COMMERCIALIZATION EFFECTS ON HOUSEHOLD INCOME, POVERTY, AND DIVERSIFICATION: A COUNTERFACTUAL ANALYSIS OF MAIZE FARMERS IN KENYA

John Olwande*and Melinda Smale[§]

*Corresponding author, Research Fellow at Tegemeo Institute of Egerton University and PhD student at the Department of Agricultural, Food and Resource Economics, Michigan State University, East Lansing, MI. Email: olwandej@msu.edu

[§]Professor at the Department of Agricultural, Food and Resource Economics, Michigan State University, East Lansing, MI. Email: msmale@msu.edu

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2014 AAEA Annual Meeting, Minneapolis, MN, July 27-29, 2014.

Commercialization effects on household income, poverty, and diversification:

A counterfactual analysis of maize farmers in Kenya

Abstract

High poverty rate persists in rural Kenya, where farming households continue to depend on agriculture for food and income, despite economic growth. Maize is the most widely grown crop, with the maize-growing smallholder population quite heterogeneous and diversified. However, less than half the growers enter the market to sell at least a portion of their harvest. Our objective in this paper is to test the effects of maize market participation on household income, poverty status, and income diversification among Kenya's smallholder maize growers. We employ a combination of propensity score matching and endogenous switching regression on household panel data covering a ten-year period. The propensity score matching results show that in the overall, participation in the maize market has a significant impact on household income and poverty, with the magnitude of impact estimates differing across segments of maize growers, while the impact on diversification of income sources is not significant in most cases. These results persist after controlling for hidden selection bias through endogenous switching regression, with the impact estimates larger. Heterogeneity effects attest to existence of underlying differences among the maize-producing households that make sellers better off than non-sellers regardless of participation in the maize market. Our findings reinforce the call for interventions to expand the capacity of smallholder maize farmers to produce for the market in efforts to raise incomes and contribute to a more widespread poverty reduction.

1. Introduction

High ppoverty rate persists in rural Kenya, where a large proportion of households continue to depend on agriculture for food and income despite improved economic growth and a steady process of decentralized urbanization. The national economy of Kenya grew at rates of 3-4 percent from the early 1990s through the end of the first decade of this century, but growth was offset by high rates of population growth (2%), with a higher rate of poverty reduction in urban as compared to rural areas (Thurlow et al. 2012). An estimated 80 percent of the people in Kenya live in rural areas, and about half of these (49.1 percent) remain poor (Republic of Kenya, 2008).

Poverty reduction in Kenya, and other agriculture-based economies of Africa south of the Sahara, is unlikely to occur without agricultural growth. How will agricultural growth occur? Despite arguments to the contrary (e.g., Collier 2008), most analysts still agree that small-scale agriculture remains the primary impetus for economic growth and any strategy aimed at diminishing rural poverty (Hazell et al. 2007; World Bank 2007; Byerlee et al. 2009). Following the global food-price crisis of 2008, Sadoulet and de Janvry (2012) referred to subsistence farming as a "safety net," and a argued that investments in small-scale agriculture are a more effective policy tool for managing price instability in poorer agricultural countries than the fiscal measures (tariffs, subsidies) most often deployed in industrialized nations.

Three-quarters of the total agricultural output in Kenya is produced by farm families on landholdings averaging 0.2-3 hectares (Republic of Kenya, 2010). Agricultural growth rates over the past few decades have been lower than overall rate of economic growth, with rapid expansion in cropland and stagnating yields for a number of crops. These include maize, the primary food staple and source of cash for the majority of Kenyan smallholders—despite high rates of adoption of maize hybrids and fertilizer (e.g., Smale and Olwande 2014; Ariga et al. 2010).

Analyses assembled by Diao et al. (2012) provide ample evidence that the economic impact of developing domestic markets for staple foods is far greater than for exported or higher value commodities in numerous countries of Africa south of the Sahara, including Kenya. Yet, in addition to stagnating maize yields, Suri (2011) demonstrated how the heterogeneity of net returns slows cumulative adoption of hybrid seed in Kenya. Data collected by Tegemeo Institute of Egerton University (2000-2010) illustrates the heterogeneity of the maize farming population with respect to maize marketing: over the decade, an estimated 21% of maize growers were autarkic (neither selling nor purchasing maize), a little over a quarter (28%) of maize growers sold, but did not purchase the crop. Another 36% purchased maize but did not sell it, and about 14% both bought and sold maize. At the same time, maize growers in Kenya typically produce and sell a range of other crops. Over that same time period, 26% of maize growers sold both maize and dairy products, and 13% sold maize alongside major industrial crops (coffee, tea, and sugar cane), of which coffee and tea are exported. The dramatic impacts of dairying on the well-being of smallholder farmers in Kenya have been extensively researched (e.g. Ngigi et al. 2010, Staal et al. 2008); tea and coffee are historical export crops in Kenya, dating to the colonial period; and sugar cane is a major industrial crop in the Western part of the country. The dynamics of agricultural growth in industrialized countries suggests that many of these maize-growing smallholders will leave agriculture for non-farm employment in the years to come, while others will become increasingly oriented toward profits, specializing in fewer farm enterprises.

Our objective in this paper is to test the effects of maize market participation on household income, poverty status, and income diversification among Kenya's smallholder maize growers. We segment our population in order to demonstrate the heterogeneity that has contextual implications for investments and policies to reduce rural poverty. A number of applied studies have addressed the constraints to market participation by smallholder farmers in Kenya (Renkow et al. 2004; Mathenge et al. 2010; ILRI 2011). To our knowledge, this study is one of few to examine the effects of market participation on household welfare indicators—a crucial link in the process of agricultural growth. Rao and Qaim (2010) examined the effects of participation of vegetable growers in supermarket chains in Kenya, and Asfaw et al. (2012) explored the relationships among input and output market participation in pigeonpea, crop diversity, and household welfare in Kenya.

Conceptually, we view farmer decision-making from the perspective of the agricultural household framework, as adapted by Key, Sadoulet and de Janvry (2000) and Barrett (2008) to market participation. A major challenge with estimating the welfare effects of market participation is the identification of the counterfactual, or the welfare status of participating farmers had they not participated. We apply a counterfactual approach (Di Falco and Veronesi 2013) to compare the effects of maize commercialization on income, poverty and income diversification. Like Asfaw et al. (2012), we apply a combination of endogenous switching regression and propensity score matching to control for the selection bias associated with the non-random decision of some farmers to participate in maize markets. To address decision-making heterogeneity, we segment the sample of maize farming households into subgroups, comparing a) all maize sellers to all maize non-sellers; b) maize farmers who sell only (exclusive sellers) to all maize non-sellers; c) maize farmers who both sell and buy to all maize non-sellers;

d) maize sellers who also sell dairy to non-sellers of maize who sell dairy; e) sellers of both maize and industrial crops to non-sellers of maize who sell industrial crops. In addition, we exploit the information in four years of panel data, using a Correlated Random Effects estimation approach to control for time-invariant, unobserved heterogeneity among farmers.

2. Conceptual framework

Our empirical model is anchored in the well-known agricultural household model. The core model depicts a farming household that maximizes utility over consumption goods produced on the farm or purchased from the market, subject to a cash income and/or credit constraint. Profit-maximizing behaviour is a special case of the model, in which consumption and production decisions can be considered separate because prices are determined exogenously in perfect markets. The defining feature of the non-separable model is that the prices guiding farmer decisions (decision prices) are endogenously determined not only by observed market prices but also by factors that influence the transactions costs associated with participating in input and output markets. Most importantly, these are household-specific prices that are unobserved and heterogeneous across smallholder farmers. Given the heterogeneity of our maize farming population, we invoke the non-separable model, in which prices are endogenous to decision-making and determined by transactions costs.

In Barrett's (2008) application of the model to market-related decisions, the household chooses whether or not to participate as a seller (a vector **M** of indicator variables equal to one for market entry, 0 otherwise) based on decision prices that depend on vectors of crop-and-household specific transactions costs. Decision prices depend on vectors of public goods and services, and household characteristics that affect search costs, such as household assets and liquidity. Barrett (2008) also differentiates distinct layers of transactions costs that are household-specific and crop-or location-specific, and among interlinked local, regional and international markets. Transaction costs create a wedge, which result in price bands (market prices plus/minus transaction costs), which result in kinked demand and supply functions with diminished responsiveness to changes in price (Key at al., 2000).

Consistent with our context, we can express the reduced form equation of Barrett's (2008) model as:

$$M_i = M_i\,(P, \mathbf{Z}, \mathbf{A}, \mathbf{G} \mid \mathbf{\Theta}) \qquad (1),$$

where M_i reflects the decision to sell maize or not by household *i,* P is the maize price as observed in the market, and the remainder of the vectors refer to the determinants of transactions costs. The vector **Z** is composed of variables representing the education, age, and gender composition of the household. Here, we add prime-age mortality, which has been shown to influence decision-making among households surveyed in some circumstances (Yamano and Jayne 2004; Chapoto et al. 2012). The vector **A** includes the total value of household assets and land size, as well as the ownership of household-specific assets that influence transactions costs (radio, transport equipment). Credit provision and relative liquidity is measured by the share of households receiving credit in the village, and also by the proportion of households in the village who are members of farmer groups and/or cooperatives. Public provision of market infrastructure (**G**) is measured at several geographical scales, drawing on secondary, geo-referenced data sources in order to introduce inter-linkages among markets (travel time to towns, population density; see also Staal et al. 2002). We also differentiate by the production potential of the agroecological zone and rainfall conditions at the nearest metereological site where precipitation is gauged (**Θ)**.

While we may be interested in explaining the decision to sell maize, our primary concern is the effect of maize market participation on the income, poverty, and income diversification status of smallholder maize farmers, and the way these effects may vary by subgroups of maize sellers (sell only; buy and sell; also sell dairy; also sell industrial crops). In section 3, we lay out our empirical strategy.

3. Estimation Strategy

Below, we begin by defining our indicators of well-being. We then present the rationale for our choice of econometric models. First, to estimate the effect of selling in the maize market on well-being, we apply propensity score matching (PSM), employing both kernel and nearest neighbor algorithms for robustness. We eliminate the observations outside the common support, in order to estimate the overall average treatment effect on the treated (ATT). In view of the limitation that propensity score matching controls for selection bias only on the basis of observed covariates, we then conduct Rosenbaum tests (2002) to gauge the sensitivity of the estimated treatment effects to hidden bias.

