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

Mary K. Mathenge & David L. Tschirley

Version: March 2008

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 educational attainment and the initial economic position of households on their 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 without at least a secondary school education. Similarly, there is evidence of income persistence for those in the low potential areas and a (weak) convergence towards the average for those in the high potential areas. 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 reduce income persistence for the very poor and those in the low potential areas but also enhances convergence for those in the high potential areas. 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. Notably, there is no conclusive evidence of the pattern of income growth or the role of education for non-poor households, implying that such may be less susceptible to long-term effects of income shocks in either direction.

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

JEL Classifications: O12 O55 I21 I32 C23

---

* Paper presented at the CSAE 2008 Conference on 'Economic Development in Africa' at St. Catherine's College, Oxford University, UK: March 16-18 2008.
Correspondence Information: Department of Agricultural Economics, Michigan State University, East Lansing, Michigan 48823 USA
E-mail: mathenge@msu.edu

1. Introduction and Problem Overview

The first Millennium Development Goal (MDG) calls for reducing by half the proportion of people living in extreme poverty and hunger by 2015. Studies of household level poverty dynamics in developing countries offer the prospect of contributing to the attainment of this goal by identifying *how many* people are mired in poverty, *who* these people are, and *what might be done* to pull them out of poverty and set them on a growth path. Yet, reliably answering any of these questions poses substantial analytical challenges. The main one among these has been lack of relevant panel data especially in rural Africa[1], where poverty is immense.

Poverty dynamics studies in Africa have used different measures of economic well-being. Studies based on expenditure include Dercon and Krishnan (2000), Grootaert and Kanbur (1995), and Grootaert et al. (1997). Baulch and Hoddinott (2000) provide a detailed review of these and other studies in developing countries. Recent studies in Africa that have used income include Gunning et al. (2000), Carter and May (2001), Fields et al. (2003a, 2003b) and Woolard and Klasen (2005). All of these studies are based on short panels (two periods at most), thus limiting the extent to which relevant factors can be controlled for. In this paper, we take advantage of a unique, broadly representative three-period panel data set from rural Kenya to study the pattern of income growth and explore any existence of long term poverty.

In Kenya, studies on poverty dynamics have mainly focused on analyzing poverty transitions and/or determinants of poverty status using discrete income-based measures of poverty (Gamba and Mghenyi, 2004; Kirimi and Sindi, 2006; Muyanga et al., 2006).

[1] Most poverty dynamics studies in Africa use the Cote d'Ivoire Living Standards Survey (CILSS) for Ivory Coast and the Kwazulu-Natal Income Dynamics Survey (KIDS) data set for South Africa.

While knowledge of factors associated with movements into and out of poverty have 'great value in the design of safety net policies', an understanding of how and why households increase their well-being relative to others is important for the design of policies that promote equitable growth (Baulch and Hoddinott, 2000). Also, use of transitions may provide only relative rankings and potentially ignore the life-cycle phenomenon (Baulch and Hoddinott, 2000). A more recent study by Burke et al. (2007) explores movement into and out of poverty using an asset based measure. In some ways, the current study parallels Barrett et al. (2006) who uses non-parametric approaches to draw inferences regarding poverty traps. This study was however carried out in a much localized context both in scale and geographical scope, potentially limiting widespread application and generalization.

This paper explores the key factors that cause changes in the economic wellbeing of rural households in Kenya. We specifically determine the effect of educational attainment and the initial income of households on subsequent 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, the study explores whether education and educational policies could be used to break income persistence and/or enhance recovery from negative shocks, especially for the 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 also explored.

The justification for using household income as a measure of economic wellbeing has been the perceived inability of poor households to smooth their consumption over time, especially within households facing liquidity constraints (Morduch, 1995;

Townsend, 1995) or with a 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'. In addition, Fields et al. (2003b) indicate that data from India and China does not find consumption to be clearly superior to income as an indicator of long term economic wellbeing. This evidence implies failure of consumption smoothing in some cases, thus confirming the relevance of income dynamics studies within the broader context of poverty reduction research. 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. The 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 useful information (Jenkins, 2000; Ravallion, 1996).

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. Access to skilled jobs, credit facilities and markets also helps in the accumulation process. At the lower end of the income distribution, cumulative disadvantage may 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). This is related to the notion of poverty traps whereby some households suffer a succession 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. This argument parallels that of the Permanent Income Hypothesis (PIH) of Friedman which assumes that transitory shocks (both positive and negative) are serially uncorrelated, thus leading household incomes to regress to their expected level in the subsequent period. A more plausible hypothesis would allow for partial correlation between successive transitory shocks, resulting in gradual convergence of household incomes towards their mean level (Fields et al., 2003a).

While the theory of cumulative advantage implies targeting those who are economically disadvantaged, to set them on a positive growth path, it may also be true that some important income shocks are independent, especially for rural households who mainly rely on agriculture, thus permitting quick recovery. Information about the extent to which each process is taking place 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 access to various endowments and market- and financial institutions to increase future incomes (while disadvantaged households are

unable to do so) and convergence towards the mean given large and uncorrelated transitory shocks, especially from agriculture.

