

EDUCATION AND INCOME DIVERSIFICATION STRATEGIES IN KENYA

Anthony Wambugu

University of Nairobi, School of Economics, P. O. Box 30197-00100, Nairobi, Kenya

ABSTRACT

Using a representative national household survey and multinomial logit regression, this study analyses the relationship between education and household choice of economic activity combination among four economic activities: wage work, wage work and farming, wage work and family enterprise, and wage work, farming and family enterprise. The results reveal pervasive combination of income generating activities among rural and urban households in Kenya. Completion of primary education and completion of secondary education are a key driving force encouraging income generation from multiple sources. Education attainment beyond secondary level tends to drive households into specializing in wage work. The results suggest that, education plays an important role in the process by which both rural and urban households generate income and sustain livelihoods, by influencing access to different income generating activity combinations in Kenya. There are also notable lifecycle and gender differences in income source diversification.

Keywords: Education, activity combination, multinomial logit, Kenya

JEL classification: O15, I26, J22, R20

1. INTRODUCTION

Models of labor mobility (for example Lewis, 1954; Harris and Todaro, 1970) envisaged labor mobility from rural subsistence sector with surplus labor, to an urban industrial sector. Economic transformation from subsistence agriculture to modern industrial urban economy would not change marginal product of rural labor or increase urban wages. Subsequent work showed that the implied constant urban wage might not hold (Harris and Todaro, 1970), growth may take place in the agricultural sector (Ranis and Fei, 1961) and differential returns to capital can induce capital mobility (Corden and Findlay, 1975).

Further, labor mobility takes other forms. For example, Bigsten (1996) found that rural to urban migration in Kenya may be circular. Other studies show that rural off-farm employment absorbs surplus labor such that households can allocate labor to multiple economic activities without migrating to urban areas (Reardon, 1997). Moreover, urban households also diversify income sources (Ellis, 1998). For example, Bigsten and Kayiizi-Mugerwa (1992) found that in Kampala, Uganda households adjusted to the 1970s and 1980s economic crisis by diversifying income sources.

A large literature exists on drivers of income diversification in developing countries (Reardon, 1997; Ellis, 1998; Reardon et al, 2007). However, the focus is on rural households (e.g. Barrett et al, 2005; Bigsten, 1985; Bigsten and Kayiizi-Mugerwa, 1995; Dercon and Krishnan, 1996; de Janvry and Sadoulet, 2001; Adams, 2002; Lay et al, 2008; Yunez-Naude and Taylor, 2001). Given that increased urbanization in Africa has not been accompanied by adequate wage employment growth, it is important to investigate household income strategies separately for rural and urban households. Such evidence can offer additional insight to policy makers on how labor allocation takes place.

An important focus of the development literature has been to examine the role of education as a driver of economic outcomes. For example, there is a vast literature on the earnings benefits of education in developing countries (Appleton, et. al, 1996; Barouni and Broecke, 2014). However, comparatively less is known about the relationship between education and household economic activity combinations. This paper contributes to the literature by examining the structure of household economic activities and how education attainment and other factors shape this structure in both rural and urban Kenya. Another contribution of the paper is that it examines not only the role of education of the household head but also that of other household members. Many household level studies of returns to education focus on the education attainment of household head. There is some evidence for Ghana (Jolliffe, 1998) that using education of household head is restrictive.

The paper is organized as follows. Section 2 presents the analytical framework of economic activity combination, while the first part of Section 3 describes the data to establish the extent of economic activity combination. The second part of Section 3 presents the econometric analysis of the relationship between activity combination behavior and education among other determinants. Section 4 concludes and summarizes the paper.

2. ANALYTICAL FRAMEWORK

Household n chooses among J economic activity combinations and derives a certain utility level, U_{nj} from each activity combination. The household knows the utility level but the researcher

does not. The household chooses the activity combination with highest utility. Therefore, economic activity combination i is chosen if and only if $U_{ni} > U_{nj} \quad \forall j \neq i$. The researcher can observe some household characteristics, X_{nj} and specify a linear function, $V_{nj} = X_n \beta_j$ that relates the characteristics to the household's utility where β_j is a vector of parameters for a specific activity combination. Because some aspects of household utility are not observed by the researcher, $V_{nj} \neq U_{nj}$. Therefore, household utility is decomposed as $U_{nj} = V_{nj} + \varepsilon_{nj}$ where ε_{nj} is a random term that represents factors that influence utility but are not in V_{nj} .

