Income Growth and Mobility of Rural Households in Kenya: Role of Education and Historical Patterns in Poverty Reduction

Mary K. Mathenge & David L. Tschirley[*]

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Portland, OR, July 29-August 1, 2007

[*] Mary Mathenge (mathenge@msu.edu) and David Tschirley (tschirle@msu.edu) are Ph.D candidate and Professor respectively at the Dept of Agricultural Economics, Michigan State University, East Lansing Michigan, 48824.

Abstract

This paper explores the key factors that cause changes in the economic wellbeing of rural households in Kenya. We specifically determine the relationship between the initial economic position of households and education on income growth and mobility. We use a three-period panel dataset to estimate a dynamic panel data model of full income. Results show strong evidence of (low) income persistence for the poor and those in the low agricultural potential areas that lack higher education. The low income persistence for the poor and uneducated may be evidence of cumulative dis-advantage and possible existence of poverty traps. As expected, higher education seems to eliminate the low income persistence for these vulnerable groups and allow convergence of incomes towards their average. This indicates the potential role of education in not only breaking the cycle of poverty for those trapped in it, but also its ability to allow increased recovery from income shocks.

Keywords: Income growth; income persistence; convergence; education; Kenya

Poverty remains among the major challenges facing the world today. Although world poverty has generally fallen in the last 40 years, progress in Sub-Saharan Africa (SSA) has been slow and uneven. The number of people reported as living on less than a dollar a day (the internationally agreed definition of absolute poverty) has doubled over the past 20 years (World Bank, 2004a). This has left many questions as to the best strategies that should be used to deal with the problem, spurring numerous research interests and massive donor funds to be used. The fight against poverty however remains an elusive goal.

In Kenya, high incidence and depths of poverty coupled with stagnating or declining income growth are the two major challenges facing the country today. Close to 46 percent of the total population and nearly half of the rural population live below the poverty line (Republic of Kenya, 2007) with meager incomes incapable of sustaining any meaningful livelihood. Even worse, poverty rates in some regions have been on the increase since the second half of the 1990s. Why? What are the alternative pathways out of poverty, and where should the limited resources be allocated? The answer to this question lies in understanding the causes of poverty and how policy can be used to break the cycle.

Household incomes are an important measure of a household's economic well-being and are key to any poverty reduction strategy. A household's total income, together with other known welfare measures such as consumption, expenditure, assets, and nutritional status, are some of the most commonly used metrics in analyzing poverty and economic mobility (Baulch and Hoddinott, 2000).

Though income dynamics studies are common in developed world (Jenkins, 2000), very few have been carried out in developing countries, especially rural Africa where poverty is immense. The main limitation has been paucity of relevant panel data in this region. The few existing studies have been carried out in Ivory Coast (using the Cote d'Ivoire Living Standards Survey (CILSS)) and South Africa (based on the Kwazulu-Natal Income Dynamics Study (KIDS) panel data set). Specific studies in Africa that have used income include Gunning et al. (2000), Fields et al (2003a, 2003b) and Woolard and Klasen (2005). Baulch and Hoddinott (2000) provide a detailed review of other earlier studies in developing countries.

Empirical studies on poverty dynamics in Kenya have mainly focused on analyzing poverty transitions and/or determinants of poverty status, thus utilizing discrete measures of poverty. While an understanding of factors associated with movements into and out of poverty has great value in the design of safety net policies, in the long run, design of policies that promote equitable growth requires information of how and why households increase their well-being relative to others (Baulch and Hoddinott, 2000). It is also important to note that use of transitions provide only relative rankings and potentially ignore the life-cycle phenomenon. A more recent study by Burke et al. (2007) explores movement into and out of poverty using an asset based measure with special reference to the importance of livestock. The current study adds to the existing literature by carrying out an in depth analysis of economic mobility using income as a continuous variable, thus utilizing much of the available information, some of which is usually masked when using discrete poverty measures.

Various theories offer alternative predictions regarding the evolution of the economic wellbeing of households over time. The theory of cumulative advantage posits that the economic wellbeing of the initially better-off households becomes better while that of the initially disadvantaged worsens (Fields et al., 2003a). This is based on the premise that wealthier households are endowed with both physical and human capital assets, whose further investment (presumably in high return activities) results in higher incomes. However, at the lower end of the income distribution, cumulative disadvantage seems to be at work whereby households without a 'minimum level of human, physical and social assets are confined to a life in poverty' (Fields et al., 2003a pp 68). This is related to the notion of poverty traps whereby some households suffer a successive run of negative shocks that forces them into destitution.