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 sub-Saharan Africa, where poverty is immense. Second, the use of a three period panel enables us to control for the household's initial economic position and still benefit from panel data methods, unlike similar studies that have relied on two period panels. The ability to account for both historical patterns and unobserved factors may provide more reliable estimates of individual effects. Of major importance here is to explore the main factors causing the economic mobility of households, especially the initially poor, over the study period in comparison with their wealthier counterparts. Third, unlike any of the studies mentioned earlier, we disaggregate the results by household poverty status and by the region's agricultural potential, thus allowing us to identify differential patterns of income growth. This is potentially 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 lagged income difference in a dynamic panel data setting, a problem either commonly assumed away or not dealt with exhaustively in earlier studies.

The rest of the paper is organized as follows. Section two discusses the conceptual approach, empirical model and estimation methods. The data used in the study and other specification issues are presented in section three. Section four presents the empirical results and discussion, and section 5 provides a summary of the findings and main conclusions.

2. Methods

2.1. 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:

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

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[3]. The maximized profits from the farm are a function of farm wages (w_f), input prices (P_Z), output prices (P_Q), human capital variables (H) and locational and other socio-economic characteristics of the household (G):

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

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:

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

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

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

which indicates that the full income of a household is dependent on performance at the farm level, 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.
[3] Refer to Singh et al. (1986) and Huffman (1991) for a detailed exposition of this model.

2.2. Empirical Model

An 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) consisting of the exogenous variables in the system and other relevant variables.

In linking equation (4) and the estimated model, we proxy the prices of farm inputs and outputs with distance and market access variables and other exogenous factors that affect the earnings from the farm such as weather. We specifically include distance to the major input suppliers (e.g., fertilizer) as an indication of access to farm-productivity enhancing technology. The amount of land cultivated and the number of livestock owned are used to give an indication of the intensity and orientation in farming. Access to extension service is also accounted for.

In considering the role of participation in the off-farm labor market in income growth, we take account of both the supply and demand factors that enable participation in that market. The ability of a household to engage in the off-farm labor market can be facilitated through enabling access, which further depends on the availability of a vibrant economy within the region and good transportation. This is based on the incentives facing the household in the form of public assets, e.g. good infrastructural development, provision of electricity and telecommunication services, public transport or the proximity of a shopping centre.

The measures of access to the off-farm labor market used in this study include access to electricity and telephone services, the proportion of adults working off the farm and the number of months (in a given year) the head of household was home, among

others. The proportion of adults working off the farm can be an indication of the availability of employment opportunities in the area, while the number of months the head was *not* at home could give an indication of the role of migratory labor in rural income growth. Village level distance variables could proxy for the level of urbanization and development in the area. We specifically pick electricity and public telephone as these facilities are only likely to be present in growing rural economies. In addition, the two are likely to have substantial impacts on off-farm employment and growth, such as in small manufacturing and service activities. Other variables such as the presence and condition of the tarmac road, and distance to the market or shopping centre are good measures of connection to the rest of the country/region and a vibrant economy, and also facilitate input delivery and output sale for agricultural activities.

The ability of a household to engage in profitable income earning activities is determined by its endowments of human and physical capital assets. Of major importance here are the human capital assets as proxied by education, experience and household composition. The education variables proxy for individual household's ability to gain access to the off-farm labor market. The education of head of household is important as he is the most likely to seek off-farm employment.

Some distance variables in the model may be endogenous. Probable in this category, are distance to electricity and phone service, since wealthier households may be in a position to bring these services closer to them. While possession of these facilities at the household level may be a function of ones' income, the specific distance variable in this case may be less endogenous as only about five percent of the sampled households had electricity at their homestead while less than one percent had telephone service,

implying that such amenities are beyond the means of most of these rural households. Other variables such as distance to tarmac road and shopping center are less likely to be endogenous, since households would have to move to those services, and such movement is relatively limited in rural Kenya[4].

Simple regression models of these distance variables on income support these views; income is not only insignificantly associated with distance to tarmac road and shopping centre, but also with distance to electricity and telephone services (see Table A1 in the appendices). As expected though, income did have a significant effect on whether a household had electricity or telephone at home. Nevertheless, to reduce any potential bias in using these distance variables, we use their respective village level means which can be considered relatively exogenous to the household. Note that distance to farm input suppliers may reflect response by input suppliers to the forces of demand which in turn could depend on general incomes in the area and is thus treated likewise.

The empirical specification of the income model, accounting for the initial income level is given by:

$$\mathbf{INC_{it} = \alpha_0 + INC_{it-1}\alpha_1 + X_{it}\delta + G_{it}\lambda + E_{it}\zeta + \varepsilon_{it}} \qquad \mathbf{i = 1,\ldots..,n \quad t=1, \ldots\ldots,T} \qquad \mathbf{(5)}$$

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, G represents locational

[4]The low attrition rate per survey (discussed later in the paper) may attest to this fact. In addition, land purchases/sales are minimal in these rural setups and most land is passed on through generations as inheritance. See Low (1986) for a review of the traditional attitudes and non-market benefits of land in much of rural Africa.

and other socio-economic characteristics of the household, while E include variables that proxy for the state of the local economy.

