1469
DCH	0.0724	0.0475
DHS	0.0696	0.0375
KIDH	-0.1070	-0.1406
RENT	0.0310	0.0239
RSON	3.0451	3.5310
Constant	-11.5578	-9.7903

Total sample size (N) = 378; N_1 = 268, N_2 = 110

U-Statistic = 0.7480 degrees of freedom = 8,1,376

Approximate F = 15.54* degrees of freedom = 8,369

*significant at 1 percent level

Classification Matrix

	NOMOVE	MOVE
NOMOVE	253	15
MOVE	63	47

Q-statistic = 130.38

APPENDIX III

RATES OF RETURN TO EDUCATION IN SOME STUDIES

Author	Year	Country	Secondary	Elementary
Carnoi	1967	Mexico	36.5	21.1
Danielsen	1971	Japan	11.4	
Gounden	1967	India	13.7	16.8
Krueger	1972	Turkey	21-23	
Hanoch	1967	U.S.A.	16.0	
Hansen	1963	U.S.A.	13.7	
Hines	1970	U.S.A.	19.5	

APPENDIX IV

TESTS FOR MULTICOLLINEARITY

A. Equation 14 of Table 5.

DEPENDENT VARIABLE: LHW

VARIABLE	COEFFICIENT		T-RATIO
YS^2	0.0082		11.32
INT	0.0062		6.78
OT	0.0204		2.54
DS	0.0219		1.40
INTERCEPT	-0.238		-3.38
R^2		0.4953	
D.F.		264	

DETERMINANT OF CORRELATION MATRIX = 0.779
CHI-SQUARE (6) = 66.059

VARIABLE	F(3, 265)
YS^2	19.68
INT	2.70
OT	10.76
DS	12.67

PATTERN OF INTERDEPENDENCE

	YS^2	INT	OT	DS
YS^2	0.18	0.67	4.54	5.39
INT	0.04	0.02	2.09	-1.85
OT	0.26	0.12	0.10	0.85
DS	0.31	-0.11	0.05	0.12

B. Mincer's Specification on the Sample

DEPENDENT VARIABLE: LHW

VARIABLE	COEFFICIENT	T-RATIO
YS	-0.011	-0.26
YS^2	0.011	4.28
EXP	0.066	3.28
EXP^2	-0.001	-3.05
INT	0.002	1.43
INTERCEPT	-0.46	-2.42
R^2	0.5067	
D.F.	263	

DETERMINANT OF CORRELATION MATRIX = 0.001862
CHI-SQUARE (10)=******

VARIABLE	F(4,264)
YS	*****
YS^2	*****
EXP	966.5
EXP^2	434.4
INT	241.9

PATTERN OF INTERDEPENDENCE

	YS	YS^2	EXP	EXP^2	INT
YS	0.95	47.17	-8.75	3.72	13.54
YS^2	0.94	0.94	4.86	-1.99	-8.48
EXP	-0.47	0.28	0.93	25.76	18.07
EXP^2	0.22	-0.12	0.84	0.86	-6.84
INT	0.64	-0.46	0.74	-0.38	0.78

BIBLIOGRAPHY

1. Adams, J.S., "Directional Bias in Intra-Urban Migration," Economic Geography, October, 1969.

2. Bailey, D. and Shotta, C., "Private and Social Rates of Return to Education of Academians," American Economic Review, March 1972.

3. Beals, Ralph E., Statistics for Economists, Chicago: Rand McNally & Co., 1972.

4. Becker, G.S., and Chiswick, B.R. "Education and the Distribution of Earnings," American Economic Review, May 1966.

5. Bowles, S. and Levin, H., "The Determinants of Scholastic Achievement. An Appraisal of Some Recent Evidence," Journal of Human Resources, Winter, 1968.

