Impact Assessment of Agricultural Commercialization on Food Security Among Smallholder Farmers in Kenya: An Application of Correlated Random Effects

G. Muricho[1]; D. Kulundu[2]; F. Sule[2]

1: ICRISAT, Nairobi Kenya, , Kenya, 2: University of Naiobi, School of Economics, Kenya

Corresponding author email: G.Muricho@cgiar.org

Abstract:

Welfare implication of agricultural commercialization in developing countries is not clear. Particularly not clear is its effect on household food security. Using panel data collected from smallholder farmers in Kenya, we analyze the impact of agricultural commercialization on household food security by fitting endogenous switching regression model in a correlated random effects framework. The results show that agricultural commercialization significantly improves household food security. Food security probability of commercialized and non-commercialized households was 62% and 32%, respectively. This 30% food security gap between the two groups of households could be reduced by 39 percentage points (12% gap) if non-commercialized households could be as efficient as commercialized household in their resource use. The other 61% percentage points (18% gap) emanated from differences in resource amounts between commercialized and non-commercialized households. The implication of these findings is that policies that stimulate and enhance agricultural commercialization are critical in improving household food security.

Acknowledegment: This study was supported by the Australian Centre for International Agricultural Research (ACIAR) and Australian International Food Security Research Centre (AIFSRC) through International Maize and Wheat Improvement Center (CIMMYT) led Sustainable Intensification of Maize-Legume Cropping Systems in Eastern and Southern Africa (SIMLESA) program, and Adoption Pathways Project. The views expressed here are those of authors and do not necessarily reflect the views of the donor or authors' institution. The usual disclaimer applies

JEL Codes: Q13, Q01

#1670

Impact Assessment of Agricultural Commercialization on Food Security among Smallholder Farmers in Kenya: An Application of Correlated Random Effects

Abstract

Welfare implication of agricultural commercialization in developing countries is not clear. Particularly not clear is its effect on household food security. Using panel data collected from smallholder farmers in Kenya, we analyze the impact of agricultural commercialization on household food security by fitting endogenous switching regression model in a correlated random effects framework. The results show that agricultural commercialization significantly improves household food security. Food security probability of commercialized and non-commercialized households was 62% and 32%, respectively. This 30% food security gap between the two groups of households could be reduced by 39 percentage points (12% gap) if non-commercialized households could be as efficient as commercialized household in their resource use. The other 61% percentage points (18% gap) emanated from differences in resource amounts between commercialized and non-commercialized households. The implication of these findings is that policies that stimulate and enhance agricultural commercialization are critical in improving household food security.

Key Words: Smallholder; Agricultural Commercialization; Food Security; Kenya; Switching Regression; Correlated Random Effects

1. Introduction

Global economic and demographic changes in terms of rising incomes and increasing urbanization, respectively, have provided smallholder farmers in developing countries with immense opportunity to commercialize their agricultural activities. This opportunity has further been fueled by recent increasing demand for non-staple western dietary habits of the growing urban population, rising household incomes, foreign investment in food markets, emergence of supermarkets and vertical integration of agricultural production and retail activities (Sharma *et al.,* 2012). Considering these changes, many sub-Saharan Africa (SSA) countries like Kenya have identified agricultural commercialization as one of the crucial pillars of their economic growth and development agenda (Republic of Kenya, 2004; Republic of Kenya, 2007; Republic of Kenya, 2010). This commercialization process is expected to reduce food insecurity and poverty among rural smallholder farmers who mainly depend on agriculture to earn their livelihood. Food security is of importance to rural households because existing empirical evidence shows that about 47% and 50% of the national and rural populations in Kenya, respectively, are food insecure compared to 40% of the urban population (World Bank, 2009). These disappointing statistics have been attributed to rapid population growth and declining agricultural productivity. To address the declining agricultural productivity problem, agricultural commercialization has been promoted for many years in the country. However, to date, empirical evidence to show that agricultural commercialization process can indeed reduce or eradicate rural food insecurity is minimal despite doubts cast on the transformation process earlier on (Republic of Kenya, 1981).

Conceptually, markets provide increased incomes to participating households who in turn can enhance their overall consumption than it could have been under subsistence orientation (Pingali, 1997; Timmer 1997). However, agricultural commercialization might also compromise food security as farmers divert most of their production resources toward pure cash cropping at the expense of food crops (Strasberg *et al.,* 1999). These two competing schools of thought on smallholder agricultural commercialization invited several empirical studies in developing countries (von Braun 1995; Govereh *et al.,* 1999; Strasberg *et al.,* 1999; Govereh and Jayne, 2003).

However, these past studies were either based on a single crop or just a few selected crops. This kind of analysis is bound to give partial and inconclusive information on the impact of agricultural commercialization compared to a study that is based on a comprehensive commercialization index constructed from all crop enterprises on the farm. Again, most of the studies were based on cross sectional data that is limited in generating information on intra and inter household differences as it could have been with panel data. Methodologically, these past studies were also based on pooled regression models to analyze the impact of agricultural commercialization. Pooled regressions assume that sampled households are similar in all aspects except that some are commercialized, and others are not – with the difference in the outcome variable being solely attributed to the treatment effect. This latter assumption might not be right if the commercialized and non-commercialized households have systematic differences that could be correlated with the outcome variable.

Also, in the literature, there is no consensus on the definition of agricultural commercialization (von Braun *et al.*, 1994; Pingali and Rosegrant 1995; Pingali 1997; Gebremedhin and Jaleta 2010). In this study we adopt the definition by Gebremedhin and Jaleta, (2010), that is, produce offered for sale and use of purchased inputs in the production process. However, the later component of this definition (use of purchased inputs) is beyond the scope of this study due to data limitations. Based on this adopted definition, a comprehensive household commercialization index (HCI) that incorporates all crop enterprises on the farm is developed and used in this study. On the other hand, there is also no consensus on the definition of "welfare". However, according to the World Bank (2000), there are three aspects of welfare i.e. poverty, inequality and vulnerability. The current study focuses on poverty which is defined as whether households or individuals have enough resources or abilities to meet their needs (Coudouel *et al.,* 2002; Haughton and Khandker, 2009). Although this poverty concept has numerous measures, this study focuses on food security whose widely used working definition provided by FAO (2003) states that "food security exists when people at all times have physical, social and economic access to sufficient, safe and nutritious food which meets their dietary needs and food preferences for an active and healthy life". To capture this subjective food security aspect of "preferred food" in the definition, a subjective assessment of household food security following Mallick and Rafi (2010) and Kassie et al., (2014a) is adopted in this study.

