EGERTON UNIVERSITY

TEGEMEO INSTITUTE OF AGRICULTURAL POLICY AND DEVELOPMENT

WPS 51/2012

THE IMPACT OF MAIZE HYBRIDS ON INCOME, POVERTY, AND INEQUALITY AMONG SMALLHOLDER FARMERS IN KENYA

Mary K. Mathenge, Melinda Smale, and John Olwande[1]

Funding for this document was provided by the American people, via the Food Security III Cooperative Agreement (GDGA-00- 000021-00) between Michigan State University and the United States Agency for International Development, Bureau for Food Security, Office of Agriculture, Research, and Technology.

Mathenge is Director, Tegemeo Institute of Egerton University, Smale is a professor of International Development at the Department of Agricultural, Food, and Resource Economics, Michigan State University, and Olwande is a research fellow with Tegemeo Institute of Egerton University.

THE IMPACT OF MAIZE HYBRIDS ON INCOME, POVERTY, AND INEQUALITY AMONG SMALLHOLDER FARMERS IN KENYA

By

Mary K. Mathenge, Melinda Smale, and John Olwande

WPS 51/2012

Tegemeo Institute of Agricultural Policy and Development

P.O Box 20498, 00200, Nairobi, Kenya

Tel: +254 20 2717818/76; Fax: +254 20 2717819

E-mail: egerton@tegemeo.org

Tegemeo Institute

Tegemeo Institute of Agricultural Policy and Development is a Policy Research Institute under Egerton University with a mandate to undertake empirical research and analysis on contemporary economic and agricultural policy issues in Kenya. The institute is widely recognized as a centre of excellence in policy analysis on the topical agricultural issues of the day, and in its wide dissemination of findings to government and other key stakeholders with a view to influencing policy direction and the decision making process. Tegemeo's consistently good quality empirically-based analytical work, and its objective stance in reporting and disseminating findings has over the past decade won the acceptance of government, the private sector, civil society, academia and others interested in the performance of Kenya's agricultural sector.

Published February, 2014

Kindaruma Lane, off Ngong Road

P.O Box 20498, 00200, Nairobi, Kenya

Tel: +254 20 2717818/76; Fax: +254 20 2717819

E-mail: egerton@tegemeo.org

URL: http://www.tegemeo.org

Tegemeo Institute acknowledges the resources support for its research programmes from key partners including the United States Agency for International Development (USAID), Michigan State University (MSU), and Egerton University, Njoro Kenya. Others include the World Bank, European Union, Department for International Development (DFID), Food and Agriculture Organization of the United Nations (FAO).

Acknowledgments

This study represents a joint collaboration between the Tegemeo Institute of Egerton University and Michigan State University.

Tegemeo Institute acknowledges the resources support for its research programmes from key partners including the United States Agency for International Development (USAID), Michigan State University (MSU), and Egerton University, Njoro, Kenya. Others include the World Bank, European Union, Department for International Development (DFID), and Food and Agriculture Organization of the United Nations (FAO).

The authors thank Nicole Mason for her comments on an earlier draft, and Patricia Johannes for her assistance in editing and formatting the paper.

Table of Contents

List of Tables

1.0 Introduction

For decades, Kenya has been depicted a maize "success story" in Sub-Saharan Africa, known for rates of hybrid maize adoption during the 1960s and 70s that paralleled those of the U.S. Corn Belt thirty years earlier (Gerhart, 1975; Byerlee and Eicher, 1997; Smale and Jayne, 2010). Over the past few decades, however, a general perception of stagnating adoption and production has been supported by FAO data and a rising maize import bill. Replacement of older hybrids by newer releases appears to have been slow (Hassan 1998; Smale et al., 2012), dampening yield potential on farms. For example, a hybrid released in 1986 and derived from this first hybrid still dominates the maize fields of Kenya, despite the dramatic increase in the number of hybrids and breadth of seed suppliers as seed markets liberalized (Swanckaert, 2012).

A number of in-depth studies of maize seed adoption have been conducted in specific regions of Kenya (Ouma et al., 2002; Salasya et al., 2002; Wekesa et al., 2002). Based on large-scale surveys, analyses of seed adoption and maize productivity have been implemented by Kenya Agricultural Research Institute (KARI) and the International Maize and Wheat Improvement Center (CIMMYT) in 1992, 2002, and 2010 (Hassan, 1998; De Groote et al., 2005). Using data collected in 1997, 2000, and 2004 by Tegemeo Institute, Egerton University and Michigan State University from a panel of households in maize-growing areas in Kenya, Suri (2011) applied a correlated random coefficient model to demonstrate that heterogeneous net returns to hybrid seed explain adoption rates. She compared distributions of hybrid and non-hybrid maize yields, illustrating that not only were mean yields higher, but the variance of yields was much lower among hybrid seed growers. Recently, Jones et al. (2012) tested the effect of hybrid use on the mean, variance, and skewness of maize yields with a stochastic production function applied to survey data collected by Tegemeo Institute during the 2006/07 cropping season. The authors found that hybrids enhance mean yields, and also reduce the exposure of smallholders to extremely low yields. In the recent past, overall adoption rates have risen compared to those reported by Suri (2011) in part as a reflection of changes in seed-to-grain price ratios and the progress of seed liberalization (Smale et al., 2012).