Second, we apply an endogenous switching regression model (ESR). This model enables us to a) control for unobserved heterogeneity, which cannot be addressed by propensity score matching because it is based on observable variation; b) estimate each component of the counterfactual (effects of participation on the participants; effects of participation on non-participants; and the heterogeneity effects for participants and non-participants); c) illustrate how the factors that influence economic status at the margin differ among maize-selling subgroups according to whether or not they also sell dairy products or industrial crops. These approaches are augmented by use of the Mundlak (1978)-Chamberlin (1984) device to handle time-invariant, unobserved heterogeneity in the context of a censored variable specification applied to panel data.

3.1. Impact outcomes

The outcome variables we consider are household income, poverty and income diversification. We measure net household income in nominal terms. Income comprises earnings from crops (gross value of crop production less input costs); livestock (gross value of livestock products plus sales of live animals less purchases of live animals plus input costs); salaries earned by all household members; businesses for all household members; informal labor employment; and remittances, pension and share dividends received by all household members.

We measure poverty by the Foster-Greer-Thorbecke (FGT) index:

$$FGT_{\alpha} = (1/n)\sum_{i=1}^{h}\left(\frac{z-y_i}{z}\right)^{\alpha} \qquad (2),$$

computing all three measures of poverty as defined by α: headcount ratio, where $\alpha = 0$; poverty gap, where $\alpha = 1$; and poverty severity, where $\alpha = 2$. In equation (2), the variable z represents the poverty line, n the number of households in the sample, h is the number of poor households (those with incomes at or below z), and y_i is household income per adult equivalent. The headcount ratio is the fraction of households below the poverty line. The poverty gap, also known as depth of poverty, is interpreted as the amount of income it would take to raise people in poverty up to the poverty line. This indicator measures the extent to which individuals fall below the poverty line as a proportion of the poverty line.

Poverty lines for each survey year were calculated by adjusting the 2006 value with the corresponding consumer price index (CPI). The official rural poverty line for 2006 was established by the Government of Kenya (Kenya shillings (KES) at 1562 per adult equivalent per month. In nominal Kenyan shillings per capita per month, the resulting poverty lines were: 1009 (2000), 1336 (2004), 1629 (2007), 2144 (2010). Since these poverty lines are expressed in per adult equivalent terms per month, we also converted the annual household income into income per adult equivalents per month.

We measure income diversification as the Simpson index, computed as:

$$D_i = 1 - \sum_{1=1}^{N} \left(s_{j,i}\right)^2 , \qquad (2),$$

where $D_i \in [0,1)$ is the diversification index for household *i*, $s_{j,i}$ refers to share of income source *i for* household *j*, $\sum_{1=1}^{N}\left(s_{j,i}\right) = 1$, and N is the total number of income sources for household i. The second term on the right side of equation (2) is Herfindahl index (HI) of concentration. Also known as the Herfindahl-Hirschman index, the HI has been extensively used by economists (e.g. Hirschman, 1964) to analyze the extent of competition among firms in an industry, calculated in terms of market shares. The index assigns a heavier weight to firms with more market power, and can thus be used as an indicator of the concentration of sales in analysis of monopolistic or oligopolistic behavior in anti-trust lawsuits.

In our analysis, households with more diversified income sources have a lower Herfindahl index and higher index of diversification, and vice-versa. A value of zero for D_i implies complete specialization in one source of income while a value towards one indicates that a household is highly diversified in terms of income sources. Income sources were categorized into five: crop, livestock, businesses & informal labor activities, and salary and remittances.

Summary statistics on dependent and outcome variables is presented in Table 1a.

3.2. Econometric Models

Participation in the maize market by the smallholder farmers is non-random. Non-randomness poses a well-known dilemma of missing data, with potential for selection bias. Simply stated, our research interest is the average effect of participating in the market on the participants, or the average effect on the 'treated' (ATT), which can be written as [$E(Y_{i1} - Y_{i0} | M_i = 1)$], where $M_i = 1$ if the i^{th} household participates and 0 otherwise. We can observe the outcome for the i^{th} household (Y_{i1}) if it participates in the maize market, but not the outcome (Y_{i0}) if it does not. Likewise, we observe non-participants only when they do not participate. Thus, the counterfactual state is observed for neither group.

We can estimate the average effects of participation by comparing outcomes between participants and non-participants, but there may be systematic differences among farmers that explain why some choose to sell maize in the market and others do not. Systematic difference would generate a 'selection bias' in our estimates of the effects of market participation. Selection bias can result from intentional or non-random selection of participants by a project or program, or in the case of a decision to market a crop, 'self-selection' by farmers who volunteer to participate. In Kenya, there are strong a priori reasons to believe that farmers who sell large volumes of maize have more assets, better proximity to markets, and more liquidity than those who do not. These are observable differences. Other unobservable differences are intrinsic, such as entrepreneurial skills.

Since the 1970s, propensity score matching (PSM) has been widely applied in evaluation studies as a means of generating a treatment and control group that mimics a randomized experiment (e.g., Dehijia and Wahba 2002; Jalan and Ravallion 2003; Rubin 2006). PSM generates a comparison sub-sample of non-participants (control group) that has ‘similar’ observable characteristics (covariates) as the participants (treatment group). To resolve the multi-dimensionality problem of multiple relevant covariates, Rosenbaum and Rubin (1983) proposed a uni-dimensional scalar function of the covariates, estimated most frequently with a logit or probit model. The predicted value from this regression, known as the propensity score, is interpreted as the probability of group membership conditional on the covariates. Here, we use a probit model and assume that, conditional on the vector of covariates, participation in the maize market as a seller is independent of outcomes. We use both nearest neighbour matching and kernel-based matching (Dehijia and Wahba, 2002) as an internal (within method) robustness check.

After matching on propensity scores and obtaining participants and non-participants that are similar in their observable characteristics, we estimate the impact of market participation on the participants (ATT) as the average of the differences in outcome variable between the matches as ATT= $E(Y_{i1} - Y_{i0})$. Observations outside the common support, which refers to the shared statistical distribution of the two groups, are eliminated before estimating ATT.

We evaluate the performance of the matching exercise by conducting three diagnostic tests: the balancing property test (Rosenbaum and Rubin, 1983) ; comparison of the value of pseudo R^2 before and after matching (Sianesi, 2004); and the likelihood ratio test for joint significance of the covariates before and after matching (Sianesi, 2004). Since PSM controls for selection bias only on the basis of observable covariates, following Dillon (2011) and Asfaw et al. (2012), we also conduct Rosenbaum tests (Rosenbaum 2002) to gauge the sensitivity of the estimated impact of market participation to hidden bias.

The PSM is based on the assumption that observable covariates account for the decision on whether or not to participate in the market. However, the participation decision, and selection bias, can also be caused by unobserved attributes. This suggests that market participation may be potentially endogenous, and PSM cannot correct for the endogeneity. As an external validity check, we apply endogenous switching regression (ESR) to control for the potential endogeneity of market participation decision. The ESR takes into account both observed and unobserved attributes in estimation of treatment impacts. A second advantage is that this approach permits simultaneous estimation of the participation decision and outcome equations for participants and non-participants (including marginal effects of covariates in each equation) and the calculation of actual and counterfactual expected values of outcome variables for both groups. This aspect of ESR deepens our comprehension of heterogeneity in population subgroups.

Several recent studies have applied ESR technique to address endogeneity problems in modeling effects of adopting agricultural innovations in smallholder agriculture (e.g. Asfaw et al. 2012; Di Falco and Veronesi 2013; Läpple et al. 2013; Negash and Swinnen 2013). The ESR model is a special case of an endogenous switching model (Maddala and Nelson 1975). In our case, the model is expressed by the following three equations:

$$M^{*} = Z\alpha + u, \quad with\ M = \begin{cases} 1 & if \quad M^{*} > 0 \\ 0 & otherwise \end{cases} \tag{3}$$

$$Regime\ 1: \quad Y_1 = X_1\beta_1 + \varepsilon_1 \ \text{ if } M = 1 \tag{4}$$

$$Regime\ 2: \quad Y_0 = X_0\beta_0 + \varepsilon_0 \ \text{ if } M = 0 \tag{5},$$

where M^* is the latent variable for market participation, M is the observed binary variable for market participation, and Z is a vector of observable characteristics hypothesized to influence participation decision, α is a vector of parameters to be estimated and u is a vector of error terms, which is normally distributed with mean zero and constant variance. Regime 1 and Regime 2 are outcome equations for participants and non-participants, respectively. Y_j is the outcome variable, X_j is a vector of observable household characteristics hypothesized to influence the outcome variable, β_j is a vector of parameters to be estimated and $\boldsymbol{\varepsilon_j}$ is the error term, and *j=1* for participants and *0* for non-participants.

A key assumption of the ESR is that the error terms in the three equations, i.e. $\varepsilon_1, \varepsilon_0$ and $u,$ have a trivariate normal distribution with mean zero and covariance matrix of the form:

$$\Omega = \begin{bmatrix} \sigma^2_{\varepsilon_1} & . & \sigma_{\varepsilon_1 u} \\ . & \sigma^2_{\varepsilon_0} & \sigma_{\varepsilon_0 u} \\ . & . & 1 \end{bmatrix} \tag{6},$$

where Ω is the covariance matrix, $\sigma^2_{\varepsilon_1}$ and $\sigma^2_{\varepsilon_0}$, respectively, are the variances of the error terms in the outcome equations (4) and (5), $\sigma_{\varepsilon_1 u}$ and $\sigma_{\varepsilon_0 u}$ are the covariance of the error terms in the outcome equations and the participation equation, and 1 is the error term of the participation equation. The variance of the error term in the participation equation is set to 1 because it can be estimated only up to a scalar factor (Dutoit, 2007). A household can be observed only in either of the regimes but not in both, making the covariance between $\sigma^2_{\varepsilon_1}$ and $\sigma^2_{\varepsilon_0}$ undefined (Maddala, 1983). The presence of selection bias implies that the expected values of the error terms in the outcome equations (4) and (5) are non-zero conditional on market participation. The conditional expectations of the error terms are given by:

$$E(\varepsilon_{1i}|M_i = 1) = \sigma_{\varepsilon_1 u}\,\frac{\phi(Z_i\alpha)}{\Phi(Z_i\alpha)} = \sigma_{\varepsilon_1 u}\lambda_{1i} \tag{7}$$

$$E(\varepsilon_{0i}|M_i = 0) = \sigma_{\varepsilon_0 u}\,\frac{-\phi(Z_i\alpha)}{1-\Phi(Z_i\alpha)} = \sigma_{\varepsilon_0 u}\lambda_{0i} \qquad (8),$$

where ϕ is the density function of the standard normal and Φ its cumulative distribution function, $\lambda_{1i} = \frac{\phi(\boldsymbol{Z_i\alpha})}{\Phi(\boldsymbol{Z_i\alpha})}$ and $\lambda_{0i} = \frac{-\phi(\boldsymbol{Z_i\alpha})}{1-\Phi(\boldsymbol{Z_i\alpha})}$. Following Maddala (1983), equations (7) and (8) can thus be written as:

$$Y_1 = X_1\beta_1 + \sigma_{\varepsilon_1 u}\lambda_{1i} + \nu_1 \text{ if } M = 1 \tag{9}$$

$$Y_0 = X_0\beta_0 + \sigma_{\varepsilon_0 u}\lambda_{0i} + \nu_0 \text{ if } M = 0 \qquad (10),$$

where $\nu_1 = \varepsilon_1 + \sigma_{\varepsilon_1 u}\lambda_{1i}$ and $\nu_0 = \varepsilon_0 + \sigma_{\varepsilon_0 u}\lambda_{0i}$ are the new error terms with zero conditional mean.

