Any views expressed in this paper are those of the author. They should not be interpreted as reflecting the views of the Institute for Development Studies or of the University of Nairobi.
The purpose of this paper is to investigate the factors which account for differences among individuals in hourly earnings in the major urban center in Kenya. There are two principal reasons for conducting such an investigation. First, there is an emerging consensus among economists that the general level of wages in the modern sector of the ‘typical’ lesser developed country is too high to permit the efficient allocation of resources. Specifically, (a) high wages promote urban-rural imbalance, thus causing high rates of migration to cities even in the face of serious urban unemployment and/or unemployment; (b) rapid rates of increase of real wages in the modern sector have lowered the rate of growth of employment relative to output, thus exacerbating the unemployment problem; and (c) high wages in the modern sector reduce domestic resources available for capital accumulation, both public and private, and hence lower the rate of economic growth below its potential maximum. Knowledge of the determinants of the structure of wages is essential to understanding why the general level is too high, which is in turn necessary for the formulation of a workable incomes policy.

* The author, Visiting Research Fellow, Institute for Development Studies, University of Nairobi and Associate Professor of Economics, The University of Michigan, has benefited from the comments and suggestions of a number of persons, including O. Ashenfelter, G. Fields, J. Morgan, L. Smith, J. Stiglitz, and W.E. Whitelaw, and is grateful to the Rockefeller Foundation for financial support.


Second, questions concerning the distribution of income are interesting in any society, and they are especially interesting in societies in which per capita income is very low. What are the income advantages of the educated, of union members, of government employees? To what extent are females discriminated against in the labor market? Answers to such questions are important in their own right as well as providing insights into the operation of the aggregate labor market.

It is quite difficult to attack these questions without reliable microeconomic data (i.e., data which provide the joint distribution of income, sex, education, union membership status, etc. across individuals), but such data are quite rare in lesser developed countries. This paper, however, is able to exploit such a data set, that from the Household Survey of the Nairobi Urban Study. Although it was primarily concerned with specific urban problems such as housing, shopping patterns, and the like, this survey contained a battery of questions on basic socio-economic information. The paper is divided into four sections: In Section I a model of individual wage differences is set out; in Section II the model is tested on the survey data; in Section III some further tests on special hypotheses are reported, and in Section IV various implications of the results concerning labor market behavior in lesser developed countries are discussed.

I. The Earnings Function

In the absence of institutional impediments to full employment the wage rate paid each worker will depend upon (a) his personal product...

3 - I am especially indebted to Prof. W.E. Whitelaw, Director of the I.D.S. Urban Study, for permission to use these data.
ivity and (b) the nonpecuniary attributes of his particular job. In most countries the actual earnings of an individual will be influenced (c) by the type of employment activity in which he is engaged, unionized, government, nonunion private, or self employment, reflecting monopoly powers and/or non-maximizing behavior on the part of some employers. Finally, in many societies there is (d) labor market discrimination against certain subgroups in the society, and the resultant earnings advantages or disadvantages (of being female, a Negro in the United States, a Catholic in Northern Ireland, from a particular tribe in an African country, etc.) must be accounted for in the earnings function. We now examine each of these sets of factors in detail.

(a) Personal Productivity

An individual's personal productivity or stock of human capital can be represented by (i) his formal education, (ii) post-formal training, both on and off the job, and (iii) innate ability. Most microeconomic data, including those used in this study, do not contain (ii), and economists are not really sure what (iii) is. To a certain extent, however, the increment to the stock of human capital due to experience and other forms of post-formal training can be represented by an individual's position in the life cycle. In general, the earnings of a particular individual relative to other individuals in the work force, holding other factors constant, first rise at an increasing rate with age after he obtains his first job, then rise at a declining rate, reach a maximum somewhere between age 45 and 55, and decline thereafter. This is usually attributed to the incidence of post-formal training, high at early and

4 For an extensive discussion of the methods used to explain earnings levels, see Giora Hanoch, Personal Earnings and Investment in Schooling, Unpublished Ph.D. dissertation, University of Chicago, 1965.
low at later ages. 5

The effect of education on the earnings potential of an individual can in principle be manipulated to obtain a rate of return to various increments of schooling, either a private or a social rate. 6 However, there are two problems associated with inferring a rate of return from cross-sectional data on earnings in the typical lesser developed country. First, if there is substantial unemployment over the range of education under consideration, the probability of obtaining employment within various periods of time must be included in the calculation of the private rate of return, and the social rate of return to educating a redundant person is negative. 7 Second, with increasing supplies of educated labor in most lesser developed countries it is very likely that the kinds of positions filled by individuals with a certain level of education a generation (or even a few years) ago are now being filled by persons with much more education. This would imply that the observed age-income profile would overstate the effect of age on potential earnings, for to some extent it represents this "seniority effect" instead of human capital accumulation after completion of formal education.


6 - The rate of return to the e\textsuperscript{th} year of schooling is that r which satisfies

\[ \sum_{t=1}^{T} (W(E, t) - W(E-1, t))(1+r)^{-t} - C_E - W(D-1, D) = 0, \]

where \( C_E \) is the cost of that year of schooling (to the individual for the private return, to the society for the social return) and T is the number of years the individual has remaining in the labor force. If the difference in earnings remains constant with age at \( \Delta W \), then the rate of return is roughly equal to \( (\Delta W/N) / (1+r) \), where \( W = W(D-1, D) \).

7 - For a thorough analysis of these issues see G. Fields, "Private and Social Returns to Education in Labour Surplus Economies," Discussion Paper No. 104, Institute for Development Studies, Nairobi, April 1971.
There are, of course, other aspects of an individual's stock of human capital. First, it would be very useful to have some measure of the quality of an individual's education. This can be partially represented by an index of how good the schooling system the individual attended was, say by expenditure per pupil. Related to this is the quality of education at home, which would be related to the education or occupation of the individual's parents. In most developing countries—certainly most African countries, however—workers' parents rarely had any formal education at all, so this variable would probably not show up as significant until the next generation of workers. Second, the rapid growth of urban areas in most underdeveloped countries and the relatively recent emergence of the modern sector implies that a large fraction of the urban work force grew up in rural areas and migrated to cities at some point. Presumably the general skills necessary for successful participation in the modern sector will be greater the longer the individual has lived in the urban area, so it should be useful to include a measure of the length of time the individual has lived in the urban area in the explanation of wage differences.

