WORKING PAPER

8/2015

Income diversification and working children

Elin Vimefall
Economics

ISSN 1403-0586

http://www.oru.se/esi/wps
Örebro University Swedish Business School
701 82 Örebro SWEDEN

Income diversification and working children

Elin Vimefall

Swedish Business School at Örebro University

70182 Örebro, Sweden

e-mail: elin.vimefall@oru.se

September 2015

Abstract
During recent years (2004-2008) the proportion of working children in sub-Saharan Africa has increased (Diallo et al. 2010). At the same time, there has been a shift in the patterns of livelihoods, whereby households rely more on sources of income from outside their own farms. When the adult in the household diversifies away from production on their own farm, this is expected to influence the children's time allocation in several ways. In this paper, we investigate how households' income diversification strategy influences children's probabilities of working and going to school among children living in farming households in Kenya, using data from the Kenya Integrated Household Budget Survey 2005/2006. Furthermore, we also analyse child labor at the intensive margin by investigating if the households income diversifications strategy influences the number of hours worked. We find that children living in households that rely solely on the production of their own farms are about 3 percent points more likely to work and about 2 percent points less likely to be in school than children from more diversified households. Furthermore, children living in household who rely only on farming for their income does also work more hours than other children. However, we do not find any differences in the proportions of working children across a number of different income diversification strategies.

Keywords: Income diversification, Child labor, Work, Schooling, Kenya

JEL Classification: O12; O55; O15; I25; J13

1. Introduction

About 10 percent of the children in the world are involved in child labor (ILO 2013), and the highest rate of child labor is found in sub-Saharan Africa (21 percent). Child labor has several negative aspects. It can directly harm a child's physical and mental health, but it can also negatively affect their education, which may in turn reduce their adult earnings (Emerson and Souza 2011). It can thus damage the national economy more broadly, since education and health are important factors in the development of a country.

Even though the rate of child labor has declined across the world, in sub-Saharan Africa child employment increased, both in absolute number and as a percentage, between 2004 and 2008 (Diallo et al. 2010).

Most working children are found in the agricultural sector, which in sub-Saharan Africa is the main source of income for a large proportion of the population. However, as the countries in sub-Saharan Africa are developing, other income sources are growing in importance. When the adult in the household diversifies away from production on their own farm, this is expected to influence the children's time allocation in several ways. Households will choose to diversify if the marginal value of any household member's time spent on off-farm work is larger than the marginal value of their time spent working on their own household's farm. Thus, when a household chooses to diversify, this will have a positive effect on its total income. Assuming that a child not having to work is viewed as a luxury good by the household, a rise in income will decrease the rate of working children. At the same time, decreasing the amount of adult labor on the household's own farm will increase the return on the child's work at the farm, and thereby the opportunity cost of schooling. This effect is expected to increase the rate of working children. Since these two effects give different predictions, it is up to empirical research to determine the relationship between income diversification and the rate of working children.

In this paper, we analyze how the household's income diversification strategy influences both the probability that a child is working and the probability that the child is going to school, focusing on farming households in the rural parts of Kenya. In addition, we analyze whether the household's income diversification strategy influences the number of hours worked by the child. To our knowledge, this is the first paper to analyze the connection between income diversification and working children and income diversification and schooling, in terms of both the extensive (probability of working and of going to school) and the intensive (number of hours worked) margin.

We analyze the work decision in conjunction with the decision to go to school, using a bivariate probit model together with data from the Kenya Integrated Household Budget Survey from 2005/2006. Then, a Tobit model is used to examine the child work supply at the intensive margin, by analyzing the number of hours worked.

We find that children living in households that rely solely on the production of their own farms are about 3 percentage points more likely to work as their main activity and about 2 percentage points less likely to be in school than other children. They also work more hours than children in diversified households.

The remainder of this paper is organized into six sections. The next section discusses the theoretical background on the relationship between income diversification and the child's time allocation. Section 3 presents the institutional setting and Section 4 presents the data. Section 5 presents the empirical strategy and section 6 show the results. The last section concludes.

2. Background and related literature

2.1 Potential mechanisms

There are several mechanisms that could explain a connection between income diversification and the probability that a child is working. First, we consider a farming household that does not

sell any labor. In the first period, the household's income is given only by its own farm's production. In the next period, the household can continue to survive only on the production of its own farm or choose to diversify its income. We assume that children can only work on their own farm. In our sample, about 94 percent of the working children stated that they worked on their own family's farm.

The household will choose to diversify if the marginal value of any household member's time spent on off-farm work is larger than the marginal value of their time spent working on their own household's farm. Therefore, if the household chooses to diversify, it is expected to increase their income.[1] A positive connection between diversification and income has been supported by empirical studies. Non-agricultural income has been shown to increase earnings, and diversification has been identified as an important way out of poverty (e.g. Barrett et al. 2005, Kristjanson et al. 2010, Freeman et al. 2004, Bigsten and Tengstam 2011, Hoang et al. 2014).[2] As a consequence of an increase in income, we would expect the rate of working children to decrease.

Why do we expect a negative relationship between income and the rate of working children? Basu and Van (1998) build their seminal paper on the assumption that parents only let their children work if their income would otherwise be under some subsistence level. This means that children not having to work is treated as a luxury good that poor households cannot afford to consume, but will choose to consume as soon as household income rises above subsistence level. Thus, in their model, a rise in income has no effect unless it takes income above subsistence level.

Even if parents do not mind letting their children work, we would still expect to observe a negative relationship between income and the child's probability of working, because of the

[1] This is somewhat of a simplification since households can have other reasons for diversification than to increase income.
[2] There is no consensus in the literature on how to define income diversification, and different studies use different measurements. However, most studies look at diversification into non-agricultural work.

decreasing marginal utility of income. Therefore, when family income increases, the utility gained from the income generated by children is reduced, making it less probable that parents will send their children to work. Furthermore, a higher income makes it possible for a family to afford substitutes for their child's work, such as a household water source removing the need to walk long distances to get water. Higher income could also be used to purchase items that might increase the child's productivity in other activities (for example, books that would increase the child's productivity in schooling) (Edmonds and Pavcnik 2005). Using data from Tanzania, Beegle et al. (2006) found that an income loss caused by a crop loss increased child labor (hours worked), while enrolment in education decreased. Edmonds (2006) found a decrease in hours of work for children when the social pension was extended to include black South African families at the end of apartheid, and Edmonds (2005) found that the rate of child labor in Vietnam decreased when expenditure per capita increased.

The income effect: Due to a higher level of income when a household diversifies, we expect to observe a lower rate of working children in diversified households.

Since the rates of return are expected to vary among income diversification strategies, we expect the rate of working children to differ both between undiversified and diversified households, and among different diversification strategies. For example, we expect that non-agricultural wage employment could provide a relatively large return for the household, while agricultural wage employment is often viewed as a low-return activity.

Next, if a household chooses to diversify it will use less labor on its own farm. This will increase the marginal productivity for a working child (and the opportunity cost of schooling). Especially when hired labor might require supervision, family labor might be preferable.

Therefore, we expect that children's work will be used as a substitute for the parents' work on the farm. This will have a positive effect on the rate of working children.

The substitution effect: Due to the increased marginal productivity of children's work on the farm when a household diversifies, we expect to observe a higher rate of working children in diversified households.

If the income effect is larger than the substitution effect, income diversification will produce lower rates of working children. If the substitution effect is larger than the income effect, then income diversification will lead to higher rates of working children.

Furthermore, households with different income diversification strategies might differ in their expectations about the rate of return of schooling. In a household where all members work on the farm, a child might be supposed to follow this tradition. Thereby, the expected rate of return from education might be lower than if the child is supposed to work in the labor market. In addition; the household might expect to achieve a positive effect on future income by letting a child work on the farm and learn from his/her parents.

Relaxing the assumption that children only work on their own households' farms could provide more connections between children's probability to work and income diversification. For example, there has been some concern that microcredits, by increasing the number of small family enterprises, could potentially increase the amount of child labor, since this is a sector where child labor is common. It could also be the case that some diversification strategies make it easier for parents to bring their children with them to work (for example to work on some other farm).