An alternative theory is based on the notion of convergence of incomes towards the average, thus enabling initially disadvantaged households to become better off and vice versa. The convergence argument is based on the assumption that income shocks do not persist and are not correlated over time. While the theory of cumulative advantage implies targeting those who are economically disadvantaged so as to set them up on a positive growth process, it may also be true that some important income shocks especially for rural households who mainly rely on the farm may be independent, thus permitting quick recovery. This kind of information would especially be insightful when disaggregated at regional level in explaining why households in some regions remain disadvantaged over time. Suffice it to say that both these dynamic processes do potentially take place: cumulative advantage as a result of using these advantages[1] to

[1] access to various endowments, market and financial institutions

build on future incomes (while the disadvantaged ones are unable to do so) and convergence towards the mean given large and uncorrelated transitory shocks.

This paper explores the key factors that cause changes in the economic wellbeing of rural households in Kenya. We specifically determine the relationship between the initial economic position of households and education on income growth and mobility. Evidence of *income persistence* for the poor would be consistent with the notion of cumulative (dis)advantage and possible existence of poverty traps. Further, we explore whether and how policies in education could be used to break income persistence for the very poor. Given the wide variation in poverty across and within regions, differences in the above impacts across income groups and regions of the country are explored.

The justification in using household income as opposed to consumption has been the perceived inability of poor households to smooth their consumption over time especially within households facing liquidity constraints or limited asset base. Using evidence from Ivory Coast and Thailand, Deaton (1997) finds consumption profile to be closely linked to the income profile and argues that the 'life-cycle model overstates the degree to which consumption is in fact detached from income over the life cycle'. This evidence implies failure of consumption smoothing in some cases, thus confirming relevance in the studies of income dynamics within the broader context of poverty reduction. An understanding of household income dynamics is fundamental to understanding the dynamics of household economic wellbeing (Baulch and Hoddinott, 2000). According to Fields et al. (2003b), the rise and fall of income and consumption experienced by households are the most direct indicators of who benefits from economic development.

Our choice of income as opposed to a discrete poverty measure is based on the advantages that come with analyzing income as a continuous variable as opposed to categorizing using an arbitrary poverty line, thus losing out on a lot of information (Jenkins, 2000, Ravallion 1996).

This study contributes to the existing body of literature in the following ways: First, it adds to the limited empirical studies on income dynamics in SSA where poverty is immense. Secondly, the use of a three period panel data enables us to control for historical patterns and still benefit from use of panel data methods unlike similar studies that have relied on two period panels (Grootaert et al., 1997, Fields et al., 2003a, 2003b, Woolard and Klasen, 2005, Glewwe and Hall, 1998). The ability to account for both historical patterns as well as unobserved factors may provide more reliable estimates of individual effects. Of major importance here is the ability to determine the economic mobility of households especially the initially poor over the study period in comparison with their wealthier counterparts. Third, unlike any of the other studies mentioned above, we disaggregate the results by poverty status and agricultural potential, thus allowing the pattern of income growth for each to unfold. This is indeed important for policy design and targeting. Fourth, we look at how policies in education can be used to break income persistence especially for those trapped in a cycle of poverty. Finally, we deal with the potential endogeneity of the lagged income difference in a dynamic panel data setting, a problem either commonly assumed or not dealt with exhaustively in earlier studies.

Methods

Conceptual Approach

The analytical framework used in this study is adopted from an agricultural household model where we assume that households are maximizing utility from consumption of goods and leisure subject to a cash income constraint[2] given by:

$$Y=\pi_f + w_o L_o + N \qquad (1)$$

where Y is cash income, π_f is net farm profits, w_oL_o is net off-farm earnings and N represents other non-labor income. The maximized profits from the farm are however a function of farm wages (w_f), input prices (P_Z), output prices (P_Q), human capital variables (H) and other household and locational characteristics of the household (G):

$$\pi^*_f = f(w_f, P_Z, P_Q, H, G) \qquad (2)$$

Off-farm wages w_o depend on the human capital assets of the household (mainly education and experience) and nature of the rural economy (E) such that:

$$w_o = f(H, E) \text{ and } H = f(\text{education, experience}) \qquad (3)$$

Combining (1), (2) and (3) above, and accounting for the value of total household time (T), we can write the full income production function of the household Y^*as:

$$Y^* = f(w_f, P_Q, P_Z, H, E, G, N) \qquad (4)$$

which indicates that the full income of a household is depended upon performance at the farm, endowments and characteristics of the household and the state of the local economy.

---

[2] Among other constraints e.g. the production technology and time constraints.