To complement research on hybrid seed adoption, a number of studies have been conducted on the maize seed industry and seed supply. Wangia et al. (2004) concluded that after the process of market liberalization that began during the mid-1980s, private sector participation at all nodes in the marketing system increased substantially, resulting in an smoother flow of inputs to many parts of Kenya. A seed sector study conducted by Nambiro et al. (2004) in Trans Nzoia District found some impact of the liberalization of the seed industry on the distribution side, where private retailers had broken the previous monopoly of the Kenya Farmers' Association. However, according to the authors, the impact of seed liberalization on maize production was minimal. At that time, it was estimated that Kenya Seed Company (KSC) provided 97% of the seed, much of which was dominated by one variety.

There is some more recent evidence that liberalization has led to entry of new seed companies in the maize market. Swanckaert (2012) reports that while KSC was the only maize seed company prior to 1992, there are currently 11 companies with varieties registered to their names. In addition, the plant variety registry of the Kenya Plant Health Inspectorate (KEPHIS) lists 164 varieties released from 1964 up to 2009, with 85 percent of these registered since 2000. The numbers of improved maize varieties and hybrids grown on farms has also increased tremendously. While Hassan (1998) found only 12 hybrids grown by farmers in 1992, Tegemeo data indicate that the number of hybrids on farms was 33 in 2004 and 50 in 2010. Nonetheless, Swanckaert (2012) concludes that although competition in the seed market has intensified, the impact of new seed companies on market concentration has been smaller than expected.

To our knowledge, a missing link in existing research on maize hybrids in Kenya has been a rigorous analysis of the impacts of seed adoption on farmer welfare. Mwabu et al. (2007) applied a bivariate probit model to explore the relationship of adoption of improved varieties to poverty of households in rural districts of Laikipia and Suba, based on a single year of data. The authors found a negative correlation between poverty and adoption of improved maize seed.

The objective of this study is to produce an initial assessment of the impact of maize hybrids on the welfare of smallholder farmers in Kenya. Our indicators of farmer welfare are household income, income inequality, and poverty. We contribute to the existing body of knowledge on maize hybrids in Kenya by documenting these impacts across different agricultural production environments. While the districts studied by Mwabu et al. (2007) are largely found in areas with lower agricultural potential, we use a representative dataset collected across major agro-ecological zones. We also apply instrumental variables with panel data methods to test the endogeneity of hybrid seed adoption.

2.0 Data

In collaboration with Michigan State University (MSU), Tegemeo Institute of Egerton University has implemented a five-year panel survey (1997, 2000, 2004, 2007, and 2010) in rural Kenya. The motivation for the survey has been to evaluate income, poverty, and developmental pathways, providing both routine policy advice and longer-term, in-depth analyses to Kenyan national decision-makers. Although crop-specific analyses have been viewed as secondary, Tegemeo and MSU have emphasized policy research on the maize value chain, along with dairy and horticulture.

The sampling frame for the panel survey was prepared in 1997 in consultation with the Central Bureau of Statistics (CBS), currently the Kenya National Bureau of Statistics (KNBS). The process is described by Argwings-Kodhek et al (1999). Census data were used to identify all non-urban divisions in the country, and these were assigned to one or more agro-ecological zones (AEZ) based on the 1990 Census, District Development Plans and the Farm Management Handbook. Within each AEZ, two or three divisions were chosen based on their importance (size of population). In each selected division, villages were randomly selected. Households were selected within selected villages with systematic sampling, and a random start. A total of 1,578 households were selected in 24 districts. For the purposes of analysis, the selected households were then grouped into 9 agro-regional zones, which represented a combination of AEZ and administrative boundaries.

The data for 1997 is excluded in this analysis owing to lack of comparability in some of the key variables of interest. We employ a balanced panel of 1243 households, all of whom grow maize. The attrition rate (after excluding the Northern Arid agro-regional zone which was dropped from the survey in 2000 due to nomadic way of life of the households, which made tracing them difficult) for the panel was 13% in 2010 compared to the initial survey, conducted in 1997.

3.0 Methods

3.1 Conceptual Approach

We view impact pathways through the lens of the agricultural household, in which family members organize their labor to maximize utility over consumption goods and leisure. They produce goods for consumption or sale, and cash constraints are relaxed primarily through farm sales and off-farm employment of family members, since the production of maize, a staple food crop, generally is not financed through credit. Growing higher-yielding hybrid maize can contribute to household income by releasing land to produce other crops or by generating cash from maize sales. Over a series of years, if farmers use hybrids successfully each season, we would expect that incremental increases in income could be capitalized into farm assets, some as equipment or livestock, contributing to capacity of the farm household to generate higher levels of farm income. Thus, households growing hybrids may move out of poverty, and by ranked among the better off in their communities. On the other hand, in an increasingly diversified agricultural economy, as households move away from reliance on maize production and meet consumption needs through other income sources, we might expect the impact of hybrid seed to matter less.

Our conceptual approach follows the class of analytical approaches generally known as "treatment models," described in-depth in an extensive body of literature that addresses the statistically-based measurement of the social and economic impacts of public programs (Ravallion, 1994; Angrist and Krueger, 2001; de Janvry et al., 2010). The motivation for these approaches is understood as essentially one of missing data (Ravallion 1994). That is, we observe the values of outcome variables, such as indicators of income or poverty, for the group of households who are targeted by a program or policy post facto, as well as those who are not. We cannot, however, observe the values of outcome variables for targeted households had they not been targeted (op. cit.).