(b) Nonpecuniary Advantages

In principle one would like to include a set of nonpecuniary attributes associated with each individual's job in the explanation of wage differences. These would include the cleanliness of the work place, an index of the required degree of physical exertion, the social status of the occupation, and the like. Any variable which increases the overall

8 Each of these variables, expenditure on schooling and father's education, was found to be quite significant in the explanation of earnings for a recent national probability sample in the U.S. See G. Johnson and F. Stafford, "Social Returns to Quantity and Quality of Schooling," University of Michigan working paper, December 1970.
attractiveness of the job would reduce the wage which the employer would have to pay in order to attract a work force of the desired quantity and quality. In practice, however, it is rather difficult to sort out those variables which should be included in the list, and no attempt is made to include any such variables in this study. A major source of difficulty in this respect is the likelihood that individuals have different tastes for each particular nonpecuniary attribute—the people do not mind a high degree of physical exertion while others do. But this would imply that individuals would avoid the types of jobs for which they have especial aversions, and this in turn suggests that nonpecuniary attributes may not be terribly important in the determination of wages.

(c) **Type of Employment Activity**

If the general level of wages in a particular country is too high to permit full employment in the modern sector of the typical lesser developed country, we would like to know why. To set up the problem for testing we will distinguish between three sets of hypotheses concerning the reason for the high wage level: (1) unionism, (2) government as employer, and (3) legal and social minimum wages.

1) **Unionism.** If a significant fraction of the work force is organized by trade unions, the union movement is to a significant extent interested in the short run economic welfare of its members, and the

---

government permits relatively free collective bargaining between unions and managements, it is very likely that the wage rate in unionized industries will be driven up above that prevailing in nonunionized industries, and the union wage level will obviously be greater than the market clearing wage in the modern sector. Whether this is sufficient to generate unemployment in the Harris-Todaro sense depends on the effect of the high union wage on the nonunion wage level. If employers of nonunion labor do not perceive that they are threatened by unionization of their workers and are not otherwise forced or motivated to pay a higher wage than is justified by market conditions, the wage in nonunion industries of the modern sector will fall below the agricultural wage (after adjustment has been made for differences in the cost of living), and full employment, with a "distorted" wage structure, will still prevail after a period of adjustment. On the other hand, there are some reasons for supposing that the level of the union wage may influence the nonunion wage directly. One of these is the "threat effect" hypothesis. As the union/nonunion wage differential increases, the probability that a majority of the workers

10 - Ignoring differences in living costs, nonpecuniary attributes, and hours of work, labor market equilibrium will be established when

\[ W_u Q + W_n (1 - Q) = W_a, \]

where \( W_u \), \( W_n \), and \( W_a \) are the wage levels for union, nonunion, and agricultural workers, respectively, and \( Q \) is an index of the probability, appropriately discounted, of obtaining employment in the unionized sector. If \( W_u \) is set greater than then \( W_n \) it then follows that the equilibrium wage in the nonunion sector will be less than the wage in agriculture, so long as \( Q > 0 \). \( Q \) will depend positively on the rate of growth of total employment and the rate of labor turnover in unionized industries as well as the relative size of the unionized branch of the modern sector. For an exposition of this approach see my "The Structure of Rural-Urban Migration Models," East African Economic Review, forthcoming June 1971.

in the typical nonunion firm will be motivated to change their status will also increase. A profit maximizing employer will therefore attempt to keep his wage offer in line with the union wage, less but not too much so. This implies, of course, that the wage level in the nonunion branch of the modern sector will not fall to that level which would equate the demand for and supply of labor in the modern sector.

(2) Government as Employer. In most lesser developed countries a fairly large fraction of wage earners are government employees, and it is quite possible that the government arbitrarily sets a wage level for its employees which is higher than is dictated by market conditions. Civil servants compose an articulate and potentially troublesome minority from the point of view of the incumbent government, and it is not difficult to imagine that the government will try to "bribe" them into submission and cooperation. In addition, decisions concerning government salary levels may be made by the bureaucrats themselves, which would be unlikely to result in systematic biases on the low side of the market wage.

(3) Legal and Social Minimum Wages. To the extent that the government enforces its code of statutory minimum wage levels, a floor may be established on nonunion wage levels, a floor may be established on nonunion wage levels which will have the effect of attracting too many people to urban areas. There may be other reasons why the nonunion wage level does not adjust to a very low level. First, the individual productivity of workers may depend on either the real wage the employer pays (through its effect on the worker's health and nutrition) or the wage the employer offers relative to offers by other firms (because of "morale" effects). By either version of the efficiency wage hypothesis the firm may find it profitable to increase its wage offer to a point...
beyond that at which it can just fill all its positions. In fact, the profit maximizing wage is determined by setting the elasticity of labour efficiently with respect to the wage equal to unity. By the health-productivity variant of the model variant of the model, firms are likely to establish minimum real wages which are independent of the wage level prevailing in the union or government sectors. By the morale effect variant the wage in the nonunion branch of the modern sector should be geared to that prevailing in union and government employment, which is roughly the same as is predicted by the threat effect model.

Second, decision makers in firms may derive utility from paying wages which are high in either a relative or an absolute sense, which implies that they are willing to sacrifice some of their profits in order to pay higher wages. This behaviour would most likely apply to firms which are owned and managed by foreigners, for foreigners may not want to be known as "exploiters" of local labour. If firms have the same (well behaved) trade off between profits and wages, it can be shown that firms in which the production process is more capital intensive will be more likely to pay high wages than more labor intensive firms (because the cost of high wages is lower for the former). This paternalism factor makes much the same suggestion as the efficiency model, namely that the wage level in the nonunion branch of the modern sector may not fall to its market clearing level.