Estimation of equations (9) and (10) by OLS would lead to biased and inconsistent estimates of the parameters β_j because doing so would mean omitting the variable $\sigma_{\varepsilon_j u}\lambda_j$ (Lapple et al., 2013). Therefore, equations (8), (9) and (10) are estimated simultaneously by full information maximum likelihood (FIML) (Lokshin and Sajaia 2004; Di Falco and Veronesi 2013; Läpple et al. 2013), which yields consistent estimates. The logarithmic likelihood function is of the following form:

$$\ln L_i = \sum_{i=1}^{n} \left(M_i \left[ln\, \phi \left({\varepsilon_{1i}}/{\sigma_{\varepsilon_1}} \right) - \ln \sigma_{\varepsilon_1} + \ln \Phi\,(\psi_{1i}) \right] + (1 - M_i) \left[\ln \phi \left({\varepsilon_{0i}}/{\sigma_{\varepsilon_0}} \right) - \ln \sigma_{\varepsilon_0} + \ln\big(1 - \Phi(\psi_{0i})\big) \right] \right) \quad (11),$$

where, $\psi_{1i} = \left(\mathrm{Z}_i \alpha + \rho_j \varepsilon_{ji} / \sigma_{\varepsilon_j} \right) \Big/ \left(\sqrt{1 - \rho_j^2} \right)$, *j=0,1.* ρ_1 is the correlation coefficient between the error term ε_1 and u, and ρ_0 the correlation coefficient between the error term ε_0 and u.

Estimating the system of equations requires that an exclusion restriction be imposed for the model to be identified. We accomplish this by including in the participation equation three variables that influence participation decision but have no effect on the outcome variable (Wooldridge, 2010). We then test for validity of these variables as instruments by testing if the variables are significant in the participation equation but are insignificant in the outcome equation for non-participants (Di Falco et al. 2013).

After estimating the parameters by FIML, we calculate conditional and unconditional expectations of outcome variables for participants and non-participants as follows:

$$\mathrm{E}(\, Y_{1i} | M_i = 1) = X_{1i}\beta_1 + \sigma_{\varepsilon_1 u}\lambda_{1i} \quad (12)$$

$$\mathrm{E}(\, Y_{0i} | M_i = 0) = X_{0i}\beta_0 + \sigma_{\varepsilon_0 u}\lambda_{0i} \quad (13)$$

$$\mathrm{E}(\, Y_{0i} | M_i = 1) = X_{1i}\beta_0 + \sigma_{\varepsilon_0 u}\lambda_{1i} \quad (14)$$

$$\mathrm{E}(\, Y_{1i} | M_i = 0) = X_{0i}\beta_1 + \sigma_{\varepsilon_1 u}\lambda_{0i} \quad (15)$$

Equations (12) and (13) compute the actual expected value of the outcome variable observed in the sub-samples for participants and non-participants, respectively. Equations (14) and (15) compute the expected values of the outcome variable in the counterfactual scenarios for participants and non-participants, respectively. The effect of participation on the participants, i.e. average treatment effect on the treated (ATT) is the difference between (12) and (14):

$$ATT = \mathrm{E}(Y_{1i}|M_i = 1) - \mathrm{E}(Y_{0i}|M_i = 1) = X_{1i}(\beta_1 - \beta_0) + (\sigma_{\varepsilon_1 u} - \sigma_{\varepsilon_0 u})\lambda_{1i} \qquad (16)$$

The effect of participation on non-participants, or the average treatment effect on the untreated (ATU), is computed as the difference between (15) and (13):

$$ATU = \mathrm{E}(Y_{1i}|M_i = 0) - \mathrm{E}(Y_{0i}|M_i = 0) = X_{0i}(\beta_1 - \beta_0) + (\sigma_{\varepsilon_1 u} - \sigma_{\varepsilon_0 u})\lambda_{0i} \qquad (17)$$

Following Di Falco and Veronesi (2013), we also compute heterogeneity effects for both participants and non-participants as follows:

$$AHT = \mathrm{E}(Y_{1i}|M_i = 1) - \mathrm{E}(Y_{1i}|M_i = 0) = (X_{1i} - X_{0i})\beta_1 + \sigma_{\varepsilon_1 u}(\lambda_{1i} - \lambda_{0i}) \qquad (18)$$

$$AHU = \mathrm{E}(Y_{0i}|M_i = 1) - \mathrm{E}(Y_{0i}|M_i = 0) = (X_{1i} - X_{0i})\beta_0 + \sigma_{\varepsilon_0 u}(\lambda_{1i} - \lambda_{0i}) \qquad (19),$$

where *AHT* and *AHU* are base heterogeneity for participants and non-participants, respectively. A positive value for *AHT* means that market participants would perform better than non-participants even if non-participants participated in the market. In the same way, a negative value for *AHU* means participants would perform better than non-participants even if they had not participated. These measures are important in understanding whether or not there is some heterogeneity between the groups that may make them different irrespective of their maize market participation status.

In all regressions, we use the Correlated Random Effects (CRE) estimation to control for time-invariant heterogeneity as proposed by Mundlak (1978) and Chamberlain (1984). As discussed above, to explore observed heterogeneity within maize-growing households in terms of market position, we estimate the impact of selling maize with respect to five subgroups: 1) all maize sellers (sell, purchase and sell) compared to all maize growers; 2) maize growers who sell only relative to all maize growers; 3) maize growers who purchase and sell relative to all maize growers; 4) maize sellers who also sell dairy relative to dairy sellers who do not sell maize; 5) maize sellers who also sell industrial crops relative to sellers of industrial crops who do not sell maize.

3.3. Data Source

We use a five-year (1997, 2000, 2004, 2007, and 2010) panel survey dataset of rural households in the major maize-growing areas of Kenya, collected by Tegemeo Institute and Michigan State University. The sampling frame for the panel survey was prepared in 1997 in consultation with the Kenya National Bureau of Statistics (KNBS). The sampling process is described in detail by Argwings-Kodhek et al (1999). Specific topics addressed in the survey vary over the years but contain a core set of variables that document aspects of farm household livelihoods. Detailed information on crops and livestock production and marketing, as well as off-farm earnings, were collected, enabling computation of household incomes and diversification.

Our analysis is based on a balanced panel of 1,243 households that covers the 1999/00, 2003/04 2006/07 and 2010/11 cropping years (hereafter referred to as 2000, 2004, 2007, and 2010, respectively). The 1997 data, collected in the first year of research, do not include all variables of interest.

We also exploit secondary, geo-referenced database to link farm household decision to higher geographical scales of analysis. We extracted data on population estimates and travel time from the Global Rural-Urban Mapping Project (GRUMP)[1]. We computed village population densities by dividing the population estimate by village land area. The population estimate is the average population count per pixel for all pixels within a 10km radius, while travel time is the time in hours it takes to travel from a village to a town with 75,000 plus inhabitants.

3.4. Explanatory Variables

The definition and summary statistics of both the outcome and explanatory variables are presented in Tables 1a and 1b. An estimated 42% of all maize growing households surveyed sold maize in the overall sample, pooled over the four years studied (Table 1a). Slightly over one quarter (28%) sold but did not purchase maize—we refer to these as commercialized maize growers. Over the years, 14% both sold and purchased maize. The subgroup of smallholders who sold both maize and dairy products is 26%, while those who sold both maize and industrial crops are smaller in share, averaging 13% over the pooled sample.

We hypothesize that female and male headed households differ in terms of their capacity to participate in the maize market (Table 1b). Virtually all female heads are widows. We measure human capital by average years of education of adults in a household and labor quantity by number of young and mature adults. We include a dummy variable indicating whether a household experienced prime-age mortality between one survey year and another to capture effects of a shock. We have no hypothesized direction of effect for this variable; recent findings

[1] The population and travel time data from GRUMP was provided by Jordan Chamberlain.

reported by Chapoto et al. (2012) differ from findings in earlier years reported by Yamano and Jayne (2004). Ownership of radio is expected to facilitate access to market information while transport equipment (bicycle, motor cycle, motor vehicle) are expected to reduce the cost of transporting output to the market. Land size and value of farm assets are included to represent resource endowments for production. We include variables measuring the proportion of village households that belong to farmer groups or cooperatives and the share that received credit as indicators of social capital and credit access, respectively. Farmers receive production and/or marketing services from the groups and cooperatives through information sharing and collective marketing. Cooperatives are mostly common among coffee and tea farmers where they engage mainly in collective marketing. Main season rainfall amount and dummies for the agroecological potential of the production zone are included to control for agroecological effects. We also include year dummies to control for temporal effects.

Reflecting the Barrett (2008) concept of interscale market linkages and Staal et al. (2008), we use travel time to town with 75000 plus inhabitants as an indicator of market access, district median maize price and village population density as exclusion variables to identify the selection equation in the endogenous switching regression. For our purposes, these variables are suitable as instruments in part because they are measured at a different scale of analysis and in part because they are pre-determined and derived form a different dataset.