Therefore, this paper contributes to the existing literature of agricultural commercialization by empirically analyzing the impact of agricultural commercialization on household food security using a more comprehensive HCI based on all crop enterprises found on the farm. Also, unlike past studies that used descriptive statistics or pooled regressions, the current study uses a more robust analytical framework by fitting a switching regression model on panel data in correlated random effects (CRE) framework. This CRE approach computes fixed effects estimators while at the same time allowing time invariant variables to be used as explanatory variables (Wooldridge, 2010). The rest of the paper is organized as follows: - Section 2 presents the methods and data used in this paper. The results are presented in section 3 starting with descriptive statistics and then the econometric regression before delving into the treatment effects of agricultural commercialization on household food security. In section 4, we present the summary, conclusions and policy implications.

2. Methods and data

Estimating the impact of agricultural commercialization on household food security outcome using non-experimental data is challenging because of lack of the counterfactual i.e. it is not possible to observe the food security outcome of a household that commercialized had it not commercialized and the vice versa. In experimental studies, this problem can be addressed by randomly enabling the treated households to be commercialized and the untreated group of households to be in the control status (non-commercialized). However, commercialization among the sampled households in this current study is not a randomly assigned treatment because households select themselves into commercialized and non-commercialized groups. This self-selection means that there could be systematic differences between these two groups of households. Therefore, evaluating the impact of agricultural commercialization on household food security by estimating a single outcome equation with a dummy variable of commercialization as one of the explanatory variable in a pooled regression might yield biased estimates. There could be an interaction between commercialization decision dummy variable and other explanatory variables (covariates).

Therefore, if explanatory variables have different effects on household food security outcome of commercialized and non-commercialized households, then separate food security outcome functions for each group of households need to be specified. This approach accounts for endogeneity because commercialization decision is potentially endogenous to the household. An endogenous switching regression (ESR) model that accounts for both endogeneity and sample selection becomes the best option to handle these estimation problems. Also, ERS allows interactions between commercialization decision and other explanatory variables in outcome function. The ESR model captures such interactions by estimating two separate equations (one for commercialized and another for non-commercialized households) along with the selection equation.

2.1 Theoretical model

Following de Janvry *et al.,* (2011), household agricultural commercialization can be modeled in a random utility framework. The farm household maximizes its utility subject to both endogenous and exogenous constraints. A farm household is expected to commercialize if commercialization is beneficial compared to otherwise. This means that the difference between utility derived from commercialization (U_{1it}) and non-commercialization (U_{0it}) denoted as G_{it}^{*}, is greater than non-commercialization utility (see Eq. 1).

$$\mathrm{G_{it}^{*}} = \mathrm{U_{1it}} - \mathrm{U_{0it}} > 0 \qquad \mathrm{Eq.\,(1)}$$

Where:

U_{1it} = Utility derived from commercialization: $U_{1it}(X_{it}, D_{it}, G_{it}^{*} = 1)$

U_{0it}= Utility derived from not commercializing in the market: $N_{0it}(X_{it}, D_{it}, G_{it}^{*} = 0)$

G_{it}^{*} = Difference between utility derived from commercialization and that derived from non - commercialization in the market.

Subscript *i* and *t* stands for the i^{th} household at time *t*.

The factors affecting the choice to commercialize or not to commercialize can then be estimated using several variants of selection models in which selection into the treatment (commercialization) is made based on expected utility (Bellemare and Barrett 2006; Alene *et al.*, 2008). The expected utility of commercialization for household *i* at time *t* is determined by two sets of variables, that is, those that are observable to the researcher (X_{it}) and those that are not (D_{it}).

Since these utilities are unobservable, then they can be expressed as a function of observable characteristics (X_{it}) and the error term (η_{it}) in the following latent variable model (see Eq. 2): -

$$T_{it}^{*} = \alpha_{it} X_{it} + \eta_{it}; \quad \text{With } H_{it} = \begin{cases} 1 \text{ if } & T_{it}^{*} > 0 \\ 0 \text{ Otherwise} \end{cases} \qquad \text{Eq. (2)}$$

Where:

T_{it} = Binary indicator variable for agricultural commercialization that equals to 1 if a household is commercialized and 0 if otherwise

α_{it} = Vector of parameters to be estimated

X_{it} = Vector of observable explanatory variables

η_{it} = Error term

2.2 Empirical model

The first-step of the two-step ESR involves modeling household's binary decision to commercialize using the probit model that can structurally be represented as shown in Eq. 3.

$$T_{it} = f(DC_{it}; PF_{it}; SC_{it}; TC_{it}) \qquad \text{Eq. (3)}$$

Where:

Subscript *i* and *t* indexes household and time, respectively

T = Binary indicator of commercialization i.e. 1 if commercialized and 0 if otherwise

DC = Demographic characteristics of the households

PF = Physical and financial endowments of the households

SC = Social capital proxies of the household

TC = Transaction costs variables of the household

The food security outcome functions conditional on commercialization decision is written in an endogenous switching regression regime model as follows: -

Regime 1: $Y_{1it} = \beta_1 X_{1it} + \varepsilon_{1it}$: If $T_{it} = 1$ Eq. (4a)

Regime 2: $Y_{0it} = \beta_0 X_{0it} + \varepsilon_{0it}$: If $T_{it} = 0$ Eq. (4b)

Where:

Y_{1it} = Outcome indicator variables of agricultural commercialization (food security) for commercialized households

Y_{0it} = Outcome indicator variables of agricultural commercialization (food security) for non–commercialized households

X_{1it} = Observed vectors of covariates determining agricultural commercialization outcome i.e. food security probability for commercialized households

X_{0it} = Observed vectors of covariates determining agricultural commercialization outcome i.e. food security probability for non–commercialized households

β_1 and β_0 = Vectors of parameters to be estimated

ε_{1it} and ε_{0it} = Error terms with zero mean and constant variance

Following Kassie *et al.*, (2014) and other impact evaluation literature, for ESR model to be identified, then X_{it} variables in Eq. 2 should contain a selection instrument, that is, variable/s that affect directly the endogenous selection variable (commercialization) but not the outcome variables (food security). In this study, all the transaction costs outlined in Eq. 3 were instrument candidates subject to verification to ascertain their suitability. The choice of transaction costs as instrument variables was informed by a combination of economic theory and findings of past empirical studies (Goetz 1992; Key *et al.*, 2000).