The joint probability density of the random vector $\varepsilon_n = \{\varepsilon_{n1}, \dots, \varepsilon_{nJ}\}$ is $f(\varepsilon_n)$. Based on this the researcher can make probabilistic statements about the household activity combinations. Specifically, the probability that household n chooses combination i is

$$\begin{aligned}
 P_{ni} &= \text{Pr ob}(U_{ni} > U_{nj}) \quad \forall j \neq i &= \text{Pr ob}(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj}) \quad \forall j \neq i \\
 &= \text{Pr ob}(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj}) \quad \forall j \neq i & \tag{1}
 \end{aligned}$$

The probability is a cumulative distribution, that is, the probability that each random term $(\varepsilon_{nj} - \varepsilon_{ni})$ is below the observed quantity, $(V_{ni} - V_{nj})$. Given the random utility model, a discrete choice model is the suitable methodology to study household income activity combinations. A multinomial logit (MNL) model (Greene, 2008) was estimated. The focus is on the probability that household n chooses activity combination i , p_{ni} controlling for several covariates. The multinomial logit response probabilities take the form

$$\text{Pr ob}(y_n = i | X_n) = \frac{e^{X_n \beta_i}}{\sum_{k=1}^J e^{X_n \beta_k}} \tag{2}$$

The model is identified by normalizing β_j to zero for $j=1$. Hence, the choice probabilities are

$$\text{Pr ob}(y_n = 1 | X_n) = \frac{1}{1 + \sum_{k=2}^J e^{X_n \beta_k}} \tag{3}$$

$$\text{Pr ob}(y_n = j | X_n) = \frac{e^{X_n \beta_j}}{1 + \sum_{k=2}^J e^{X_n \beta_k}} \tag{4}$$

The model in three sets of explanatory variables: human capital (education); other characteristics of the labor force (size, age, gender); and geographic characteristics (region and area of residence). To measure the impact of education, six dummy variables for the highest level of household head's education, and six dummy variables for the highest education level among other labor force members are included. The second set of dummy variables is included because even where the head is not educated there may be another adult who is educated. The average age of labor force and average age squared are proxies for accumulated experience and household's position along the lifecycle.

Households may be constrained in choice of activity combinations by quantity of labor. Therefore, household labor force, measured as number of household members 15 to 65 years old and not in school is controlled for. A larger labor force can widen the scope for engaging in multiple activities through reduction in indivisibilities problem that arises due to inadequate labor (Bigsten, 1996). Gender differences and household responsibilities are allowed for by adding a dummy for male-headed households. The share of women in the labor force is also included to take account of the importance of women labor in the households. The two gender variables can also reflect differences in preferences towards activity combinations.

Income generating opportunities are likely to differ across Kenyan regions. To capture regional differences, six dummy variables for province of residence are included. The variables can also control for regional differences in output and inputs prices, state of infrastructure and availability of public services (e.g. health care and school facilities).

To derive unique probabilities for the multinomial logit model, coefficients for wage only households are normalized to zero. The coefficients of other categories are interpreted relative to this category.

The signs on the coefficients show how risk-ratio, $p(y_n=j)/p(y_n=0)$, the ratio of probability of a household engaging in activity mix j , relative to wage work changes when a covariate changes.

The magnitudes of the coefficients do not reflect partial changes. A positive coefficient would not imply that a rise in the associated variable increases p_{ni} , the probability of household n choosing i . The probability of another outcome may rise relatively more such that, p_{ni} falls. Thus the direction of change in $p(y_n=j)$, for a change in an explanatory variable is not clear from the sign of associated coefficient. The change in probability of choosing a given alternative from a change in an independent variable depends on values of all explanatory variables and coefficients of each every alternative in the choice set. Consequently, the estimated coefficients and associated partial effects can have different signs (Long, 1997). For continuous variables the change in probability is

$$\frac{\partial \Pr(y = i | x)}{\partial x_k} = P(y = i | x) \left[\beta_{ki} - \sum_{j=1}^J \beta_{kj} \Pr(y = j | x) \right] \tag{5}$$