4. Results

4.1 Descriptive Statistics

Table 2 illustrates the range of the variation among Kenyan smallholders in terms of maize quantities sold. Considering the full sample of farmers in all years studied (2000-2010), the average amount sold was 1.8 mt, but ranged from 4.5 to 103,500 kgs. The share of production sold averaged 42%. The farmers we describe as "commercialized" sold 2.4 mt on average, but again amount sold was as small as 7.8 kg and the share of production sold averaged 46%. As expected, mean amount sold by maize farmers who both sold and purchased maize was less than half the overall mean (0.8 mt) and the share of marketed production was also lower (35%). As is the case in Table 1, statistics for sellers of both maize and dairy products are roughly similar to those of commercialized maize growers, while those that represent sellers of both maize and industrial crops parallel the group that both sells and purchases maize. Statistics shown in Table 1 and Table 2 indicate that over the past decade, smallholder maize production in Kenya remained oriented toward home consumption, with under half of the producers engaged in marketing and the share of production marketed also under half. The statistics presented in Table 3 show that maize sellers and non-sellers generally differ significantly in outcome indicators and covariates. For example, market participants have significantly higher income and lower poverty rates and have more diversified income sources than non-participants. They are also more educated on average and are more endowed in land and other physical assets. Clearly, comparing the two groups in terms of the outcome indicators when the distributions of their observable characteristics are significantly different could be misleading. The data support the need to utilize other statistical approaches. We address this problem in the next section through PSM and endogenous switching regression, following the econometric strategy outlined above.

4.2. Propensity Score Matching

The probit equation that estimates propensity scores is presented in Table 4[2]. Assets, land size and being in medium or high agricultural potential significantly increase the probability of participation in the maize market, holding other factors constant. Probability of participation also rises with population densities, but also with travel time to towns with over 75000 inhabitants. This latter effect, though counterintuitive, is consistent with the findings of other studies conducted in Kenya. The density of maize traders in villages has risen over time in the major maize-producing areas of the country, and the overall distance to maize selling points has declined (Chamberlin et al. 2009). The finding also reflects the distribution of population in Kenya, where higher potential maize-growing areas are actually more remote with respect to major towns than lower potential maize-growing regions. On average, higher maize prices are associated with a lower propensity to sell. This is consistent with the non-separable model of the agricultural household, which predicts that the higher the market price of a food staple in a locality, the greater the incentive for net-consuming households to withhold it from the market in order to avoid purchase later in the season.

Only 17 observations fell outside of the common support when participants and non-participants were matched with a caliper size of 0.05, representing 25% of the sample standard deviation of propensity scores. Distributions are portrayed graphically in Appendix A (Figure 1A), also illustrating the similarity in results between nearest neighbour (5 neighbors) and kernel algorithms for matching. Off-support observations were excluded from subsequent analyses. Tests of balance between the two groups provide strong evidence that matching participants and non-participants on propensity scores eliminates a large share of the bias that results from comparing univariate outcomes between the two groups (Table 1A). That is, difference-of-mean

[2] We present regression results only for the overall sample, and not for the various sub-groups of maize growers

tests between the two groups are no longer statistically significant. The Pseudo-R^2 fell from 0.144 to 0.003 after matching, while the value of the log-likelihood ratio declined from 950.74 to 19.05 (the p-value of the Chi-squared statistic rose from 0.000 to 0.518).

The ATTs for all subsamples of maize sellers and all outcome indicators are shown in Table 5. Overall, participation in the maize market has a significant impact on household income and poverty, raising income by 37 percentage points on average and reducing poverty headcount and gap by nine and five percentage points, respectively. Poverty severity also declines significantly when maize farmers sell their crop, by three percentage points. These results generally hold when the kernel-based matching algorithm is employed, attesting to robustness of the results to matching algorithm.

Looking at the sub-category of commercialized growers (i.e. those who sell but do not purchase), we see a larger and significant impact of market participation on income and poverty. Participation raises income by 43 percentage points, and lowers poverty head count and gap by 12 and six percentage points, respectively. On the other hand, the impact of market participation on income and poverty among maize farmers who remain more subsistence-oriented (both sell and purchase) is nearly half that of commercialized maize farmers. For this subgroup, income and poverty head count diminishes to 23 and five percentage points, respectively. However, the impact on poverty gap also diminishes to four percentage points at the mean.

For maize growers who sell both maize and dairy products, participation in the maize market raises income by 33 percentage points and reduces poverty head count and gap by seven and three percentage points, respectively. Poverty severity also reduces by two percentage points. The magnitudes of effects are similar for maize sellers who also sell cash crops (coffee, tea, and/or sugar cane), with an average income gain of 33 percentage points, although the effect on poverty headcount and poverty gap is nearly twice as great at the mean. The poverty headcount, gap and severity decrease by 13, six and three percentage points, respectively when maize farmers who sell industrial crops also sell maize.

Participation in the maize market has no significant impact on the diversification of income sources among any of the subgroups. This finding may reflect the persistent centrality of maize in livelihood strategies of most Kenyan smallholders, but also the fact that most Kenyan smallholders have diversified across crops and farm-nonfarm income sources.

The sensitivity analysis results (Γ) based on the Rosenbaum bounds are presented in the last column of Table 5. The value of Γ indicates how sensitive the estimated impact is sensitive to hidden bias, and shows the critical level at which the impact estimate may be questioned. A value of Γ close to 1 indicates that the impact estimate is highly sensitive to hidden bias while a larger value of Γ indicates less sensitivity of the estimate to hidden bias. We observe considerable variation of robustness to hidden bias by the four outcome variables across the subgroups of maize sellers. Poverty severity, gap and headcount ration in that order are the least sensitive to hidden bias. For income, the critical level at which we would question our conclusion about the impact of market participation ranges between 1.5 and 2.4, while for income diversification our conclusion about impact of market participation would be questioned for Γ values of 1.1-1.2, suggesting our conclusion is highly sensitivity to hidden bias. Results from ESR, which addresses the problem of hidden bias, are discussed next.

4.3. Endogenous Switching Regression

Diagnostic tests confirm that the estimated coefficients of the three instrumental variables (travel time, district median maize price and village population density) are jointly insignificant ($F_{3,2763}$, p-value= 0.451) in the income equation for non-participants (regression shown in Table B1) while they are in fact individually significant in the market participation equation (see Table 4[3]). This supports their use to identify the outcome equations.

For maize sellers and non-sellers, we estimate the expected income and income diversification index under actual and counterfactual conditions as explained in equations (12) – (15). We then use the estimated expected income to compute poverty headcount ratio, gap and severity under the actual and counterfactual conditions for subgroups, because these variables are censored. The average effects of market participation on income, poverty and income diversification are then computed according to equations (16)-(19).

The estimated income equation, including significance and magnitude of marginal effects of hypothesized determinants, differs (Wald test, at 8% overall significance) between maize sellers and non-sellers (Table 6a[4]). This is apparent in the higher negative effect of female headship, which is offset by the positive effect of education, among non-participants. Membership in farmer cooperatives or associations also has a larger effect among non-sellers than sellers. By contrast, among maize sellers, location in the high potential zone for maize production is a highly significant determinant of variation in net income, as are the value of assets and land size. The number of mature adults (labor supply) in the household also has a greater positive effect on net income among sellers than non-sellers. Ownership of radios and transport vehicles raises income for both groups. The estimated income diversification equations for maize sellers and non-sellers is presented in Table 6b.

[3] We present regressions only for the overall sample and not for the sub-groups of maze growers
[4] We present regressions only for the overall sample and not for the sub-groups of maze growers

Estimated impacts of selling maize, as generated by ESR, are presented in Table 7a for overall sample of maize producers and Tables 7b and 7c for the four sub-samples. Among all maize producers combined, ESR results also demonstrate that selling maize has positive and significant impacts on income and poverty reduction (Table 7a). Considering all maize growers, households that sold maize would have earned 49 percentage points less had they not sold maize; their poverty headcount, gap and severity would have increased by 22, nine and five percentage points, respectively. Had non-participants chosen to sell maize, they would have increased their income by 17 percentage points and their poverty headcount, gap and severity would have declined by eight, four and two percentage points. We also observe a positive and significant impact of maize sales on diversification of household income sources, which was not evident in the PSM results. Heterogeneity effects show that even if participants had chosen not to sell maize, they would have earned higher income and benefited from lower poverty levels (by all the three measures of poverty) than non-participants, though they would have been less diversified in terms of income sources. In parallel, even if they had sold maize, non-participating households would have earned less income from more diverse source, and remained poorer than those who actually sold maize. These findings attest to existence of underlying differences among the maize-producing households that persist regardless of participation in the maize market.

Considering the sub-sample of commercialized maize producers, selling maize increased their income by 59 percentage points and their poverty headcount, gap and severity reduced by 27, 11 and six percentage points, respectively (Table 7b). A positive and significant impact of maize sales on diversification of household income sources is also observed for this group. The impact of maize sales on income and poverty diminishes for the sub-group of maize growers that both sold and purchased maize. Selling maize increased their income by 32 and reduced poverty headcount by 31 percentage points. Poverty gap and severity reduced by 17 and four percentage points, respectively. These results indicate the differential impact of exclusive and partial commercialization of maize on smallholder welfare.

The impact of maize sales on income and poverty for maize selling households who also sold dairy or industrial crops (coffee/tea/sugarcane) is presented in Table 7c. For maize sellers who also sold dairy, maize sales increased income by 45 percentage points. Poverty head count, gap and severity reduced by 19, six and three percentage points, respectively. Maize sales also had a positive and significant impact on income diversification. Compared to other dairy-selling households that did not sell maize, heterogeneity effects show that maize selling households would have had higher income, lower poverty levels and less diverse income sources even if those households had sold maize, suggesting existence of heterogeneous differences between maize sellers and non-sellers that make sellers better of within this group of households. These results also generally persist in the sub-group of maize producers who also sold industrial crops, where we observe maize sales having a significant impact on income, poverty and diversification of income sources, and maize selling households doing better in terms of income and poverty than their non-selling counterparts even if these had sold maize.

5. Conclusion

High poverty rate persists in rural Kenya, where farming households continue to depend on agriculture for food and income, despite economic growth. Maize is grown by most households in the major maize-growing areas of Kenya, but in the 2000-2010 period covered by our panel data, only about 42 percent entered the maize market to sell at least a portion of their harvest and only about a quarter sold but did not purchase maize. The maize-growing smallholder population is heterogeneous and diversified, with about one quarter selling both maize and dairy products and about 1 in 8 selling maize as well as industrial crops.