6. Brown, L. and Holmes, J. "Intra-Urban Migrant Lifelines: A Spatial View" Review of Regional Studies, Spring, 1971.

7. Caplow, T. "Incidence and Direction of Residential Mobility in a Minneapolis Sample," Social Forces, May, 1949.

8. Carnoy, M. "Rates of Return to Schooling in Latin America," Journal of Human Resources, 2, 1967.

9. Clark, D. and Fong, H. "Returns to Schooling and Training in Singapore," Malayan Economic Review, Oct., 1970.

10. Danielsen, A.L. and Okachi, K., "Private Rates of Return to Schooling in Japan," Journal of Human Resources, Summer, 1971.

11. de Leeuw, Frank, "The Demand for Housing: A Review of Cross-Section Evidence," The Review of Economics and Statistics, February, 1971.

12. Denison, E.F., "The Sources of Past and Future Growth," in Wykstra, ed. 1971.

13. Eckaus, R.S., "Returns to Education with Standardized Incomes," Quarterly Journal of Economics, Feb., 1973.

14. ________, Estimating the Returns to Education: A Disaggregated Approach, Berkeley, The Carnegie Foundation for the Advancement of Teaching, 1973.

15. Farrar, D.E. amd Glauber, R.R., "Multicollinearity in Regression Analysis: The Problem Revisited," Review of Economics and Statistics, February, 1967.

16. Ford, R. G., "Population Succession in Chicago," American Journal of Sociology, March, 1950.

17. Freund, J.E., Mathematical Statistics, Englewood Cliffs: Prentice-Hall, 1971.

18. Gounden, A.M.N., "Investment in Education in India," Journal of Human Resources, Summer, 1967.

19. Griliches, Z., and Mason, W. "Education, Income and ABility," Journal of Political Economy, May 1972.

20. Gustman, A.L., "On Estimating the Rate of Return to Education," Applied Economics, June, 1973.

21. Hanoch, G., "An Economic Analysis of Earnings and Schooling," Journal of Human Resources, 2, 1967.

22. Hansen, W.L., "Total and Private Rates of Return to Investment in Schooling," Journal of Political Economy, April 1963.

23. Harbison, F.H. and Myers, C.A., Education, Manpower and Economic Growth, New York, McGraw-Hill Book Company, 1964.

24. Harris, R., Tolley, G., and Harrel, C. "The Residence Site Choice," Review of Economics and Statistics, May, 1968.

25. Hause, J.C., "Earnings Profile: Ability and Schooling," Journal of Political Economy, May/June 1972, Part 2.

26. Hines, F., et al., "Social and Private Rates of Return to Investment in Schooling by Sex-Race Groups and Regions," JHR, Summer, 1970.

27. Johnson, G.E., The Determinants of Individual Hourly Earnings in Urban Kenya: Methodology and Preliminary Results, Institute for Development Studies, University of Nairobi, 1971 (mimeographed).

28. Johnson, G.E. and Stafford, F.P., "Social Returns to Quantity and Quality of Schooling," Journal of Human Resources, Spring, 1973.

29. Johnston, J., Econometric Methods, McGraw-Hill, 1972, New York.

30. Juster, F.T., ed. Education, Income, and Human Behavior, Carnegie Commission on Higher Education. New York: McGraw-Hill, 1975.

31. Kain, J.F., "The Journey-to-Work as a Determinent of Residential Location," Papers of the Regional Science Association, 1962.

32. Krueger, A.O., "Rates of Return to Turkish Higher Education," Journal of Human Resources, Fall, 1972.

33. Ladd, G., "Linear Probability Function and Discriminant Functions," Econometrica, October, 1966.

34. Lansing, J.B. and Blood, D.M., "A Cross-Section Analysis of Non-business Air Travel," Journal of the American Statistical Association, Vol. LIII, 1958.

35. Lee, T.H., "Demand for Housing: A Cross-Section Analysis," The Review of Economics and Statistics, Vol. 45, 1963.

36. Link, C.R., "The Quantity and Quality of Education and Their Influence on Earnings, The Case of Chemical Engineers," Review of Economics and Statistics, May, 1973.