In the case of hybrid seed use among smallholders in Kenya, we observe values of outcome variables for adopters and non-adopters, but not for adopters had they not adopted. Unlike households targeted by a program or policy, adopters in Kenya choose to grow hybrid seed, or "self-select" into the treatment group. From a large body of previous empirical research on the adoption of agricultural technologies, we know that adopters are generally those who are wealthier in terms of various types of human capital and have more access to "soft" (information and financial services) and "hard" (roads, vehicles, and marketplaces) market infrastructure. Thus, any estimate of the impact of hybrid seed use on outcome variables that does not take this into account will exhibit a bias due to the underlying effects of these factors. This selection bias, attributable in our case to self-selection through seed choice rather than explicit targeting, reflects the fact that adopters are better off than non-adopters even before they adopt.

Experimental and quasi-experimental methods have been proposed to address selection bias. Experimental approaches include randomized treatments or randomized controlled trials. These approaches are not feasible in our case, where Kenyan farmers have grown maize hybrids since they were first released in 1964. Quasi-experimental approaches consist of instrumental variables regression, propensity score matching, and difference-in-difference estimation. Propensity score matching involves estimating the probability that a farmer plants hybrid seed as a function of a set of observed explanatory variables, and comparing outcome variables for adopters and non-adopters who have high likelihood of adoption. The implicit assumptions of this approach are that 1) only the factors that matter in adoption are those specified in the regression equation, and that 2) all relevant determinants are observable. We know that these assumptions are difficult to justify when we have a limited number of observed explanatory variables and we know that certain intrinsic, unobservable attributes influence the seed choices made by Kenyan farmers.

While feasible in our case, the matching approach is not well-suited to analyzing the impact of this project because specific socio-economic groups were not targeted for an intervention. In addition, according to Handa and Maluccio (2010), matching methods are more promising in evaluating easily measured outcomes, such as those related to child schooling and health, than it is for more complex outcomes, such as expenditures (or income).

Panel data methods are also designed to control for unobserved heterogeneity that is correlated with observed variables, through fixed effects or first-differencing models (De Janvry et al., 2010; Ravallion, 1994; Wooldridge, 2002). Difference-in-difference models are the most common in research design for policy analysis with panel data by comparing the change in outcome variables between the sub-population that received a treatment (or self-selected into the group) and the sub-population that did not. In the case of a panel with only two years of observations, fixed effects estimation produces similar results to first-differencing

From the foregoing and of the feasible approaches to quantitative assessment of social and economic impacts, we consider instrumental variables combined with panel data methods as the best suited to our data-generating process and research hypotheses. The instrumental variable approach relies on econometric methods to separate the effects of belonging to a group (through targeting or choice) from those of other factors that influence impact. Identifying valid instrumental variables is the major challenge associated with this method. In our case, valid instrumental variables are those that determine whether or not a farmer uses hybrid seed in maize production, but only influence outcome variables through hybrid seed use.

In summary, we hypothesize that either because of observables (farm size, education, labor supply) or unobservables (e.g., intrinsic management ability, unmeasurable soil quality), or simultaneity/feedback process, the decision by Kenyan smallholder farmers to grow hybrid maize seed is endogenous in income. In the presence of endogeneity, which results from non-random selection of analytical units, the correlation of independent variables with error terms, or a chain of causality among independent and dependent variables, estimators generated by ordinary least squares are biased.. Our approach estimates the local average treatment effect (LATE): the effect of self-selection into the group of farmers that uses maize hybrids on income, inequality, poverty, identified through instrumentation. Next we define the outcome variables, present the estimation procedure, and describe the explanatory variables.

3.2. Estimation Strategy

3.2.1. Impact Outcomes

We considered three outcome variables, namely household income, income inequality, and poverty. Household income, expressed in current, nominal terms, is comprised of net crop income (gross value of crop production less input costs); net livestock income (gross value of livestock products plus sales of live animals less purchases of live animals plus input costs); salaries for household members; net business income for household members; income from informal labour employment for household members; and remittances, pension and share dividends received by all household members

For poverty, we employed a binary outcome variable measuring whether or not the respective household income fell below the poverty line. While we can use Foster-Greer-Thorbecke (FGT) indices to compare users and non-users of hybrid seed, the choice of the binary poverty outcome was based on the need to construct the indicator at the household level to enable the kind of estimation envisioned in this study[2]. We applied the official poverty lines established by the Government of Kenya, for each survey year, in nominal Kenyan shillings (KES) per month: 1009 (2000), 1336 (2004), 1629 (2007), 2144 (2010). Since these poverty lines are expressed in per capita terms per month, we also converted the annual household income into income per adult equivalents per month for comparison purposes.

As with the FGT indices, popular measures of inequality, such as the Gini or Theil indices, are also calculated over a distribution of households or individuals. People often compare themselves with others in their immediate reference group, such as a village, rather than with the whole society (Yitzhaki, 1979). Based on the observation and their analysis of the effects of migration on households in Mexico, Stark and Taylor (1989) proposed an inequality index calculated at the individual or household level and given as:

$$RD(Y_i) = AD(Y_i) * P(Y_i) \quad (5)$$

Relative deprivation $RD(Y_i)$ was calculated for each household *i*, taking the remaining households in the sample as the reference group. $AD(Y_i)$ is the mean income of households in the sample richer than a given household *i* and $P(Y_i)$ is the proportion of households in the sample that are richer than a given household *i* (Stark and Taylor, 1989). To construct the index, households were ranked by income from lowest to highest. As is the case with other indicators of inequality, relative deprivation is typically calculated with income data, although it can also be computed with other variables, such as land. Conforming to other outcome variables, we computed the index over current nominal income. The higher the value of the index, the greater is the deprivation of the household relative to other households in sample.