In this paper, we have used a combination of propensity score matching (PSM) and endogenous switching regression (ESR) to evaluate the effects of maize sales on income, poverty, and the diversity of income sources among smallholder maize farmers. Using both methods provides a robustness or validity check. The underlying conceptual framework, which dictates the exogenous covariates in the regression models, is the non-separable case of the agricultural household model. We measure poverty according to the Foster-Greer-Thorbecke index and income diversification with the Simpson index, which is related to (1 minus) the Hirschman index of concentration.

We have addressed their heterogeneity in several ways. First, we defined maize selling households according to five segments: 1) all maize sellers; 2) farmers who sell maize but do not purchase it; 3) farmers who both sell and purchase maize; 4) maize sellers who also sell dairy products; and 5) maize sellers who also sell industrial crops. Second, unlike PSM, ESR controls for the potential endogeneity of market participation decision through simultaneous estimation of the market participation decision and household income equations for participants and non-participants. The method enables us to calculate not only the ATT, but also the average treatment effect on the untreated (ATU) and treatment heterogeneity (ATH) of market participation. The treatment heterogeneity compares the effect of participation for households that actually participated and those that did not if they had (ATT-ATU). Finally, in all regressions, we utilize the Mundlak-Chamberlin device (Correlated Random Effects approach) to control for time-invariant heterogeneity.

Results from the PSM have shown that in the overall, participation in the maize market has a significant impact on household income and poverty. Across the segments of maize growers, the impact is larger for commercialized growers and those that also sell dairy products and smaller for more subsistence-oriented growers and those that also sell industrial crops. Participation in the maize market, however, has no significant impact on diversification of income sources among any of the subgroups, reflecting the persistent centrality of maize in livelihood strategies of most Kenyan smallholders, but also the fact that most Kenyan smallholders have diversified across crops and farm-nonfarm income sources.

After controlling for hidden selection bias through ESR, the estimated impact of market participation on income and poverty is larger for the overall sample of maize sellers and for each of the sub-groups. We also observe a positive and significant impact of selling maize on diversification of household income sources, which was not evident in the PSM results. Heterogeneity effects attest to existence of underlying differences among the maize-producing households that make sellers better off than non-sellers regardless of participation in the maize market.

We conclude that maize is not only the most important staple food in Kenya but is a major income-generating enterprise for smallholders who produce and sell. Sellers appear clearly advantaged relative to their counterparts who do not sell (although this depends on the other cash-earning farm enterprise), an indication that the smallholder maize market in Kenya may be concentrated in the hands of more well-off farmers. The study findings reinforce the call for interventions to expand the capacity of a broad base of smallholder maize farmers to produce for the market for a broader distribution these benefits. Increased smallholder maize market participation will not only raise incomes and contribute to a more widespread poverty reduction, as demonstrated here, but will also contribute to improving food security.

References

Argwings-Kodhek, G., Jayne, T., Nyambane, G., Awuor, T., Yamano, T., 1998. How Can Micro-level Household Survey Data Make a Difference for Agricultural Policy Making? Nairobi: Egerton University/Tegemeo Institute of Agricultural Policy and Development. Available at: http://www.tegemeo.org/viewdocument.asp?ID=28

Ariga, J. and T. S. Jayne. 2010. Private-sector responses to public investments and policy reforms: The case of fertilizer and maize market development in Kenya. In D. Spielman and R. Pandya-Lorch (eds.), *Proven Successes in Agricultural Development: A Technical Compendium to Millions Fed*. International Food Policy Research Institute (IFPRI), Washington, DC.

Asfaw, S., L. Lipper, T. J. Dalton, and P. Audi. 2011. Market participation, on-farm crop diversity and household welfare: micro-evidence from Kenya. *Environment and Development Economics* 17: 579-601.

Asfaw, S., Kassie, M., Simtowe, F., Lipper, L., 2012. Poverty Reduction Effects of Agricultural Technology Adoption: A Micro-evidence from Rural Tanzania. *Journal of Development Studies*. 48, 1288–1305.

Barrett, C.B. 2008. Smallholder market participation: Concepts and evidence from eastern and southern Africa. *Food Policy* 33: 299-317.

Byerlee, D., A. de Janvry, J. Van Wassenhove, and D Barry. 2009. Smallholders unite. *Foreign Affairs*. March/April 2009.

Caliendo, M. and Kopeinig, S., 2005. Some practical guidance for the implementation of propensity score matching, IZA Discussion Papers, No. 1588

Chamberlain, G. 1984. Panel Data. In Handbook of Econometrics, Vol. 2, ed. Z. Griliches, and M.D. Intriligator. Amsterdam, North Holland: Elsevier.

Chamberlin, J., and T.S. Jayne. 2009. Has Kenyan Farmers' Access to Markets and Services Improved? Panel Survey Evidence, 1997–2007. Working paper, Nairobi, Kenya: Tegemeo Institute of Agricultural Policy and Development.

Chapoto, A., L. Kirimi and S. Kadiyala. 2012. Poverty and prime-age mortality in Eastern and Southern Africa: Evidence from Zambia and Kenya. *World Development* 40(9): 1839-1853.

Collier, P. 2008. The politics of hunger. *Foreign Affairs* November/December.

Dehejia, R. H., and S. Wahba. 2002. Propensity Score Matching Methods for Non-experimental Causal Studies. *Review of Economics and Statistics* 84: 151-161.

Diao, X, J. Thurlow,. S. Benin and S. Fan (eds.). 2012. *Strategies and Priorities for African agriculture: Economywide Perspectives from Country Studies*. International Food Policy Research Institute (IFPRI), Washingon, DC.

Di Falco, S., Veronesi, M., 2013. How Can African Agriculture Adapt to Climate Change? A Counterfactual Analysis from Ethiopia. *Land Econ*. 89, 743–766.

Dillon, A., 2011. The Effect of Irrigation on Poverty Reduction, Asset Accumulation, and Informal Insurance: Evidence from Northern Mali. *World Dev*. 39, 2165–2175.

Dutoit, L.C., 2007. Heckman's Selection Model, Endogenous and Exogenous Switching Models, A survey.

Hazell, P. C. Poulton, S. Wiggins, and A. Dorward. 2007. The Future of Small Farms for Poverty Reduction and Growth. 2020 Discussion Paper 42. International Food Policy Research Institute (IFPRI), Washington, D.C.

Hirschman, A. O. 1964. The Paternity of an Index. *The American Economic Review* (American Economic Association) **54** (5): 761

International Livestock Research Institute (ILRI). 2011. Towards priority actions for Market Development for African Farmers: Proceedings of an International Conference. 13-15

May 2009, Nairobi, Kenya. AGRA (Alliance for a Green Revolution in Africa) and ILRI (International Livestock Research Institute), Nairobi, Kenya.

Jalan, J., R de Janvry A avallion, M., 2003. Estimating the Benefit Incidence of an Antipoverty Program by Propensity-Score Matching. *J. Bus. Econ. Stat.* 21, 19–30.

Key N, Sadoulet E. and de Janvry. 2000. Transaction costs and agricultural household supply response. *American Journal of Agricultural Economics* 82(2), 245–259.

Läpple, D., Hennessy, T., Newman, C., 2013. Quantifying the Economic Return to Participatory Extension Programmes in Ireland: an Endogenous Switching Regression Analysis: An Endogenous Switching Regression Analysis. *J. Agric. Econ*. 64, 467–482.

Lokshin, M., Sajaia, Z., 2004. Maximum likelihood estimation of endogenous switching regression models. *Stata J*. 4, 282–289.

Maddala, G.S. and Nelson, F.D., 1975. Switching regression models with exogenous and endogenous switching. Proceeding of the American Statistical Association (Business and Economics Section), pp. 423–426.

Mathenge, M., Place, F., Olwande, J., and Mitoefer, D. 2010. Participation in Agricultural Markets among the Poor and Marginalized: Analysis of Factors Influencing Participation and Impacts on Income and Poverty in Kenya. Unpublished Study Report Prepared for the FORD Foundation.

Mundlak, Y. 1978. On the pooling of time series and cross section data. *Econometrica* 46: 69-85.

Negash, M., Swinnen, J.F.M., 2013. Biofuels and food security: Micro-evidence from Ethiopia. *Energy Policy* 61, 963–976.

Ngigi, M., M.A. Ahmed, S. Ehui and Y. Assefa. 2010. Smallholder Dairying in East Africa. In S. Haggblade and P. B.R. Hazell, Successes in African Agriculture: Lessons for the Future. International Food Policy Research Institute and Johns Hopkins University Press, Washington, DC, and Baltimore.

Rao, E.J.O. and Qaim, M. 2010. Supermarkets, farm household income and poverty: Insights from Kenya. Contributed Paper presented at the Joint 3rd African Association of Agricultural Economists (AAAE) and 48th Agricultural Economists Association of South Africa (AEASA) Conference, Cape Town, South Africa, September 19-23, 2010.

Ravallion, M. (1994). Poverty comparisons. *Fundamentals in Pure and Applied Economics* Volume 56. Harwood Academic Publishers.

Renkow, M., Hallstrom, D.G., Karanja, D.D., 2004. Rural infrastructure, transactions costs and market participation in Kenya. Journal of Development Economics 73 (3), 349–367

Republic of Kenya, 2008. Well-Being in Kenya: A Socio-Economic Profile. The Kenya National Bureau of Statistics

Rosenbaum, P.R., Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70, 41–55.

Rosenbaum, P.R., 2002. *Observational Studies*. Second Edition. New York, NY: Springer.

Rubin, D.B. 2006. *Matched Sampling for Causal Effects*. Cambridge University Press, Cambridge.

Sadoulet, E. and A. de Janvry. 2012. Subsistence farming as a safety net for food-price shocks. In M. Cohen and M. Smale (eds.), Global Food-Price Shocks and Poor People: Themes and Case Studies. Development in Practice Books, Routledge, Oxon, UK.

Sianesi, B., 2004. An evaluation of the Swedish system of active labour market programs in the 1990s. *The Rev. Econ. and Stat*., 86(1): 133–155.