Another possibility concerns the impact of potential labor mobility on the desired wage level of the typical firm. In most underdeveloped economics the bulk of the labor force has very strong ties with rural areas, and very often an individual will quit a good job in the modern
sector in order to go back to his home area for a year or two. In order to counter this problem and minimize total labor costs (wages plus training costs) firms may offer very high wage rates compared to rural wages, and this can result in substantial urban unemployment even though there are no institutional imperfections such as unions.  

To the extent that this phenomenon is important in the modern sector, it also complicates the interpretation of union/nonunion wage differentials, for the costs of labor turnover may be different for union and nonunion firms.

The Murky Sector. Even if a migrant cannot find employment in either the high or relatively high tranches of the modern sector, there is still a large class of opportunities for small-scale self employment in the urban area. Much of this activity in the "murky" sector requires no capital and is not subject to any barriers to entry, so one would expect the forces of competition to work here. To the extent that a new urban in-migrant has some chance of eventually obtaining employment in the modern sector, the returns to the labor of petty capitalists should be less than the equivalent in agriculture. On the other hand, if there is only a very small chance that an individual can crack into the modern sector, returns to the urban self employed and to agricultural labor (after adjusting, of course, for differences in labor quality) should be about the same.

12 - J. Stiglitz is exploring some of the implications of this model in a paper in progress.

Testable Implications. By including the type of employment activity of the individual as an explanatory variable it is possible to estimate the relative pecuniary advantages accruing to those working in a unionized private industry or for the government relative to each other and those who work in private employment in the modern sector. It is also possible to estimate the returns to the self employed relative to other groups. The results of such an exercise will provide some insight into the validity of the various hypotheses concerning the structure of wages we have set out. For example, if government workers receive more than union workers in private industry - holding qualifications constant - the government as employer hypothesis would be supported, the unionism hypothesis refuted.

A convenient method for performing tests of this sort is to set up the regression model in the form

$$\log W = \alpha_U + \alpha_G + \alpha_{UG} + \alpha_S + \ldots,$$

where $W$ is the wage rate the individual receives, $U$ a dummy variable which is one if the individual is a union member and zero if he is not, $G$ is a dummy variable reflecting government employment, $S$ a dummy variable reflecting self employment status, and where, of course, all other variables influencing wages are included. The interaction term, $U^G$, would be included to allow for the possibility that unions influence the wages of government workers to a different extent than those who are employed in private industries. The coefficient $\alpha_U$ represents the logarithm of the difference between the wages of union and nonunion workers in private employment. Hence, $\frac{W_{U}}{W_{N}} = e^{\alpha_U}$, and the proportionate wage advantage due to union membership status is
The wage of government workers relative to unionized workers in private industries is \( a_{1} \), of the self employed relative to nonunion workers in private industries \( a_{2} \), and so forth.

To test the set of hypotheses which predict that there is a floor to the wage rate in nonunion industries in the modern sector, one would see if \( a_{2} \), the coefficient on self employment status, is significantly less than zero. If \( a_{2} = 0 \), i.e. \( W_{n} = W_{u} \), it is then fairly clear that there are no barriers to the nonunion private wage level falling to its competitive level. This test, however, is open to the "alibi" that self employment yields a nonpecuniary advantage to individuals because of the associated independence from externally imposed hours, pace of work, and the like. This argument would also imply a negative value of \( a_{3} \).

(d) Discrimination

In most societies females are at a disadvantage in the labor market relative to males, in part because most employers are males. If this is so, females will receive lower wages than males, other factors held constant. For the typical African country the other major possibility for discrimination is on ethnic grounds, i.e. by tribe.

To test for any such possibilities the regression model can include one-zero dummy variables for sex and tribal affiliation, say

\[
\log W = b_{1} F + b_{2} T_{1} + b_{3} T_{2} + \ldots,
\]

where \( F \) indicates female, \( T_{1} \) membership in a particular tribe and \( T_{2} \) membership in another, and where again, the other variables are also included in the regression. The estimated wage of females relative
to males is \( b_1 \), of tribe 1 relative to tribe 2 \( b_2 \). There are other possible interpretations of these coefficients besides pure labor market discrimination. A significantly negative value of \( b_1 \), for example, is also consistent with the hypothesis that females are simply not productive as males even when educational and other qualifications are identical. In order for labor market discrimination to take place, all employers must derive utility from employing one group over another — in this case they must be willing to sacrifice some of their profits in order to employ males instead of females, members of one tribe rather than another. If there is a fairly large class of employers who do not have such tastes and desire only to maximize profits, a negative value of \( b_1 \) would not represent labor market discrimination, for in the absence of quality differences among groups in the work force the forces of competition would drive relative wages to equality.

There is another type of discrimination in the labor market, that due to discrimination by employees rather than employers. This seems most relevant in the African case, for members of one tribe are somewhat reluctant in many instances to work with members of another tribe. Employers, on the other hand, may not really care if an employee is from one tribe or another. If the tribes are mixed in a particular workplace, however, productivity may fall because of friction or poor communication. The employer therefore hires only members of a single tribe. Marginal firms will hire members of that tribe with the lowest wage rate, and this would tend to equalize wages by tribe.