Given the level of skewness of the distribution of the level income variable, we preferred to use its natural logarithm (Figures A.1, A.2). The use of natural logarithms has an added advantage in that it negates the need to deflate nominal incomes. According to Wooldridge (2003: 430), the use of natural logarithms in combination with year dummies achieves the same results as using real incomes. The distributions of Stark's inequality index are far less skewed than that of the income variable used to construct it (Figure A.3), although they are not smooth. For this reason, and because the index is originally defined (Stark and Taylor, 1989) in terms of levels, we did not transform the inequality indicator.

3.2.2 Econometric Models

Separate regressions were estimated for each outcome variable of interest described in section 3.2.1. The model of interest is as given below:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \gamma Z_{it} + \alpha_i + \varepsilon_{it} \qquad i=1,..,N \; t=1,..,T \qquad (1)$$

Where Y is the outcome variable of interest, X is a vector of exogenous explanatory variables, Z is hybrid seed use, α_i represents time-invariant unobserved factors that affect the outcome variable and ε is the random error term. We also added year effects to the estimation (as part of X) to control for other time-varying, unobservable factors that do not vary across households.

Model (1) can be estimated using the usual panel data regression methods depending on the construction of the outcome variable.Panel data methods help to control for time-invariant unobservable factors among farm households that could be correlated with the included explanatory variables (Wooldridge 2002). Often, this unobserved heterogeneity is considered to capture intrinsic features that are not easily measured (such as farm management capability, unmeasurable soil quality, etc). In addition and as we have argued previously, we have strong conceptual and empirical grounds to expect that our variable of interest, Z (hybrid seed use), is endogenous in the outcome models due to self-selection, simultaneity, or measurement error. Estimating Model (1) without controlling for endogeneity would result in biased estimates, thus misrepresenting the impacts of hybrid seed use on outcome variables.

To control for the potential endogeneity of the binary variable for hybrid seed use, we applied Instrumental Variables (IV) method with panel data (Schaffer 2010) for the continuous outcome variables. This is a two-stage least squares procedure, with a binary variable for hybrid seed use in the first stage regression. Angrist (2001) and Angrist and Krueger (2001) argue that even in the case of a dichotomous variable in the first of the two equations, two-stage least squares produces consistent estimators that are less sensitive to assumptions on functional form.

Although we also run a first difference IV model and observed that the basic result is maintained, we report results for the Fixed Effects two-Stage Least Squares (FE2SLS) method. The test for endogeneity is by itself inbuilt in these models and test results reported with the estimates. We estimated all models with robust variance-covariance matrices.

Given that both the poverty outcome variable and the suspected endogenous variable, hybrid maize adoption, are binary and non-linear, a probit-2SLS method is inappropriate because it implies that, in the second stage, a nonlinear function of an endogenous variable is replaced with the same nonlinear function of fitted values from a first-stage estimation (Wooldridge 2002:236). The Control Function Approach (CFA) however enables us to test and account for endogeneity bias in such non-linear models. As in a 2SLS model, the control function approach requires use of instrumental variables in the first stage, reduced form estimation of hybrid seed use. In the second stage, however, the structural model is estimated with the observed endogenous variable and the residual from the first stage added as explanatory variables. The test of endogeneity is the statistical significance of the coefficient of the residual, when the regression is estimated with bootstrapped standard errors. The control function approach is described in early work by Smith and Blundell (1986).

Given the non-linearity of the binary poverty outcome variable, we apply probit with Correlated Random Effects (CRE). As proposed by Mundlak (1978) and Chamberlain (1984), the CRE model helps to control for unobserved heterogeneity and its correlation with observed factors in a non-linear models. Application of the model requires that the means of time-varying explanatory variables are included in the regression.

Standard model diagnostics include tests of a) the relevance of the instrument set; b) model identification; and c) endogeneity of the adoption variable. Model diagnostics for (a) include a) the evaluation of the joint F-test for excluded instruments in the first stage regression; b) the Hansen's J test for overidentifying restrictions; and c) the Kleibergen-Paap statistic which provides a test for the weakness (underidentification) of instruments. Rejection of the null hypothesis supports evidence that instruments are correlated with the endogenous regressor. Failure to reject the null hypothesis in the Hansen-J test indicates that the 'extra' instrumental varialbes are exogenous in the structural equation, thereby supporting the validity of the instruments, Finally, the endogeneity test is defined as the difference between the Hansen-J statistics with and without the instruments that are hypothesized to identify the endogenous variable (Schaffer 2010). The results of these tests are reported in section 4.

3.2.3 Explanatory Variables

Explanatory variables are defined and summary statistics reported in Table 1.

Our explanatory variables include only those that vary over time, excluding dummies for agricultural potential, such as agroecological zone, and administration (district), which are highly correlated with other independent variables. We also include year dummies. To address the crucial importance of agroecological factors, we estimate separate regressions for sub-samples based on two classifications: 1) areas with high potential maize productivity and other maize-growing areas; 2) areas with high and low productivity potential.