Empirical Model

A full income production function based on equation (4) is estimated to determine the key factors that cause changes in the economic well being of rural households in Kenya. In this study, we use the reduced form version of equation (4) comprising of all the exogenous variables in the system and other relevant variables. The underlying assumption of this model is that real household income is a function of the household's endowments or stock of assets (X_{it}) and the economic environment (Z_{it}) in which these assets become productive and an error term (ε):

$$Y_{it} = f\,(X_{it}, Z_{it}, \varepsilon_{it}) \qquad (5)$$

The empirical specification of the income model, accounting for historical patterns is given by:

$$\mathbf{INC_{it} = \alpha_0 + INC_{it-1}\alpha_1 + X_{it}\delta + Z_{it}\lambda + \varepsilon_{it} \qquad i = 1,\ldots,n \quad t=1,\ldots,T \qquad (6)}$$

where: INC is the real value of income. Included in X are variables related to the household's endowments of physical, social and human capital, while the Z's include locational and other socio-economic characteristics of the household.

The inclusion of a lagged dependent variable helps to account for historical patterns and may also serve as control for some omitted variables. While an indication of the pattern of income growth is undeniably relevant, it would also be of policy importance to assess how education affects these income growth rates and persistence. This implies that the coefficient of the lagged income variable may vary across households with different educational levels. We therefore add an interaction term for the lagged income variable and some measures of human capital to determine how income persistence differs by education. The education of the head of household is used given that in most cases, the heads are responsible for making decisions for the entire household regarding use of the available physical and human assets.

Accounting for income persistence, delineating the education variable and including the respective interaction term, model (6) above becomes:

$$\mathbf{INC_{it} = \alpha_0 + INC_{it-1}\alpha_1 + Ed_{it}\,\beta_1 + INC_{it-1}{*}Ed_{it}\,\alpha_2 + X_{it}\delta + Z_{it}\lambda + \alpha_i + \mu_{it}} \quad \textbf{(7)}$$

where: Ed is the education variable, and X and Z are as earlier defined. To control for any omitted unobserved factors that may potentially correlate with the above variables or other included explanatory variables, we have explicitly accounted for them in the above model: α_i represents the time invariant unobservable effects and μ_{it} is a purely random component.

Estimation

The **dynamic panel data model** (7) has implications on the estimation methods often used. First, the unobserved effects are most likely correlated with the lagged dependent variable (LDV), thus rendering OLS inconsistent. Secondly, even though we could get rid of the unobserved effects through differencing or fixed effects, it is logical that future values of the LDV are potentially correlated with the idiosyncratic error term (Cov (INC_{is}, μ_{it}) $\neq 0$, for s>t) implying that the within estimation is also inconsistent. This problem also bedevils Generalized Least Squares (GLS) since it requires strict exogeneity of the regressors. The most viable solution to this problem has been to take first differences to eliminate the unobserved effects and then instrument for the lagged difference variable (Ahn and Schmidt, 1995; Wooldridge, 2002).

$$\mathbf{\Delta INC_{it} = \Delta INC_{it-1}\,\alpha_1 + \Delta Ed_{it}\,\beta_1 + \Delta(INC_{it-1}{*}Ed_{it}\,)\,\alpha_2 + \Delta X_{it}\delta + \Delta Z_{it}\lambda + \Delta\,\mu_{it}} \qquad \mathbf{(8)}$$

Taking first differences of the data does clearly help to eliminate the time-invariant unobservable factors, but this comes at a cost of reducing variation in the regressors. This problem is however minimized in this case since our panel has more than a year's gap, resulting in somewhat longer differences (3-4 years between periods). The first difference approach also helps us to explain changes in the economic wellbeing of households.

In this study, and to ensure consistency of the estimated parameters, equation (8) above is thus estimated using First Difference Two-stage Least Squares (FD-2SLS) so as to account for the endogeneity of the lagged income difference (LID) in the model. Following Anderson and Hsiao (1982) and Wooldridge (2002), we use previous lags of income level (INC_{it-2}) as instruments for the (LID) variable. Since we effectively can only use one such instrument from our data, we also use lagged mean rainfall deviation (R_{it-2}) as another potential instrument. The rainfall variable provides over-identifying restrictions to allow testing the validity of the instrument set. To account for the potential lack of strict exogeneity of the interaction term with education, we use the respective lagged interaction term ($INC_{it-2}*Ed_{it-2}$) as an instrument. It is however important to note that the use of previous income levels as potential instruments is only legitimate when there is no serial correlation in the errors (Arellano and Bond, 1991; Wooldridge, 2002). This is nevertheless not applicable in this case as we effectively only have a cross-section of differenced data after accounting for historical patterns and unobserved heterogeneity.

The use of 2SLS is appropriate since the effects of multicollinearity (some level of which cannot be denied in these models) are less serious in 2SLS than in IV estimation. The estimation and inference are further made robust to heteroscedasticity and serial correlation.

Data and Sample Area

The data used in this study come from the Tegemeo Agricultural Monitoring and Policy Analysis Project (TAMPA) data set which consists of a three-period panel collected over a period of seven years. The household surveys were carried out in 1996/97, 1999/00 and 2003/04 cropping seasons. The specific sample used in this study consists of a total of 3972 households (1324 for each year). The panel contains data on economic, demographic and other social characteristics of the households. Table 1 presents the description of the variables used in this study including their means and standard deviations.