14 For the original analysis of the economic implications of various types of discrimination see G.S. Becker, *The Economics of Discrimination*, Chicago, 1977.
II. Empirical Results

In the Spring of 1971 an intensive household survey of Africans in Nairobi was conducted by the Institute for Development Studies under the auspices of the Nairobi City Council. The sample was confined to low and middle income areas of the city, which does not create much of a bias because most residents of high income areas are Asians and Europeans. The survey instrument contained a battery of questions on invariable basic socio-economic which provided sufficient information to estimate an earnings function of the sort discussed in Section I. The sample used in this paper consists of all respondents who received some wage income in Nairobi during 1970, and the dependent variable is the logarithm of the wage rate (measured in Cents Kenya) per hour of work for the respondent's modal work activity in 1970. The principal reason that we chose to explain wage rather than income differences among individuals is that a large part of the sample worked less than full time, and income earned per month is of course very sensitive to this. The wage rate, on the other hand, is a measure of an individual's earning capacity or stock of human capital. It would be very interesting for a country like Kenya to explain differences in incomes for a year for the entire adult population, for this would provide information on the determinants of unemployment and underemployment as well as wages. This, however, would require a national probability sample, for a large fraction of individuals who desire employment in the modern sector are forced to spend much of their time in rural areas.

---

15 - Wage incomes are adjusted to include direct allowances, but no attempt was made to adjust for the value of free housing in the cases in which it was provided.

The specification of the earnings function is

$$\log W = \alpha E' + \beta_1 A + \beta_2 A^2 + \beta_3 \text{ARR} + \beta_4 U + \beta_5 G + \beta_6 U \times G$$

$$+ \beta_7 S + \delta F + Y \text{TRB} + \delta G + \epsilon,$$

where

- $E'$: Vector of educational dummy variables
- $A$: Age of respondent
- $\text{ARR}$: Age arrived in Nairobi
- $U$: Union member
- $G$: Government employee
- $S$: Self-employed
- $F$: Female
- $Y \text{TRB}$: Vector of tribal dummy variables.

The level of educational attainment is entered in categoric rather than metric form to allow for the possibility that the effect of education on wages may not be "smooth." The principal cutoff points in Kenya are seven (primary) and eleven (secondary) years of schooling, and one might expect the marginal effect of a year of schooling on earnings potential to be much greater for, say, the seventh year than for the fifth or sixth. The coefficients on age and age squared, $\beta_1$ and $\beta_2$, should be positive and negative, respectively, reflecting the earlier discussion of human capital accumulation over the life cycle. The coefficient on $\text{ARR}$ should be negative, for it is the negative of the annual increment to an individual's stock of human capital due to residing in Nairobi. Expectations concerning the signs and values of $\beta_4, \beta_5, \beta_6,$ and $\delta,$ the coefficients on the type of employment activity variables, were discussed fairly thoroughly in the previous section, as was the interpretation of the coefficient on $F$. Four tribes, Kikuyu, Kamba, Luo, and Luhya, were selected for inclusion in the regression equation, for they together represent 93 per cent of the sample. The coefficient on each should be interpreted as the proportionate deviation
Table 1.
Determinants of Hourly Earnings in Urban Kenya, 1970

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1) Education Categoric</th>
<th>(2) Education Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Educ. 1-3</td>
<td>.155 (.058)</td>
<td>-</td>
</tr>
<tr>
<td>&quot; 4-6</td>
<td>.317 (.050)</td>
<td>-</td>
</tr>
<tr>
<td>&quot; 7</td>
<td>.550 (.054)</td>
<td>-</td>
</tr>
<tr>
<td>&quot; 8-10</td>
<td>1.013 (.066)</td>
<td>-</td>
</tr>
<tr>
<td>&quot; 11</td>
<td>1.535 (.078)</td>
<td>-</td>
</tr>
<tr>
<td>&quot; 12+</td>
<td>1.818 (.208)</td>
<td>-</td>
</tr>
<tr>
<td>E</td>
<td>-</td>
<td>.010 (.014)</td>
</tr>
<tr>
<td>E²</td>
<td>-</td>
<td>.011 (.001)</td>
</tr>
<tr>
<td>A</td>
<td>.083 (.009)</td>
<td>.086 (.009)</td>
</tr>
<tr>
<td>A²</td>
<td>-.00079 (.00011)</td>
<td>-.00080 (.00011)</td>
</tr>
<tr>
<td>ARR</td>
<td>-.013 (.002)</td>
<td>-.013 (.002)</td>
</tr>
<tr>
<td>U</td>
<td>.274 (.041)</td>
<td>.265 (.041)</td>
</tr>
<tr>
<td>G</td>
<td>.165 (.053)</td>
<td>.153 (.053)</td>
</tr>
<tr>
<td>UMG</td>
<td>-.178 (.071)</td>
<td>-.164 (.070)</td>
</tr>
<tr>
<td>S</td>
<td>-.717 (.064)</td>
<td>-.725 (.063)</td>
</tr>
<tr>
<td>F</td>
<td>-.238 (.054)</td>
<td>-.208 (.054)</td>
</tr>
<tr>
<td>KIK</td>
<td>-.112 (.065)</td>
<td>-.091 (.065)</td>
</tr>
<tr>
<td>KAM</td>
<td>-.036 (.071)</td>
<td>-.001 (.073)</td>
</tr>
<tr>
<td>LUO</td>
<td>.008 (.073)</td>
<td>.024 (.073)</td>
</tr>
<tr>
<td>LUH</td>
<td>.112 (.076)</td>
<td>.133 (.076)</td>
</tr>
<tr>
<td>Const.</td>
<td>2.83 (.185)</td>
<td>2.75 (.184)</td>
</tr>
</tbody>
</table>

R² = .490

a - Estimated standard errors of regression coefficients in parentheses beside estimated coefficients.
of the earnings of the particular tribe from those of "others" (minor Kenya tribes and a scattering of individuals from Tanzania, Uganda, Ethiopia, the Congo, and Sudan), other things equal.

The initial results for a sample of 1254 individuals are presented in Table 1, and the tightness of the fit indeed confirms the notion that the wage structure in urban Kenya can be explained by a set of socio-economic variables. The results are best interpreted one set of variables at a time.