Since there are several potential mechanisms by which children's' probability to work and a household's income diversification strategy could be connected, it is up to empirics to shed light on the connection.

2.2 Earlier literature

An extensive literature exists on the determinants of child labor, nicely summarized in, for example, Basu and Tzannatos (2003), Bhalotra and Tzannatos (2003), Edmonds (2007) and, more recently, Congdon Fors (2012).

Even though there is no previous literature on the relationship between income diversification and child labor or income diversification and schooling, a related literature analyzes the connection between access to microcredits and child labor. This literature is based on the concern that microcredits increase the number of small household enterprises, where child labor is common. Wydick (1999) argues that access to credit for household enterprises could have both a positive and a negative effect on schooling (and child labor). First, access to credit could allow the household to use hired labor instead of the labor of children. This would increase schooling and decrease the rate of child labor. Second, when the household enterprise gets more capital, the marginal product of child labor (and the opportunity cost of schooling) increases. This effect would instead increase the rate of child labor and decrease schooling. Putting his model to the test, Wydick (1999) find that the positive effect of credit seems to dominate (i.e. access to credit increases schooling in the sample tested). However, in family enterprises with a high risk of moral hazard connected to hired labor, households seem to prefer to let their children work, and the positive effect on schooling becomes lower. Shimamura and Lastarria-Cornhiel (2010) find some evidence that credit uptake reduces the probability of being enrolled in education among young girls in rural Malawi, and Hazarika and Sarangi (2008) show that access to microcredit increases the child's probability of working during peak season. However,

they find that the increase is mostly caused by children taking over domestic chores while their parents work in the household enterprise. The authors find no effect on the probability of being in school.

To our knowledge, the only previous literature about determinants of child labor in Kenya is provided by Moyi (2011) and Buchmann (2000). Moyi (2011) uses a definition of child labor that includes household chores. He finds that, especially in the rural parts of Kenya, many children combine labor and school. Furthermore, children in wealthier households have a lower probability of doing labor (or combining labor and school). Children who live in households where the head has some education and who are the biological children of the head of household have a lower probability of only doing labor (compared to only being in school). That many children combine school and work is confirmed by Buchmann (2000). Furthermore, she finds that, even though wage work is uncommon among children in Kenya, children do help on the family farm and with household chores.

3. Institutional setting

Even though only 15 percent of Kenya's land is suitable for farming, agriculture dominates the economy, and in 2006 about 75 percent of the labor force worked in that sector. Tea, horticultural products and coffee are the main cash crops, while maize is the main food staple. The semi-arid areas in the north and east are instead dominated by livestock.

In general, farming households in Kenya do diversify their income, and about 45 percent of the income comes from outside the agricultural sector.[3] The percentage of income that comes from farming is lower in the richest group of households, while the part that comes from non-agricultural wages is higher (Figure 1).

[3] Own calculation, using data from the RIGA database.

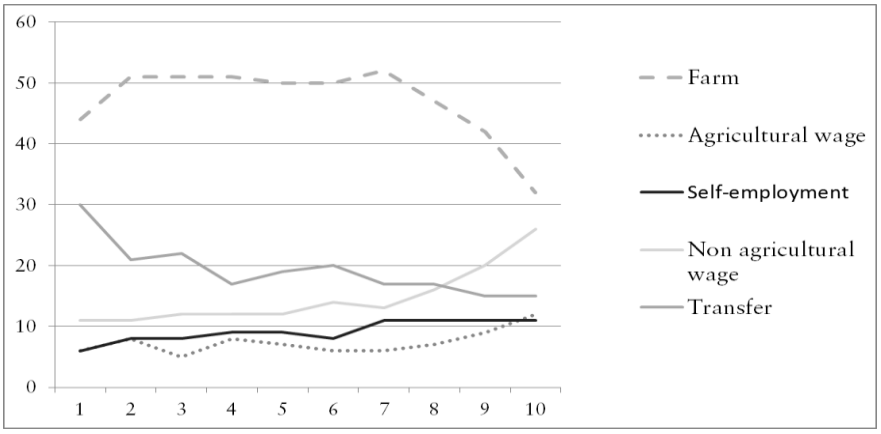

Figure 1 Source of income and income share, by expenditure decile

Source: Own computation, using data from the RIGA database.

The school system consists of eight years of free primary education, four years of secondary education and four years of college/university education. A child is supposed to start school at the age of six and primary education is free of charge. Most children get some education and in 2009 the enrollment rate in primary education was about 82 percent (World Bank data). At the end of grade 8, children take the Kenya Certificate of Primary Education examination that determines the school to which the child can move on to for his or her secondary education. In 2009, the enrollment rate in secondary education was 50 percent (World Bank data).

Most working children in Kenya work in agriculture, and many children combine work with school. Children's employment is regulated by the Employment Act (2007) and the Children's Act (2007). The minimum age for employment is 16 years. However children aged 13-16 are allowed to do light work that is not harmful to their health, development or education. The age restriction does not apply when working for another family member or in the agricultural sector (unless the work is dangerous to the child's life, health, or moral development). According to the Children's Act, every child should be protected from any work that interferes with his or her education.

4. Data and descriptive statistics

4.1 Work and schooling

Data are extracted from the Kenya Integrated Household Budget Survey 2005/2006, which covers 1343 clusters of 10 randomly selected households, giving a sample of 13430 households. Our analysis uses data only on children of 6-14 years old, which is the age interval when they are supposed to be in primary education. Since we are interested in what happens to the rate of working children when farming households diversify, we only include households living in the rural parts of the country that participate in farming. This gives a final sample of 10847 children.

All children were asked about their main activity during the past seven days. A child was defined as working if their main activity was working for pay, working in their own family's business, or working on their own family's agricultural holding.[4] Since child employment is illegal in Kenya, we might have a problem with under-reporting of children working for pay. It is also important to note that we only capture those children who actually live in the household. Furthermore, this measure could be dependent on when the interview was conducted, since for example, we would expect more children to state that they had worked as their main activity during harvest.

All children were also asked if they were currently attending school.[5] In our sample, 7 percent of the children reported work as their main activity (Table 1), while 90 percent attended school. A larger proportion of boys were working as their main activity (7.93 percent) than girls (5.83 percent). Among those who worked, 94 percent reported working on a family agricultural holding, 2 percent in the family business and 4 percent for pay. 66 percent of the working children did attend school.

[4] Mothers or guardians answered for children under 10 years of age.
[5] Again, mothers or guardians answered for children under 10 years of age. If school was not in session at the time, the respondent was asked if he/she had attended school in the session just completed and planned to attend in the next session.

Table 1: Rate of children working and children attending school, n = 10847

Age	Working (%)	Schooling (%)
6	3.28	81.47
7	3.32	89.07
8	5.26	89.49
9	5.27	93.28
10	8.45	91.32
11	7.06	93.97
12	10.61	91.31
13	8.93	92.56
14	10.59	91.42
Total	6.89	90.27

There is a large degree of variation in the definitions used when analyzing child labor/work, which presumably could have a large impact on the results. Several studies highlight the importance of including household chores when analyzing, while others point to the importance of how the questions are asked (Dillon et al. 2012). Because of data limitations, we only include information about the child's main activity during the past seven days. This gives a lower threshold for the rate of working children in Kenya.[6] However, since these children work as their main activity, this is presumably the group for which work comes at the expense of schooling and can be considered especially harmful for the child's future development. Therefore, it is particular important to understand why these children work.

Children who did not attend school stated that their "parents did not let me" as the main reason, followed by having no money and working to help at home (Table 2). However, we do not have information about why parents do not let children attend school. One possible reason could be that they need their children to help in the household or on the farm.

[6] Only children who worked as their main activity were asked more detailed questions about their work. Even though children who did not work as their main activity stated whether they had worked for anyone outside the household during the previous 12 months, they did not state whether they had worked on the household's farm. However, as a robustness test, we re-estimate our models using a broader definition of child labor, including all children who stated that they had worked at any time during the previous 12 months.