For dummy variables, the difference in probability when x_k jumps from x_s to x_e is given by

$$\Delta \Pr(y = i | x) = \Pr(y = i | x, x_k = x_e) - \Pr(y = i | x, x_k = x_s) \tag{6}$$

The model is based on the assumption that each of the random components in the utilities is independently identically distributed with Type I extreme value distribution. This assumption gives rise to the independence from irrelevant alternatives (IIA) property (Long, 1997). The MNL restriction on agents' choices that the IIA assumption implies would require that the probability of holding a given activity combination be independent of other combinations. The Hausman test of IIA is used to test this restriction on choices. The test compares maximum likelihood estimates based on full sample with maximum likelihood estimates in which one combination is dropped, while households that actually had the combination are dropped. Under the null hypothesis the two sets of estimates should be close.

3. ESTIMATION RESULTS AND DISCUSSION

To determine the extent of economic activity combination by households, we analyze data from the National Household Welfare Monitoring and Evaluation Survey [WMSII) collected by the Kenya Central Bureau of Statistics. It is the most extensive nationally representative survey available. A wide range of information on household characteristics, expenditures, incomes and economic activities was collected. The analytic sample for this study contains 9,183 households. Table 1 shows the distribution of households over seven economic activity combinations.

Table 1: Percentage distribution of household economic activity combinations

Economic activity combinations	Total	Rural	Urban
Wage work only	18(17)	13(12)	52(53)
Farming only	1(1)	2(1)	0(0)
Own business only	0(0)	0(0)	0(0)
Wage work and farming	49(49)	54(55)	13(13)
Wage work and own-business	7(7)	4(3)	26(26)
Farming and own business	0(0)	0(0)	0(0)
Wage work, farming, and own business	23(25)	26(28)	8(8)

Note: Figures in () are for the analytic sample while figures outside the () are for data as they are in the data file. Column figures may not add to 100 because a small percentage of households have no earnings in listed activities.

The results show that income activity combination is widespread in both urban and rural Kenya. Little empirical evidence exists on the issue in urban areas. Table 1 shows that forty-nine per cent of the sample households combined farming with wage work; the proportion is higher in rural areas (55 per cent), than in urban areas (13 per cent). A quarter of sample households engaged in all the three activities; the proportion in rural areas (28 per cent) is almost four times that in urban areas. Seventeen per cent of the households engaged in wage work only; the proportion of households in this category is larger in urban areas (53 per cent) than in rural areas (12 per cent). Similarly, while a quarter of urban households engaged in both wages and own business, less than 5 per cent of rural households had this mix.

Maximum likelihood estimates of the MNL model are reported in Table 2. The activity combination model (2) was estimated on the full, rural and urban samples respectively. Because households' activity combination choices within a cluster may be related, the standard errors are adjusted for cluster effects. Also, because estimation requires adequate sample size in each category of the dependent variable, analysis is restricted to the four most frequent activity combinations: wage work only; farming and wage work; wage work and own business; and wage work, farming and own business.¹ The corresponding partial effects of estimated parameters of education variables computed following equations (5) and (6) are reported in Tables 4. The partial effects of the education variables for the urban sample are mostly statistically insignificant except at post-secondary education level. The signs of the partial effects also contrast sharply with those of the rural sample.

¹Cases of own business only, farming only, and farming and business only are few (less than a percent of sample).