**Education**

The single most important variable explaining the logarithm of an individual's hourly earning is his level of educational attainment (the simple correlation coefficient between log W and E, measured in years, is .427). Two regressions are presented in Table 1, (1) which enters education in categoric form and (2) which enters education in metric form with a quadratic. There is little to choose between the alternative forms, for the categoric case appears to be well represented by the second degree polynomial in E. When education is entered in both categoric and metric forms, a regression which is not reported, all of the coefficients associated with education become insignificant due to multicollinearity, which one would expect from (1) and (2). The coefficients on the other variables are little affected by the way education is entered into the regression.

We will first discuss the categoric results, equation (1), for the coefficients, which represent logarithmic deviations of the particular educational category from a zero level of formal educational attainment are more easily interpreted. The results indicate an increasing marginal
proportionate effect of education on wages as the educational level is increased.\(^{18}\) Table 2 presents the percentage earnings advantage due to moving from a particular educational category to each higher one.\(^{19}\) (The category 12+ was excluded from this table because of the paucity of observations in this range.) The underscored figures are the percentage increases due to moving to the next highest category. Taking the midpoint of each category (2 for 1-3, etc.), the average percentage increment due to a year of schooling are 8.5 for 0 to 2, 6 for 2 to 5, 13.5 for 5 to 7, 29.5 for 7 to 9, and 34.0 for 9 to 11. As pointed out in fn. 6, this is a rough measure of the conventional rate of return to schooling when out-of-pocket (or out-of-treasury, for the social rate) costs are ignored. It should be pointed out that if the conventional rate

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>17</td>
<td>38</td>
<td>73</td>
<td>175</td>
<td>366</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>18</td>
<td>49</td>
<td>136</td>
<td>297</td>
<td>2</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>27</td>
<td>101</td>
<td>267</td>
<td>5</td>
<td>R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>59</td>
<td>169</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>68</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^{18}\) Increasing returns to education, at least through secondary school, are not uncommon for underdeveloped countries. See, for example, Martin Carnoy, "Earnings and Schooling in Mexico," Economic Development and Cultural Change, July 1967, pp. 408-19.

\(^{19}\) For example, the estimated value of \(\log W_{11} - \log W_0 = 1.54\), so \(W_{11}/W_0 = e^{1.54} = 4.66\). The percentage increment due to moving from 0 to 11 years of schooling is then \(100 \times (4.66 - 1) = 366\).
of return to schooling had any especial significance in Kenya it
would be incumbent upon us to work out more sophisticated estimates
of the structure of returns. This would involve obtaining figures for
private and social costs at various levels of \( E \) and equating present
values by iteration. However, as was noted earlier, the probability
of obtaining a job in the modern sector versus either a position in
agriculture or unemployment, as well as estimates of the effect of
education on productivity in agriculture, are necessary to calculate
meaningful estimates. The figures in Table 2 are thus interesting
from the point of view of the distribution of income rather than educational
policy.

Increasing marginal returns to schooling are also indicated
by equation (2). The estimated effect of education on wages is given
by \( \frac{\delta \log W}{\delta E} = 0.10 + 0.022E \). The \( t \) test for the hypothesis that
the second derivative of \( \log W \) with respect to \( E \) is positive is simply
twice the coefficient on \( E^2 \) divided by its standard error, which yields
a \( t \) value in excess of 20. In discussion of the results on the other
determinants of wages in this section, the estimates from equation (2)
will be used.

The enormous wage advantage accruing to employed persons with some
secondary schooling helps to explain the serious unemployment problem
Kenya is currently facing with secondary school leavers. On the one
hand, the huge wage advantage of secondary over primary school graduates,
169 per cent, means that it is well worth the student's time to finish
his studies even if there is a very good chance that he will not obtain
the kind of job for which he is qualified. Similarly, since education is

---

20 - On this problem see P. Kinyanjui, "The Education, Training and Unemploy-
the principal route to success in the society, there is also strong pressure on the government by the citizenry to increase the capacity of the school system. On the other hand, the high wage differential will lower the expansion (in absolute terms) of the number of new jobs in the economy suitable for school leavers. An employer will require a secondary over a primary school graduate only if the efficiency of the former is at least 161 per cent greater than the latter's. In a capital-poor country like Kenya, there cannot be expected to be many such positions. Thus, unless the wage differential between educational categories can be reduced, the problem of unemployed school leavers is likely to get worse.  

Age and Age Arrived in Nairobi

The coefficients on age and age squared are each quite significant, having the expected positive and negative signs, respectively, and the coefficient on the age at which the individual arrived in Nairobi is also significant in the expected negative direction. For a person who arrived in Nairobi at a certain age, the effect on the potential wage rate is given by \( \frac{3}{\log W} / \partial A = .073 - .00160A \), for the coefficient on ARR must be added to the coefficient on A in order to give the total effect of age. This implies that wage rates reach a maximum in cross-sectional terms at about age 46. The predicted difference in the logarithm of the wage rate of a worker 46 years old and one 22 years old is .446, which means that the older makes about 56 per cent more than the younger worker. This is very

21 - It is a very interesting question whether the existence of large unemployment rates for school leavers will reduce the demand for education in the future. It is also interesting to speculate why unemployed school leavers do not return to agriculture, at least during peak periods of farming activity, as those with less schooling do. One possibility, which was suggested by my colleague L.D. Smith, is that unemployed secondary school leavers are in debt to their families or other persons in the villages and cannot yet admit failure.
The coefficient on ARR implies that for each year a worker has resided in Nairobi his potential wage increases by 1.3 per cent. This result probably reflects two factors. First, to some extent the coefficient picks up the "acculturation" effect discussed in Section I, which was the a priori justification for including the variable in the regression in the first place (and for including the question in the survey instrument). Second, and perhaps more important, the length of time an individual has been in Nairobi represents the number of years the individual has been able to search for the best jobs available to him, and, hence, in a probabilistic sense it is a measure of his success. Furthermore, in line with the discussion of the probable effects of education in Section I, increased supplies of educated labor have caused employers to increase hiring standards of jobs with given attributes, which makes it increasingly difficult for successive series of new entrants into the labor force to obtain high-paying positions given their qualifications.