Table 1: Variable definition and summary statistics

Variable	Construction	2000		2004		2007		2010		Pooled	
		Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Dependent											
Income	Sum of gross income (KES) from crops, livestock, salaries, remittance, business and informal labor activities	170,234	195,802	186,535	215,563	201,659	210,659	278,463	386,998	209,223	267,240
Income inequality	Stark's inequality index	126,050	42,049	138,712	46,114	147,604	49,905	209,054	66,045	126,050	42,049
Poverty	1=poor (income per adult-equivalent per month) < official poverty line; 0 else	0. 27	0.44	0.30	0.46	0.31	0.46	0.32	0.45	0.30	0.46
Explanatory											
Hybrid adoption	1=Grow hybrid seed in the year, 0 otherwise	0.68	0.47	0.61	0.49	0.73	0.44	0.82	0.38	0.71	0.45
De jure headship	1=Reported head of household is female, 0 otherwise	0.12	0.32	0.20	0.40	0.23	0.42	0.27	0.44	0.21	0.40
Education	Average educational attainment (years) of all adults in household	7.12	2.86	7.21	2.99	7.20	3.02	7.59	3.00	7.28	2.97
Young adults	Number of adults 15-24 years	2.07	1.62	1.96	1.63	1.91	1.60	1.81	1.53	1.94	1.60
Mature adults	Number of adults 25-64 years	2.35	1.26	2.26	1.30	2.10	1.26	2.06	1.34	2.19	1.29
Extension	Km to extension service	5.45	5.93	5.27	5.87	4.42	5.23	5.16	5.29	5.08	5.60
Tarmac	Km to tarmac road	7.67	7.87	7.59	7.81	7.61	7.48	7.11	7.23	7.50	7.60
Land size	Total land area (acres), last survey	6.04	9.71	6.06	8.48	6.12	8.98	5.82	8.81	6.01	9.00
Rain	Total mm rain in the growing season associated with the survey year	583.91	271.77	685.24	298.61	611.59	195.99	413.89	200.76	573.66	265.11
Seed access	Km to nearest seller of certified maize seed	5.59	7.46	3.85	7.38	3.41	4.42	4.13	5.70	4.25	2.70
Seed-to-grain price ratio	Predicted value of farmgate seed-to-grain price ratio (village, year, distance to fertilizer source, own transport equipment)	11.17	2.14	11.25	2.034	11.78	2.03	7.44	2.10	10.42	6.42

Source: Authors.

Other independent variables include labor quality and quantity, measured in terms of the average education of all adults in the household, and the numbers of young and mature adults in the household. While household size and the number of children, in particular, is often argued to be endogenous, we consider the number of working-age adults to be exogenous in the short-term. We include whether the recognized household head is female or male. In fact, most women household heads in the Tegemeo panel are widows, and there are strong reasons to believe that they differ from households headed by men, most of whom have spouses. Land size is included to represent endowments of physical capital, access to other resources, and scale of the farm operation. Unfortunately, while we do have soil quality data, these are measured in terms of categorical or binary variables that cannot be included in the fixed effects estimation. Distances to extension services and the tarmac road are included as explanatory variables in both income and hybrid seed regressions.

Included in the hybrid seed category are all hybrids as indicated by the farmer/respondent or recognized by the enumerator. We did not include improved open-pollinated varieties, which represent a very small percentage of improved maize types grown by Kenyan farmers, as do recycled hybrids.

Theory suggests that the ratio of seed-to-grain prices, and the distance to the nearest seller of certified maize seed are potentially good instruments for identifying the effect of hybrid seed use on household income (Heisey et al. 1998). We hypothesize that distance to the nearest seller of certified maize seed is strongly related to hybrid seed use, but not necessarily to overall household income. Virtually all of households surveyed rely on a range of farm and non-farm income sources, in addition to maize.

Seed price, calculated here as seed costs per kg at the farm gate, is observed only for farmers who purchase seed (most of which is hybrid maize), even though all farmers face prices. The theoretic framework of the agricultural household dictates that both household and market characteristics, as well as market prices, are determinants of decision prices, which vary by household. We predicted the seed-to-grain price ratio using village dummies, year effects, distance to nearest seller of fertilizer, and a binary variable measuring ownership of transport equipment.

4.0 Results

4.1 Descriptive Statistics

Use rates for hybrid seed are shown in Table 2, by year and agroecological zone. The results suggest that overall rates climb from about two-thirds of farmers in the early 2000s to over four-fifths a decade later. Rates have exceeded 80% in the highlands since the beginning of the period, and have varied most over the years in the lowlands.

Table 2: Percentage of households growing hybrid seed, by agroecological zone and year

	2000	2004	2007	2010
Coastal lowland	28.4	1.3	37.8	37.7
Lowland	73.5	50.0	61.8	91.2
Lower midland 3-6	26.5	17.2	38.8	56.3
Lower midland 1-2	75.2	75.0	86.1	89.6
Upper midland 2-6	79.7	68.2	77.4	92.6
Upper midland 0-1	87.1	85.3	88.4	91.1
Lower highland	85.2	85.2	91.0	95.1
Upper highland	92.3	92.3	100.0	100.0
All zones	68.1	61.1	73.1	82.0

Source: Authors.

Comparisons of hybrid users and non-users on outcome indicators are presented in Tables 3 through 5. With respect to household income, hybrid maize growers have statistically higher income on average than farmers who grew local or improved open-pollinated varieties (Table 3).

Table 3: Household income by use of hybrid maize seed and year

Year	No hybrid maize seed		Hybrid maize seed	
	Mean	Std. Deviation	Mean	Std. Deviation
2000	108,175	112,909	199,419	218,840
2004	135,820	164,140	219,240	237,350
2007	127,596	126,258	229,442	229,039
2010	160,571	230,835	306,483	410,571

Source: Authors.
Note: T-tests show adopters have significantly higher income than non-adopters in each year at 1% significance. Household annual income reported in nominal KES.