Since Y_{1it} and Y_{0it} are not observed simultaneously, the covariance between ε_{1i} and ε_{0i} is not defined (Maddala, 1983). Important implication of this error structure is that because the error term of the selection model (Eq. 2) is correlated with the error terms of the welfare outcome models (Eq. 4a and Eq. 4b), the expected values of ε_{1i} and ε_{0i} conditional on the sample selection are non-zero.

$$E[\varepsilon_{Pit} \backslash H_{it} = 1] = \delta_{\varepsilon_{P\eta it}} \frac{\phi(\beta X_{it})}{\Phi(\beta X_{it})} = \delta_{\varepsilon_{P\eta it}} \lambda_{Pit} ---------Eq.(5)$$

and:-

$$E[\varepsilon_{Nit} \backslash H_{it} = 0] = -\delta_{\varepsilon_{N\eta it}} \frac{\phi(\beta X_{it})}{1 - \Phi(\beta X_{it})} = \delta_{\varepsilon_{N\eta it}} \lambda_{Nit} ---------Eq.(6)$$

Where:

$\phi\ (.)$ = Standard normal probability density function

$\Phi\ (.)$ = Standard cumulative density function

$\lambda_{Pit} = \frac{\phi(\beta X_{it})}{\Phi(\beta X_{it})}$

$\lambda_{Nit} = -\frac{\phi(\beta X_{it})}{1-\Phi(\beta X_{it})}$

λ_{Pit} and λ_{Nit} are the IMR computed from the selection equation

The computed IMR is included in Eq. 4a and Eq. 4b to correct for selection bias in the two-step estimation procedure i.e. endogenous switching regression. If the estimated covariance $\delta_{\varepsilon_{1\eta it}}$ and $\delta_{\varepsilon_{0\eta it}}$ are statistically significant, then the decision to commercialize and the food security probability outcome variables are correlated, that is, there is evidence of endogenous switching and the null hypothesis of absence of sample selectivity bias will be rejected.

Therefore, Eq. 4a and Eq. 4b are used to estimate the average counterfactual food security probability distribution i.e. what could have been the food security probability outcome of the commercialized households had they not commercialized and the vice versa. Following the wage decomposition literature pioneered by Oaxaca (1973), this analytical framework is also used to decompose the food security probability gap between commercialized and non-commercialized households into the portion that is caused by differences in the amount of resources held by the two groups of households (quantity or level effect) and that component due to differences in the resource use efficiency (efficiency or return effect). The actual expected food security probability outcomes for commercialized and non-commercialized households are computed using Eq. 7a and Eq. 7b, respectively. On the other hand, the counterfactual expected food security outcome probability outcomes are estimated using Eq. 8a and Eq. 8b for commercialized and non-commercialized households, respectively.

Actual scenarios (observed from the sample data):

Commercialized: $E(Y_{1it} \backslash T = 1; X) = \beta_1 X_{1it} + \gamma_{1\varepsilon} \lambda_{1it}$ $Eq.(7a)$

Non-commercialized: $E(Y_{0it} \backslash T = 0; X) = \beta_0 X_{0it} + \gamma_{0\varepsilon} \lambda_{0it}$ $Eq.(7b)$

Counterfactual scenarios:

Commercialized if didn't commercialize: $E(Y_{1it} \backslash T = 0; X) = \beta_1 X_{0it} + \gamma_{1\varepsilon} \lambda_{0it}$ $Eq.(8a)$

Non-commercialized if commercialized: $E(Y_{0it} \backslash T = 1; X) = \beta_0 X_{1it} + \gamma_{0\varepsilon} \lambda_{1it}$ $Eq.(8b)$

Applying these conditional expectations and using agricultural commercialization as the treatment variable, decomposition of the observed food security gap between commercialized and non-commercialized households (Eq. 7a less Eq. 7b) is computed as shown in Table 1. The difference in food security probability outcome of commercialized households emanating from their differences in efficiency of use of their currently held resources compared to the efficiency of non-commercialized households is obtained by subtracting Eq. 8a from Eq. 7a. Similarly, the difference in food security probability outcome of non-commercialized households emanating from their differences in efficiency of use of their currently held resources compared to the efficiency of commercialized households is obtained by subtracting Eq. 7b from Eq. 8a (Table 1). On the other hand, the difference in food security probability outcome of commercialized households as a result of their differences in the amount of resources held compared to the amount of resources held by non-commercialized households, holding efficiency constant, is obtained by subtracting Eq. 7b from Eq. 8a. Finally, the difference in food security probability outcome of non-commercialized households originating from their differences in the amount of resources held by commercialized household, holding their resource use efficiency constant, is obtained by subtracting Eq. 8b from Eq. 7a (Table 1).

Table 1. Conditional expectations, treatment effects and heterogeneity effects

Household type	Market participating households' response to characteristics	Non-market participating households' response to characteristics	Returns effects (difference caused by difference in resource use efficiency)
Commercialized households	(8a) $E(Y_{1i}/T=1)$	(9a) $E(Y_{0i}/T=1)$	(8a) – (9a)
Non-commercialized households	(9b) $E(Y_{1i}/T=0)$	(8b) $E(Y_{0i}/T=0)$	(9b) – (8b)

Level effect (difference caused by differences in resource quantities)	LE_0 = (8a) – (9b)	LE_1 = (9a) – (8b)	(8a) – (8b)

The first-step of the two-step ESR based on Eq. 3 is estimated using a probit model. On the other hand, the second step of the ESR is based on Eq. 4a and Eq. 4b where, like in the first step, the dependent variable in both equations (household food security probability) is binary. The independent variables were grouped into four categories: - i) demographic characteristics; ii) physical and financial asset endowments; iii) social capital; and iv) transaction costs as shown in Table 2. As already mentioned, transaction costs explanatory variables were assessed for their suitability as instrument variables using a simple falsification test following Di Falco *et al.*, (2011) and Kassie *et al.*, (2014). Those transaction costs variables that passed the suitability test were used accordingly i.e. excluded in the second stage of estimating the treatment outcome equations.