Table 1.Summary Statistics of Variables Used in the Models

Variable Description	Variable Name	Unit	Mean	Std Deviation
Household Income	Income	Ksh(‘000)	165.29	192.98
Distance to tarmac road	dtmrd	km	7.72	7.90
Mean dist to elect and phone	mdist	Km	4.52	4.86
Age of head	agehd	years	54.98	13.63
Education of head	educhd	years	6.36	4.89
Gender (male headedness)	malehd	1/0	.84	.37
% adults working off farm	padofe	years	.36	.27
Land cultivated	ldcult	acres	13.70	16.44
Number of livestock owned	nlvstok	count	18.21	42.49
Number of adult males	nmalad	count	2.49	1.58
Number of adult females	nfemad	count	2.41	1.41
Months head at home	month	months	10.45	3.53
Group Membership	gpmem	1/0	.77	.42
Completed primary school	primo	1/0	.53	.49
Completed high school	seco	1/0	.21	.41
Some college education	cole	1/0	.17	.38

No. of Observations=2648

Empirical Findings and Discussions

Table 2a presents the parameter estimates of the first difference model given by equation (8) in the methods section. For comparison purposes, five different models are estimated representing different treatments of the lagged dependent variable. Only model (5) involves IV estimation. Model (1) represents the most basic type of estimation possible with a two-year panel though it ignores the role of historical patterns in income determination.

Model (2) shows the results when we include lagged income as a level variable in a differenced model. Three things are particularly noteworthy here. One is that the inclusion of the lagged income variable is not in built in the original formulation of the model and thus does not in itself go through the differencing procedure. Secondly, there is a significant increase in the coefficient of determination from the first model, thus indicating the importance of accounting for income persistence in an income model. Thirdly, the results show evidence of convergence of income towards the mean, a result that is consistent with earlier studies in Africa that followed a similar econometric approach, namely Grootaert et al. (1997) and Fields et al. (2003a). The reliability of these results may however be in question as the estimation fails to account for the endogeneity of the lagged income variable. Fields et al. (2003b) and Woolard and Klasen (2005) use a similar procedure but also instrument for the endogenous lagged income variable. Fields et al. (2003b) find mixed results with the IV method and alludes to the sensitivity of results to the treatment of the income variable. On the other hand, Woolard and Klasen (2005) indicate that convergence is maintained with the IV estimation but the coefficient is greatly reduced for the rural areas.

Table 2a.Determinants of Income Growth

Model	1	2	3	4	5
Dependent Variable = Δ Income					
Δ dtmrd	0.51	0.03	0.01	-0.17	-1.29
	(0.27)	(0.02)	(0.01)	(0.10)	(0.55)
Δ mdist	9.23**	9.55***	9.16***	8.88***	6.88
	(2.54)	(3.49)	(2.90)	(2.87)	(1.57)
Δ agehd	2.62	3.63*	4.23*	3.70*	-0.69
	(0.97)	(1.95)	(1.84)	(1.80)	(0.15)
Δ agehd2	-0.02	-0.03*	-0.03*	-0.03*	-0.01
	(0.84)	(1.79)	(1.65)	(1.65)	(0.19)
Δ educhd	2.42	4.86***	1.61	6.31*	40.48**
	(1.22)	(3.29)	(1.07)	(1.87)	(2.23)
Δ malehd	-5.69	30.80***	8.33	15.06	58.37
	(0.31)	(2.66)	(0.39)	(0.84)	(1.29)
Δ padofe	102.71***	87.05***	78.52***	79.67***	96.78***
	(5.91)	(6.61)	(5.10)	(5.24)	(4.29)
Δ ldcult	1.32	-0.17	1.49	1.21	-0.89
	(0.69)	(0.16)	(1.06)	(0.95)	(0.60)
Δ nlvstok	1.26***	0.73***	1.09***	1.06***	0.94***
	(3.61)	(3.07)	(4.12)	(4.28)	(3.56)
Δ nmalad	9.01	10.88**	9.83*	11.01**	19.21*
	(1.37)	(2.42)	(1.90)	(2.26)	(1.95)
Δ nfemad	12.03**	10.44***	12.70***	12.30***	9.19
	(2.38)	(3.08)	(2.96)	(2.88)	(1.51)
Δ month	-2.99	-2.17	-2.17	-2.34	-3.83
	(1.54)	(1.28)	(1.49)	(1.61)	(1.55)
Δ grpmem	11.83	2.77	11.03	10.27	5.14
	(0.98)	(0.34)	(1.10)	(1.08)	(0.37)
year dummy	-29.90***	109.54***	-12.88*	-8.74	14.87
	(3.54)	(9.85)	(1.69)	(1.28)	(0.43)
Lagged Income		-0.67***			
		(14.24)			
LID			-0.57***	-0.47***	0.49*
			(6.91)	(3.83)	(1.82)
LID* educhd				-0.02	-0.15*
				(1.64)	(1.96)
Observations	1324	1324	1324	1324	1324
R-squared	0.09	0.51	0.42	0.43	