The inclusion of a lagged dependent variable helps to account for historical patterns and may also serve as control for some omitted variables (Wooldridge, 2003 pp 300). The coefficient on this variable indicates the extent to which income changes in one year persist in future years. While an indication of the pattern of income growth is undeniably relevant, it would be of additional policy importance to assess how education affects income persistence, especially for low income households. For the model to capture such an effect, the coefficient of lagged income must be allowed to vary across households with different educational levels. 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.

In addition, the inclusion of the lagged dependent variable enables us to allow for any potential feedback process in some of the included variables. Such a feedback process is plausible, given that income received during the previous period could be invested in starting or expanding some income earning activity off the farm or could be used to improve access to public assets such as telephone service. Thus the inclusion of a lagged dependent variable as an explanatory variable in (5) not only serves to account for historical patterns but also allows flexibility in the assumptions made on the regressors.

Accounting for income persistence, delineating the education variable and including the respective interaction term, model (5) above becomes:

$$\mathbf{INC_{it} = \alpha_0 + INC_{it-1}\alpha_1 + Ed_{it}\,\beta_1 + INC_{it-1}*Ed_{it}\,\alpha_2 + X_{it}\delta + G_{it}\lambda + E_{it}\zeta + \eta_i + \mu_{it}} \qquad \mathbf{(6)}$$

where: Ed is the education variable and X, G and E are as earlier defined. To control for any omitted time invariant 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 as η_i . μ_{it} is a purely random component.

2.3. Estimation

The **dynamic panel data model** (6) 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. Second, 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 this method 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).

$$\Delta INC_{it}=\Delta INC_{it-1}\alpha_1 + \Delta Ed_{it}\beta_1 + \Delta(INC_{it-1}*Ed_{it})\alpha_2 +\Delta X_{it}\delta +\Delta G_{it}\lambda +\Delta E_{it}\zeta +\Delta\mu_{it} \quad (7)$$

Taking first differences of the data eliminates the time-invariant unobservable factors, but this comes at a cost of reducing variation in the regressors (Wooldridge, 2002). This problem is however minimized in this case since our panel has three- to four years between each round. The first difference approach also helps to explain changes in the economic wellbeing of households.

In this study, and to ensure consistency of the estimated parameters, equation (8) 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 take advantage of the three period panel and use previous lags of income level (INC_{it-2}) as instruments for the LID variable. Since we can only use one such instrument from our data, we also use lagged mean rainfall deviation 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 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.

3. Data and Other Specification Issues

3.1. 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 seven years. The household surveys cover the 1996/97, 1999/00 and 2003/04 cropping seasons. The specific sample used in this study consists of a total of 3972 observations (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	Unit	Mean	Std Deviation
Household Income	Ksh ('000)	165	193
Household Demographics			
Education of head	years	6.3	4.9
Age of head	years	55	13.6
Gender of head (male headedness)	1/0	0.84	0.37
% adults working off farm	years	0.36	0.27
Number of adult males	count	2.5	1.6
Number of adult females	count	2.4	1.4
Head completed primary school	1/0	0.53	0.50
Head completed high school	1/0	0.21	0.41
Head with some college education	1/0	0.17	0.38
Public Infrastructure			
Distance to tarmac road	Km	7.72	7.90
Mean distance to input seller	Km	5.93	6.22
Distance to shopping centre	Ksh	9.09	13.3
Distance to extension service	Km	5.39	4.79
Mean dist to electricity & phone	Km	4.53	4.86
Other Variables			
Land cultivated	acres	13.7	16.4
Number of livestock owned	count	18.2	42.5
Months head at home	months	10.4	3.53
Group Membership	1/0	0.78	0.42

No. of Observations=2648

3.2. Specification Issues

Sample Attrition/ Selection Bias

Biases can occur in estimation with incomplete panels. Our survey retained 86% of all households over the three survey periods. Some of the reasons given for exclusion were household dissolution, relocation, refusal to interview, and lack of contact, among others. There would be no selection bias if we can assume that the attrition was a purely random occurrence, but in practice, it is possible that households that were not surveyed could be significantly different from those that remained in the panel. However, previous studies using the same data set do not find significant evidence of any selection bias (Burke et al., 2007). This evidence, and the fairly low attrition rate (about 7% for each survey) allows us to consider attrition in this case as a fairly random occurrence and the nature of those who left the sample unlikely to cause any substantial bias in the results.

Measurement Error

A major source of bias with panel data comes from measurement error, due to the relationship between the explanatory variables and the composite error that includes measurement error. Measurement error could arise when computing final measures of wellbeing. Other sources include a downward reporting bias for income, inaccurate interviewing by enumerator, misreporting by respondents, or data entry mistakes. Many of these types of measurement error could be purely random and so become part of $\mu_{it;}$ others could be constant over time and hence eliminated during differencing.

The most serious estimation biases emanate from systematic measurement errors. For instance, households with high incomes may tend to underreport them (to avoid

exclusion from any future income generating interventions) while those with very low incomes may overestimate (to portray an average household and avoid dishonor). However, consistent with the other income dynamics literature in Africa, these data show little mobility at the two extremes of the income distribution compared to middle income households[5]. This evidence suggests that errors among high and low income households could mostly be eliminated through differencing. It is also plausible to expect minimal misreporting among middle income groups (no incentive since they fall within the average household). Further, we believe that thorough training of enumerators and close supervision of data collection and data entry has minimized errors from such sources.