Smale, M. and J. Olwande. 2014. Demand for Maize Hybrids and Hybrid Change on Smallholder Farms in Kenya. *Agricultural Economics* 45: (in press). Article first published online : 23 DEC 2013 DOI: 10.1111/agec.12095

Staal, S.J., A. Nin Pratt, and M. Jabbar. 2008. Dairy development for the resource poor. Part 2: Kenya and Ethiopia dairy development case studies. Pro-Poor Livestock Policy Initiative Working Paper 44-2. International Livestock Research Institute (ILRI), Nairobi, Kenya.

Suri, T. 2011. Selection and Comparative Advantage in Technology Adoption. *Econometrica* 79 (1): 159-209.

Thurlow, J., J. Kiringai and M. Gautum. 2012. Kenya. In X. Diao, J. Thurlow,. S. Benin and S. Fan (eds.), *Strategies and Priorities for African agriculture: Economywide Perspectives from Country Studies*. International Food Policy Research Institute (IFPRI), Washington, DC.

Wooldridge, J. M., 2010. Econometric Analysis of Cross Section and Panel Data, Second Edition. MIT Press, Cambridge, MA.

World Bank. 2007. *Agriculture for Development. World Development Report 2008*. Washington, D.C., the World Bank.

Yamano, T., and T.S. Jayne. 2004. Measuring the Impacts of Working-age Adult Death on Rural Households in Kenya. *World Development* 32: 91-119.

Table 1a: Summary statistics: dependent and outcome variables

Variable	Construction	2000 (N=1190)		2004(N=1222)		2007(N=1221)		2010(N=1212)		Pooled(N=4845)	
		Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Sell maize	Household sold maize (1=yes)	0.40	0.49	0.45	0.50	0.47	0.50	0.37	0.48	0.42	0.49
Sell maize, no purchase	Household sold but did not purchase maize (1=yes)	0.24	0.43	0.30	0.46	0.33	0.47	0.25	0.43	0.28	0.45
Sell and purchase maize	Household sold and purchased maize (1=yes)	0.16	0.36	0.16	0.37	0.14	0.35	0.12	0.32	0.14	0.35
Sell both maize & dairy	Household sold maize and dairy products	0.24	0.42	0.28	0.45	0.31	0.46	0.23	0.42	0.26	0.44
Sell both maize and industrial crops	Household sold maize and industrial crops	0.14	0.35	0.13	0.33	0.14	0.35	0.12	0.33	0.13	0.34
Income	Sum of net income from crops, livestock, salaries, remittance, business and informal labor activities (‘000 KES)	161,917	191,491	170,256	196,535	184,226	194,287	278,765	390,157	198,873	261,711
Poverty head count	FGT index (see text)	0.26	0.44	0.30	0.46	0.32	0.46	0.31	0.46	0.30	0.46
Poverty gap	FGT index (see text)	0.11	0.22	0.13	0.24	0.12	0.22	0.12	0.22	0.12	0.22
Poverty severity	FGT index (see text)	0.06	0.15	0.07	0.17	0.06	0.14	0.06	0.15	0.06	0.15
Income diversification	Simpson index (see text)	0.47	0.17	0.48	0.16	0.51	0.16	0.49	0.16	0.49	0.16

Table 1b: Summary statistics: explanatory variables

Variable	Construction	2000 (N=1190)		2004(N=1222)		2007(N=1221)		2010(N=1212)		Pooled(N=4845)	
		Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Female head	Gender of household head (1=female)	0.12	0.32	0.20	0.40	0.24	0.42	0.27	0.44	0.21	0.40
Education	Education of adults (years)	7.14	2.86	7.25	2.96	7.20	3.02	7.58	3.01	7.29	2.97
Young adults	No. of adults 15-24 years	2.08	1.62	1.94	1.57	1.92	1.60	1.80	1.53	1.94	1.58
Mature adults	No. of adults 25-64 years	2.35	1.27	2.25	1.26	2.09	1.25	2.06	1.33	2.19	1.29
Mortality	Household experienced prime-age mortality (1=yes)	0.06	0.24	0.06	0.23	0.06	0.24	0.05	0.22	0.06	0.23
Radio	Ownership of radio (1=yes)	0.83	0.38	0.88	0.33	0.90	0.29	0.86	0.34	0.87	0.34
Transport equipment	Ownership of transport equipment (1=yes)	0.45	0.50	0.48	0.50	0.51	0.50	0.48	0.50	0.48	0.50
Assets	Value of household assets (KES)	42,431	76,699	172,809	348,350	222,122	404,755	260,217	433,305	175,079	356,267
Land	Land size (acres)	6.07	8.56	6.15	9.02	5.81	8.84	5.16	8.62	5.80	8.77
Group	Proportion of village households in farmer groups/cooperatives	0.79	0.19	0.76	0.22	0.75	0.22	0.71	0.24	0.75	0.22
Credit	Proportion of village households that received credit	0.48	0.29	0.33	0.32	0.52	0.28	0.56	0.26	0.47	0.30
Rainfall	Rainfall amount in main season (mm)	593	268	690	295	611	197	417	200	578	263
Medium potential	Medium potential dummy	0.24	0.43	0.25	0.43	0.25	0.43	0.25	0.43	0.25	0.43
High potential	High potential dummy	0.47	0.50	0.47	0.50	0.46	0.50	0.45	0.50	0.46	0.50
Travel time	Travel time to nearest town with 75000+ inhabitants (hours)	3.07	2.38	3.08	2.37	3.07	2.35	3.10	2.37	3.08	2.37

Variable	Construction	2000 (N=1190)		2004(N=1222)		2007(N=1221)		2010(N=1212)		Pooled(N=4845)	
		Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Maize price	District median maize grain price (log) (KES/kg)	12.57	1.22	13.02	1.59	12.63	1.39	20.87	1.61	14.78	3.82
Population density	Village population density (persons/km^2)	313	178	356	206	385	222	411	239	366	215

Source: Authors

Table 2: Quantity (kgs) and share (%) of production of maize sold, all years, by subgroup

Subgroup	Mean (kg)	Std. Dev.	Min	Max	% of production sold
Sell maize	1875	5223	4.50	103500	41.96
Sell maize, no purchase	2416	6287	7.75	103500	45.61
Sell and purchase maize	835	1457	4.50	14400	34.94
Sell both maize and dairy	2368	5791	4.50	103500	45.91
Sell both maize and industrial crops	803	1363	6.75	12750	35.59

Source: Authors

Table 3: Summary statistics of outcome and explanatory variables between participants and non-participants in unmatched sample

Variable	Maize sellers		Maize non-sellers		Difference in mean	t-statistic
	Mean	Std. Dev.	Mean	Std. Dev.		
Income	254,368	333,872	158,135	181,798	96232.9***	(12.860)
Poverty head count	0.21	0.41	0.36	0.48	-0.155***	(-11.81)
Poverty gap	0.07	0.17	0.15	0.25	-0.0798***	(-12.47)
Poverty severity	0.04	0.11	0.09	0.17	-0.0501***	(-11.54)
Income diversification	0.49	0.16	0.48	0.17	0.0123***	(2.580)
Female head	0.19	0.39	0.22	0.41	-0.0264***	(-2.25)
Education	7.75	2.83	6.96	3.02	0.795***	(9.300)
Young adults	2.09	1.66	1.82	1.52	0.268***	(5.830)
Mature adults	2.23	1.25	2.15	1.31	0.0784***	(2.100)
Mortality	0.05	0.23	0.06	0.24	-0.00409	(-0.61)
Radio	0.90	0.30	0.85	0.36	0.0571***	(5.850)
Transport equipment	0.55	0.50	0.43	0.49	0.122***	(8.450)
Assets	202,235	383,607	155,145	333,445	47089.8***	(4.550)
Land	8.05	11.62	4.15	5.27	3.898***	(15.670)
Group	0.74	0.22	0.77	0.22	-0.0301***	(-4.75)
Credit	0.42	0.29	0.51	0.31	-0.0879***	(-10.10)
Rainfall	625	243	544	273	80.90***	(10.680)
Medium potential	0.27	0.44	0.23	0.42	0.0357***	(2.840)
High potential	0.55	0.50	0.40	0.49	0.157***	(10.950)
Travel time	3.02	2.32	3.12	2.40	-0.103	(-1.49)
Maize price	14.14	3.76	15.25	3.79	-1.108***	(-10.09)
Population density	348	208	380	220	-31.53***	(-5.04)

t statistics in parentheses
* p<0.1, ** p<0.05, *** p<0.01

Table 4: Probit estimation results of propensity scores for maize market participation

Variable	Coefficient	P-value
Female head	0.034	0.508
Education	0.026*	0.092
Young adults	0.030	0.114
Mature adults	0.002	0.931
Mortality	0.025	0.756
Radio	0.072	0.256
Transport equipment	-0.054	0.226
Assets	0.041***	0.004
Land	0.263***	0.000
Group	0.275	0.112
Credit	-0.149	0.294
Rainfall	0.103	0.153
Medium potential	0.274***	0.001
High potential	0.390***	0.000
Year 2004 dummy	0.006	0.938
Year 2007 dummy	0.011	0.899
Year 2010 dummy	0.238	0.197
Travel time	0.038***	0.000
Maize price	-0.996***	0.001
Population density	0.002**	0.012
Constant	1.699	0.061
Observations	4845	
Log likelihood	-2788.466	
Prob > chi2	0.000	
Pseudo R2	0.155	

* p<0.1, ** p<0.05, *** p<0.01

Table 5: Average Treatment Effect (ATT) of maize market participation, by subgroup