Table 2: Reasons for not attending school, n = 799

	(%)
Parents did not let me	27.03
No money	20.03
Working to help at home	18.90
Not interested	8.64
Own illness	7.38
School too far	6.76
School conflicts with beliefs	4.01
Family illness	2.13
Other	5.17

There was a large dispersion in the number of hours worked per week (Figure 2), with some children working so much that participation in school would have been impossible.[7] Others worked very few hours but still reported this as their main activity.

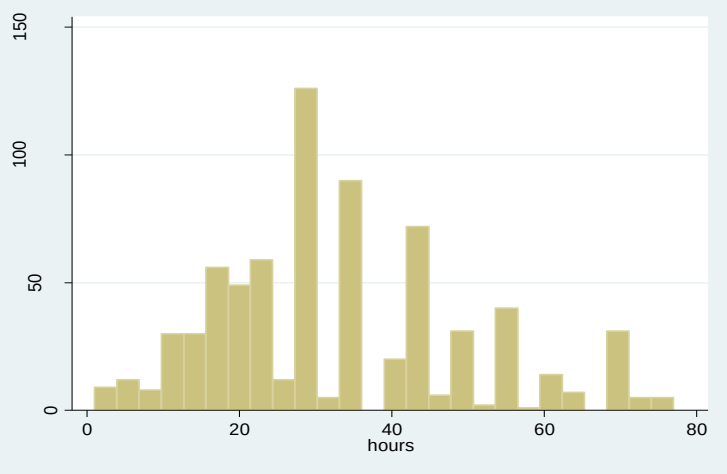

Figure 2 Hours of work

4.2 Income diversification strategy

The main explanatory variable is the household's income diversification strategy. This is traditionally defined based on the income-generating activities engaged in by all members of the household, including the child work contribution. This makes the variable endogenous in the child work and school equation. To control for this, in this paper we define the household's

[7] Children who stated that they worked more than 80 hours per week have been excluded.

income diversification strategy based on the activities engaged in by the adults in the household. An adult is defined as an individual aged 16 or over, since this is when employment becomes legal in Kenya.

Incomes are reported from four different sources: one's own farm, agricultural wage labor, non-agricultural wage labor, and non-agricultural family business. We define a household as having a farm if any member of the household has been engaged in farming during the previous 12 months (self-employed or as a tenant), or if any member has raised or owned livestock, poultry, fish or bees during the past 12 months. Since we do not have information about which household members engaged in these activities, we have to assume that no child is solely responsible for farming. An individual (aged 16 or over) is defined as doing wage labor if he or she gave "paid employee" as her or his employment status. The wage category is further divided into agricultural wage labor and non-agricultural wage labor. This division is based on ISIC (REV2) codes where 1000-1999 are defined as agricultural, and 2000 and above as non-agricultural. A household is defined as having a non-agricultural business if anyone (aged 16 or over) in the household operated any non-agricultural income-generating enterprise (self-defined) that produced goods or services, or if anyone (aged 16 or over) owned a shop or operated a trading business in the period in question.

Since all households in our sample get some income from production from their own farms, the households can be combined into seven groups: Full-time farmer (F), Farmer and Agricultural wage worker (FA), Farmer and Non-agricultural wage worker (FN), Farmer and family Business (FB), Farmer, Agricultural wage worker and family Business (FAB), Farmer, Non-agricultural wage worker and Agricultural wage worker (FNA), and Farmer, Non-agricultural wage worker and family Business (FNB)[8]. In our sample, 56 percent of the children lived in a full-time farmer household (Table 3).

[8] No household reported combining all 4 income sources (FNAB).

Table 3: Income diversification strategies (%), n = 10847

F	56.06
FA	7.42
FN	12.64
FB	17.03
FAB	1.31
FNA	1.34
FNB	4.20

F= Full-time farmer, FA= Farmer and Agricultural wage worker, FN= Farmer and Non-agricultural wage worker, FB= Farmer and family Business, FAB = Farmer, Agricultural wage worker and family Business, FNA = Farmer, Non-agricultural wage worker and Agricultural wage worker, FNB= Farmer, Non-agricultural wage worker and family Business

We expect that one of the main mechanisms for a possible connection between income diversification and child work/schooling works through income effects. Evaluating the poverty levels in different household types, we find that poverty is greatest among the full-time farming households, while it seems to be lowest among the households that diversify into non-agricultural wage labor. [9]

Table 4: Poverty level by diversification strategy, n = 10847

	Poverty level (%)
F	58.42
FA	55.54
FN	43.73
FB	48.12
FAB	45.54
FNA	38.82
FNB	32.19

F= Full-time farmer, FA= Farmer and Agricultural wage worker, FN= Farmer and Non-agricultural wage worker, FB= Farmer and family Business, FAB = Farmer, Agricultural wage worker and family Business, FNA = Farmer, Non-agricultural wage worker and Agricultural wage worker, FNB= Farmer, Non-agricultural wage worker and family Business

[9] Poverty is evaluated using the national poverty line of Kenya.

Even though this could be viewed as support for the hypothesis that income diversification increases income, it is important to note that the mechanism also could work the other way around, i.e. that better-off households have a larger probability of diversifying.

5. Econometric specification

5.1 Work and schooling

Our main interest lies in the decision of the parents to send the child to work or not, which gives a binary dependent variable. Another aspect is the relationship between school and work. A probit or logit model with school as an explanatory variable for the work decision would lead to simultaneity bias. Thus, following Canagarajah and Coulombe (1997) and Skyt-Nielsen (1998), we use a bivariate probit model that treats the work and school decisions as two interdependent decisions. A household will choose to let a child work (go to school) if the net utility from doing so is positive. However, since the utility level is unobservable, we model this using a latent variable approach:

$$y^*_{1i} = \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_{1i} \qquad y_{1i}=1 \quad \text{if } y^*_{1i} > 0, \text{ otherwise } 0$$

$$y^*_{2i} = \beta_3 x_{1i} + \beta_4 x_{2i} + \varepsilon_{2i} \qquad y_{2i}=1 \quad \text{if } y^*_{2i} > 0, \text{ otherwise } 0$$

$$\varepsilon_i = (\varepsilon_{i1,} \varepsilon_{i2})' \sim N(0,\sum)$$

$$\sum = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

y^*_{1i} and y^*_{2i} are the unobserved net utilities from letting the child work (y^*_{1i}) and go to school (y^*_{2i}). If y^*_{1i} is positive, the child will work, and we will observe y_{1i} =1, (otherwise y_{1i} = 0). If y^*_{1i} is positive, the child will be sent to school, and we will observe y_{2i}=1 (otherwise y_{2i} = 0).

Since a child can work as their main activity, but still attend school, y_{1i} and y_{2i} could both be one for the same individual.

$\mathbf{x}_1$ are a set of dummy variables indicating the household's income diversification strategy and $\mathbf{x}_2$ are control variables explaining the work and school decision. ρ is the coefficient of correlation of the error terms, measuring the correlation between outcomes after controlling for the explanatory variables (Greene 2003). The model is estimated with maximum likelihood, and the error terms are clustered at the district level.

5.2 Hours worked

To analyze the number of hours worked, we use a Tobit model, censored from below, since it is only possible to work a positive number of hours. Since we do not have any information about the amount of time spent on education, we focus here on the time spent on work. As in the previous section, we use a latent variable approach:

$$y^*_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i$$

$$y_i = y_i^* \quad \text{if } y_i^* > 0$$

$$y_i = 0 \quad \text{if } y_i^* \leq 0$$

$$\varepsilon_i \sim N(0, \sigma^2)$$

where y^*_i is the optimal number of hours the household would let the child work. However, since it is not possible to work a negative number of hours, if $y_i^* \leq 0$ the household will choose to not let the child work and we will observe $y_i = 0$. If $y_i^* > 0$, we will observe $y_i = y_i^*$. $\mathbf{x}_1$ are a set of dummy variables indicating the household's income diversification strategy, and $\mathbf{x}_2$ are control variables. The error terms are clustered at the district level.