4.0 Empirical Findings and Discussions

4.1. Characteristics and Mobility of Households across Income Groups[6]

The economic wellbeing of a household is defined not only by the income earning ability but more so the physical and human capital assets that each is endowed with. This trend is evidenced in Table 2, which shows the endowments of the households by income quintile. It is clear that each of the physical and human capital assets increases with income, implying that the wealthy are not only able to earn higher incomes, but are also endowed with higher levels of other assets including education. Land, livestock, and other household and farm assets are likely to be simultaneously determined with income. The education of the household head, however, is minimally influenced by current income since rarely are any of these heads in the school-going age, and participation in

[5] See Table 3 whose contents are presented and discussed later in the paper.
[6] For a detailed descriptive analysis of the characteristics of these households, see Burke at al. (2007).

adult education in Kenya remains low (Republic of Kenya, 2007). Such causality may however be established through income persistence over generations and inheritance.

Table 2.Household Endowments by Income Quintiles

Year	Income Quintile	Mean Incomes (Ksh)	Total land cultivated (acres)	Education of head of household (years)	Value of livestock owned (Ksh)	Value of household and farm assets (Ksh)
2000	Lowest	25246	9.1	4.3	19017	43761
	2	64936	14.2	5.1	26314	72538
	3	108562	12.2	6.1	31984	87056
	4	176770	14.1	6.9	47210	189702
	Highest	421816	21.6	8.2	84143	316794
2004	Lowest	25739	7.3	4.2	17691	66861
	2	68064	9.9	4.9	30980	83463
	3	115561	13.6	6.4	44585	114666
	4	189119	15.3	8.1	63833	207529
	Highest	455967	19.5	9.2	128628	386362
Total	Lowest	25493	8.2	4.3	18354	55311
	2	66500	12.1	5.0	28647	78000
	3	112062	12.9	6.3	38285	100861
	4	182945	14.7	7.5	55522	198616
	Highest	438891	20.5	8.7	106385	351578

Source: Author's calculation

Interestingly, except for the education of the head, there is a sharp jump in the incomes and other assets of the household between the 4th and 5th income quintiles. This reveals the existence of some very wealthy households in these predominantly poor rural communities. Noteworthy also, is the fact that the very poor as defined by the lowest income category are limited in all their endowments[7] and have least education. This is consistent with observations by Baulch and Hoddinott (2000) that 'poverty reflects a conjunction of low endowments, low returns to those endowments and vulnerability to shocks'.

Further evidence can be seen in Table 3 which gives the income quintile mobility matrix for these households between 2000 and 2004. The table shows that 48 percent of households in the lowest quintile in 2000 remained in the same quintile in 2004, while another 25 percent moved up only one quintile. On the other hand, about half of the households that were in the highest income quintile in 2000 remained there in 2004 while another 26 percent moved just one quintile lower. In between though, there is evidence of substantial mobility as earlier mentioned, a fact that could be explained by the ability of these households to move in both directions, unlike the highest and lowest income households. Overall, about 75 percent of all households remained relatively immobile between 2000 and 2004 (38% maintained their status quo while 37% moved only one quintile) with less than 10 percent moving across 3 or 4 quintiles (see Table A2 in the appendices).

[7] Possibly due to negative shocks whose impacts tend to persist or lack of a minimum economic empowerment adequate to push them into a positive growth process.

Table3. Mobility of Households across Income Quintiles

		2004 Income Quintiles				
		1	2	3	4	5
2000 Income Quintiles	1	48	25	15	8	4
	2	23	30	21	17	9
	3	11	25	32	20	12
	4	12	12	21	29	25
	5	4	9	12	26	49

Source: Author's calculation

N/B: Figures represent the distribution of households (%) from each income quintile in 2000 across the five 2004 income quintiles.

4.2. Econometric Results

General Results: Table 4a presents the parameter estimates of the first difference model given by equation (7). For comparison purposes, five different models are estimated, representing different treatments of the lagged dependent variable. Only model 5 involves IV estimation. Model 1 is a first difference (FD) model that represents the most basic type of estimation possible, especially with a two-period panel. This model 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 noteworthy here. One is that the lagged income variable is not differenced. Second, there is a significant increase in the coefficient of determination from the first model, thus indicating the importance of accounting for the initial economic position in an income model. Third, 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) find that reversion to the mean is maintained with the IV estimation (i.e., the coefficient on LDV is negative) but the coefficient is greatly reduced for rural areas. While this may be an accurate reflection of the KIDS data, it is also possible that the inability to difference the lagged income variable and/or the validity of the instrument set may be introducing additional bias to the results. Of concern here is the use of assets and number of household earners as valid instruments for the lagged dependent variable in an income model.

Model 3 also accounts for historical patterns by including the 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. As in Model 2, the coefficient of the lagged income difference is negative and significant. 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.