**Type of Employment**

The results on U, G, UG, and S indicate that the type of employment activity in which the individual is engaged is an important determinant of his earning capacity. Each of the coefficients on these variables should be interpreted as the logarithmic deviation of the wage for that type of employment relative to other types.

---

22 - The estimated wage of a 46-year-old worker relative to that of a 22-year-old worker was estimated to be 1.73 in the U.S. from the regression equation in Johnson and Stafford, op. cit., which is somewhat higher than the estimate from equation (2). However, when the effect of age is estimated separately for the employed and the self-employed, as is done in Section III, the estimated age 46/age 22 differential becomes 1.74, which is virtually identical with the U.S. result.
of activity from that of nonunion workers in private industries, the
excluded category, after other things are taken into account. The
estimated advantage of each group over each other and the associated
standard error are shown in Table 3, and each of these differences
except that between unionists in public and private employment is
significantly different from zero at conventionally acceptable confidence
levels.

Table 4 presents indexes of estimated hourly earnings by type
of employment activity in relation to nonunion employees in private
industries and the percentage of the sample in each category. These

23 - Since \( \log W = \beta_{1}U + \beta_{2}G + \beta_{4}U \times G + \ldots \), the estimated wage advantage of, say, unionists in the private sector over nonunion government workers is \( \hat{\beta}_{4} - \hat{\beta}_{5} \), and its estimated standard error is

\[
\{ \nabla(\hat{\beta}_{4}) \nabla(\hat{\beta}_{5}) \}^{1/2}
\]
Table 4.

Indices of Estimated Hourly
by Type of Employment

<table>
<thead>
<tr>
<th>Employment Type</th>
<th>Index</th>
<th>% of Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private Nonunion</td>
<td>100</td>
<td>25.1</td>
</tr>
<tr>
<td>Union</td>
<td>130</td>
<td>39.0</td>
</tr>
<tr>
<td>Government Nonunion</td>
<td>116</td>
<td>9.7</td>
</tr>
<tr>
<td>Union</td>
<td>129</td>
<td>16.7</td>
</tr>
<tr>
<td>Self Employed</td>
<td>48</td>
<td>9.5</td>
</tr>
</tbody>
</table>

Results demonstrate that there is a large union/nonunion wage differential in Kenya, a wage advantage of 30 per cent in private industries and 11 per cent in government employment. It is interesting to note that estimates of the union wage advantage, for private industries from similar data sets range from 28 to 34 per cent, and these are within a single standard error of the estimate for Kenya. Nonunion workers in government earn 16 per cent more than their counterparts in private industries. Now all this suggests that both the unionism and government-as-employer hypotheses concerning the high wage level in the modern sector are relevant for Kenya, with the former perhaps being more important.

At the same time, wages of nonunion workers in private industries are more than twice the hourly earnings of the self employed, which is consistent with the different wage models which predict a floor to the wages received by nonunion employees: threat effects, legal minimum wages, productivity effects, and paternalistic wage setting. It is, unfortunately, not possible with this data set to test between these, for they each

predict the same thing, i.e. $W_n > W_s$. It is also rather difficult
to test the prediction that the wage level for the self employed will
be no greater than the wage level in agriculture, for, although there
are no official statistics, the cost of living is much greater in Nairobi
than in the countryside. The predicted hourly earnings of a typical
(i.e., the mean values of the other variables: $A = 33$, $E = 5$, $ARR = 20$,
and from the "other" tribal classification) male in self employment is
46 cents ($K$), and the customary wage in agriculture is about 38 cents. 25
Real wages are then equal if the urban cost of living is 21 per cent greater
than the rural cost of living, and it is generally believed that 21 per
cent understates the difference. 26 Thus, although we can make no claims
for the precision of the supportive statistics, it appears that the self
employed do not earn much more than individuals in agriculture.

Discrimination

Females in Nairobi earn 19 per cent less than males (males earn
22 per cent more), observable qualifications held constant. To what
extent this is due to discrimination and to what extent it is due to lower
productivity on the part of females is an open question. In the United States

---

25 - For comparative purposes, the predicted wage of this mean individual
in private nonunion employment is 97 cents and 126 cents in private
union employment. In 1970 100 cents ($K$) = $0.14$ (US) = $5.8$ p (UK).
26 - There are many problems in performing urban-rural wage comparisons,
some of which I have gone into in my "Wages, Employment, . . ." op.
cit. The major ones of these include the seasonality of work
activity in agriculture, nonpecuniary differences with respect to
location and different types of work, and the fact that a fairly
substantial portion of the typical urban income recipient's earnings
are remitted to his home area. W.E. Whitelaw
in the estimation of a "remittances function" with the data set
employed in the present study.
the wage disadvantage of females is between 25 and 30 per cent, depending on age and education. This is to a large extent explained by the fact that females generally drop out of the labor force (or are expected to do so) in order to raise families during the period of maximum human capital formation. This type of labor supply behavior is not so common in Kenya. An additional source of earnings differentials between females and males concerns the differential incidence of the various types of employment. For this sample, 35.3 per cent of females were union members as compared with 58.3 per cent of males, and 26.6 per cent of females were self employed as compared with 6.8 per cent of males. (The incidence of \( U \) and \( U^*G \) was essentially the same for the sexes.) Now the regression equation accounted for the influences of union membership and self employment status on wages, and it is interesting to find out how much additional 'discrimination' can be accounted for by these two differences. One rather crude way is to multiply the difference between the proportions of the sexes in each category times the appropriate estimated effect of that category on wages from equation (2). (It is rough because it is not possible to calculate standard errors of the total earnings differential estimated in this fashion.) This implies that because of the institution of unionism the difference between the logarithm of female and male wages is changed by \( .265 \times (1.353 - -.583) = -.061 \). Similarly, the greater incidence of self employment changes the estimated value of \( \log W_f - \log W_m \) by \( -.725 \times (.266 -.068) = -.144 \). These effects - other just about equal the direct effect of \( F \) on \( \log W \), and adding these to to the direct coefficient implies that the earnings disadvantage of females relative to males is about 40 per cent.