5.3 Level of income

Since there are several practical problems with measuring income in developing countries it is common to instead use household expenditure (Deaton 2000). We follow this approach and use per adult equivalent expenditure to measure the income level in the household. However, since a working child contributes to the household income, the variable would be endogenous in the work equation. This simultaneity would create a positive bias in the coefficient for income. Therefore we use the instrument approach. To find a good instrument for income is not an easy task. We use three instruments measuring the quality of housing: (1) whether the household has a toilet within the dwelling, (2) whether it has electricity and (3) whether the floor in the house is made of cement. Since these goods represent relatively large investments, it is not likely that they will have been financed by child work. Even though these instruments might not be perfect, we argue that they are still an improvement compared to using expenditure directly. To use the instrument approach we build on the model in Section 5.1 and introduce an equation for the household's expenditure including the instruments. This gives a three-equation mixed-process model (Roodman 2011):

$$y^{*}_{1i} = \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 y_{3i} + \varepsilon_{1i} \qquad y_{1i}=1 \quad \text{if } y^{*}_{1i} > 0, \text{ otherwise } 0$$

$$y^{*}_{2i} = \beta_3 x_{1i} + \beta_4 x_{2i} + \beta_3 y_{3i} + \varepsilon_{2i} \qquad y_{2i}=1 \quad \text{if } y^{*}_{2i} > 0, \text{ otherwise } 0$$

$$y_{3i} = \beta_3 x_{1i} + \beta_4 x_{2i} + \beta_5 x_{3i} + \varepsilon_{3i}$$

$$\varepsilon_i = (\varepsilon_{i1}, \varepsilon_{i2}, \varepsilon_{i3})' \sim N(0, \sum)$$

$$\sum = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & \sigma_{33} \end{bmatrix}$$

where y_{3i} is the household expenditure and x_{3i} is our instrument. The coefficient of correlation between expenditure and work, ρ_{13}, measures the endogeneity of expenditure in the work equation, and the coefficient of correlation between expenditure and schooling, ρ_{23}, the endogeneity of expenditure in the school equation. The model is estimated using maximum likelihood, and the error terms are clustered at the district level.

5.4 Control variables

The child's time allocation will be affected by the marginal rate of return from different activities. If labor productivity increases as a child gets older, older children will have a comparative advantage in working. However, the use of a child's time is not necessarily determined by the actual productivity, but rather by what the parents assume it to be. Among mothers interviewed in Kenya, 26 percent stated that boys were smarter than girls, and 24 percent thought that the job market was worse for women than for men (Buchmann 2000). Daughters of parents who thought that the job opportunities were worse for women had a lower probability of being enrolled in school. The belief that boys were smarter than girls did not influence the probability of girls being in school. To control for this, we include both the child's age and gender.

The child's marginal productivity from working on the family farm will furthermore be determined by the amount of land relative to the amount of labor. In a larger family, the child's probability of working is thereby expected to be smaller (holding land area constant). Bhalotra and Heady (2003) argue that, even though an increase in land holdings can increase income and thereby decrease the amount of child labor, it can also have the opposite effect by creating an incentive to employ the children. Since the marginal productivity is expected to increase with land size, the effect becomes stronger as land holdings increase. Basu et al. (2010) show that the relationship between child labor and landholdings takes the form of an inverted U, whereby

child labor first increases and then decreases. However, in their setting (Vietnam), it seems that the relationship works through an increase in hours spent on domestic chores. To control for this, we include a dummy capturing whether the household owns land, the area of land farmed, the area of the land squared, and the amount of labor in the household.

The probability of working or going to school is not only expected to depend on the amount of labor in the household, but also on the number of children. Having more children will increase the subsistence level of income. Therefore, the quality and quantity of children may come at the expense of each other. Higher quality costs more, which reduces the number of children that can be supported (Becker and Tomes1976). If having a larger family raises the required subsistence income, then having more siblings will reduce a child's probability of being in school and increase their probability of working. However, it could be important to take the activities of the siblings into consideration (Patrinos and Psacharopoulos1997). For example, having siblings too young for school or work means that someone has to take care of them. Therefore, we control for whether a child has any siblings in the following age intervals: under the age of 6, 6-14, and 15-18. These groups correspond to the ages when a child is too young to be in school, and is supposed to be in primary education and secondary education, respectively.

The work versus school decision can also be influenced by the parents' level of education, with more educated parents expected to have a lower probability of sending their children to work and a higher probability of sending them to school. The coefficient for education in the child work equation thus represents the parents' attitudes to work, aspirations for the child's future, and time preferences (Bhalotra and Tzannatos 2003). Children of more educated parents have been shown to have a lower probability of working and a higher probability of being in school (e.g. Emerson and Souza 2003). To control for this we include the educational level of the head of household.

The work versus school decision can also differ among different groups in society, because of different norms and culture. Therefore, we include a variable capturing the household's ethnolinguistic background.[10]

The probabilities of working and of being in school will furthermore be influenced by the access to, cost of and quality of education. To control for this, we include the mean education cost (by cluster) in Kenyan Shillings for students in government primary schools (divided by 1000). The quality of a school is measured as the percentage of students in the second class who can write. Distance to school is measured as the percentage of households that are more than 5 km from the nearest public primary school. Do to data limitations, the last two measures are at the district level.

Kenya is a highly diversified country, and school and work opportunities are expected to differ among different parts of the country. Therefore, all equations include a set of province dummies. For definitions of all explanatory variables, see Table A1 in the appendix, and for descriptive statistics, see Table A2.

6. Results

In Section 2 we concluded that there exists a potential mechanism that could produce either a positive or a negative relationship between income diversification and work/schooling. In this section, we turn to the data to shed some light on the relationship.

6.1 Income diversification strategy

First, we want to know whether the household's overall income diversification strategy is correlated with the probabilities that a child is working and attending school, respectively.

[10] This information comes from information about which language the household used when answering the survey, which was available in eleven local languages and English and Swahili. Since there is a strong connection between language and ethnicity in Kenya we argue that this variable is a good proxy for the household's ethnolinguistic background.

Estimates of the marginal effects from the bivariate probit model, using only the household's income diversification as an explanatory variable, are presented in Table 5.

Table 5: Marginal effects of income diversification strategy estimated at the mean. Three-equation mixed-process model without controls, n=10847

	Work		Schooling		Work		Schooling	
Diversified	-0.042	***	0.080	***				
	(0.013)		(0.022)					
FA					-0.033	**	0.060	***
					(0.017)		(0.022)	
FN					-0.047	***	0.089	***
					(0.017)		(0.022)	
FB					-0.035	**	0.073	***
					(0.014)		(0.023)	
FAB					-0.039		0.090	**
					(0.030)		(0.038)	
FNA					-0.049	*	0.116	***
					(0.027)		(0.045)	
FNB					-0.074	**	0.124	***
					(0.031)		(0.035)	
ρ	-0.497	***			-0.495	***		
Log-pseudolikelihood	-5891				-5883			
Controls	No		No		No		No	

Notes: *** 1% significance level, ** 5% significance level, * 10% significance level. Standard errors clustered at the district level. FA= Farmer and Agricultural wage worker, FN= Farmer and Non-agricultural wage worker, FB= Farmer and family Business, FAB = Farmer, Agricultural wage worker and family Business, FNA = Farmer, Non-agricultural wage worker and Agricultural wage worker, FNB= Farmer, Non-agricultural wage worker and family Business. The comparison group is Full-time farmers (F).

In the first specification, we compare children from households that have a diversified income portfolio to children from households that obtain all of their income from production from their own farm. We find that if a child comes from a full-time farming household, the probability that he or she work as main activity increases by about 4 percentage points, and the probability that he or she is in school decreases by about 8 percentage points. Looking at the different forms of diversification, it seems to be the case that children from households that diversify into non-agricultural wage labor (alone, or together with other activities) have the lowest probability of working. However, the difference is not statistically significant. To see whether these results

hold after controlling for other factors, we re-estimate the model including all control variables (Table 6).