Table 4a.Determinants of Income Growth

Model	1 FD-No LDV	2 FD-with LDV	3 FD-with LID	4 Full-No IV	5 Full - 2SLS
Lagged Income Variable (LDV)		-0.68***			
		(14.31)			
Lagged Income Difference (LID)			-0.58***	-0.48***	0.49*
			(7.11)	(3.99)	(1.80)
LID* Education				-0.02	-0.15*
				(1.62)	(1.95)
Δ Distance to tarmac road	0.66	0.61	-0.00	-0.17	-1.16
	(0.34)	(0.38)	(0.00)	(0.10)	(0.48)
Δ Mean distance to input seller	-1.08	1.00	1.07	0.76	-2.32
	(0.97)	(1.05)	(0.58)	(0.46)	(1.29)
Δ Distance to shopping centre	-0.13	-0.73**	0.48	0.41	-0.34
	(0.30)	(2.19)	(0.96)	(0.88)	(0.66)
Δ Distance to extension service	0.81	1.01	3.67	3.32	-0.34
	(0.34)	(0.57)	(1.46)	(1.40)	(0.12)
Δ Mean dist to electricity & phone	7.30**	4.69**	5.40*	5.48*	6.79**
	(2.21)	(2.28)	(1.78)	(1.87)	(2.05)
Δ Age of head	2.83	3.51*	4.26*	3.78*	-0.36
	(1.05)	(1.87)	(1.89)	(1.86)	(0.08)
Δ Age of head squared	-0.02	-0.02*	-0.03	-0.03*	-0.01
	(0.90)	(1.67)	(1.64)	(1.66)	(0.26)
Δ Education of head	2.41	4.78***	1.54	6.09*	40.50**
	(1.22)	(3.24)	(1.03)	(1.83)	(2.23)
Δ Gender of head (male headedness)	-4.57	29.28**	10.14	16.57	59.47
	(0.25)	(2.50)	(0.50)	(0.94)	(1.31)
Δ % adults working off farm	102.60***	89.17***	74.96***	76.49***	97.93***
	(5.82)	(6.69)	(4.82)	(4.97)	(4.14)
Δ Land cultivated	1.37	-0.14	1.45	1.19	-0.82
	(0.71)	(0.13)	(1.06)	(0.95)	(0.54)
Δ Number of livestock owned	1.27***	0.73***	1.15***	1.12***	0.91***
	(3.93)	(2.95)	(5.41)	(5.65)	(3.73)
Δ Number of adult males	8.89	11.10**	9.74*	10.87**	19.02*
	(1.34)	(2.46)	(1.88)	(2.22)	(1.94)
Δ Number of adult females	12.38**	10.96***	13.11***	12.71***	9.45
	(2.44)	(3.18)	(3.03)	(2.95)	(1.54)
Δ Months head at home	-2.92	-2.08	-2.19	-2.34	-3.73
	(1.53)	(1.23)	(1.51)	(1.62)	(1.51)
Δ Group Membership	11.83	3.66	11.49	10.67	4.69
	(0.98)	(0.45)	(1.16)	(1.12)	(0.34)
year dummy	-32.08***	98.60***	-9.15	-5.73	11.66
	(3.47)	(8.55)	(1.06)	(0.68)	(0.32)
Observations	1324	1324	1324	1324	1324
R-squared	0.09	0.51	0.43	0.44	

Dependent Variable = Δ Income
Robust t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Source: Author's calculation

As shown from the results, the parameters from models 4 and 5 show broadly similar patterns, but with important exceptions for the LID. Without accounting for the endogeneity of the LID (model 4), its coefficient is negative and significant and does not vary significantly with education of the head. This implies that households are recovering from income shocks, such that worse off households become better in future periods and vice versa. This result is consistent with earlier studies which find overwhelming support for the convergence of household incomes towards the mean, and suggests that serially uncorrelated gains/losses are a substantial proportion of the full income of households.

The above results however change when we account for the weak exogeneity of the LID (model 5). The coefficient of the LID turns positive and significant and varies significantly with education of the household head. The combined effect however remains positive but insignificantly different from zero at mean education changes (0.43 increase in years of education between 2000 and 2004; see Table 4b). The positive and significant coefficient of the LID provides evidence of income persistence for those with less education. It is noteworthy that a 3.3 increase in years of education fully eliminates income persistence, after which the coefficient of the LID turns negative. This indicates possible recovery of income shocks for those with higher education.

Table 4b. Combined Effects of LID and Education at Mean Levels

Table	Model	Variable	Combined effects	F- ststistics	p-value
4a	4	LID	-0.48	19.67	.0000
		Education	5.47	3.37	.0668
	5	LID	0.43	1.59	.2073
		Education	39.88	4.98	.0258

Source: Author's calculation

The difference in the results given by Models 4 and 5 may not be very surprising and could be explained by looking at the estimation methods applied to both models. The method of 2SLS applied to Model 5 uses only the predicted portion of the suspect endogenous variable; this prediction can be viewed as the permanent income component of full income. On the other hand, the LID variable in model 4 consists of both the permanent and transitory components, the latter part possibly generating the 'reversion towards the mean' phenomenon. The differences in these results justify the use of appropriate estimation methods to enable the drawing of relevant conclusions. 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. An additional benefit of the approach is that, by eliminating the transitory component of income, model 5 allows us to focus on whether long-term income persistence is present. This approach is fairly consistent with that in Barrett et al. (2006).

The results of the over identification test for the validity of the instrument set and the first stage regressions are given in Table A3 in the appendices. 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 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 policy recommendations requires further analysis as to the level of education that can achieve the required results. Table A4 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 education. Results show that attainment of a primary school education made no significant contribution to household income and also failed to significantly reduce income persistence or enhance convergence. Among households whose head had attained either a secondary or some college education, the respective education variable was positive and highly significant on its own, and also caused large 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.