As measured by Stark's index, average relative deprivation is significantly higher among non-hybrid growers than growers of maize hybrids, in each year (Table 4). Relative to all other households in the national sample, incomes are higher and more widely distributed among maize growers who plant hybrid seed. However, a quick review of the figures in Table 4 does not suggest that the coefficient of variation[3] is greater among maize farmers who grow hybrids. Kolmogorov-Smirnov tests indicate that the distribution of inequality based on time-pooled data among hybrid maize farmers is lower-valued overall and also significantly different at 1%.

Table 4: Relative deprivation (income) by use of hybrid maize seed and year

Year	No hybrid maize seed		Hybrid maize seed	
	Mean	Std. Deviation	Mean	Std. Deviation
2000	142,639.6	34,930.5	118,255.2	42,885.5
2004	152,727.9	40,111.7	129,637.8	47,372.7
2007	168,657.7	39,723.4	139,742.3	51,033.0
2010	241,426.0	50,872.9	201,379.4	66,950.9

Source: Authors.
Note: Inequality measured by Stark's index of relative deprivation (defined in text) with respect to income in nominal KES. T-tests show non-adopters have significantly higher mean relative deprivation than adopters in each year at 5% significance.

The relationship of hybrid use to poverty status, as measured by FGT indices[4] described above, is summarized in Table 5. The proportion of households falling below the poverty line is around twice as high for farmers who do not grow hybrid maize in each year. The average depth of poverty is also significantly greater for poor households that do not grow hybrid maize, relative to hybrid maize growers in all years but 2010. Similarly, the severity of poverty, or the variation among the poor, is less for hybrid users than non-users in three out of the four survey years.

[3] The standard deviation divided by the mean.

[4] The formula for the FGT index is: $FGT_\alpha = (1/n)\sum_{i=1}^{h}(\frac{z-y_i}{z})^\alpha$. When α=0, the FGT is a headcount; when α=1, the FGT is the poverty gap; when α=2, the FGT is poverty severity.

Table 5: Use of maize hybrids by poverty status and year

	Headcount		Poverty depth		Poverty severity	
	Grow maize hybrid		Grow maize hybrid		Grow maize hybrid	
Year	No	Yes	No	Yes	No	Yes
2000	41.67	19.63	0.468	0.354	0.283	0.181
2004	42.14	21.96	0.450	0.386	0.267	0.206
2007	46.20	25.77	0.416	0.359	0.228	0.175
2010	54.70	25.43	0.417	0.389	0.229	0.199

Source: Authors. .
Note: Statistical tests show adopters were less poor by any of the three indices in all years except 2010 at 5% significance, when poverty depth and severity was not significantly different between adopters and non-adopters.

4.2 Econometric Results

In this section, we present regression results for each of the three impact outcomes. Regression results are shown in the main text for models estimated on the combined sample and the sample disaggregated by maize productivity. Results for separate models estimated by agricultural productivity potential are presented in the annex (Tables A.2, A.3 and A.4.)The principal hypotheses tested in both linear and non-linear regressions concern the endogeneity of the decision to grow hybrid maize, generated by simultaneity or selection bias, and the impact of growing maize hybrids on income-related outcomes (total household income, income inequality, and poverty).

The regressions presented in Table 6 confirm that growing maize hybrids has a positive influence on total household income, and a greater impact in the higher potential maize-growing areas than in other environments. As expected based on economics principles, the marginal payoff (incentive) to smallholder households who grow higher-yield maize seed is larger in magnitude in areas that are better suited to maize.

Among other explanatory factors, the impact of female headship is strongly negative in the high potential maize-growing areas, although this effect is lost in other areas and in the overall sample (Table 6). One interpretation of this finding is that where more resources are allocated to maize maize production, as would be expected in the higher potential zone, the loss of the male head may have serious consequences for labor resources in the household, diminishing not only the investments in production of maize and other crops but the capacity to allocate the time of family members to nonfarm income sources that can generate substantial earnings in Kenya's economy.

The importance of adult labor is reflected in the significance of the coefficients on the number of both young and mature adults.

Not surprisingly, as has been demonstrated in many studies of rural Kenya and other developing countries, education consistently raises incomes. This variable is significant across all of the country, but the effect is less visible in the disaggregated regressions (Table A.2 shows weak significance in areas with low overall agricultural productivity, however). The scale of the farm, and the access to other resources that land provides, has a positive effect on incomes, particularly in the areas where maize productivity, and agricultural productivity potential in general, are lower. The fact that neither distance to the tarmac road, nor extension, affect income is also of interest. Market infrastructure has steadily improved in Kenya (Chamberlin and Jayne 2009), while some studies have questioned the efficacy of extension (Gautam 2000; Marenya and Barrett 2009). Moreover, public extension may not have as much of a role in the household incomes of rural Kenyans as in the past, simply because nonfarm sources are increasingly more important and other information sources are increasing available to farmers. To say that extension has an insignificant impact on overall household income does not imply, however, that it has no effect on the profitability of certain farm enterprises. The weak negative and significant effect of total annual rainfall is somewhat surprising, although it may be an artifact of the relationship between income, year, and rainfall in the underlying data. The lowest rainfall among all years was in the year with the highest income, when product prices were especially high. Positive year effects are most prominent in 2004 and 2007 overall and in areas with lower productivity potential, but stronger in 2010 in the high potential maize-growing areas.