Robust t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Source: Authors calculation

Model (3) also accounts for historical patterns by including a lagged income variable but this time in the initial formulation of the model such that it is differenced with the other variables. This is only possible with at least a three year panel. The coefficient of the lagged income variable however remains negative and significant whether the level or differenced form is used. The rest of the coefficients also remain stable across the two models except for the education variable. Models (4) and (5) both interact the LID with education; in the latter, we instrument for the LID and its interaction with education as discussed in the methods section. As shown from the results, the parameters from the two models show fairly similar patterns but with a few exceptions.

A major difference though and one that is of main interest in this paper is the coefficient of the LID. Without accounting for the endogeneity of the LID, the coefficient is negative and significant and does not vary significantly with education of the head. This implies that households are recovering from income shocks and worse off households become better in future periods and vice versa. This result is consistent with findings from earlier studies which find overwhelming support for the convergence of household incomes towards the mean. This scenario may exist when a larger proportion of the full income of a household consist of transitory gains/losses which are less persistent allowing quick recovery of shocks or de-cumulation of gains. This result may not be surprising given that without using instrumental variables methods, the lagged income variable consists of both the permanent and the transitory components.

The above results however change when we account for the weak exogeneity of the LID. The coefficient of the LID turns positive and significant and seems to vary significantly with education of head. It is however noteworthy that after 3.3 years of education, the combined effect turns negative but remains insignificant even at mean education level (6.36 years) as shown in Table 2b. This implies possible recovery of income shocks for those with higher levels of education, but a cycle of low income persistence for those with less education especially the 30% with less than 4 years of education. For these households, results point to the existence of poverty traps and cumulative disadvantage which is thankfully broken at higher levels of education.

Table 2b. Combined Effects of LID and Education at Mean Levels

Table	Model	Variable	Combined effects	F- ststistics	p-value
	4	LID	-0.59	18.28	.0000
3		Education	3.63	3.48	.0623
	5	LID	-0.46	1.64	.2002
		Education	20.38	4.96	.0261

Source: Authors calculation

The difference in the results given by Models (4) and (5) especially in regard to income persistence justifies the use of appropriate estimation methods to enable the drawing of relevant conclusions. These differences may be explained by looking at the procedure of IV estimation used. The method of 2SLS applied to Model (5) implies that the final estimation uses the predicted portion of the suspect endogenous variable which can be viewed as the permanent income component of full income while the LID in model (4) consists of both the permanent and transitory components. In this paper, we take Model (5) as representing the most reliable parameter estimates based on the appropriateness of the estimation procedure that not only accounts for the endogeneity of the LID, but its interaction with education as well. The results of the over identification test for the validity of the instrument set and the first stage regressions are given in Table A1. The instruments are both individually and jointly significant in the first stage regressions of the two endogenous variables. There is also strong evidence of failure to reject the exogeneity of the instrument set.

Though the result of how education can break the cycle of low income persistence for the poor is interesting by itself, deriving relevant policy recommendations require further analysis as to the level of education that can achieve the required results. Table A2 in the appendices presents regression results of model (5) with four different specifications for the education variable: continuous (as in the original model) and three binary variables indicating completion of primary school, completion of secondary school, and at least some college/professional training. The results show that attainment of at least a primary school education made no significant contribution to household income and also failed to significantly reduce income persistence or enhance convergence. This result is contrary to households whose head had attained either a secondary or some college education. In these cases, the respective education variable was highly significant and also caused great reductions in the positive coefficient for the LID. This is an indication of the role of post-primary education in feeding income growth and also in breaking (low) income persistence for the poor.

Based on the above findings, Table 3 presents the regression results of the models with secondary education disaggregated by poverty status. These results show that for households that are below the poverty line, those whose heads have less than secondary educations are locked up in a cycle of low income persistence and cumulative disadvantage. This is in contrast to their counterparts with at least a secondary education who show strong evidence of convergence of incomes towards the average. However, for those households that are above the poverty line, there is no conclusive evidence of the pattern of income growth nor the role of education, implying that such households may be less susceptible to long-term effects of income shocks in either direction[3]. On its own, attainments of secondary education does not have a significant influence on income growth for those who are non-poor, but has a very large and significant positive effect for the poor.

.

[3] It is possible that incomes for such non-poor households is driven by their (higher) asset holdings.