Results by Poverty Status: Based on the above findings, Table 5 presents the regression results of the models with the binary variable for secondary education, disaggregated by poverty status. For poor households, the coefficient on LID is positive and significant, while the coefficient on its interaction with education is negative and significant. These results show that poor households whose heads have less than a secondary education are locked in a cycle of low income persistence, which is evidence of cumulative disadvantage and possible existence of poverty traps. This is in contrast to their counterparts with secondary education that manage to break this cycle and recover from negative shocks and hopefully into a positive growth process. However, for households above the poverty line, there is neither evidence of income persistence, nor the role of education in breaking that persistence; these households may be less susceptible to long-term effects of income shocks in either direction[8]. On its own, attainment of secondary

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

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.

Table 5.Models Disaggregated by Poverty Status

	1 **General**	2 **Below Poverty**	3 **Above Poverty**
LID	0.13	0.32***	-0.01
	(0.53)	(3.02)	(0.01)
LID*Secondary education	-1.43*	-3.07***	-1.15
	(1.93)	(4.36)	(0.84)
Δ Distance to tarmac road	0.44	5.23*	-3.32
	(0.21)	(1.82)	(0.86)
Δ Mean distance to input seller	-0.47	-1.15	2.85
	(0.34)	(0.83)	(0.69)
Δ Distance to shopping centre	-0.17	-0.73	0.35
	(0.39)	(1.42)	(0.23)
Δ Distance to extension service	1.81	-1.23	13.35
	(0.66)	(0.51)	(1.55)
Δ Mean dist to electricity & phone	5.71*	2.78	10.46
	(1.75)	(0.87)	(0.95)
Δ Age of head	2.03	-0.95	13.66*
	(0.77)	(0.18)	(1.70)
Δ Age of head squared	-0.02	-0.01	-0.12
	(1.09)	(0.22)	(1.62)
Δ Secondary education	354.60**	539.79***	475.07
	(2.19)	(4.32)	(1.18)
Δ Gender of head (male headedness)	42.11	87.62**	49.15
	(1.36)	(2.31)	(0.55)
Δ % adults working off farm	108.10***	92.60***	137.76
	(5.05)	(3.71)	(1.34)
Δ Land cultivated	-0.13	-2.01	-2.21
	(0.16)	(1.47)	(1.40)
Δ Number of livestock owned	1.21***	0.73**	2.42**
	(5.39)	(2.14)	(2.42)
Δ Number of adult males	14.54**	1.56	49.77***
	(2.05)	(0.23)	(2.59)
Δ Number of adult females	10.46*	12.12*	19.65
	(1.91)	(1.85)	(0.93)
Δ Months head at home	-2.62	-3.75*	1.42
	(1.31)	(1.75)	(0.34)
Δ Group Membership	5.61	2.48	15.42
	(0.48)	(0.13)	(0.46)
year dummy	-8.60	-27.78**	52.02
	(0.45)	(2.16)	(0.62)
Observations	1324	935	389

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

Source: Author's calculation

As for the other determinants of income growth, it seems that having a male head of household positively influences income growth for poor households, but has no significant effect for the non-poor. This could possibly 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 endowed with less land and other assets hence 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 rural income growth. These findings are consistent with those of Giles (2006) in studying China's rural labor market.

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 for the poor, 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 effects 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 is 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 may cause an insignificant result in an otherwise significant variable.

Results by Agricultural Potential: Considering the pattern of income growth by agricultural potential (Table 6), we observe strong evidence of income persistence for those households in the lower agricultural potential areas whose heads had no college training. This persistence is however broken for households with post secondary training (as indicated by the negative coefficient of -0.89 on the interaction term), 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 6.Models Disaggregated by Agricultural Potential

	1 General	2 Low Potential	3 High Potential
LID	0.04	0.16***	-0.03
	(0.20)	(2.65)	(0.12)
LID*college education	-1.19*	-0.89**	-1.11*
	(1.94)	(2.41)	(1.84)
Δ Distance to tarmac road	0.19	-0.54	0.63
	(0.10)	(0.22)	(0.24)
Δ Mean distance to input seller	-0.68	0.86	-1.61
	(0.52)	(0.80)	(0.50)
Δ Distance to shopping centre	-0.16	-0.00	0.15
	(0.37)	(0.01)	(0.22)
Δ Distance to extension service	1.44	0.91	5.89
	(0.57)	(0.44)	(0.99)
Δ Mean dist to electricity & phone	6.52**	2.55	11.45***
	(2.24)	(0.70)	(2.59)
Δ Age of head	1.32	0.80	3.67
	(0.54)	(0.26)	(1.07)
Δ Age of head squared	-0.01	-0.01	-0.03
	(0.75)	(0.21)	(1.30)
Δ College Education	321.63**	200.32**	329.15**
	(2.33)	(2.42)	(2.28)
Δ Gender of head (male headedness)	30.55	19.41	47.60
	(1.18)	(0.81)	(1.49)
Δ % adults working off farm	105.78***	122.79***	88.78***
	(5.32)	(4.62)	(3.18)
Δ Land cultivated	0.38	-0.36	0.99
	(0.42)	(0.63)	(0.50)
Δ Number of livestock owned	0.78*	1.01*	0.78
	(1.77)	(1.84)	(1.62)
Δ Number of adult males	12.21*	10.16	10.94
	(1.90)	(1.21)	(1.32)
Δ Number of adult females	9.37*	16.20**	3.28
	(1.77)	(2.30)	(0.48)
Δ Months head at home	-3.01*	-1.52	-4.17*
	(1.69)	(0.68)	(1.85)
Δ Group Membership	1.71	8.79	-5.29
	(0.17)	(0.53)	(0.37)
year dummy	-9.38	17.91	-11.85
	(0.51)	(1.41)	(0.40)
Observations	1324	430	894