Table 6: Marginal effects of income diversification strategy estimated at the mean. Three-equation mixed-process model with controls, n=10847

	Work		Schooling		Work		Schooling	
Diversified	-0.029	***	0.017	*				
	(0.008)		(0.008)					
FA					-0.035	**	0.026	*
					(0.016)		(0.014)	
FN					-0.034	**	0.021	*
					(0.014)		(0.013)	
FB					-0.020	*	0.009	
					(0.011)		(0.011)	
FAB					-0.031		0.010	
					(0.032)		(0.033)	
FNA					-0.042		0.034	
					(0.029)		(0.036)	
FNB					-0.049	*	-0.016	
					(0.028)		(0.021)	
Expenditure	-0.028	**	0.056	***	-0.026	**	0.056	***
	(0.012)		(0.017)		(0.012)		(0.018)	
ρ_{12}	-0.755	***			-0.752	***		
ρ_{13}	0.280	***			0.270	**		
ρ_{23}	-0.336	***			-0.335	***		
Log-likelihood	-22208				-22172			
Controls	Yes		Yes		Yes		Yes	

Notes: *** 1% significance level, ** 5% significance level, * 10% significance level. Standard errors clustered at the district level. FA= Farmer and Agricultural wage worker, FN= Farmer and Non-agricultural wage worker, FB= Farmer and family Business, FAB = Farmer, Agricultural wage worker and family Business, FNA = Farmer, Non-agricultural wage worker and Agricultural wage worker, FNB= Farmer, Non-agricultural wage worker and family Business. The comparison group is Full-time farmers (F).

Including all control variables, we still find that children from full-time farming households have a larger probability of working as their main activity and a lower probability of being in school. However, the magnitude of the effect has decreased.

As previously discussed, we expect that household expenditure is endogenous in the work and school equations. Therefore, we introduce the instruments presented in Section 5.3. Using the instruments, we find that we can reject the hypothesis that expenditure is exogenous in the work and school equations, supporting the view that instruments are needed. All the instruments

are relevant in the first-stage regression (at the 1% level of significance). Furthermore, Hansen's over-identification test cannot reject the hypothesis that the over-identification restrictions are valid. Comparing the estimates with and without the instruments shows that the estimated coefficients for expenditure increase in magnitude (Table A3, appendix). In the work equation, the coefficient goes from statistically insignificant to significant, supporting the hypothesis that poverty is a reason for child labor. It does also change the estimated coefficients for the income diversification strategies, indicating that expenditure is important in explaining the differences in rate of working children among income diversification strategies.

Looking at the control variables, we find that older children and boys have a higher probability of working than others (Table A4, appendix). If the head of household has an education the probability of a child working decreases. Somewhat surprisingly, the biological children of the head of household have a higher probability of working, and the probability that a child is working decreases when the cost of education increases. The effect of having a head of household with at least some secondary education seems to work through the household income. Excluding expenditure from the equation makes the estimated coefficient for secondary education both statistically significant and larger in magnitude.

Dropping the children who are six years old does not affect the results. However, changing the sample to children aged 10-14 increases the effect of the income diversification strategy, by increasing the magnitude of the coefficients (Table A5, appendix). Next, we change the dependent variable to include children who did not work during the previous seven days, but who had a job to return to, or who had been employed or had participated in casual labor at any time during the year. This increases the rate of working children in our sample to 8.83 percent. Re-estimating our model using this definition of a working child only has a minor impact on the results (Table A6, appendix).

The previous results suggested that diversification into non-agricultural work might have the largest impact on the probability of a child working. To further test this hypothesis, we run our model without the income diversification strategy dummies, instead using three dummy variables capturing whether any adult in the household participated in non-agricultural wage labor, agricultural wage labor or a non-agricultural business (Table A7, appendix). These results show that, after the controls are added, only the coefficients for non-agricultural wage labor and business are statistically significant, confirming our expectations that these are the forms of income diversification that are most important for the rate of working children.

6.2 Hours worked

In Section 6.1 we investigated how income diversification influences the probability to work and be in school. Now we turn to the intensive margin and analyze how income diversification affects the number of hours worked. The dependent variable is the number of hours spent working during the last seven days.[11]

As discussed earlier, we expect that household expenditure is endogenous in the work equation. However, in the Tobit regression we cannot reject the hypothesis that household expenditure is exogenous. Therefore, we do not use the instrumental approach but instead include household expenditure directly in the model.

Table 7 shows that children from households that fully specialize in production from their own farms work about 22 more hours per week than children in diversified households. Including the controls makes the magnitude decrease and shows that children who live in undiversified households work about 17 hours more per week than children in diversified households.

[11] When cleaning the data we exclude individuals that stated that they had worked more than 80 hours in a week.

Comparing different income diversification strategies does not show any clear pattern. However, children living in households that combine farming with non-agricultural wage labor (and other income-generating activities) seem to work fewer hours than other children.

Table 7: Tobit regression, hours worked, censored at 0.

	Hours worked							
	779 uncensored observations							
Diversified	-21.889	***	-17.266	***				
	(5.387)		(3.689)					
FA					-18.132	**	-16.685	**
					(7.795)		(6.763)	
FN					-22.740	***	-19.013	***
					(7.445)		(6.218)	
FB					-19.245	***	-12.895	***
					(6.779)		(4.883)	
FAB					-16.426		-17.126	
					(15.187)		(13.259)	
FNA					-27.064	**	-26.032	*
					(13.703)		(13.208)	
FNB					-41.458	***	-34.663	***
					(14.324)		(12.779)	
Expenditure			-0.198				-0.030	
			(1.238)				(1.216)	
Controls	No		Yes		No		Yes	

Notes: *** 1% significance level, ** 5% significance level, * 10% significance level. Standard errors clustered at the district level. FA= Farmer and Agricultural wage worker, FN= Farmer and Non-agricultural wage worker, FB= Farmer and family Business, FAB = Farmer, Agricultural wage worker and family Business, FNA = Farmer, Non-agricultural wage worker and Agricultural wage worker, FNB= Farmer, Non-agricultural wage worker and family Business. The comparison group is Full-time farmers (F).

As a robustness test we exclude the children who stated that they worked more than 60 hours a week (Table A8, appendix). We find that this does not influence the results when we compare diversified to non-diversified households. When we compare different income diversification strategies it does affect the magnitude of the effect for some coefficients. However, the overall conclusions remain the same.

7. Conclusion

This paper seeks to estimate the relationship between household income diversification strategies and the rate of working children and schooling. We find that the increase in the number of households that diversify their income by working outside of their own farms seems to decrease the proportion of working children. Focusing on children that work as their main activity, we find that children from households that are engaged in full-time farming are about 3 percentage points more likely to work as their main activity and about 2 percentage points less likely to be in school than children from diversified households. Children from undiversified households have both a larger probability of working, and work longer hours than children from diversified households.

However, since we focus solely on children who work as their main activity, more research is needed to understand how income diversification influences the child's total time allocation. Furthermore, to be able to better understand the specific mechanism by which income diversification and the child's probability to work is related, we would need to use panel data. This could, for example, help us to analyze the potential income effect, by determining whether income increases when households diversify or whether households that diversify are already better off.

Taken together with previous literature that highlights diversification as an important way to increase income, our results support the conclusion that diversification is good for the household, and for the children. Therefore, a recommended policy would be to increase the farming households' access to the labor market.

References

Barrett C. B., Bezuneh M., Clay D. C., and Reardon T. (2005). Heterogenous Constraints, Incentives and Income Diversification Strategies in Rural Africa. Quarterly Journal of International Agriculture 44(1), pp 37-60.

Basu, K., Das, S. and Dutta, B. (2010). Child labor and household wealth: Theory and empirical evidence of an inverted-U. Journal of Development Economics, 91, 8-14.

Basu, K. and Tzannatos, Z. (2003). The Global Child Labor Problem: What Do We Know and What Can We Do? World Bank Economic Review, 17(2), 147-173.

Basu, K. and Van, P. H. (1998). The Economics of Child Labor. American Economic Review, 88, 412-427.

Becker, G. and Tomes, N. (1976). Child Endowments and the Quantity and Quality of Children. The Journal of Political Economy, 84(4), 143-162.

Beegle, K., Dehejia, R. and Gatti, R. (2006). Child labor and agricultural shocks. Journal of Development Economics, 81, 80-96.

Bhalotra, S. and Heady, C. (2003). Child Farm Labor: The Wealth Paradox. World Bank Economic Review, 17(2), 197-227.

Bhalotra, S. and Tzannatos, Z. (2003). Child Labor: What Have We Learnt? Social Protection Discussion Paper No.317, Human Development Department, World Bank.