2.3 Estimation strategy for the endogenous switching regression (ESR)

Given that this study is using a two wave panel data, a correlated random effects (CRE) approach is employed using the *Mundlak–Chamberlain* Device (Mundlak, 1978; Chamberlain 1982). One major advantage of panel data is its ability to control for time invariant unobserved heterogeneity. Traditional approaches of panel data (random and fixed effects) have some inherent weakness. On one hand, the random effects models assume no correlation between the unobserved heterogeneities and the observed explanatory variables in the model. If this random effect assumption holds, then across-sectional analysis employing OLS estimation would also consistently estimate the model parameters (Wooldridge, 2010). On the other hand, while fixed effects approach looks attractive as it assumes arbitrary correlation between the unobserved heterogeneity and observed explanatory variables, it does not include the time invariant observed explanatory variables. Therefore, CRE approach that preserves the advantages of fixed approach while at the same time enabling the inclusion of time invariant explanatory variables (Wooldridge, 2010) is more attractive and it is adopted in this study.

Structurally, as demonstrated by Wooldridge (2010), the CRE approach allows for the correlation between unobserved heterogeneity (C_i) and the vector of explanatory variables across all time periods (X_{it}). In this framework, the assumption is that there is a linear relationship between the unobserved time varying individual heterogeneity and the observed explanatory variables that can be modeled as follows: -

$$C_i = \varphi + \bar{X}_i\lambda + a_i ------------------------------ Eq.(9)$$

Where: -

φ is a scalar

$\bar{X}$ is the averages of time varying explanatory variables

λ is a vector of coefficients to be estimated

a_i is the error term assumed to have zero mean conditional on the entire history of the covariates (X_{i1}, X_{i2}, ------, X_{iT}) i.e. a_i is uncorellated with X_{it} for all t and therefore X_i

The reduced form of the model in which φ is absorbed into the intercept term and y are added to the set of explanatory variables including time invariant variables is estimated as follows:-

$$Y_{it} = a_t^* + X_{it}\beta + \bar{X}_i\lambda + Z_i\gamma + a_i + \varepsilon_{it} -------------------- Eq.(10)$$

Where:

Y_{it} is the outcome variable

Z_i is a vector of time invariant explanatory variables

Following Schunck (2013) and Burke and Jayne (2014), β are estimated parameters in the model that are interpreted as "within-household" cluster or household effect. It is important to note that these "within-household" estimates are similar to the FE estimates, that is, these coefficients are the effect of a given time varying variable's effect of deviation from its overall average or "permanent" level. Therefore, logically, these coefficients can be interpreted as the effect of a deviation within a household (Burke and Jayne 2014). On the other hand, λ and γ are estimated model parameters that are interpreted as "between" cluster or household effects. These variables are constant for each household across the panel period and therefore they only represent "between-household" effect. This means that time varying covariates (X_{it}) can be decomposed into "within" and "between" cluster or household effects (Schunck 2013; Burke and Jayne 2014). Therefore, to construct the *Mundlak-Chamberlain* devise, panel average variables of selected time varying based on the two panel periods were added in the selection and food security outcome models as additional explanatory variables.

The efficient method to estimate ESR models is by full information maximum likelihood (FIML) estimation (Lokshin and Sajaia, 2004). An alternative estimation method is fitting one equation at a time by either 2SLS or maximum likelihood estimation. However, these later methods are less efficient than FIML because they require some potentially cumbersome adjustments to derive consistent standard errors (Lokshin and Sajaia, 2004). The 2SLS or maximum likelihood estimation also shows poor performance in case of high multicollinearity between covariates of the selection model (Maddala, 1983). The FIML approach relies on joint normality of error terms in the selection and outcome equations and thus more efficient (Lokshin and Sajaia, 2004).

2.4 Data

The study is based on balanced two wave panel data collected from 457 rural farming households (914 observations). The first-round data was collected in January – April 2011 and the second round collected in August – November 2013. The surveyed households were randomly selected from villages in Bungoma and Siaya districts in western Kenya and Embu, Imenti South and Meru South districts in eastern Kenya. A semi-structured questionnaire was used by trained enumerators to collect data on household socioeconomic characteristics, crop production and utilization including marketing, household own self-assessment of food security in the last twelve months preceding the survey, total household cash expenditure on food and non-food items and sources of other household incomes including credit and savings among many more variables.

3. Results and discussion

Descriptive and econometric results of ESR model using correlated random effects (CRE) approach are presented and discussed in this section. Also, results of the treatment (commercialization) under actual and counterfactual frameworks are presented and discussed in detail before conclusions and policy implications are drawn.

3.1 Descriptive statistics

About 75% of the surveyed households were commercialized and their average commercialization intensity was 37% though unconditional commercialization intensity was about 27% (Table 2). Food secure households were more commercialized than their food insecure counterparts. While no causality is implied, these descriptive statistics show positive correlation between household agricultural commercialization and food security.

Household demographic variables, physical and financial assets, social capital and transaction costs variables also showed some marked differences across food secure and food insecure households (Table 2).

Table 2. Descriptive statistics for the Endogenous Switching Regression Model: Food security

Variable label	Food secure (N=497)		Food insecure (N=417)		Total (N=914)		Difference
	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	
Commercialized households	0.8491	0.3583	0.6211	0.4857	0.7451	0.4361	0.2280***
Proportion of value of crop produced sold	0.3449	0.2711	0.1915	0.2352	0.2749	0.2664	0.1534***
Demographic characteristics:							
Household head sex	0.8632	0.3440	0.8106	0.3923	0.8392	0.3676	0.0526**
Household head age	50.6519	13.9085	50.8561	13.5423	50.7451	13.7355	-0.2042
Household head education	8.1362	3.7991	7.1583	3.7460	7.6900	3.8043	0.9779***
Household size	4.7801	2.1998	5.4062	2.3346	5.0658	2.2825	-0.6261***
Dependency ratio	0.8126	0.7005	1.0869	0.9060	0.9377	0.8119	-0.2742***
Physical and financial assets:							
Per capita owned farm size	0.2617	0.2395	0.2037	0.2056	0.2353	0.2264	0.0580***
Tropical livestock units (TLU)	1.7567	2.0382	1.5151	1.6229	1.6465	1.8632	0.2416*
Mean soil fertility score	2.1808	0.5710	2.0024	0.6480	2.0994	0.6135	0.1784***
Total annual non-farm income (1000 KSh)	130.4707	272.6154	58.6484	106.1973	97.7027	216.3209	71.8222***