Table 2: Multinomial logit estimates of household activity combinations

Explanatory variables	Rural Kenya			Urban Kenya		
	Farm/Wage	Business/Wage	Farm/Business/Wage	Farm/Wage	Business/Wage	Farm/Business/Wage
Age (years)	0.13*** (4.85)	0.07 (1.09)	0.13*** (4.39)	0.03 (0.34)	0.13** (2.04)	0.30** (2.05)
Age squared	-0.0012*** (3.50)	-0.0013 (1.47)	-0.0013*** (3.39)	-0.0001 (0.11)	-0.0017* (1.87)	-0.0040* (1.81)
Male household head	0.21 (1.61)	0.14 (0.63)	0.30** (2.16)	0.11 (0.32)	0.16 (0.52)	0.65 (1.47)
Education: Household head						
Some primary	0.65*** (4.90)	0.67*** (2.81)	0.79*** (5.52)	-0.23 (0.55)	0.24 (0.79)	0.04 (0.07)
Full primary	0.38** (2.13)	0.35 (1.22)	0.69*** (3.58)	0.27 (0.63)	-0.21 (0.63)	0.07 (0.13)
Some secondary	0.09 (0.42)	0.55* (1.67)	0.58** (2.57)	-0.08 (0.18)	-0.21 (0.68)	-0.54 (0.91)
Full secondary	0.22 (1.10)	0.48 (1.59)	0.60*** (2.84)	-0.10 (0.23)	-0.21 (0.69)	-0.00 (0.01)
Post secondary	-0.57** (2.03)	0.24 (0.51)	-0.19 (0.61)	-0.33 (0.55)	-0.39 (1.05)	-2.36* (1.96)
University	-1.61*** (3.12)	0.26 (0.33)	0.06 (0.11)	-0.25 (0.40)	0.06 (0.12)	-1.87 (1.40)
Education: non-household head						
Some primary	1.03*** (7.10)	0.61** (2.36)	1.17*** (7.21)	0.21 (0.61)	0.18 (0.74)	0.66 (1.36)
Full primary	1.19*** (7.07)	0.65** (2.24)	1.23*** (6.61)	-0.56 (1.50)	0.16 (0.54)	0.65 (1.20)
Some secondary	1.07*** (5.36)	0.69* (1.81)	1.09*** (5.14)	-0.32 (0.84)	0.10 (0.33)	0.52 (0.98)
Full secondary	1.04*** (5.61)	0.73** (2.12)	1.01*** (4.90)	-0.19 (0.41)	0.46 (1.46)	0.74 (1.49)
Post secondary	1.20** (2.48)	1.11 (1.44)	1.46*** (2.85)	-0.16 (0.27)	-1.04* (1.96)	0.07 (0.08)
University	0.68 (1.32)	-0.40 (0.31)	0.08 (0.15)	-1.63 (1.41)	-0.47 (0.70)	0.80 (0.57)
Share of women	1.12*** (5.96)	0.08 (0.21)	0.99*** (4.73)	-0.11 (0.25)	0.38 (1.02)	1.38** (2.23)
Log Labor force	0.38*** (3.37)	-0.35 (1.40)	0.70*** (5.35)	0.75** (2.18)	-0.27 (1.02)	0.33 (0.74)
Constant	-3.54*** (6.05)	-2.34** (2.06)	-4.37*** (6.69)	-2.79* (1.90)	-3.35*** (2.95)	-10.53*** (4.09)
Sample size	7751	7751	7751	1304	1304	1304
Pseudo R ²	0.09			0.08		
Log-Likelihood	-7405.97			-1390.78		

Note: Robust z-statistics within parentheses. * Significant at 10%; ** significant at 5%; *** significant at 1%. Regressions include dummy variables for region of residence.

Table 3: Hausman test statistic of IIA assumption for multinomial logit models

	Rural			Urban		
	Chi-square	p-value	Evidence	Chi-square	p-value	evidence
(a)	3.47(46)	1	For Ho	-0.95(46)	1	For Ho
(b)	0.22(46)	1	For Ho	0.24(46)	1	For Ho
(c)	1.98(46)	1	For Ho	0.15(46)	1	For Ho

Note: The omitted combination in (a) farming and wage; (b) business and wage; (c) wage, business and farming. Degrees of freedom in parentheses

Table 4: Multinomial logit estimates, partial effects at means of explanatory variables

Education	Rural			Urban		
	Farm/Wage	Business/Wage	Farm/Wage/Business	Farm/Wage	Business/Wage	Farm/Wage/Business
Household head						
Some primary	0.005***	0.001***	0.044***	-0.028	0.055	-0.001
Full primary	-0.040**	-0.002	0.076***	0.035	-0.049	0.005
Some secondary	-0.093	0.008*	0.105**	0.001	-0.031	-0.021
Full secondary	-0.065	0.004	0.086	-0.003	-0.038	0.003
Post-secondary	-0.114**	0.021	0.054	-0.015	-0.048	-0.055*
University	-0.382***	0.035	0.272	-0.019	0.032	-0.048
Non-household head						
Some primary	0.026***	-0.009**	0.053***	0.011	0.017	0.035
Full primary	0.046***	-0.010**	0.033***	-0.053	0.033	0.041
Some secondary	0.042***	-0.007*	0.025***	-0.034	0.019	0.031
Full secondary	0.050***	-0.005**	0.014***	-0.035	0.087	0.038
Post-secondary	-0.012**	-0.003	0.078***	0.006	-0.156*	0.017
University	0.139	-0.014	-0.091	-0.087	-0.077	0.081