Subgroup	N	Outcome variable	Participants	Non-participants	ATT	S.E.	t-statistic	Critical level of hidden bias (Γ)[a]
All maize sellers (Comparison group: all non-sellers)	Participants = 2034 Non-participants = 2794	Log of income	11.98	11.61	0.37***	0.05	7.02	2.0
		Poverty headcount	0.21	0.31	-0.10***	0.02	-4.98	1.8
		Poverty gap	0.07	0.13	-0.05***	0.01	-5.54	2.8
		Poverty severity	0.04	0.07	-0.03***	0.01	-5.22	3.6
		Income diversification index	0.49	0.49	0.01	0.01	0.74	1.2
Sell maize, no purchase (Comparison group: all non-sellers)	Participants = 1338 Non-participants = 2794	Log of income	12.18	11.75	0.43***	0.06	7.18	2.4
		Poverty headcount	0.13	0.25	-0.12***	0.02	-5.82	2.8
		Poverty gap	0.04	0.10	-0.06***	0.01	-5.82	4.3
		Poverty severity	0.02	0.05	-0.03***	0.01	-5.11	5.6
		Income diversification index	0.49	0.49	0.00	0.01	0.23	1.1
Sell and purchase maize (Comparison group: all non-sellers)	Participants = 701 Non-participants = 2794	Log of income	11.61	11.38	0.23***	0.06	3.86	1.5
		Poverty headcount	0.36	0.41	-0.05**	0.02	-2.19	1.1
		Poverty gap	0.14	0.17	-0.04***	0.01	-3.09	1.6
		Poverty severity	0.07	0.10	-0.03***	0.01	-3.25	2.0
		Income diversification index	0.49	0.48	0.01	0.01	0.86	1.2
Sell both maize & dairy (Comparison group: non-sellers of maize but sellers of dairy)	Participants = 1297 Non-participants = 1310	Log of income	12.27	11.94	0.33***	0.07	4.91	2.0
		Poverty headcount	0.12	0.19	-0.07***	0.02	-2.98	2.6
		Poverty gap	0.04	0.07	-0.03***	0.01	-3.46	3.6
		Poverty severity	0.01	0.03	-0.02***	0.01	-3.18	4.6
		Income diversification index	0.53	0.52	0.01	0.01	0.73	1.2
Sell both maize & industrial crops (Comparison group: non-sellers of maize but sellers of industrial crops)	Participants = 661 Non-participants = 1095	Log of income	12.19	11.86	0.33***	0.06	5.90	2.0
		Poverty headcount	0.11	0.24	-0.13***	0.02	-5.68	3.1
		Poverty gap	0.03	0.09	-0.06***	0.01	-5.78	4.5
		Poverty severity	0.01	0.04	-0.03***	0.01	-4.89	6.1
		Income diversification index	0.47	0.46	0.00	0.01	0.28	1.1

Significant at: * <0.1, ** p<0.05, *** p<0.01

[a] Critical level of hidden bias evaluated at significance level of 5%

Note: Nearest neighbour algorithm used.

Table 6a: Determinants of income among maize sellers and non-sellers (FIML endogenous switching regression)

Variable	Participants			Non-participants		
	Coef.	Robust S. E.	P-value	Coef.	Robust S. E.	P-value
Female head	-0.203***	0.042	0.000	-0.319***	0.068	0.000
Education	0.008	0.016	0.629	0.040**	0.018	0.023
Young adults	0.016	0.016	0.333	-0.001	0.029	0.970
Mature adults	0.100***	0.018	0.000	0.060***	0.022	0.008
Mortality	-0.031	0.078	0.694	-0.041	0.070	0.555
Radio	0.124**	0.059	0.037	0.225***	0.073	0.002
Transport	0.155***	0.038	0.000	0.158***	0.052	0.002
Assets	0.057***	0.017	0.001	0.016	0.013	0.199
Land	0.345***	0.049	0.000	0.126	0.088	0.154
Group	0.331**	0.143	0.021	0.763***	0.270	0.005
Credit	0.008	0.121	0.946	-0.112	0.213	0.599
Rainfall	0.012	0.062	0.849	-0.003	0.087	0.975
Medium potential	0.085	0.060	0.160	-0.080	0.089	0.367
High potential	0.195***	0.050	0.000	-0.019	0.069	0.781
Year 2004 dummy	-0.012	0.056	0.824	-0.011	0.087	0.902
Year 2007 dummy	0.070	0.053	0.186	0.271***	0.072	0.000
Year 2010 dummy	0.456***	0.068	0.000	0.524***	0.091	0.000
Constant	8.035***	0.459	0.000	9.419***	0.382	0.000
Observations	4828					
σ	0.716*	0.076	0.002	1.157**	0.072	0.019
ρ	0.093*	0.083	0.263	-0.033	0.031	0.282
Wald test of independence of equations	Chi-square=3.03		P-value = 0.0818			

* p<0.1, ** p<0.05, *** p<0.01

Note: Dependent variable is natural logarithim of net income. σ is the standard deviation of the error term of the equation and ρ is the correlation coefficient between the respective outcome equation and the selection equation. A statistically significant correlation coefficient implies presence of selection bias.

Table 6b: Determinants of income diversification among maize sellers and non-sellers (FIML endogenous switching regression)

Variable	Participants			Non-participants		
	Coef.	Robust S. E.	P-value	Coef.	Robust S. E.	P-value
Female head	-0.026**	0.012	0.037	-0.010	0.009	0.248
Education	-0.003	0.006	0.534	-0.003	0.003	0.320
Young adults	-0.005	0.004	0.293	-0.004	0.003	0.246
Mature adults	0.002	0.005	0.739	0.012***	0.004	0.002
Mortality	-0.023	0.021	0.257	-0.006	0.014	0.670
Radio	0.005	0.015	0.763	0.018*	0.010	0.073
Transport equipment	0.000	0.011	0.987	0.001	0.007	0.911
Assets	0.000	0.010	0.968	0.003	0.002	0.187
Land	-0.056	0.037	0.134	-0.011	0.013	0.390
Group	-0.093*	0.053	0.083	0.013	0.029	0.657
Credit	0.073**	0.033	0.027	0.035	0.023	0.130
Rainfall	-0.031	0.020	0.127	-0.006	0.010	0.554
Medium potential	-0.064	0.097	0.510	0.000	0.020	0.991
High potential	-0.074	0.104	0.477	0.001	0.021	0.953
Year 2004 dummy	-0.002	0.016	0.907	0.010	0.011	0.341
Year 2007 dummy	-0.001	0.022	0.968	0.025**	0.011	0.025
Year 2010 dummy	0.013	0.020	0.528	0.009	0.012	0.444
Constant	1.223*	0.646	0.058	0.355***	0.078	0.000
Observations	4828					
σ	0.231***	0.111	0.002	0.165	0.018***	0.000
ρ	-0.968	0.174	0.453	-0.110	0.362	0.764
Wald test of independence of equations	Chi-square=0.72 P-value = 0.397					

* p<0.1, ** p<0.05, *** p<0.01

Note: σ is the standard deviation of the error term of the equation and ρ is the correlation coefficient between the respective outcome equation and the selection equation. A statistically significant correlation coefficient implies presence of selection bias.

Table 7a: Estimated impact of selling maize on income, poverty and diversification

Outcome variable		Decision		Participation effect
		Sell maize	Do not sell maize	
Log of income	Participants	11.983	11.496	0.487***
		0.015	0.014	0.021
	Non-participants	11.643	11.452	0.191***
		0.014	0.014	0.020
	Heterogeneity effect	0.340***	0.044**	0.296***
		0.021	0.020	0.007
Poverty headcount	Participants	0.174	0.397	-0.224***
		0.008	0.011	0.014
	Non-participants	0.328	0.419	-0.091***
		0.009	0.009	0.013
	Heterogeneity effect	-0.154***	-0.022	-0.132***
		0.013	0.014	0.011
Poverty gap	Participants	0.047	0.139	-0.092***
		0.003	0.005	0.006
	Non-participants	0.108	0.151	-0.043***
		0.004	0.004	0.006
	Heterogeneity effect	-0.061***	-0.012*	-0.049***
		0.005	0.006	0.004
Poverty severity	Participants	0.019	0.066	-0.047***
		0.001	0.003	0.003
	Non-participants	0.048	0.072	-0.024***
		0.002	0.003	0.003
	Heterogeneity effect	-0.030***	-0.007*	-0.023***
		0.003	0.004	0.002
Income diversification index	Participants	0.489	0.460	0.029***
		0.001	0.001	0.001
	Non-participants	0.857	0.480	0.377***
		0.001	0.001	0.001
	Heterogeneity effect	-0.368***	-0.020***	-0.348***
		0.001	0.001	0.001

*** p<0.01, ** p<0.05, * p<0.10

Table 7b: Estimated impact of selling maize on income, poverty and diversification

Outcome variable		Sell maize, no purchase			Sell and purchase maize		
		Decision stage		Participation effect	Decision stage		Participation effect
		Participate	Not participate		Participate	Not participate	
Log of income	Participants	12.058	11.474	0.585***	11.816	11.502	0.314***
		0.014	0.014	0.020	0.013	0.014	0.020
	Non-participants	11.789	11.452	0.337***	11.777	11.452	0.325***
		0.013	0.014	0.019	0.013	0.014	0.019
	Heterogeneity effect	0.269***	0.021	0.247***	0.039**	0.050**	-0.011*
		0.020	0.020	0.008	0.019	0.020	0.006
Poverty headcount	Participants	0.142	0.412	-0.271***	0.225	0.394	-0.169***
		0.008	0.011	0.013	0.009	0.011	0.014
	Non-participants	0.248	0.419	-0.171***	0.257	0.420	-0.162***
		0.008	0.009	0.012	0.008	0.009	0.012
	Heterogeneity effect	-0.106***	-0.007	-0.099***	-0.032**	-0.026*	-0.007
		0.012	0.014	0.012	0.013	0.014	0.011
Poverty gap	Participants	0.037	0.145	-0.108***	0.059	0.137	-0.078***
		0.003	0.005	0.005	0.003	0.005	0.006
	Non-participants	0.076	0.151	-0.075***	0.071	0.151	-0.080***
		0.003	0.004	0.005	0.003	0.004	0.005
	Heterogeneity effect	-0.039***	-0.006	-0.033***	-0.012***	-0.014**	0.003
		0.004	0.006	0.004	0.004	0.006	0.004
Poverty severity	Participants	0.014	0.069	-0.055***	0.023	0.064	-0.042***
		0.001	0.003	0.003	0.002	0.003	0.003
	Non-participants	0.032	0.072	-0.040***	0.027	0.072	-0.045***
		0.002	0.003	0.003	0.001	0.003	0.003
	Heterogeneity effect	-0.018***	-0.003	-0.015***	-0.005**	-0.008**	0.003
		0.002	0.004	0.003	0.002	0.004	0.002
Income diversification index	Participants	0.487	0.449	0.038***	0.508	0.830	-0.322***
		0.001	0.001	0.001	0.001	0.000	0.001
	Non-participants	0.854	0.480	0.374***	0.431	0.472	-0.041***
		0.001	0.001	0.001	0.001	0.001	0.001
	Heterogeneity effect	-0.367***	-0.031***	-0.335***	0.077***	0.358***	-0.281***
		0.001	0.001	0.001	0.002	0.001	0.002