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

Source: Author's calculation

Summary and Conclusions

The results from the study suggest differences in the role of initial economic position of households on rural income growth and mobility, across agricultural potential areas and across poverty status. Overall, rural households in Kenya show evidence (at 10% level) of income persistence, which is broken at higher levels of education. As discussed earlier, this result somehow deviates from earlier findings from Africa where overwhelming support for convergence of incomes has been indicated, possibly due to the differences in the econometric procedures employed. These differences underscore the importance of using appropriate estimation methods. Disaggregation of these results by poverty status and agricultural potential does however provide some answers and also reveal differences that could be important for policy.

Households below the poverty line and whose head does not have a secondary education show strong evidence of income persistence, which is clearly broken for households whose head had a secondary education. No clear pattern emerges for non-poor households with or without secondary education; the direction of effects does however show faster recovery for those with higher education. The existence of income persistence for the 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 on a positive growth process. Overall, results indicate 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.

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 the form of college training is required to break this cycle of low income persistence; the need for such a high level of education is likely related to 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. Government made primary education free for all children four years ago. While this was a policy step of major importance, our analysis suggests that it is not sufficient to improve income growth, nor is it adequate to break the cycle of persistent low incomes 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.

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." *Review of Economic Studies* 58 (2): 277-297.

Barrett, C. B., P. P. Marenya, J. Macpeak, B. Minten, F. Muriithi, W.O. Kosura, F. Place, J. C. Randrianarisoa, J. Rasambainarivo and J. Wangila. 2006. "Welfare Dynamics in Rural Kenya and Madagascar." *Journal of Development Studies* 42 (2):248-277.

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.

Carter, M. R. and J. May. 2001. "One Kind of Freedom: Poverty Dynamics in Post-Apartheid South Africa." *World Development* 29 (12):1987-2006.

Dercon, S. and P. Krishnan. 2000. "Vulnerability, Seasonality and Poverty in Ethiopia." *Journal of Development Studies* 36 (6): 25-53.

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.

Gamba, P. and E. Mghenyi. 2004. "Rural Poverty Dynamics, Agricultural Productivity and Access to Resources." Working Paper No. 21. Nairobi, Kenya: Egerton University, Tegemeo Institute.

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.

Grootaert, C. and R. Kanbur 1995. "The Lucky Few amidst Economic Decline: Distributional Changes in Cote d' Ivoire as seen through Panel Data Sets, 1985-88." *The Journal of Development Studies* 31 (4): 603-619

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.

Kirimi, L. and K. Sindi. 2006. "A Duration Analysis of Poverty Transitions in Rural Kenya." Selected paper presented at the American Agricultural Economics Association Annual Meeting, July 23-26, Long Beach, California.

Low, A. 1986. *Agricultural Development in Southern Africa: Farm-Household Economics and the Food Crisis*. James Currey Ltd: London.

Morduch, J. 1995. "Income Smoothing and Consumption Smoothing." *Journal of Economic Perspectives* 9 (3): 103-114.

Muyanga, M., M. Ayieko and M. Bundi. 2006. "Transient and Chronic Rural Household Poverty: Evidence from Kenya." Paper Presented at the 2007 CSAE Conference, Oxford Univeristy.

Ravallion, M. 1996. "Issues in Measuring and Modeling Poverty." *Economic Journal* 106: 1328-1343.

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

Townsend, R. 1995. "Consumption Insurance: An Evaluation of Risk-bearing Systems in Low-Income Countries." *Journal of Economic Perspectives* 9 (3): 83-102.

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.

Wooldridge, J. 2003. *Introductory Econometrics: A Modern Approach 2e.* Thomson, South-Western.

World Bank. 2000. Global poverty report (July 2000). http://www.worldbank.org/html/extdr/extme/G8_poverty2000.pdf

Appendices

TableA1. Simple Regression Results of Distance Variables on Household Income

Distance variable	Coefficient	p-value
Distance to tarmac road	0.00062	0.516
Mean distance to input seller	0.0010	0.778
Distance to shopping centre	0.0012	0.288
Distance to extension service	-0.00024	0.837
Distance to electricity	0.00023	0.719
Distance to telephone	0.000033	0.934
Mean dist to electricity & phone	0.00062	0.617
Household had electricity	.000110	0.000
Household had telephone	.000042	0.000

Source: Author's calculation

Table A2. Mobility of Households across Income Quintiles

		2004 Income Quintiles				
		1	2	3	4	5
2000 Income Quintiles	1	128 (9.7)	65 (4.9)	39 (2.9)	22 (1.7)	10 (.8)
	2	62 (4.7)	79 (6.0)	55 (4.2)	45 (3.4)	24 (1.8)
	3	29 (2.2)	65 (4.9)	84 (6.3)	54 (4.1)	33 (2.5)
	4	33 (2.5)	33 (2.5)	56 (4.2)	76 (5.7)	67 (5.1)
	5	12 (.9)	23 (1.7)	31 (2.3)	68 (5.1)	131 9.9)

Author's calculation
N/B: Figures in parenthesis are percent of households in total sample.