Had contacts with extension staff	0.4930	0.5005	0.4988	0.5006	0.4956	0.5003	-0.0058
Household got agricultural credit	0.1751	0.3804	0.0959	0.2948	0.1389	0.3461	0.0791***
Social capital:							
Household belongs to APN	0.5694	0.4957	0.4436	0.4974	0.5120	0.5001	0.1258***
Number of dependable relatives in village	5.9195	10.5353	6.0600	10.7790	5.9836	10.6416	-0.1404
Trusts in grain traders	0.7545	0.4308	0.7170	0.4510	0.7374	0.4403	0.0375
Transaction costs:							
Owns mobile phone	0.8873	0.3165	0.7722	0.4199	0.8348	0.3716	0.1151***
Transport cost to main market	47.7143	34.1472	52.2974	35.7174	49.8053	34.9279	-4.5831**
Own local transport means	0.7243	0.4473	0.5707	0.4956	0.6543	0.4759	0.1536***
Regional dummy	0.6258	0.4844	0.3381	0.4736	0.4945	0.5002	0.2876***

Statistical significance: *** at 1%; ** at 5%; * at 10%

3.2 Econometric results

Results from ESR model in correlated random effects framework were as presented in Table 3. Determination of the validity of instruments used was based on theory and empirical evidence. Theoretically, all transaction cost variables used in the model were assumed to affect agricultural commercialization more directly than food security outcomes. Empirically, Goetz (1992) and Key *et al.*, (2000) have demonstrated that transaction costs determine agricultural market participation (commercialization). Therefore, following Lokshin and Glinskaya (2009), Di Falco *et al.*, (2011) and Kassie *et al.*, (2014b), this study tested the validity of the identified instrument variables (all transaction cost variables) using a simple falsification procedure. The results showed that only household ownership of cellphone met this criterion and was used as an instrument in the ESR model.

Since this paper focuses on assessing the impact of agricultural commercialization on household food security outcome, only important policy variables from the selection model are discussed here. Policy variables that were significant in determining agricultural commercialization included farm size, soil fertility and credit access. Households with bigger farm sizes, more fertile soils and accessed agricultural input credit were likely to be commercialized (Table 3). Also, households that belonged to agricultural production networks/groups were likely to commercialize. Lastly, smallholder farmers who owned cellphones were likely to be commercialized while those that were in more remote areas (had higher transport costs to the nearest main market) were unlikely to commercialize (Table 4).

Econometric results of the determinants of household food security probability showed that education level of the household head had a positive and significant "between-household" effect on food security of commercialized households (Table 3). This finding highlights the importance of formal education in enabling commercialized households to access modern agricultural production technologies and market information needed to produce surplus and participate in markets. Secondly, dependence ratio had a negative and significant "between-household" effect on food security of non-commercialized households (Table 3). Non-commercialized households with higher dependency ratio were likely to be food insecure because of increased burden on active working members to provide food for non-productive members.

Econometric results from physical and financial asset variables showed that the "within-household" effects were more important in explaining food security probability of non-commercialized than commercialized households (Table 3). These "within-household" effects were mainly positive thus highlighting the importance of household wealth in explaining food security among non-commercialized households. For example, the tropical livestock units (TLU) had a positive and significant "within-household" effect on food security probability among non-commercialized households implying that non-commercialized households might be depending on selling off some of their livestock to buy food. On the other hand, household soil fertility score for all operated plots had a positive and significant "within-household" effect on food security probability among non-commercialized households (Table 3). This means that non-commercialized households are more dependent on own produced foods and therefore increasing fertility of their plots is likely to increase productivity of their own foods (Govereh and Jayne 2003, Mather *et al.*, 2011).

Table 3. Endogenous Switching Regression: Impact of agricultural commercialization on household food security outcome

Variable label	Selection model: Determinants of household agricultural commercialization		Outcome models: Determinants of household food security outcome			
			Commercialized households		Non-commercialized households	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Demographic characteristics:						
Household head sex	-0.0903	0.1546	0.0874	0.1693	-0.1211	0.3126
Household head age	0.0089	0.0144	0.0054	0.0148	0.0143	0.0275
Household head education	0.0342**	0.0167	0.0414**	0.0169	0.0385	0.0318
Household size	0.0265	0.0627	-0.0305	0.0704	-0.0397	0.1169
Dependence ratio	-0.0539	0.1280	-0.0719	0.1334	0.4422	0.2858
Physical and financial capital:						
Owned livestock size	0.0147*	0.0080	-0.0021	0.0087	0.0264*	0.0161
Per capita owned land	3.8204**	1.6995	-1.2190	1.7845	-1.0003	3.7695
Per capita owned land squared	-1.7225**	0.8657	1.3666	1.1780	0.0909	1.7762
Soil fertility score	-0.1511	0.1283	0.1275	0.1400	0.6040**	0.2441
Annual non-farm income	-0.0093*	0.0049	0.0073	0.0046	0.0166*	0.0102
Got agricultural input credit	0.0569	0.2492	0.2192	0.2232	-0.5613	0.5707
Contacts with extension	0.2137	0.1554	-0.2202	0.1593	-0.0324	0.3129
Ox-plough ownership	na	na	-0.2410	0.3156	-0.1811	0.6678
Social capital:						

Membership to APNs	0.5515***	0.1153	0.2838**	0.1198	0.5145**	0.2334
Dependable relatives in village	0.0010	0.0051	-0.0088*	0.0050	-0.0098	0.0150
Trust grain traders	0.1804	0.1203	0.2764**	0.1251	0.3372	0.2413
Transaction costs:						
Own mobile phone	0.8913***	0.2171	na	na	na	na
Transport to nearest main market	-0.0156**	0.0078	-0.0111*	0.0067	-0.0300	0.0194
Own transport means	0.1327	0.1125	0.3787***	0.1173	0.6914***	0.2263
Regional dummy	0.9400***	0.1405	0.7489***	0.1393	-0.4514	0.3329
Mundlak - Chamberlain device:						
Household head age	-0.0116	0.0151	-0.0064	0.0156	-0.0152	0.0294
Household size	-0.0342	0.0692	-0.0333	0.0776	0.0403	0.1320
Dependence ratio	0.1346	0.1500	-0.2047	0.1603	-0.9026**	0.3542
Owned livestock size	-0.0063	0.0107	0.0080	0.0115	-0.0100	0.0205
Per capita owned land	-1.6208	1.7661	1.5496	1.7241	0.2693	3.9939
Per capita owned land squared	0.3784	0.9639	-1.0921	1.0211	0.5232	2.2847
Soil fertility score	0.4152**	0.1753	0.1460	0.1776	-0.4520	0.3642
Annual non-farm income	0.0078	0.0064	-0.0035	0.0060	-0.0204	0.0139
Got agricultural input credit	0.7084**	0.3317	-0.0402	0.3091	2.0555***	0.7076
Contacts with extension	-0.1432	0.2137	0.5442**	0.2200	-0.2803	0.4359
Own mobile phone	-0.2894	0.2895	na	na	na	na
Ox-plough ownership	na	na	0.1337	0.3972	0.0328	0.7307
Constant	-1.7167***	0.4261	-1.2708***	0.4695	-0.9832	0.6994