Bigsten A. and Tengstam S. (2011). Smallholder Diversification and Income Growth in Zambia. Journal of African Economies, 20(5), 781-822.

Buchmann, C. (2000). Family Structure, Parental Perceptions, and Child Labor in Kenya: What Factors Determine Who is Enrolled in School? Social Forces, 78(4), 1349-1379.

Canagarajah, S. and Coulombe, H. (1997). Child labor and schooling in Ghana. World Bank Working Paper No.1844.

Childrens Act (2007). Laws of Kenya. Revised Edition 2010. Published by the National Council for Law Reporting.

Congdon Fors, H. (2012). Child Labour: A review of recent theory and evidence with policy implications. Journal of Economic Surveys, 26(4), 570-593.

Deaton, A. (2000). The Analysis of Household Surveys: A Microeconometric Approach to Development Policy. World Bank. The Johns Hopkins University Press, Baltimore and London

Diallo, Y., Hagemann, F., Etienne, A., Gurbuzer, Y. and Mehran, F. (2010). Global child labor developments: Measuring trends from 2004 to 2008. International Labour Organization.

Dillon, A., Bardasi, E., Beegle, K. and Serneels, P. (2012). Explaning variation in child labor statisitcs. Journal of Development Economics 98, pp 136-147.

Edmonds (2007). Child labor. Handbook of Development Economics, 4, 3607-3709.

Edmonds, E. (2005). Does Child Labor Decline with Improving Economic Status? The Journal of Human Resources, 40(1), 77-99.

Edmonds, E. (2006). Child labor and schooling responses to anticipated income in South Africa. Journal of Development Economics, 81(2), 386-414.

Edmonds, E. and Pavcnik, N. (2005). Child Labor in the Global Economy. The Journal of Economic Perspectives, 19(1), 199-220.

Emerson, P. and Souza, A. (2003). Is There a Child Labor Trap? Inter-Generational Persistence of Child Labor in Brazil. Economic Development and Cultural Change, 51(2), 375-398.

Emerson, P. and Souza, A. (2011). Is Child Labor Harmful? The Impact of Working Earlier in Life on Adult Earnings. Economic Development and Cultural Change, 59(2), 345-385.

Employment Act (2007). Laws of Kenya. Revised Edition 2012. Published by the National Council for Law Reporting with the Authority of the Attorney-General.

Freeman H. A., Ellis, F. and Allison, E. (2004). Livelihoods and Rural Poverty Reduction in Kenya. Development Policy Review, 22(2), 147-171.

Greene, W. (2003). Econometric Analysis,5th Ed. Prentice Hall, New York.

Hazarika, G. and Sarangi, S. (2008). Household Access to Microcredit and Child Work in Rural Malawi. World Development, 36(5), 843-859.

Hoang, T. X., Pham, C. S. and Ulubasoglu M. A. (2014). Non-Farm Activity, Household Expenditure, and Poverty Reduction in Rural Vietnam: 2002-2008. World Development, 64, 554-568.

ILO (2013). Global child labour trends 2008 to 2012. International Labour Office, International Programme on the Elimination of Child Labour (IPEC), Geneva.

Kristjanson, P., Mango, N., Krishna, A., Radeny, M. and Johnson, N. (2010). Understanding Poverty Dynamics in Kenya. Journal of International Development, 22, 978-996.

Moyi, P. (2011). Child Labor and School Attendance in Kenya. Educational Research and Reviews, 6(1), 26-35.

Patrinos, H. A. and Psacharopoulos, G. (1997). Family Size, Schooling and Child Labor in Peru: An empirical analysis. Journal of Population Economics, 10, 387-405.

Roodman, D. (2011). Fitting fully observed recursive mixed-process models with cmp. The Stata Journal, 11(2), 159-206.

Shimamura, Y. and Lastarria-Cornhiel, S. (2010). Credti Program Participation and Child Schooling in Rural Malawi. World Development, 38(4), 567-580.

Skyt-Nielsen, H. (1998). Child Labor and School Attendance: Two Joint Decisions. Working Paper 15, Centre for Labor Market and Social Research, Denmark.

Wydick, B. (1999). The effect of microenterprise lending on child schooling in Guatemala. Economic Development and Cultural Change, 47(4), 853-869.

Table A1: Variables and definitions

Variable	Definition
Main variables	
F	The household (individuals aged 16 and over)gets all of its income from the production of its own farm
FA	The household (individuals aged 16 and over) gets income from the production of its own farm and agricultural wage labor
FN	The household (individuals aged 16 and over) gets income from the production of its own farm and non-agricultural wage labor
FB	The household (individuals aged 16 and over) gets income from the production of its own farm and non-agricultural self-employment
FAB	The household (individuals aged 16 and over) gets income from the production of its own farm, agricultural wage labor and non-agricultural self-employment
FNA	The household (individuals aged 16 and over) gets income from the production of its own farm, non-agricultural wage labor and agricultural wage labor
FNB	The household (individuals aged 16 and over) gets income from the production of its own farm, non-agricultural wage labor and non-agricultural self-employment
Child	
Age	Age in years
Age squared	Age squared
Boy	Female = 0, Male = 1
Biological child	Of head of household =1, otherwise = 0
Household	
Plot size	Land area in acres of parcel used for farming
Plot size2	Land area in acres of parcel used for farming, squared
Land	Household owns land
Siblings 0-5	Child has at least one sibling aged 0-5
Siblings 6-14	Child has at least one sibling aged 6-14
Siblings 15-18	Child has at least one sibling aged 15-18
Adult	Number of adults (aged 16 and over) in the household
Education of head of household (dummies)	None (ref) Some primary Some secondary or above
Ethnolinguistic background (dummie)	Embu Kalenjin

	Kamba
	Kisii
	Kikuyu (ref)
	Luhya
	Luo
	Maasai
	Meru
	Mijikenda
	Somali
	Swahili
	English
Expenditure	
Instrument:	
Toilet	Household has toilet facility located within the dwelling =1, otherwise = 0
Electricity	Household has electricity=1, otherwise=0
Cement floor	The floor in the main dwelling is made out of cement=1, otherwise=0
Geographic	
Distance to school	Percentage of households that are more than 5 km from the nearest public primary school. Calculated by district.
Cost of education	Mean education cost in Kenyan Shillings for students in government primary schools, divided by 1000. Calculated by cluster.
Quality of education	Quality of school measured as the percentage of students in the second class that can write, calculated by district.
Province (dummies)	Central (ref)
	Coast
	Eastern
	Northeastern
	Nyanza
	Rift Valley
	Western

Table A2: Descriptive statistics

	All n = 10847	F n = 6081	FA n = 805	FN n = 1371	FB n =1847	FAB n =142	FNA n =145	FNB n =456
Age	9.88	9.89	9.86	9.91	9.89	9.75	9.82	9.77
	(2.59)	(2.59)	(2.59)	(2.56)	(2.58)	(2.64)	(2.59)	(2.68)
Male	0.50	0.51	0.49	0.52	0.49	0.51	0.44	0.49
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Bilogical child	0.80	0.79	0.82	0.83	0.80	0.77	0.75	0.80
	(0.40)	(0.41)	(0.39)	(0.38)	(0.40)	(0.42)	(0.43)	(0.40)
Not living with parents	0.15	0.17	0.12	0.11	0.14	0.13	0.10	0.13
	(0.36)	(0.37)	(0.32)	(0.31)	(0.34)	(0.34)	(0.31)	(0.34)
Embu	0.02	0.02	0.05	0.03	0.01	0.02	0.01	0.01
	(0.14)	(0.14)	(0.21)	(0.16)	(0.09)	(0.14)	(0.12)	(0.11)
Kalenjin	0.15	0.16	0.18	0.17	0.09	0.13	0.19	0.19
	(0.36)	(0.36)	(0.38)	(0.37)	(0.28)	(0.34)	(0.40)	(0.39)
Kamba	0.09	0.07	0.06	0.11	0.13	0.12	0.13	0.15
	(0.29)	(0.26)	(0.24)	(0.31)	(0.33)	(0.33)	(0.34)	(0.36)
Kikuyu	0.09	0.09	0.14	0.09	0.07	0.11	0.21	0.05
	(0.29)	(0.28)	(0.35)	(0.29)	(0.25)	(0.32)	(0.41)	(0.21)
Kisii	0.04	0.04	0.04	0.02	0.03	0.01	0.03	0.03
	(0.19)	(0.20)	(0.20)	(0.13)	(0.18)	(0.08)	(0.18)	(0.17)
Luhya	0.04	0.03	0.04	0.02	0.05	0.04	0.07	0.06
	(0.19)	(0.17)	(0.20)	(0.14)	(0.22)	(0.20)	(0.25)	(0.23)
Luo	0.10	0.07	0.10	0.10	0.19	0.13	0.10	0.10
	(0.30)	(0.26)	(0.30)	(0.29)	(0.39)	(0.33)	(0.30)	(0.30)
Maasai	0.04	0.05	0.02	0.03	0.04	0.06	0.00	0.01
	(0.20)	(0.22)	(0.16)	(0.17)	(0.20)	(0.24)	(0.00)	(0.08)
Meru	0.06	0.07	0.08	0.06	0.02	0.02	0.06	0.03
	(0.23)	(0.25)	(0.26)	(0.24)	(0.15)	(0.14)	(0.23)	(0.17)
Mijikenda	0.03	0.02	0.00	0.03	0.06	0.06	0.00	0.03
	(0.16)	(0.13)	(0.04)	(0.18)	(0.23)	(0.23)	(0.00)	(0.16)
Somali	0.04	0.06	0.00	0.01	0.01	0.00	0.00	0.00