*** p<0.01, ** p<0.05, * p<0.10

Table 7c: Estimated impact of selling maize on income, poverty and diversification (Participants: Maize sellers who also sell dairy or industrial crops sellers)

Outcome variable		Sell both maize & dairy products			Sell both maize and industrial crops		
		Decision stage		Participation effect	Decision stage		Participation effect
		Participate	Not participate		Participate	Not participate	
Log of income	Participants	12.112	11.665	0.447***	12.264	11.818	0.446***
		0.014	0.013	0.019	0.015	0.015	0.021
	Non-participants	11.756	11.632	0.124***	12.996	11.696	1.299***
		0.013	0.013	0.018	0.012	0.013	0.018
	Heterogeneity effect	0.355***	0.032*	0.323***	-0.732***	0.122***	-0.854***
		0.019	0.019	0.007	0.019	0.019	0.007
Poverty headcount	Participants	0.106	0.296	-0.190***	0.090	0.251	-0.161***
		0.007	0.010	0.012	0.006	0.010	0.012
	Non-participants	0.262	0.310	-0.049***	0.001	0.285	-0.284***
		0.008	0.009	0.012	0.001	0.009	0.009
	Heterogeneity effect	-0.155***	-0.014	-0.142***	0.089***	-0.034***	0.123***
		0.011	0.013	0.010	0.005	0.013	0.012
Poverty gap	Participants	0.026	0.089	-0.062***	0.018	0.070	-0.052***
		0.002	0.004	0.004	0.002	0.003	0.004
	Non-participants	0.076	0.096	-0.020***	0.000	0.084	-0.084***
		0.003	0.003	0.005	0.000	0.003	0.003
	Heterogeneity effect	-0.050***	-0.008	-0.042***	0.018***	-0.014***	0.032***
		0.004	0.005	0.003	0.001	0.005	0.004
Poverty severity	Participants	0.010	0.037	-0.027***	0.006	0.028	-0.022***
		0.001	0.002	0.002	0.001	0.002	0.002
	Non-participants	0.031	0.041	-0.010***	0.000	0.035	-0.035***
		0.002	0.002	0.002	0.000	0.002	0.002
	Heterogeneity effect	-0.022***	-0.004	-0.017***	0.006***	-0.006***	0.013***
		0.002	0.003	0.002	0.001	0.002	0.002
Income diversification index	Participants	0.528	0.493	0.036***	0.485	0.412	0.073***
		0.001	0.001	0.001	0.001	0.001	0.002
	Non-participants	0.823	0.543	0.280***	0.933	0.489	0.444***
		0.001	0.001	0.001	0.002	0.001	0.002
	Heterogeneity effect	-0.295***	-0.050***	-0.245***	-0.448***	-0.077***	-0.372***
		0.001	0.001	0.001	0.002	0.002	0.002

*** p<0.01, ** p<0.05, * p<0.10

Appendix A

Table A1. Covariate balance in matched sub-samples: percent reduction in bias

Covariate	Sample	Mean		% bias	% reduction in |bias|	Test of difference in mean between participants and non-participants	
		Participants	Non-participants			t-statistic	P-value
Gender of household head (1=female)	Unmatched	0.190	0.217	-6.600		-2.250**	0.025
	Matched	0.192	0.188	1.000	84.300	0.340	0.737
Education of adults (years)	Unmatched	7.752	6.958	27.200		9.300***	0.000
	Matched	7.727	7.709	0.600	97.700	0.210	0.834
No. of adults 15-24 years	Unmatched	2.090	1.822	16.800		5.830***	0.000
	Matched	2.084	2.082	0.100	99.300	0.040	0.972
No. of adults 25-64 years	Unmatched	2.233	2.155	6.100		2.100**	0.036
	Matched	2.227	2.254	-2.100	64.900	-0.690	0.488
Household experienced prime-age mortality (1=yes)	Unmatched	0.055	0.059	-1.800		-0.610	0.544
	Matched	0.055	0.054	0.600	63.900	0.210	0.836
Owns radio (1=yes)	Unmatched	0.902	0.845	17.300		5.850***	0.000
	Matched	0.902	0.904	-0.500	97.100	-0.180	0.857
Owns transport equipment (1=yes)	Unmatched	0.550	0.428	24.600		8.450***	0.000
	Matched	0.549	0.559	-2.100	91.400	-0.670	0.500
Value(log) of household assets (KES)	Unmatched	11.134	10.578	25.500		8.610***	0.000
	Matched	11.119	11.169	-2.300	91.100	-0.780	0.436
Land size (log)(acres)	Unmatched	1.806	1.388	57.200		20.030***	0.000
	Matched	1.785	1.828	-5.900	89.700	-1.680	0.093
Proportion of village households in farmer groups/cooperatives	Unmatched	0.737	0.767	-13.800		-4.750***	0.000
	Matched	0.738	0.749	-4.700	65.900	-1.570	0.117
Proportion of village households that received credit	Unmatched	0.422	0.510	-29.500		-10.100***	0.000
	Matched	0.424	0.429	-1.700	94.400	-0.540	0.587
Rainfall amount (log) in main season (mm)	Unmatched	6.332	6.123	35.100		11.840***	0.000

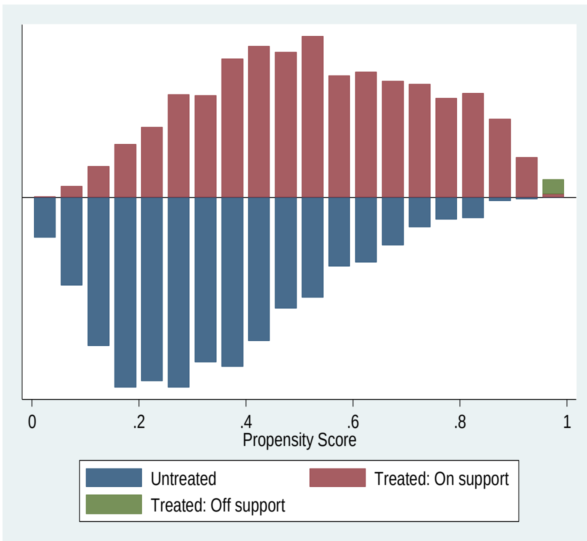
	Matched	6.331	6.299	5.200	85.100	1.910	0.056
Medium potential dummy	Unmatched	0.269	0.233	8.200		2.840***	0.005
	Matched	0.271	0.286	-3.500	57.200	-1.080	0.278
High potential dummy	Unmatched	0.553	0.397	31.800		10.950***	0.000
	Matched	0.550	0.525	4.900	84.500	1.560	0.119
Year 2004 dummy	Unmatched	0.271	0.238	7.500		2.590**	0.010
	Matched	0.270	0.257	3.200	57.900	1.000	0.318
Year 2007 dummy	Unmatched	0.281	0.231	11.500		3.970***	0.000
	Matched	0.281	0.283	-0.500	95.800	-0.150	0.882
Year 2010 dummy	Unmatched	0.216	0.275	-13.800		-4.710***	0.000
	Matched	0.217	0.218	-0.300	98.000	-0.090	0.927
Travel time to nearest town with 75000+ inhabitants (hours)	Unmatched	3.020	3.122	-4.300		-1.490	0.136
	Matched	3.008	2.914	4.000	8.600	1.240	0.215
District median maize grain price (log) (KES/kg)	Unmatched	2.618	2.696	-32.800		-11.300***	0.000
	Matched	2.619	2.618	0.700	98.000	0.210	0.837
Village population density (persons/km2)	Unmatched	348.260	379.790	-14.700		-5.040***	0.000
	Matched	349.030	342.650	3.000	79.800	0.940	0.349

* p<0.1, ** p<0.05, *** p<0.01

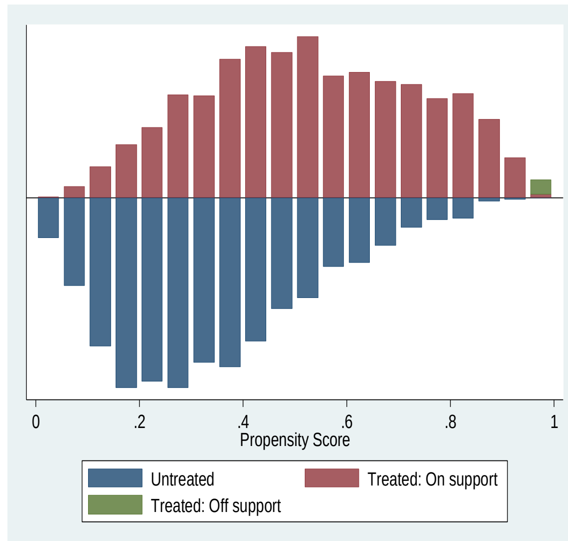
Source: Authors.

Figure A1. Common Support

Nearest neighbour matching



Kernel matching



Source: Authors.

Appendix B

Table B1. Test result for admissibility of instruments (ESR)

Variable	Coefficient	Robust Standard Error	P-value
Dependent variable: Log of household income (Non-participants sub-sample)			
Female head	-0.321	0.069***	0.000
Education	0.040	0.018**	0.024
Young adults	-0.001	0.029	0.980
Mature adults	0.059	0.022***	0.007
Mortality	-0.037	0.070	0.600
Radio	0.210	0.073***	0.004
Transport equipment	0.201	0.058***	0.001
Assets	0.020	0.012	0.108
Land	0.133	0.088	0.132
Group	0.747	0.268***	0.005
Credit	-0.105	0.216	0.628
Rainfall	0.000	0.088	0.999
Medium potential	-0.255	0.116**	0.028
High potential	-0.068	0.076	0.373
Year 2004 dummy	0.009	0.096	0.925
Year 2007 dummy	0.311	0.098***	0.002
Year 2010 dummy	0.493	0.315	0.118
Travel time	0.018	0.013	0.158
Maize price	0.161	0.529	0.761
Population density	0.000	0.001	0.481
Constant	10.090	1.177***	0.000
Observations	2794		
F (30, 2763)	56.510		
Prob > F	0.000		
R^2	0.294		
Test for joint significance of instruments			
F(3, 2763)	0.880		
Prob > F	0.451		

* p<0.1, ** p<0.05, *** p<0.01