With respect to the tribal dummies it is most interesting to look at differences among the four major tribes, for these represent 93 per cent
of the total sample. The estimated logarithmic differences in wages by tribe - after other differences are taken into account - are shown in Table 5. The most interesting result is that Kikuyus, the largest tribe in Kenya and the politically dominant one, earn significantly less than each of the other three.

There are two principle alternative (although not mutually exclusive) explanations of this. First, Kikuyus may be discriminated against by employers, for employers, who are generally Asians and Europeans and who may resent Kikuyus for their role in the independence movement of the 1950's. Another variant of this approach is based on the fact that a large number of the major firms in Nairobi were formed or expanded significantly during the 1950's, a time in which the movement of Kikuyus into Nairobi was restricted by the colonial authorities. These high wage firms were thus staffed by members of other tribes, and, because members of particular tribes prefer to work with their "own people" (the employee discrimination model), Kikuyus have on average tended to be excluded from the better jobs.

Table 5.

Estimated Differences in Logarithm of Wages by Major Tribe

<table>
<thead>
<tr>
<th>ADVANTAGE OF:</th>
<th>KAM</th>
<th>LUO</th>
<th>LEH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.090</td>
<td>.115</td>
<td>.224</td>
</tr>
<tr>
<td></td>
<td>(.045)</td>
<td>(.047)</td>
<td>(.052)</td>
</tr>
<tr>
<td></td>
<td>0.025</td>
<td>.134</td>
<td>KIK</td>
</tr>
<tr>
<td></td>
<td>(.064)</td>
<td>(.068)</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>V</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>E R</td>
</tr>
<tr>
<td></td>
<td>0.109</td>
<td>LUO</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.061)</td>
<td></td>
</tr>
</tbody>
</table>
Second, the Kikuyu areas in Kenya are slightly closer and are quite a bit more accessible to Nairobi than the Kamba areas, and they are much more so than the Luo and Luhya areas, which are in the Western part of Kenya. Since practically all the respondents have retained their rural ties, 27 this would imply that the reservation price for prolonged residence in Nairobi would be lower for Kikuyus than for Kambas than for the Luo or Luhya. In other words, a Luo who was not fortunate enough to obtain a good job in Nairobi within a certain period of time would be more likely to go back to his home area than would a Kikuyu, for the disutility of being in Nairobi is greater for the former than the latter.

Consistent with this interpretation is the fact that the ratio of the fraction of the Nairobi population which is Kikuyu to the equivalent fraction for the nation is 1.78, whereas it is 1.34 for Kambas and 1.08 and 1.17 for Luos and Luhyas, respectively. 28 It should be stressed, however, that this economic interpretation of tribal differentials does not rule out a discrimination explanation at the same time; the two are not mutually exclusive.

III. Effect of Human Capital Variables on the Returns to the Self Employed

A rather important social question in Kenya and most other LDC's concerns what to do with the educated unemployed. If large numbers of primary and secondary school leavers cannot find employment in the modern sector of the economy, they will be forced to move into agriculture or urban self employment. This suggests a question: what extent does education increase the earnings capacity of individuals outside

27 - Rural ties are retained in the sense that individuals own or expect to inherit land in their home areas; wives are left in the home area, and the urban worker remits a portion of his income to the rural area for support of his family.

the modern sector?

The IDSUS data are obviously of no help in attacking the returns to education in the agricultural sector question, but it can be employed to answer the question for petty capitalists in urban areas. By adding interaction terms between the human capital variables and the dummy variable for self employment status to the other variables in equation (2), that is

\[ \log W = \alpha_1 E + \alpha_2 L^2 + \alpha_3 EXS + \alpha_4 A + \beta_1 A^2 + \beta_2 A^3 + \beta_3 \times S + \ldots, \]

one can test whether the human capital variables, individually and in concert, influence the earnings of the self employed differently from the way they influence the earnings of those individuals employed by business-firms and the government. The results of such a test are shown in Table 5. (The coefficients on the other variables are little changed by the introduction of the interaction terms and are thus not reported.) For comparative purposes the corresponding results from equation (2) are also shown. The null hypothesis that the effect of the human capital variables on the earnings of the two groups, the self employed and employees, is the same is decisively refuted (F = 7.79), which is to say that the increase in the variance explained is worth the loss of five degrees of freedom.

The estimated effect of education on the earnings of the self employed is \( \alpha_1 + \alpha_2 + 2(\alpha_2 + \alpha_3)E = .020 + .002E, \) which is less than

29 For a very interesting attempt to estimate the effect of schooling on earnings in the agricultural sector of an [LDC see Alberto Valdes S., "Wages and Schooling of Agricultural Workers in Chile," Economic Development and Cultural Change, January 1971, pp. 313-29.}
Table 6.
Effects of Introducing Interaction between Human Capital Variables and Self Employment Status

<table>
<thead>
<tr>
<th>Variable</th>
<th>Interactive Model</th>
<th>Equation (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient on Variable</td>
<td>Coefficient on Interaction Term</td>
</tr>
<tr>
<td>E</td>
<td>.022 (.015)</td>
<td>-.001 (.006)</td>
</tr>
<tr>
<td>E^2</td>
<td>.011 (.001)</td>
<td>-.010 (.007)</td>
</tr>
<tr>
<td>A</td>
<td>.096 (.009)</td>
<td>-.074 (.029)</td>
</tr>
<tr>
<td>A^2</td>
<td>-.00093 (.00014)</td>
<td>.00095 (.0003)</td>
</tr>
<tr>
<td>ARR</td>
<td>-.010 (.002)</td>
<td>-.030 (.007)</td>
</tr>
<tr>
<td>S</td>
<td>1.47 (.36)</td>
<td>-.725 (.063)</td>
</tr>
<tr>
<td>R^2</td>
<td>.507</td>
<td>.491</td>
</tr>
</tbody>
</table>