	(0.19)	(0.24)	(0.06)	(0.11)	(0.11)	(0.00)	(0.00)	(0.00)
Swahili	0.29	0.29	0.27	0.30	0.28	0.30	0.18	0.31
	(0.45)	(0.45)	(0.44)	(0.46)	(0.45)	(0.46)	(0.38)	(0.46)
English	0.03	0.03	0.02	0.04	0.03	0.00	0.02	0.04
	(0.16)	(0.16)	(0.13)	(0.18)	(0.17)	(0.00)	(0.14)	(0.20)
Age of head	46.80	47.68	46.43	45.10	45.62	47.45	49.50	44.11
	(13.04)	(13.76)	(12.55)	(11.27)	(12.72)	(11.85)	(10.99)	(9.97)
Head no education	0.32	0.39	0.28	0.20	0.26	0.26	0.19	0.13
	(0.47)	(0.49)	(0.45)	(0.40)	(0.44)	(0.44)	(0.40)	(0.33)
Head primary education	0.47	0.46	0.54	0.42	0.51	0.50	0.47	0.48
	(0.50)	(0.50)	(0.50)	(0.49)	(0.50)	(0.50)	(0.50)	(0.50)
Head some secondary education	0.19	0.13	0.17	0.37	0.21	0.21	0.34	0.38
	(0.40)	(0.34)	(0.38)	(0.48)	(0.41)	(0.41)	(0.47)	(0.49)
Owns no land	0.30	0.33	0.31	0.30	0.21	0.19	0.26	0.23
	(0.46)	(0.47)	(0.46)	(0.46)	(0.41)	(0.39)	(0.44)	(0.42)
Size of plot	2.52	2.39	2.17	2.74	2.69	2.93	2.74	3.13
	(5.04)	(5.81)	(3.77)	(4.09)	(3.90)	(4.17)	(2.96)	(3.06)
Has siblings aged 0-5	0.22	0.20	0.25	0.24	0.24	0.24	0.15	0.24
	(0.41)	(0.40)	(0.43)	(0.42)	(0.43)	(0.43)	(0.36)	(0.43)
Has siblings aged 6-14	0.84	0.83	0.82	0.83	0.85	0.83	0.88	0.86
	(0.37)	(0.37)	(0.38)	(0.37)	(0.36)	(0.38)	(0.33)	(0.34)
Has siblings aged 15-18	0.54	0.52	0.53	0.56	0.51	0.54	0.66	0.65
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.48)	(0.48)
Number of adults	2.57	2.39	2.71	2.99	2.43	2.78	3.79	3.21
	(1.31)	(1.22)	(1.40)	(1.44)	(1.10)	(1.07)	(1.54)	(1.36)
Distance to school	56.54	56.19	55.66	55.58	59.93	61.22	53.61	51.35
	(28.22)	(28.53)	(28.61)	(29.08)	(26.02)	(27.50)	(27.58)	(28.26)
Cost of education	0.57	0.55	0.53	0.65	0.56	0.64	0.70	0.65
	(0.59)	(0.53)	(0.47)	(0.90)	(0.52)	(0.55)	(0.79)	(0.52)
Quality of education	0.70	0.69	0.69	0.72	0.72	0.72	0.73	0.74
	(0.22)	(0.22)	(0.23)	(0.23)	(0.22)	(0.19)	(0.20)	(0.21)

Coast	0.10	0.09	0.04	0.11	0.14	0.07	0.00	0.12
	(0.30)	(0.29)	(0.19)	(0.31)	(0.35)	(0.26)	(0.00)	(0.32)
Eastern	0.22	0.21	0.24	0.24	0.20	0.20	0.23	0.24
	(0.41)	(0.41)	(0.42)	(0.43)	(0.40)	(0.40)	(0.43)	(0.43)
Northeastern	0.04	0.06	0.00	0.01	0.01	0.00	0.00	0.00
	(0.19)	(0.24)	(0.06)	(0.11)	(0.11)	(0.00)	(0.00)	(0.00)
Nyanza	0.15	0.13	0.15	0.13	0.23	0.13	0.15	0.15
	(0.36)	(0.34)	(0.36)	(0.33)	(0.42)	(0.34)	(0.36)	(0.36)
Rift Valley	0.29	0.31	0.33	0.30	0.19	0.31	0.32	0.26
	(0.45)	(0.46)	(0.47)	(0.46)	(0.40)	(0.46)	(0.47)	(0.44)
Western	0.13	0.11	0.13	0.12	0.16	0.17	0.11	0.19
	(0.33)	(0.31)	(0.33)	(0.33)	(0.37)	(0.38)	(0.31)	(0.39)

F= Full-time farmer, FA= Farmer and Agricultural wage worker, FN= Farmer and Non-agricultural wage worker, FB= Farmer and family Business, FAB = Farmer, Agricultural wage worker and family Business, FNA = Farmer, Non-agricultural wage worker and Agricultural wage worker, FNB= Farmer, Non-agricultural wage worker and family Business

Table A3: Marginal effects of income diversification strategy estimated at the mean. Three-equation mixed-process model, with controls. With and without instrument for expenditure

	Work		Schooling		Work		Schooling	
	With instrument				Without instruments			
Diversified	-0.029	***	0.017	*	-0.034	***	0.025	***
	(0.008)		(0.008)		(0.009)		(0.008)	
Expenditure	-0.028	**	0.056	***	-0.0001		0.017	***
	(0.012)		(0.017)		(0.002)		(0.005)	

Notes: *** 1% significance level, ** 5% significance level, * 10% significance level. Standard errors clustered at the district level.

Table A4: Marginal effects estimated at the mean.