Statistical significance: *** at 1%; ** at 5%; * at 10%

Model description:

Descriptor	Selection model	Commercialized households	Non-commercialized households
Number of obs	914	681	233
LR chi2(30)	244.8800	141.2900	58.8000
Prob>chi2	0.0000	0.0000	0.0013
Pseudo R2	0.2360	0.1562	0.2008
Log likelihood	-396.4215	-381.6888	-116.9893

Non-farm annual household income had a positive and significant "within-household" effect on food security of non-commercialized households (Table 3). This relationship could be because of non-commercialized households using their non-farm income to buy food to make up for the shortfall in own produced food staples. This might be contrary to commercialized households who are more likely to produce surplus (Govereh and Jayne (2003).

Access to agricultural credit had a positive and significant "between-household" effect on food security of non-commercialized households (Table 3). This finding confirms that all crop enterprises on the farms of non-commercialized households are food crops and accessing this input credit goes directly to boost their food crop production. On the other hand, contacts with extension staff had a positive and significant "between-household" effect on food security of commercialized households (Table 3). Therefore, extension information is important in enabling commercialized households to produce more food crops beyond their subsistence levels.

While physical and financial assets were significant in explaining food security mainly among non-commercialized households, social capital and transaction costs were more important in explaining food security among commercialized households (Table 3). These empirical findings support the new institutional economics theory which postulates that social capital type of institutions and transaction costs variables are important in reducing the costs incurred by commercialized households in concluding their market transactions (North, 1990).

3.3 Treatment effects of agricultural commercialization on household food security

To disentangle the impacts of agricultural commercialization on household food security stemming from observed and unobserved heterogeneities between commercialized and non-commercialized households, a counterfactual analysis was built from ESR estimates using post estimation procedures (Lokshin and Sajaia, 2004). In Table 4, cell (a) and cell (b) represent the actual (observed) expected household probability of being food secure among commercialized and non-commercialized households, respectively. This means that, among commercialized households, the observed expected probability of a household being food secure was about 62% compared to 32% among non-commercialized households (Table 4). These results compare very well with descriptive statistics generated directly from the data as presented in Figure 1. These descriptive statistics in Figure 1 were found to be statistically different at 1% level of significance. A quick and direct comparison of these statistics could imply that agricultural commercialization helped commercialized households to have food security probability advantage of almost 30 percentage points. However, such simple and direct comparison could be misleading without taking into consideration the observed and unobserved characteristics of commercialized and non-commercialized households that could be correlated with respective food security outcome. This calls for building counterfactual scenarios of expected values for the two groups of households, that is, cell (c) and cell (d) in Table 4 to decompose the sources of this 30% difference in food security probability between commercialized and non-commercialized households.

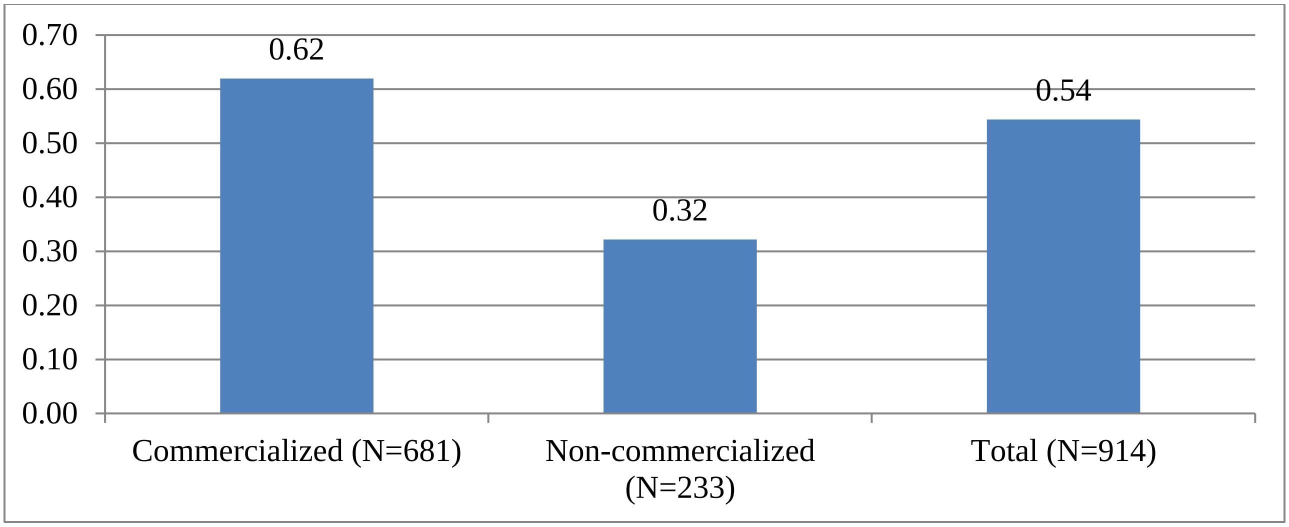

Figure 1. Proportion of households that are food secure

The counterfactual analysis and decomposing the 30% food security gap followed the Oaxaca (1973) wage decomposition. This food security gap decomposition revealed that food security probability outcome of commercialized households would be reduced significantly by almost 23% if efficiency of their currently held resources was to be the same like that of the non-commercialized households, that is, cell (a) less cell (d) in Table 4. This is what is commonly referred to as "returns effect" in recent literature (Kassie *et al.,* 2015). Similarly, the food security probability outcome for commercialized households could be reduced by about 7%, that is, cell (d) less cell (b) in Table 4, if this group of households had the same amount of resources like the ones held currently by the non-commercialized households. This is what is now commonly called the "level effect" (Kassie *et al.,* 2015). Therefore, the 30% food security probability outcome advantage that commercialized households have over non-commercialized households stems from commercialized households' superiority in resource use efficiency (21%) and advantage in resource amount (9%). These results mean that for commercialized households food security probability to come down from 32% level of non-commercialized households from their 62%, then commercialized households' resource use efficiency have to come down and at the same time their resource level or amount have to be reduced too. However, no policy maker will be interested in pursuing interventions that could result in such a scenario.