Table 7.
Estimated Derivative of log W with Respect to E for Employees, Self Employed, and Difference

<table>
<thead>
<tr>
<th>E</th>
<th>Employees</th>
<th>Self Employed</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.043 (.028)</td>
<td>.013 (.043)</td>
<td>.030 (.045)</td>
</tr>
<tr>
<td>4</td>
<td>.102 (.026)</td>
<td>.019 (.022)</td>
<td>.083 (.024)</td>
</tr>
<tr>
<td>7</td>
<td>.175 (.026)</td>
<td>.025 (.058)</td>
<td>.150 (.060)</td>
</tr>
<tr>
<td>11</td>
<td>.263 (.030)</td>
<td>.033 (.110)</td>
<td>.230 (.120)</td>
</tr>
</tbody>
</table>
the estimated effect for employees. \( \alpha_1 + 2\alpha_2 E = .021 + .022E \), for all values of \( E \). The estimated influence of \( E \) on the logarithm of the wage rate and the associated estimated standard errors for employees, the self-employed, and the differential effect of education on the two groups for selected values of \( E \) is shown in Table 7. The most interesting result is that whereas the estimate of \( \frac{3}{E}(\log W)\frac{3E}{3E} \) is significantly greater than zero for \( E > 2 \) for employees it is not so for any value of \( E \) for the self-employed. The estimated difference in the effect of education on the earnings of the two groups is significant for \( E > 3 \).

The interaction results also suggest that there is no relation between wages and age for the self-employed. The derivative of \( \log W \) with respect to \( A \) is \( -.018 + .00004A \) for the self-employed, which is negative for all relevant values of \( A \), but its standard error is always at least as great as its value. For employees, on the other hand, the estimated value of \( \frac{3}{A}(\log W)\frac{3A}{3A} \) is significantly positive for \( A < 43 \) and significantly negative for \( A > 50 \). (\( W \) is highest in a relative sense at \( A = 46 \).) The effect of the age arrived in Nairobi variable, however, is significantly greater for the self-employed than for employees. This most likely represents a selection effect. Individuals who are successful petty capitalists - for whatever the reason - will tend to stay in this occupation. Those who are less successful will search harder for a job or return to their home areas to work in agriculture.
IV. Summary and Conclusions

The principal empirical findings of this study of wages in Kenya's modern sector can be summarized as follows:

(1) There is an increasing marginal proportionate effect of education on potential earnings of individuals from 0 through 11 years of schooling. The estimated derivative of the logarithm of the wage with respect to education increases from .010 at E = 0 to .252 at E = 11. This is very difficult to justify economically since there is a surplus of educated labor in Kenya.

(2) The effect of age on earnings in Kenya is quite similar to that in developed countries. Relative wages rise with age until age 46 and fall thereafter. The age at which the individual migrated to Nairobi also has a negative effect on earnings, which probably mainly reflects the combination of seniority systems and increasing supplies of would-be entrants to the modern sector.

(3) Union wages are estimated to be 30 per cent greater than nonunion wages in private industries and 11 per cent greater in government employment, other things held constant. For nonunion workers, government employment increases wages by 16 per cent. The results appear to be most consistent with the view that unionism is the principal cause of high wages in the modern sector of Kenya.

(4) At the same time, the nonunion wage level in private industries is more than twice the hourly returns for the self-employed, which suggests in turn that even the nonunion labor market is in disequilibrium with respect to the rural wage level.

(5) The earnings of females in Nairobi are on average 1 per cent lower than those of males, other observable factors held constant. Females, however, tend to work in nonunion occupations and to be self-employed.
to a greater extent than males, both of which factors exacerbate the earnings disadvantage of females, and upon adding the effect of these to the differential, the wage disadvantage of females rises to forty per cent.

(6) Earnings of members of the largest tribe in Kenya, the Kikuyu, are significantly less than any of the other three major tribes, which is consistent with an interpretation of the existence of one or another discriminatory practices against them. This result, however, is also consistent with simple migration theory, for the home areas of the Kikuyu are closest to Nairobi of any of the tribes.

(7) The effect of the human capital variables, age and education, is systematically different for the self employed and employees. This suggests that petty capitalism is an alternative to the modern sector only for the relatively uneducated.

In conclusion, it seems appropriate to stress the usefulness of the periodic gathering of data sets comparable to this one for different countries. If such a data set had been available for Kenya in, say, 1965, we would have been able to evaluate the trends in the returns to education, the relation between earnings and age, advantages to union membership, and so forth. This would have permitted the testing of a wider range of hypotheses than was possible with this single cross-section. For example, with data for two time periods one could answer the question of whether the government wage level leads or lags the private industry union wage level. Another set of questions which could be attacked concerns the changes in educational qualifications for entry into the better jobs in the modern sector. We would also like to point out that future surveys could profitably obtain certain types of information which were not included in the IDSUS survey. First, it is important to obtain
as much information on the background of the individual as possible, the occupation of the father, the size of the community in which the individual grew up, the number of older and younger brothers and sisters, and so forth. The importance of particular indicators of background varies among countries, of course, and the selection of the "optimal" set requires knowledge of both the techniques of economic analysis and the culture of the particular country, but inclusion of such variables would aid the estimation of the effect of variables like those included in this study. Second, recent developments in human capital theory have stressed the importance of post-formal training in the accumulation of individual human capital. This is probably even more important in LDC's than in developed economies. Accordingly, future surveys should attempt to find out as much as possible about on-the-job training, occupational history, and formal technical training of individuals.

It would be very interesting to contrast the results of similar studies to this one for a number of different countries (and, where appropriate, different urban areas within specific countries). This would permit the analysis of the effects of various institutional differences - with respect to the stage of economic development, type of economic and political organization, racial and ethnic homogeneity, and other cultural delineations - among countries on labor market behavior and the personal distribution of income.