Three-equation mixed-process model. All control variables

	Work		Schooling	
Diversified	-0.029	***	0.017	*
	(0.009)		(0.009)	
Age	0.026	***	0.094	***
	(0.009)		(0.010)	
Age^2	-0.001	*	-0.004	***
	(0.0004)		(0.0004)	
Male	0.022	***	0.010	
	(0.006)		(0.007)	
Biological	0.039	**	-0.026	*
	(0.017)		(0.014)	
Not living with parents	0.032	*	-0.032	**
	(0.017)		(0.014)	
Age of head	0.000		0.000	
	(0.00)		(0.000)	
Head primary education	-0.027	*	0.060	***
	(0.015)		(0.017)	
Head some secondary education	-0.013		0.073	***
	(0.017)		(0.021)	
Expenditure	-0.028	**	0.056	***
	(0.012)		(0.017)	
Embu	0.029		-0.001	
	(0.040)		(0.051)	
Kalenjin	0.024		-0.032	
	(0.038)		(0.029)	
Kamba	0.030		0.039	
	(0.036)		(0.034)	
Kisii	0.015		-0.052	
	(0.046)		(0.043)	
Luhya	0.051		-0.040	
	(0.038)		(0.026)	
Luo	-0.065		-0.016	
	(0.053)		(0.028)	
Maasai	0.052		-0.115	***
	(0.047)		(0.031)	
Meru	-0.069	*	-0.017	
	(0.039)		(0.037)	
Mijikenda	0.063		-0.102	***
	(0.063)		(0.029)	
Somali	-0.111	**	-0.106	***
	(0.044)		(0.028)	
Swahili	0.041		-0.062	***
	(0.064)		(0.020)	
English	-0.000		-0.010	***
	(0.033)		(0.023)	
Owns land	-0.009		0.059	***
	(0.015)		(0.011)	

Plot size, acres	0.002		0.0009
	(0.003)		(0.002)
Plot size acres^2	-0.000		-0.000
	(0.000)		(0.000)
Sibling 0-5	-0.008		0.009
	(0.008)		(0.007)
Siblings 6-14	-0.010		0.011
	(0.009)		(0.010)
Siblings 15-18	-0.007		0.010
	(0.008)		(0.008)
Adults	-0.0003		-0.0003
	(0.003)		(0.003)
Distance to school	-0.0006	**	-0.0002
	(0.0003)		(0.0002)
Cost of education	0.026	**	-0.008
	(0.010)		(0.010)
Quality of education	-0.073	*	0.057
	(0.038)		(0.110)
Province dummies	Yes		Yes

Notes: *** 1% significance level, ** 5% significance level, * 10% significance level. Standard errors clustered at the district level. Reference group for ethnolinguistic background is Kikuyu. Reference group for educational level is no education.

Table A5: Marginal effects of income diversification strategy estimated at the mean. Three-equation mixed-process model with controls. Different definition of sample.

	Age 7-14, n= 9536			Age 7-14, n = 9536		
	Work		School	Work		School	
Diversified	-0.030	***	0.013				
	(0.009)		(0.008)				
FA				-0.033	**	0.026	**
				(0.017)		(0.013)	
FN				-0.036	**	0.011	
				(0.016)		(0.013)	
FB				-0.022	*	0.010	
				(0.013)		(0.012)	
FAB				-0.023		-0.015	
				(0.034)		(0.032)	
FNA				-0.040		0.030	
				(0.031)		(0.32)	
FNB				-0.061	**	0.021	
				(0.030)		(0.022)	
Controls	Yes			Yes			
ρ_{13}	0.286	**		0.273	**		
P_{23}	-0.301	**		-0.303	**		
P_{12}	-0.804	***		-0.801	***		
Log-pseudolikelihood	-19484			-19451			

	Age 10-14, n=5831			Age 10-14, n=5831		
	Work		School	Work		School	
Diversified	-0.041	***	0.014				
	(0.012)		(0.009)				
FA				-0.042	*	0.033	**
				(0.022)		(0.014)	
FN				-0.049	**	0.001	
				(0.020)		(0.013)	
FB				-0.029	*	0.014	
				(0.015)		(0.012)	
FAB				-0.011		-0.015	
				(0.045)		(0.035)	
FNA				-0.039		0.027	
				(0.035)		(0.033)	
FNB				-0.129	***	0.036	
				(0.037)		(0.028)	
Controls	Yes			Yes			
ρ_{13}	0.284	**		0.268	**		
P_{23}	-0.118			-0.128			
P_{12}	-0.856	***		-0.864	***		
Log-pseudolikelihood	-12115			-12088			

Notes: *** 1% significance level, ** 5% significance level, * 10% significance level. Standard errors clustered at the district level. F= Full-time farmer, FA= Farmer and Agricultural wage worker, FN= Farmer and Non-agricultural wage worker, FB= Farmer and family Business, FAB = Farmer, Agricultural wage worker and family Business, FNA = Farmer, Non-agricultural wage worker and Agricultural wage worker, FNB= Farmer, Non-agricultural wage worker and family Business

Table A6: Marginal effects of income diversification strategy estimated at the mean. Three-equation mixed-process model with controls. Different definition of work.

	Work as main activity n= 10847				Worked at any time during year n=10847			
Diversified	-0.030	***			-0.031	*		
	(0.010)				(0.009)			
FA			-0.035	**			-0.039	**
			(0.017)				(0.016)	
FN			-0.028	**			-0.035	**
			(0.014)				(0.014)	
FB			-0.026	**			-0.021	*
			(0.013)				(0.011)	
FAB			-0.043				-0.031	
			(0.039)				(0.031)	
FNA			-0.027				-0.042	
			(0.039)				(0.029)	
FNB			-0.042				-0.049	*
			(0.025)				(0.028)	
Controls	Yes		Yes		Yes		Yes	
ρ_{13}	0.280	***	0.270	**	0.270	***	0.266	***
P_{23}	-0.336	***	-0.335	***	-0.334	***	-0.334	***
P_{12}	-0.755	***	-0.752	***	-0.754	***	-0.734	***
Log-pseudo likelihood	-22208		-22172				-22606	

Notes: *** 1% significance level, ** 5% significance level, * 10% significance level. Standard errors clustered at the district level. FA= Farmer and Agricultural wage worker, FN= Farmer and Non-agricultural wage worker, FB= Farmer and family Business, FAB = Farmer, Agricultural wage worker and family Business, FNA = Farmer, Non-agricultural wage worker and Agricultural wage worker, FNB= Farmer, Non-agricultural wage worker and family Business. The comparison group is Full-time farmers (F).

Table A7: Participating in different income-generating activities.
Marginal effects estimated at the mean. Three-equation mixed-process model, with controls. N =10847

	Work		Schooling		Work		Schooling	
Agricultural wage	-0.026	**	0.053	***	-0.028		0.022	*
	(0.013)		(0.020)		(0.018)		(0.013)	
Non-agricultural wage	-0.043	***	0.081	***	-0.053	***	0.003	
	(0.014)		(0.020)		(0.017)		(0.012)	
Business	-0.033	**	0.066	***	-0.031	**	0.013	
	(0.013)		(0.021)		(0.012)		(0.012)	
Controls	Yes				Yes			
ρ_{13}					0.267	**		
P_{23}					-0.124			
P_{12}	-0496	***			-0.865	***		
Log-likelihood	-5886				-12095			
Controls	No		No		Yes		Yes	

Notes: *** 1% significance level, ** 5% significance level, * 10% significance level. Standard errors clustered at the district level. FA= Farmer and Agricultural wage worker, FN= Farmer and Non-agricultural wage worker, FB= Farmer and family Business, FAB = Farmer, Agricultural wage worker and family Business, FNA = Farmer, Non-agricultural wage worker and Agricultural wage worker, FNB= Farmer, Non-agricultural wage worker and family Business. The comparison group is Full-time farmers (F).

Table A8: Tobit regression, hours worked, censored at 0. Check of robustness to excluding observations of children reportedly working more than 60 hours per week.

	Upper limit 80 hours n = 10824	Upper limit 60 hours n = 10755	Upper limit 80 hours n = 10824	Upper limit 60 hours n = 10755
Diversified	-17.266 ***	-17.220 ***		
	(3.689)	(3.716)		
FA			-16.685 **	-18.375 ***
			(6.763)	(6.721)
FN			-19.013 ***	-17.856 ***
			(6.218)	(5.786)
FB			-12.895 ***	-13.512 ***
			(4.883)	(4.931)
FAB			-17.126	-21.320
			(13.259)	(15.806)
FNA			-26.032 **	-21.147 *
			(13.208)	(11.908)
FNB			-34.663 ***	-29.360 ***
			(12.779)	(11.263)
Controls	Yes	Yes	Yes	Yes

Notes: *** 1% significance level, ** 5% significance level, * 10% significance level. Standard errors clustered at the district level. FA= Farmer and Agricultural wage worker, FN= Farmer and Non-agricultural wage worker, FB= Farmer and family Business, FAB = Farmer, Agricultural wage worker and family Business, FNA = Farmer, Non-agricultural wage worker and Agricultural wage worker, FNB= Farmer, Non-agricultural wage worker and family Business. The comparison group is Full-time farmers (F).