Table 3.Models Disaggregated by Poverty Status

	(1) General	(2) Below Poverty	(3) Above Poverty
LID	0.13	0.31***	0.02
	(0.56)	(3.17)	(0.05)
LID*Secondary education	-1.40*	-3.04***	-1.24
	(1.88)	(4.42)	(0.92)
Δ dtmrd	0.31	4.98*	-3.54
	(0.15)	(1.75)	(0.90)
Δ mdist	7.26*	-1.35	9.10
	(1.81)	(0.31)	(0.74)
Δ agehd	1.84	-1.15	13.79*
	(0.71)	(0.22)	(1.70)
Δ agehd2	-0.02	-0.01	-0.12*
	(1.03)	(0.18)	(1.66)
Δ seco	348.73**	534.02***	503.56
	(2.13)	(4.36)	(1.28)
Δ malehd	40.48	86.66**	56.89
	(1.31)	(2.30)	(0.61)
Δ padofe	108.26***	86.43***	134.46
	(5.24)	(3.67)	(1.33)
Δ ldcult	-0.13	-2.02	-1.91
	(0.16)	(1.49)	(1.18)
Δ nlvstok	1.20***	0.75**	2.20**
	(4.76)	(2.26)	(2.12)
Δ nmalad	14.54**	1.61	49.76**
	(2.06)	(0.24)	(2.55)
Δ nfemad	10.08*	11.70*	18.97
	(1.85)	(1.83)	(0.89)
Δ month	-2.68	-3.70*	1.32
	(1.33)	(1.74)	(0.30)
Δ grpmem	5.83	3.84	14.26
	(0.51)	(0.20)	(0.42)
year dummy	-6.67	-22.70*	53.78
	(0.39)	(1.89)	(0.75)
Observations	1324	935	389

Robust t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Source: Authors calculation

As for the other determinants of income growth, it seems that having a male head of household positively influences income growth for those below poverty, but has no significant effect for the non-poor. This could be explained by the reduction of discriminatory practices based on gender for the non-poor female heads as compared to their poor counterparts. The proportion of adults working off the farm, which could be an indication of the availability of employment opportunities in the area, shows the same pattern: a clear positive effect for the poor but not as clearly for the non-poor. This is not surprising given that the poor are more likely to benefit from other income earning activities and is further an indication of the role of access to the off-farm labor market in breaking the cycle of poverty, a subject well advanced by Giles (2006).

A similar pattern is observed for the number of months the head stayed home. The higher the number of months the head was at home, the lower the impact on income growth, which again implies that working away from the farm for the head, resulted in positive income gains for the poor. This may indicate the role of migratory labor in rural income growth. The number of livestock owned had positive income gains for both the poor and non-poor but the amount of land cultivated had no significant influence on either. This latter result is surprising given that we observe a general increase in land cultivated with income, but seems consistent with findings from Burke et al. (2007). It is however possible (as may be the case with a few other variables) that the low variability of these variables across the years, which is only made worse by the differencing operation, may cause an insignificant result in an otherwise significant variable.

Considering the pattern of income growth by agricultural potential (Table 4), we observe strong evidence of income persistence for those households in the lower agricultural potential areas and whose heads had no college training. This persistence is however broken for households with post secondary training (as indicated by the resulting negative coefficient of -0.67), thus showing evidence of convergence towards the average for such households. This observation is plausible given the low returns to agriculture in the low potential areas and the fact that reduction of income persistence in such areas may only be realized through access to the off-farm labor market, whose entry may require more education and training beyond what a secondary school education may offer. As expected, households in the high potential areas seem to recover from shocks with or without college training. However, those with college training tend to recover faster (coefficient of -1.14) from such income shocks than do their counterparts without this training.

Table 4.Models Disaggregated by Agricultural Potential

	(1) General	(2) Low Potential	(3) High Potential
LID	0.05	0.18***	-0.03
	(0.24)	(2.91)	(0.10)
LID*college education	-1.18*	-0.85**	-1.11*
	(1.92)	(2.29)	(1.81)
Δ dtmrd	0.03	-0.31	0.76
	(0.02)	(0.13)	(0.29)
Δ mdist	8.69**	0.20	11.46***
	(2.44)	(0.05)	(2.64)
Δ agehd	1.13	0.79	3.51
	(0.46)	(0.25)	(1.01)
Δ agehd2	-0.01	-0.00	-0.03
	(0.69)	(0.18)	(1.29)
Δ cole	320.24**	191.45**	330.17**
	(2.30)	(2.34)	(2.23)
Δ malehd	29.40	18.83	48.08
	(1.14)	(0.77)	(1.51)
Δ padofe	106.20***	121.88***	91.16***
	(5.52)	(4.57)	(3.33)
Δ ldcult	0.35	-0.27	0.98
	(0.39)	(0.50)	(0.49)
Δ nlvstok	0.77*	0.92	0.76
	(1.71)	(1.41)	(1.57)
Δ nmalad	12.29*	9.59	11.45
	(1.92)	(1.14)	(1.36)
Δ nfemad	8.98*	15.47**	2.89
	(1.70)	(2.21)	(0.43)
Δ month	-3.08*	-1.29	-4.22*
	(1.72)	(0.57)	(1.86)
Δ grpmem	1.73	10.40	-5.51
	(0.17)	(0.64)	(0.38)
year dummy	-7.04	16.46	-16.98
	(0.42)	(1.46)	(0.66)
Observations	1324	430	894