Table 4. Average expected household food security probability outcomes

Type of household	Household food security probability outcome		
	Commercialized characteristics	Non-commercialized characteristics	Returns effects
Commercialized (N=681)	0.6195 (a)	0.3906 (d)	0.2289***
Non-commercialized (N=233)	0.4377 (c)	0.3218 (b)	0.1158***
Level effects	0.1818***	0.0687***	**0.2976**

Statistical significance: *** at 1%; ** at 5%; * at 10%

On the other hand, the food security probability outcome for non-commercialized households could significantly be improved by almost 12% if their resource use efficiency could be improved to that of commercialized households. This means that if non-commercialized households were as efficient as commercialized households in the use of their currently held resources, the 30% food security gap will be reduced by about 12% points to 18% due to what is commonly called the "returns" effect, that is, cell (c) less cell (b) in Table 4. The remaining 18% food security gap can be closed by improving the amount of resources held by the non-commercialized households. In other words, if the resources held by the non-commercialized households can be increased to the level of the amount of resources held by commercialized households while holding else constant (ceteris paribus), then their food security probability outcome will increase by 18% as a result of what is usually referred to as the "level" effect. Basically, these counterfactual results for non-commercialized households implies that the main source of the observed food security gap between commercialized and non-commercialized households (30%) is the differences in the amount of resources held by the two groups of households. The low amount of resources held by the non-commercialized households compared to commercialized households contributes to about 61% of the food security gap. The remaining 39% of the food security gap is due to poor efficiency in use of the resources by non-commercialized households compared to commercialized households.

3.4 Summary and conclusion

Global economic and demographic changes in terms of rising incomes and increasing urbanization, respectively, have provided smallholder farmers in developing countries with immense opportunity to commercialize their agricultural activities. However, the implication of this agricultural transformation process on the welfare of rural farming households is not clear. Particularly not clear in theoretical and empirical literature is the effect of agricultural commercialization on household food security. Most of the previous empirical studies that that analyzed the welfare impact of agricultural commercialization were based on either cross-sectional data and/or pooled regression models. However, panel data and use of more rigorous and robust models like switching regression could are more informative.

In this study, two wave balanced panel data collected from smallholder farmers in western and eastern parts of Kenya were used. The study applied endogenous switching regression using correlated random effects framework to analyze the impact of smallholder agricultural commercialization on household food security. The results showed that physical and financial assets were positively and significantly related to food security outcome among non-commercialized households. On the other hand, social capital and transaction costs were more important in determining the food security outcome of commercialized households. Therefore, while non-commercialized households were more dependent on physical and financial assets to be food secure, commercialized households were more dependent on the efficient working of markets.

Impact assessment analyzes showed that agricultural commercialization has a positive and significant effect on household food security outcome. The counterfactual analysis revealed that commercialized households could have significantly lowered their food security if they were not commercialized. On the other hand, non-commercialized households could have improved their food security significantly if they had commercialized. The observed probability of a commercialized household being food secure was about 62% while that of a non-commercialized household was 32%. This 30% food security gap was decomposed into two main sources, that is, that emanating from differences in resource use efficiency between the two groups and amount of resources held by each group. The food security of non-commercialized households would improve significantly by almost 12% if they were as efficient in using their currently held resources as commercialized household are. This difference in resource use efficiency accounted for about 39% of the observed food security gap. The other 18% gap (i.e. 61% of the gap) was due to differences in the amount of resources held by the two groups of households. This means that the food security gap between commercialized and non-commercialized households can only be fully closed if both efficiency of resource use and the amount of resources held by non-commercialized households can be improved.

References

Alene D.A., V.M. Manyong, G. Omanya, H.D. Mignouna, M. Bokanga, G. Odhiambo (2008). Smallholder market participation under transactions costs: Maize supply and fertilizer demand in Kenya. *Food Policy 33 (2008) 318–328*

Bellemare M.F. and Barrett C.B. (2006). An Ordered Tobit Model of Market Participation: Evidence from Kenya and Ethiopia. *American Journal of Agricultural Economics 88(2) (May 2006): 324 – 337)*

Burke, W. J. and T.S. Jayne (2014). Smallholder Land Ownership in Kenya: Distribution Between Households and Through Time. *Agricultural Economics 45 (2014) 185 – 198*

Chamberlain G. (1982). Multivariate Regression Models for Panel Data. *Journal of Econometrics, Vol. (18): 5 – 46*

Coudouel, A., J.S. Hentschel, and Q.T. Wodon (2002). Poverty Measurement and Analysis. In: J. Klugman (ed.). A Sourcebook for Poverty Reduction Strategies. Washington DC. World Bank.

de Janvry A., Dustan A., and Sadoulet E. (2011). Recent Advances in impact analysis methods for ex-post impact assessments of agricultural technology: Options for the CGIAR. Report prepared for the workshop: Increasing the rigor of ex-post impact assessment of agricultural research: A discussion on estimating treatment effects, organized by the CGIAR Standing Panel on Impact Assessment (SPIA), 2 October, Berkeley, California, USA. Independent Science and Partnership Council Secretariat: Rome Italy

Di Falco S., M. Yesuf, G. Kohlin and C. Ringler (2011). Estimating the Impact of Climate Change on Agriculture in Low-Income Countries: Household Level Evidence from the Nile Basin, Ethiopia. *Environmental and Resource Economics, 52 (4). pp. 457-478. ISSN 0924-6460*

FAO (2003). Trade Reforms and Food Security: Conceptualizing the linkages. Commodity Policy and Projections Service, Commodities and Trade Division, FAO, Rome http://www.fao.org/docrep/005/y4671e/y4671e06.htm Accessed on 19 November 2013

Gebremedhin B. and M. Jaleta (2010). Commercialization of Smallholders: Does Market Orientation Translate into Market Participation? Improving Productivity and Market Success (IPMS) of Ethiopian Farmers Project Working Paper 22. Nairobi, Kenya, ILRI

Goetz S.J. (1992). A Selectivity Model of Household Food Marketing Behaviour in Sub-Saharan Africa. *American Journal of Agricultural Economics, 446 – 452*

Govereh J. and T.S. Jayne (2003). Cash Cropping and food crop productivity: Synergies or trade-offs? *Agricultural Economics 28 (2003) 39 – 50*

Govereh J.; Jayne T.S. and Nyoro J. (1999). Smallholder Commercialization, Interlinked Markets and Food Crop Productivity: Cross-Country Evidence in eastern and southern Africa. *http//www.aec.msu.edu/fs2/ag_transformation/atw_govereh.PDF*

Haughton and Khandker (2009). Handbook on Poverty and Inequality. The International Bank for Reconstruction and Development / World bank. http://go.worldbank.org/7JGPK76TM0 Accessed on 24 September 2014

Jayne, T.S. and M. Muyanga (2012). Land Constraints in Kenya's Densely Populated Rural Areas: Implications for Food Policy and Institutional Reform. Contributed Paper prepared for presentation at the 86th Annual Conference of the Agricultural Economics Society, University of Warwick, United Kingdom. 16 -18 April 2012.