37. Mills, E. S., Studies in the Structure of the Urban Economy, Baltimore: John Hopkins Press, 1972.

38. Mincer, J., "On-the-Job Training: Costs, Returns and Some Implications," Journal of Political Economy, Supplement, Ocotber 1962.

39. ________, "The Distribution of Labor Incomes: A Survey," Journal of Economic Literature, March 1970.

40. ________, Schooling, Experience, and Earnings, New York, National Bureau of Economic Research, 1974.

41. Moore, E. G., "Models of Migration and the Intra-urban case," Australian and New Zealand Journal of Sociology II, April, 1966.

42. Morgenstern, R.D., "Direct and Indirect Effects on Earnings of Schooling and Socio Economic Background," Review of Economics and Statistics, May, 1973.

43. Morrison, D., "On the Interpretation of Discriminant Analysis," Journal of Marketing Research, May, 1969. Also in Akar, David A. (ed) Multivariate Analysis in Marketing: Theory and Application, Belmont: Wadsworth Publishing Co., Inc., 1971.

44. Muth, R. F., Cities and Housing: The Spatial Pattern of Urban Residential Land Use, Chicago: University of Chicago Press, 1969.

45. Perlman, R., The Economics of Education, New York, McGraw-Hill, 1973.

46. Press, S. J., Applied Multivariate Analysis, New York: Holt, Rinehart and Winston, Inc., 1972.

47. Reinmuth, J.E., Applied Regression and Discriminant Analysis, (mimeo), Eugene: University of Oregon, 1970.

48. Roberts, J., A Land Full of People: Kenya Today, Eyre and Spottiswoode, London, 1966.

49. Rossi, P. H., Why Families Move: A Study in the Social Psychology of Urban Residential Mobility, Glencoe, Illinois, 1955.

50. Schultz, T.W., Investment in Human Capital, New York, The Free Press, 1971.

51. ________, "Investment in Human Capital," American Economic Review, March, 1961.

52. Simmons, J. W., "Changing Residence in the City: A Review of Intra-Urban Mobility," Geographical Review, October, 1968.

53. Steinnes, D. N. and Fisher, W. D., "An Econometric Model of Intraurban Location," Journal of Regional Science, April, 1974.

54. Taubman, P. and Wales, T., "Mental Ability and Higher Educational Attainment in the Twentieth Century," in Juster, ed., 1975.

55. ________, "Education as an Investment and a Screening Device," in Juster, ed., 1975.

56. Theil, H., Principles of Econometrics, New York: John Wiley & Sons, Inc., 1971.

57. Warner, S., Stockastic Choice of Mode in Urban Travel: A Study in Binary Choice, Northwestern University: Northwestern University Press, 1962.

58. Watson, P.L., "Choice of Estimation Procedure for Models of Binary Choice," Regional and Urban Economics, 4, 1974.

59. Welch, F., "Black-White Difference in Returns to Schooling," American Economic Review, December, 1973.

60. Whitelaw, W.E., Nairobi Household Survey: Description of the Methodology and Guide to the Data. Nairobi: Institute for Development Studies, 1971.

61. ________, Rural-Urban Relations, Low-Income Settlements, Intraurban Mobility and the Spatial Distribution of Households in Nairobi Kenya. Working papers in Economics #6, University of Oregon, April, 1974.

62. Winger, A. R., "Housing and Income," Western Economic Journal, June, 1968.

63. Wolpert, J., "Migration as an Adjustment to Environmental Stress," Journal of Social Issues, 1966.

64. Wonnancott, R.J. and Wonnancott, T.H., Econometrics, New York: John Wiley & Sons, Inc., 1970.

65. Wykstra, R.A., ed., Education and the Economics of Human Capital, New York, The Free Press, 1971.