Notes: corresponding coefficients * significant at 10%, ** significant at 5%, and *** significant at 1%. Partial effects of region dummies and other variables are omitted for brevity.

Households headed by individuals with incomplete primary education are significantly more likely than those without education to combine economic activities. The largest partial effect is to combine the three activities. This result holds for other household members with the exception that it reduces the probability of combining business and wage work. Thus with only some primary education, a household is less unlikely to rely on wage income only. This may reflect that this level of education is unlikely to lead to well paid wage jobs.

The partial effects of full primary education attainment have a mixed pattern. Full primary education of the household head significantly reduces the probability of a household combining farming and wage work by 4 percentage points and raises the probability to combine the three activities by 7.6 percentage points. If the most educated other household member (not the head) has attained full primary education, the probability to combine farm work and wage work as well as to combine the three activities is significantly higher (4.6 percentage points and 3.3 percentage points respectively) than in households where the most educated other household member has no formal education.

Secondary education attainment also increases chances of households combining income generating activities. If household head has some secondary education, the probability of the household to engage in the three activities goes up by 10.5 percentage points. Some or full secondary education attainment of other household labor force significantly influences the probability of engaging in the three activities, but it has relatively larger impacts on engagement in farming and wage work.

Post-secondary education and university education are significant predictors of some activity combinations. In rural areas, a household whose head has post secondary education or university education is significantly less likely to combine farming and wage work relative to wage work only and the partial effects are relatively large (11.4 percentage points and 38.2 percentage points respectively). Households where the other most educated member (not the head) has post-secondary education, are less likely to combine wage work and farming, but more likely to combine the three activities. In urban areas, post-secondary education of the household head reduces the probability of combining the three activities by 5.5 percentage points. In contrast, if the other most educated household member has post-secondary education, the chance for such a household to combine business and wage work is reduced by 15.6 percentage points.

Average age of household labour force is a significant predictor of activity combinations and the relationship is concave, meaning there is a turning point. The older the household labor force the more likely rural households are to combine farming and wage work or farming, wage work and business. Urban households with older laborforce are more likely to engage in wage work and

business or farming, wage work and business. This suggests that diversification of economic activities takes time. Households develop networks overtime that helps them enter new income generating activities. Household savings overtime can help break liquidity constraints in setting up own business. Households with younger labor force engage in wage work only.

The results reveal notable differences in activity combination strategies of male headed households and female headed households. Male-headed households in rural areas are more likely to combine the three economic activities. This may signal scope to leave the primary income earning activity to other household members, while the household head supplies labor to other activities. Further, in rural areas, the share of women in the household labor force raises the probability of the household combining farming and wage work. It also raises the probability to combine the three activities for rural and urban households.

The size of the household labour force also significantly predicts the probability of some activity combinations. The larger the household labor force, the more unlikely it is that a rural household will combine farming and wage work. Instead, the more likely it is that the household will combine the three activities. A larger household labor force may help dilute the indivisibilities problem since more workers are available to take advantage of income earning opportunities. For the urban households, larger labour force raises probability of combining wage work and farming work. These may be households that own land in rural areas but reside in urban areas.

4. SUMMARY AND CONCLUSION

This study examined the nature and predictors of household income generation activity combinations in Kenya, with emphasis on the role of education attainment. In rural areas the most prevalent combination is wage work and farming, followed by a combination of wage work, farming and household business. In contrast, the most prevalent activity in urban areas is wage work only, followed by the combination of wage work and business.