Table 6: Impacts of hybrid maize on income, all households and by maize productivity

	(1) Income	(2) Income High Maize Potential Area	(3) Income Other Maize-Growing areas
Grow hybrid	2.06*	2.44*	1.92+
	(0.872)	(1.106)	(1.063)
Female head	-0.02	-0.41**	0.08
	(0.082)	(0.128)	(0.100)
Education	0.02*	0.03	0.02
	(0.010)	(0.018)	(0.012)
Young adults	0.03*	0.04*	0.03*
	(0.013)	(0.019)	(0.016)
Mature adults	0.10**	0.10**	0.10**
	(0.015)	(0.026)	(0.018)
Extension	0.00	-0.00	0.00
	(0.004)	(0.005)	(0.005)
Tarmac	-0.00	-0.00	-0.00
	(0.005)	(0.010)	(0.006)
Farm size	0.01*	-0.00	0.02**
	(0.003)	(0.004)	(0.006)
Rainfall	-0.00+	0.00	-0.00+
	(0.000)	(0.000)	(0.000)
2004	0.26**	-0.01	0.27**
	(0.075)	(0.141)	(0.101)
2007	0.16**	0.16+	0.17*
	(0.061)	(0.085)	(0.076)
2010	0.15	0.30*	0.17
	(0.163)	(0.123)	(0.214)
Observations	4,874	1,298	3,576
Number of hhid	1,239	331	908

Source: Authors.
Note: Robust standard errors in parentheses. ** p<0.01, * p<0.05, + p<0.1

As expected given the descriptive statistics reported in Table 4, growing maize hybrids has an overall effect that reduces relative deprivation with respect to income, at the 5% level of significance. When disaggregated by maize productivity potential, however, the effect is statistically weak in both models (Table 7). When grouped by agricultural productivity potential, growing maize hybrids reduces inequality only in the lower potential zones (Table A.3). Again, consistent with the general development literature and past research in Kenya, education improves the status of the average household across the nation (but not by maize-producing or agricultural potential). Meanwhile, the strong significance on the coefficient of the variable recording the number of mature adults in the household attests to the importance of labor constraints in improving the status of farm households with respect to income. Farm size counteracts relative deprivation, particularly in the lower potential areas. Interestingly, higher rainfall weakly increases inequality overall but reduces inequality in the higher productivity areas when these are considered separately. A disturbing finding for Kenyan development is that relative deprivation with respect to income appears to have risen in successive survey waves.

Table 7: Impacts of hybrid maize on relative deprivation (income), all households and by maize productivity

	(1) Relative deprivation	(2) Relative deprivation High maize potential area	(3) Relative deprivation Other maize-growing areas
Grow hybrid	-102,337.52*	-125,073.07+	-96,387.40+
	(46,975.665)	(68,439.429)	(55,574.399)
Female head	5,914.99	31,364.24**	-20.32
	(4,086.534)	(7,421.378)	(4,894.726)
Education	-1,015.55*	-1,737.95	-664.25
	(499.288)	(1,079.668)	(582.856)
Young adults	-990.01	-1,141.05	-1,017.91
	(691.703)	(1,126.517)	(870.766)
Mature adults	-5,936.50**	-5,425.43**	-5,916.21**
	(824.577)	(1,525.862)	(956.694)
Extension	-49.69	272.66	-154.47
	(202.803)	(280.405)	(288.566)
Tarmac	172.63	419.32	185.96
	(280.949)	(579.638)	(310.273)
Farm size	-521.22**	82.35	-1,353.69**
	(198.968)	(219.317)	(334.994)
Rainfall	16.99+	-13.59	11.26
	(9.262)	(27.092)	(9.184)
2004	3,307.64	11,734.25	4,170.89
	(3,971.177)	(8,288.849)	(5,206.285)
2007	24,141.48**	18,740.02**	25,981.27**
	(3,203.987)	(4,882.376)	(3,909.184)
2010	97,948.32**	82,791.41**	99,804.59**
	(8,743.827)	(7,214.781)	(11,169.868)
Observations	4,875	1,298	3,577
Grow hybrid	1,239	331	908

Source: Authors.
Note: Robust standard errors in parentheses ** p<0.01, * p<0.05, + p<0.1

The regression estimating the impact of growing hybrid maize on whether or not total household income falls below the national poverty line is shown in Table 8. Average partial effects (marginal effects) are presented for the second-stage regression with hybrid seed use and residual from the first-stage regression included as explanatory variables. Errors have been bootstrapped for 50 iterations.

The negative coefficient on growing hybrid maize in the poverty regression has strong statistical significance overall and in the lower potential maize-growing areas, but not in the high potential maize-growing environments. This is consistent with the evidence in the data that adoption rates are much higher in the higher potential areas (regardless of household income), and income is also more widely distributed. Growing hybrid maize is the norm for farmers in these environments. We postulate that it is instead the particular hybrid grown, soil fertility and other management practices that differentiate more profitable from less profitable maize production, and additional sources of income that distinguish richer from poorer farm households.

Other coefficients in this regression are of interest. First, education of adults in the household reduces the chances that the household's income will fall below the poverty line across all areas, and in separate regressions except in the high-potential maize-growing areas. There, we surmise that education levels are so high that marginal effects can no longer be as easily detected. T-tests and Kolmogorov-Smirnov tests confirm that years of adult education are higher at the mean and throughout the distribution in the high-potential maize-growing zone than elsewhere. A more surprising result is the positive relationship of numbers of adults in the household to the probability of being poor. It may be that the consumption needs of larger households offset their income-generating capacity, or that households with larger numbers of adults remaining on the farm are those whose income-generating capacity off-farm, where wages are higher, is circumscribed. As in most of the other regressions, distance to the tarmac road and to extension services have no observable effect on the likelihood of falling below the poverty line. Relative to 2000, probabilities of poverty rose in both 2007 and 2010, all other factors held constant. Interestingly, this is visible across all agroecologies, and in the higher potential agricultural areas, but not in the higher potential maize-growing areas. Thus, the pattern according to productivity potential depends very much on the classification of the agroecologies or agroregions.