Robust t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Source: Authors calculation

Summary and Conclusions

The results from the study suggest differences in the role of historical patterns on rural income growth and mobility and the potential of education in breaking the cycle of income persistence across different agricultural potential areas and poverty status. Overall though, rural households in Kenya show weak evidence (at 10% level) of income persistence especially at lower levels of education. At higher levels of education, there is evidence of convergence of incomes towards the mean though this remains insignificant even at mean levels of education of up to a primary education (6.36 years). As discussed earlier, this result somehow deviates from earlier findings from Africa possibly due to the differences in the econometric procedures employed. Disaggregation of these results by poverty status and agricultural potential does however provide some answers.

Households below the poverty line and without a secondary education show strong evidence of income persistence which is clearly broken for households whose head had at least a secondary education. There is no clear pattern of income growth for non-poor households with or without secondary education. The existence of income persistence for the very poor and uneducated is consistent with the theory of cumulative advantage and possible existence of poverty traps. This does imply the need for targeting those who are economically disadvantaged so as to set them up on a positive growth process.

A similar pattern emerges for households in the low potential areas where evidence of income persistence is observed. As expected, a much higher education in form of college training is required to break this cycle of low income persistence given high entry barriers into viable income earning activities in the off-farm sector as a substitute to the low returns from agriculture. Higher education in this case enables quick recovery from income shocks for those in the high agricultural potential areas.

Given the results of this study, the need for a comprehensive education policy cannot be overemphasized. While primary school education is important, it is not sufficient to impact positively on income growth and neither is it adequate in breaking the cycle of low income persistence for those trapped in poverty. Investments and programs in education to encourage enrollment and completion of secondary school education are therefore going to be key for future poverty reduction strategies.

The analysis of the factors that cause income changes over time is important in trying to understand the causes of poverty and hence devising appropriate pro-poor policies. While the international environment has powerful impacts on poor countries and their ability to reduce poverty, the lives of the poor are mostly affected by actions at the country and local level (Global poverty report, 2000). There is thus need for governments to design policies which are specific to their circumstances and hence the need for such localized studies cannot be overemphasized

Appendices

Table A1. Results of the First Stage Regression and Over-id test

Dependent Variable	(1) LID	(2) LID*Education
Δ dtmrd	-0.62	-13.17
	(0.43)	(1.06)
Δ mdist	-1.85	-31.26
	(0.55)	(0.90)
Δ agehd	-0.75	-37.23
	(0.34)	(0.93)
Δ agehd2	0.01	0.19
	(0.30)	(0.73)
Δ educhd	-6.63***	189.13***
	(2.63)	(7.31)
Δ malehd	29.79*	454.66*
	(1.66)	(1.86)
Δ padofe	25.90*	-102.66
	(1.65)	(0.84)
Δ ldcult	-0.82	-15.00
	(0.71)	(1.29)
Δ nlvstok	-0.29	-1.56
	(1.16)	(0.63)
Δ nmalad	2.42	75.55*
	(0.45)	(1.66)
Δ nfemad	0.43	-10.35
	(0.09)	(0.23)
Δ month	0.75	-5.67
	(0.50)	(0.38)
Δ grpmem	5.40	44.06
	(0.58)	(0.39)
year dummy	130.85***	667.05***
	(10.58)	(5.03)
Instruments		
Lagged Income level	-0.87***	-3.96***
	(12.41)	(3.70)
Lagged income*education	.042***	0.33***
	(5.34)	(3.04)
Lag mean rainfall deviation	-2.57***	-20.07***
	(3.14)	(2.97)
Joint sig of Instrument set		
F-statistics	51.90	7.11
p-value	0.0000	0.0001
Over-id Test		
Chi-square statistic	0.184	
p-value	0.6676	