Kassie M., H. Teklewold, M. Jaleta, P. Marenya and O. Erenstein (2015). Understanding the Adoption of a Portfolio of Sustainable Intensification Practices in eastern and southern Africa. *Land Use Policy 42 (2015) 400 – 411*

Kassie M., Ndiritu S. W. and Stage J. (2014a). What determines gender inequality in household food security in Keny? Application of exogenous switching treatment regression. *World Development, 56, 153 - 171*

Kassie, M., M. Jaleta and A. Mattei (2014b). Evaluating the Impact of Improved Maize varieties on Food Security in Rural Tanzania: Evidence from Continuous Treatment Approach. *Food Security DOI 10.1007/s12571-014-0332-x*

Key, N. E. Sadoulet and A de Janvry (2000). Transaction Costs and agricultural Household Supply Response. *American Journal of Agricultural Economics. Pp. 245 – 259.*

Lokshin M. and E. Glinskaya (2009). The Effect of Male Migration on Employment Patterns of Women in Nepal. *The World Bank Economic Review, Vol. 23, No. 3, pp.481 – 507*

Lokshin, M. and Z. Sajaia (2004). Maximum Likelihood Estimation of Endogenous Switching Regression Models. *Stata Journal 4(3): 282 – 289*

Maddala G.S. (1983). *Limited Dependent and Qualitative Variables in Econometrics.* New York: Cambridge University Press.

Mallick D. and Rafi M. (2010). Are female headed households more food insecure? Evidence from Bangladesh. *World Development, 38, 593 - 605*

Mather, D., D. Doughton and T.S. Jayne (2011). Smallholder Heterogeneity and Maize Market Participation in Southern and Eastern Africa: Implications for Investments Strategies to Increase Marketed Food Staple Supply. MSU International Development Working Paper 113, October 2011

Mathenge M., F. Place, J. Olwande and D. Mithoefer (2010). Participation in Agricultural Markets among the Poor and Marginalized: Analysis of Factors Influencing Participation and Impacts on Income and Poverty in Kenya. A study Report, Tegemeo Institude of Agriculturasl Policy and Development and World Agroforestry Centre. http://www.tegemeo.org/documents/other/Report-on-Market-Participation-among-the-Marginalized-groups-in-Kenya-Final-Report-Kenya-v3.pdf Accessed on 5 September 2013

Mundlak Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica, Vol. 46, No. 1 (Jan. 1978), pp. 69 – 85*

Muyanga, M., M. Ayieko and M. Bundi (2006). Determinants of Transient and Chronic Poverty: Evidence from Kenya. PMMA Network Session Paper. A paper Presented During the 5th PEP Research Network General Meeting, June 18 – 22, 2006, Addis Ababa, Ethiopia

North, D.C. (1990). *Institutions, Institutional Change and Economic Performance.* Cambridge, Cambridge University Press.

Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labour Markets. *International Economic Review, Volume 14, Issue 3 (Oct. 1973), 693 - 709*

Pingali P. L. (1997). From Subsistence to Commercial Production Systems: The Transformation of Asian Agriculture. *American Journal of Agricultural Economics, 79 (May 1997): 628 - 634*

Pingali P.L. and M.W. Rosegrant (1995). Agricultural Commercialization and diversification: processes and policies. *Food Policy, Vol. 20, No. 3, pp. 171 – 185, 1995*

Republic of Kenya (1981). *National Food Policy paper*. Sessional Paper 4 Nairobi: Government Printers

Republic of Kenya (2004). The Economic Recovery Strategy for Wealth and Employment Creation 2003 - 2007. Government Printers, Nairobi

Republic of Kenya (2007). Kenya Vision 2030. Government Printers, Nairobi

Republic of Kenya (2010). *Agricultural Sector Development Strategy 2010 – 2020*. http://www.ascu.go.ke/DOCS/ASDS%20Final.pdf Accessed on 5 September 2013

Schunck, R. (2013). Within and Between estimates in random effects models: Advantages and drawbacks of correlated random effects and hybrid model. *The Stata Journal (2013): 13, Number 1, pp. 65 - 76*

Sharma, P.V.; D. Jain and S. De (2012). Managing Agricultural Commercialization for Inclusive Growth in South Asia. Global Development Network (GDN). The Global Research Capacity Building Program.

Strasberg P.J., Jayne T.S., Yamano T. Nyoro J. Karanja D. and Strauss J. (1999). Effects of Agricultural Commercialization on Food Crop Input use and productivity in Kenya. Michigan State University International Development Working Papers No. 71. Michigan, USA

Timmer CP. 1997. Farmers and markets: The political economy of new paradigms. *American Journal of Agricultural Economics 79(2):621–627*

Von Braun J., Bouis H and Kennedy E. (1994). Conceptual Framework. In: von Braun J and Kennedy E. (eds), *Agricultural Commercialization, Economic Development and Nutrition.* John Hopkins University Press, Baltimore, Maryland, USA. Pp. 9-33

von Braun, J. (1995). Agricultural commercialization: Impacts on income and nutrition and implications for policy. *Food Policy, Vol. 20, No. 3, pp. 187 – 202, 1995*

Wooldridge J. (2010). *Econometric Analysis of Cross Sectional and Panel Data.* The MIT Press, Cambridge, Massachusetts.

World Bank (2000). World Development Report 2000/2001: Attacking Poverty. The World Bank, Washington DC

World Bank (2009). *Kenya - Poverty and inequality assessment: executive summary and synthesis report*. Washington, DC. WorldBank. http://documents.worldbank.org/curated/en/2009/04/10842664/kenya-poverty-inequality-assessment-executive-summary-synthesis-report. Accessed 7 September 2015