Table A3. Results of the First Stage Regression and Over-id test

Dependent variable	LID		LID* Education	
	Coefficient.	Std. Error	Coefficient.	Std. Error
Δ Distance to tarmac road	-0.41	1.45	-12.99	12.37
Δ Mean distance to input seller	3.19**	1.43	1.54	9.99
Δ Distance to shopping centre	-0.11	0.41	-1.51	3.58
Δ Distance to extension service	0.74	1.99	-11.17	17.43
Δ Mean dist to electricity & phone	-4.42	2.98	-21.33	25.93
Δ Age of head	-1.19	2.19	-37.98	40.40
Δ Age of head squared	0.01	0.02	0.19	0.26
Δ Education of head	-6.67***	2.51	189.22***	25.85
Δ Gender (male headedness)	27.53	17.68	447.15*	244.19
Δ % adults working off farm	-25.40*	15.76	-90.12	124.81
Δ Land cultivated	-0.88	1.16	-15.06	11.61
Δ Number of livestock owned	-0.26	0.23	-1.72	2.43
Δ Number of adult males	2.81	5.44	76.71*	45.76
Δ Number of adult females	0.48	4.65	-11.47	45.43
Δ Months head at home	0.68	1.50	-5.82	14.88
Δ Group Membership	-4.50	9.37	-45.46	111.71
Year dummy	127.46***	13.25	660.13***	133.04
Instruments				
Lagged Income level	-0.87***	0.07	-4.01***	1.10
Lagged income*education	0.04***	0.01	0.33***	0.11
Lag mean rainfall deviation	-2.64***	0.86	-20.84***	7.11
Joint sig of Instrument set				
F-statistics	55.93		6.93	
p-value	0.0000		0.0001	
Over-id Test				
Chi-square statistic		0.089		
p-value		0.7649		

Source: Author's calculation

TableA4. Comparison with different Education Levels

	1 Continuous Education	2 With primary	3 With secondary	4 With College education
LID	0.49*	-0.08	0.13	0.04
	(1.80)	(0.20)	(0.53)	(0.20)
LID*education	-0.15*	0.10	-1.43*	-1.19*
	(1.94)	(0.09)	(1.93)	(1.94)
Δ Distance to tarmac road	-1.16	0.69	0.44	0.19
	(0.48)	(0.35)	(0.21)	(0.10)
Δ Mean distance to input seller	-2.32	-0.81	-0.47	-0.68
	(1.29)	(0.50)	(0.34)	(0.52)
Δ Distance to shopping centre	-0.34	-0.07	-0.17	-0.16
	(0.66)	(0.16)	(0.39)	(0.37)
Δ Distance to extension service	-0.34	1.22	1.81	1.44
	(0.12)	(0.41)	(0.66)	(0.57)
Δ Mean dist to electricity & phone	6.79**	7.27	5.71*	6.52**
	(2.05)	(1.53)	(1.75)	(2.24)
Δ Age of head	-0.36	2.56	2.03	1.32
	(0.08)	(0.85)	(0.77)	(0.54)
Δ Age of head squared	-0.01	-0.02	-0.02	-0.01
	(0.26)	(0.59)	(1.09)	(0.75)
Δ education of head	40.50**			
	(2.23)			
Δ primary education		-22.65		
		(0.12)		
Δ secondary education			354.60**	
			(2.19)	
Δ college education				321.63**
				(2.33)
Δ Gender of head (male headedness)	57.46	-7.13	40.30	28.57
	(1.26)	(0.15)	(1.29)	(1.09)
Δ % adults working off farm	96.75***	98.67***	106.48***	103.33***
	(4.19)	(4.85)	(4.97)	(5.19)
Δ Land cultivated	-0.77	1.56	-0.11	0.39
	(0.51)	(0.64)	(0.12)	(0.42)
Δ Number of livestock owned	0.91***	1.28***	1.21***	0.78*
	(3.73)	(3.00)	(5.39)	(1.77)
Δ Number of adult males	19.02*	8.54	14.54**	12.21*
	(1.94)	(1.07)	(2.05)	(1.90)
Δ Number of adult females	9.45	12.43**	10.46*	9.37*
	(1.54)	(2.47)	(1.91)	(1.77)
Δ Months head at home	-3.73	-2.92	-2.62	-3.01*
	(1.51)	(1.64)	(1.31)	(1.69)
Δ Group Membership	4.69	11.45	5.61	1.71
	(0.34)	(0.83)	(0.48)	(0.17)
year dummy	11.66	-29.84	-8.60	-9.38
	(0.32)	(1.34)	(0.45)	(0.51)
Observations	1324	1324	1324	1324

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

Source: Author's calculation