Source: Authors calculation

TableA2. Comparison with different Education Levels

Models	1 Continuous Education	2 With primary	3 With secondary	4 With College education
LID	0.49*	-0.10	0.13	0.05
	(1.82)	(0.25)	(0.56)	(0.24)
LID*education	-0.15*	0.25	-1.40*	-1.18*
	(1.96)	(0.21)	(1.88)	(1.92)
Δ dtmrd	-1.29	0.64	0.31	0.03
	(0.55)	(0.31)	(0.15)	(0.02)
Δ mdist	6.88	10.19	7.26*	8.69**
	(1.57)	(1.40)	(1.81)	(2.44)
Δ agehd	-0.69	2.18	1.84	1.13
	(0.15)	(0.66)	(0.71)	(0.46)
Δ agehd2	-0.01	-0.01	-0.02	-0.01
	(0.19)	(0.41)	(1.03)	(0.69)
Δ educhd	40.48**			
	(2.23)			
Δ primo		-50.13		
		(0.23)		
Δ seco			348.73**	
			(2.13)	
Δ cole				320.24**
				(2.30)
Δ malehd	58.37	-6.88	40.48	29.40
	(1.29)	(0.15)	(1.31)	(1.14)
Δ padofe	96.78***	102.63***	108.26***	106.20***
	(4.29)	(4.92)	(5.24)	(5.52)
Δ ldcult	-0.89	1.54	-0.13	0.35
	(0.60)	(0.64)	(0.16)	(0.39)
Δ nlvstok	0.94***	1.29***	1.20***	0.77*
	(3.56)	(2.84)	(4.76)	(1.71)
Δ nmalad	19.21*	7.83	14.54**	12.29*
	(1.95)	(0.91)	(2.06)	(1.92)
Δ nfemad	9.19	11.96**	10.08*	8.98*
	(1.51)	(2.26)	(1.85)	(1.70)
Δ month	-3.83	-3.02	-2.68	-3.08*
	(1.55)	(1.61)	(1.33)	(1.72)
Δ grpmem	5.14	10.04	5.83	1.73
	(0.37)	(0.65)	(0.51)	(0.17)
year dummy	14.87	-32.04	-6.67	-7.04
	(0.43)	(1.25)	(0.39)	(0.42)
Observations	1324	1324	1324	1324

Robust t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Source: Authors calculation

References

Ahn, C.S., and P. Schmidt. 1995. “Efficient Estimation of Models for Dynamic Panel Data.” *Journal of Econometrics* 68:5-27.

Anderson, T.W. and C. Hsiao. 1982. “Formulation and Estimation of Dynamic Models Using Panel Data.” *Journal of Econometrics* 18:47-82

Arellano, M. and S. Bond. 1991. “Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations.” *The Review of Economic Studies* 58 (2): 277-297.

Baulch, B. and J. Hoddinott. 2000. “Economic mobility and poverty dynamics in developing countries.” *Journal of Development Studies* 36 (6):1-24.

Burke, W. J., T. S. Jayne, H. A. Freeman and P. Kristjanson. 2007. “Factors Associated with Farm Household’s Movement into and out of Poverty in Kenya: The Rising Importance of Livestock.” MSU International Development Working Paper No. 90.

Deaton, A. 1997. *The Analysis of Household Surveys: A Microeconomic Approach to Development policy*. The John Hopkins University Press: Baltimore.

Fields, G.S., P.L. Cichello, M. Menendez and D. Newhouse. 2003a. “For richer or for Poor? Evidence from Indonesia, South Africa, Spain and Venezuela.” *Journal of Economic Inequality* 1:67-99.

Fields, G. S., P. L. Cichello, S. Freije, M. Menendez and D. Newhouse. 2003b. “Household Income Dynamics: A Four-Country Story.” *Journal of Development Studies* 40 (2): 30-54.

Giles, J. 2006. "Is Life more Risky in the Open? Household Risk-Coping and the Opening of China's Labor Markets." *Journal of Development Economics* 81 (1): 25-60.

Glewwe, P. and G. Hall. 1998. "Are Some Groups More Vulnerable to Macro-Economic Shocks than Others? Hypothesis Tests based on Panel Data from Peru." *Journal of Development Economics* 56: 181-206.

Grootaert, C., R. Kanbur and G. T. Oh. 1997. "The Dynamics of Welfare Gains and Losses: An African Case Study." *The Journal of Development Studies* 33 (5): 635

Gunning, J. W., J. Hoddinott, B. Kinsey and T. Owens. 2000. "Revisiting Forever Gained: Income Dynamics in the Resettlement Areas of Zimbabwe, 1983-96." *The Journal of Development Studies* 36 (6): 131-154.

Jenkins, S. P. 2000. "Modeling Household Income Dynamics." *Journal of Population Economics* 13 (4): 529 – 567.

Republic of Kenya. 2007. Economic Survey. Kenya Bureau of Statistics, Ministry of Planning and National Development.

Woolard, I. and S. Klasen. 2005. "Determinants of Income Mobility and Household Poverty Dynamics in South Africa." *The Journal of Development Studies* 41 (5):865-897.

Wooldridge, J. 2002. *Econometric Analysis of Cross-Section and Panel Data*. The MIT Press: Cambridge, Massachssetts, London, England.

World Bank. 2004a. "Global Poverty Monitoring." Available at http://www.worldbank.org/research/povmonitor/index.htm