A key predictor of household activity combination is education attainment. But whose education matters? Most previous studies at household level focus on the education of household head only. In this study households headed by persons whose highest education attainment is primary level tend to combine farming, wage work and household business. But households headed by highly educated persons tend to specialize on wage work. Education of other household workers induces households to combine farming, wage work and own business. It also increases access to the farming and wage work combination. The results also revealed gender differences in diversification strategies.

Education has direct monetary return as previous studies show. This study shows that a benefit of education that is often not considered is that schooling is a key predictor of household activity combinations. This is probably because education enhances the ability of households to take advantage of emerging income generation activities within rural or urban areas. Education policy can continue to increase access to secondary and post-secondary education to help households' access high return wage activities. Future studies can estimate the causal effect of education on household activity combinations if valid and relevant instruments can be found. Further investigation is required to understand gender differences in incidence of activity combinations and predictors of activity combinations; the dynamics of household income diversification; and welfare implications.

REFERENCES

Adams, R.H. (2002), Nonfarm income, inequality, and land in rural Egypt. *Economic Development and Cultural Change*, 50(2): 339-364.

Appleton, S., Hoddinot, J., and McKinnon, J. (1996), Education and health in Sub-Saharan Africa, *Journal of International Development*, 8(3): 307-339.

Barouni, M. and Broecke, S. (2014), The Returns to Education in Africa: Some New Estimates, *Journal of Development Studies*, 50(12): 1593-1613

Barrett, C.B., Clark, M.B., Clay, D.C. and Reardon, T. (2005), Heterogeneous constraints, incentives and income diversification strategies in rural Africa, *Quarterly Journal of International Agriculture* 44(1): 37-60.

Bigsten, A. (1996), The Circular Migration of Smallholders in Kenya. *Journal of African Economies*, 5(1): 1-20

Bigsten, A. (1985), "What do Smallholders do for a Living? Some Evidence from Kenya", in M. Lundahl (ed.), *The Primary Sector in Economic Development*, Croom Helm, London, 1985

Bigsten, A. and Kayiizi-Mugerwa, S. (1995), Rural sector responses to economic crisis in Uganda. *Journal of International Development*, 7(2): 181-209.

Bigsten, A. and Kayiizi-Mugerwa, S. (1992), Adaptation and distress in the urban economy: A study of Kampala households. *World Development*, 20(10): 1423-1441

Corden, W. and Findlay, R. (1975), Urban unemployment, intersectoral capital mobility and development policy. *Economica*, 42(165): 59-78.

de Janvry, A. and Sadoulet, E. (2001), Income strategies among rural households in Mexico: The role of off-farm activities. *World Development*, 29(3): 467-480.

Dercon, S. and Krishnan, P. (1996), Income portfolios in rural Ethiopia and Tanzania: Choices and constraints. *Journal of Development Studies*, 32(6): 850-875.

Ellis, F. (1998), Household strategies and rural livelihood diversification. *Journal of Development Studies*, 35(1): 1-38.

Greene, W. H., (2008), *Econometric Analysis*, Sixth edition, Upper Saddle River, New Jersey: Prentice-Hall,

Harris, J. and Todaro, M. (1970), Migration, unemployment and development: A two sector analysis, *American Economic Review*, 60(1): 126-42

Jolliffe, D. (1998), Skills, schooling, and household income in Ghana. *The World Bank Research Observer*, 12(1): 81-104

Lay, J., Mahmoud, T.O. and M'Mukaria, G.M. (2008). Few opportunities, much desperation: the dichotomy of non-agricultural activities and inequality in western Kenya, *World Development* 36(12): 2713–2732.

Ranis, G. and Fei, J. (1961), A theory of Economic development, *American Economic Review*, 51(4): 533-565.

Reardon, T. (1997), Using evidence of household income diversification to inform study of rural nonfarm labor market in Africa. *World Development*, 25(5): 735-747.

Reardon, T., Stamoulis, K. and Pingali, P. (2007), Rural nonfarm employment in developing countries in an era of globalization, *Agricultural Economics*, 37(1): 173-183.

Long J (1997), Regression models for categorical and limited dependent variables, Sage Publication Thousand Oaks, California.

Yunez-Naude, A. and Taylor, J.E. (2001), The determinants of nonfarm activities and incomes of rural households in Mexico with emphasis on education. *World Development*, 29(3): 561-572.