Table 8: Impact of hybrid maize on poverty, control function with CRE

	(1) Poor All areas Marginal effects	(2) Poor High potential maize area Marginal effects	(3) Poor Other maize-growing areas Marginal effects
Grow hybrid	-0.256***	-0.232	-0.266***
	(0.0549)	(0.224)	(0.0753)
Residual	0.102*	0.00781	0.107
	(0.0553)	(0.214)	(0.0776)
Female head	0.00794	0.151**	-0.0281
	(0.0287)	(0.0685)	(0.0358)
Education	-0.0159***	-0.0150	-0.0164***
	(0.00427)	(0.0104)	(0.00578)
Young adults	0.0284***	0.0329***	0.0265***
	(0.00599)	(0.00919)	(0.00660)
Mature adults	0.0138**	0.0285**	0.0107
	(0.00642)	(0.0135)	(0.00683)
Extension	-0.00157	-0.00150	-0.00119
	(0.00161)	(0.00273)	(0.00225)
Tarmac	-0.00695	-0.00560	-0.00590
	(0.00459)	(0.00818)	(0.00477)
Farm size	-0.00147	-0.000581	-0.00172
	(0.00271)	(0.00536)	(0.00311)
Rainfall	6.77e-05	-5.49e-05	5.85e-05
	(5.18e-05)	(0.000221)	(7.14e-05)
2004	0.00823	0.0509	0.00536
	(0.0208)	(0.0633)	(0.0196)
2007	0.0576***	0.0292	0.0663***
	(0.0168)	(0.0382)	(0.0187)
2010	0.104***	0.0714	0.103***
	(0.0201)	(0.0622)	(0.0322)
Observations	4,844	1,289	3,555

Robust standard errors in parentheses ** p<0.01, * p<0.05, + p<0.1
Source: Authors.
Note: Coefficients of means of time-varying variables not presented. Errors bootstrapped 50 iterations.

Diagnostic statistics for the reported models of the three outcome variables are shown in Table 9. The joint F-test statistics are all significant confirming the strong relationships between the instrumental variables and growing hybrid maize. The results of the endogeneity test indicate rejection of the null hypothesis of exogeneity implying that the adoption variable is endogenous in all the outcome models. The results of the Hansen-J statistic indicate strong evidence of failure to reject the exogeneity of the instrument set, thus giving credence to their validity. In the control function application to estimate poverty outcomes, the significance of the coefficient on the residual in the second stage, combined with the F-test on the instruments in the first stage, is our only test of endogeneity. The significance of the residual leads us to reject the null hypothesis that growing hybrid maize is exogenous in poverty outcomes across all survey areas, although this is not the case when separate regressions are estimated according to agricultural potential or maize productivity (Tables 8, A. 4).

Table 9: Table 9. Summary of diagnostic tests for instrumental variables

Model	F -test	Endogeneity test	Hansen J Statistic
Income			
General	5.10 (0.006)	9.68 (0.002)	0.307 (0.5796)
High maize productivity	4.96 (0.007)	7.380 (0.007)	0.491 (0.4836)
Other maize-growing areas	3.06 (0.0468)	5.53 (0.019)	0.288 (0.5913)
Inequality			
General	5.11 (0.006)	8.957 (0.0028)	1.891 (0.1691)
High maize productivity	4.96 (0.007)	4.979 (0.0257)	0.043 (0.8361)
Other maize-growing areas	3.06 (0.0468)	5.890 (0.0152)	1.007 (0.3156)
Poverty			
General	-80.31 (0.005)	0.102 (0.0.553)	--
High maize productivity	-8.56634(0.01)	0.0078(0.214)	--
Other maize-growing areas	-62.5514	0.107 (0.0776)	--

Source: Authors. Notes: Value of test statistic (p-value).

5.0 Conclusions

To our knowledge, despite more than a half-century of improved maize production in Kenya, and numerous, in-depth adoption studies, the impact of hybrid maize seed on the welfare of smallholders has not been estimated with rigorous quantitative methods. The objective of this paper is to initiate that work. We applied an instrumental variables, fixed effects model and a CRE probit model with the control function to data collected from a balanced panel of 1243 households in four waves (2000, 2004, 2007, and 2010). The impacts we considered were income, income inequality, and poverty status. Diagnostic statistics provide evidence of the endogeneity of hybrid use in all of these outcome variables, due either to self-selection bias or simultaneity.

Comparisons of sample statistics confirm that total household income is higher for maize growers who plant hybrid seed relative to those who do not. The poverty status of maize-growing farm families who do not grow hybrids is substantially higher, in terms of the Foster-

Greer-Thorbecke indices of headcount ratios, the mean poverty gap, and mean severity of poverty. In addition, mean inequality, as measured by Stark's index of relative deprivation, is also higher among non-hybrid maize growers.

Regression results confirm that growing hybrid maize seed is endogenous in income and poverty outcomes as we have defined them, either through self-selection bias, simultaneity, or other underlying factors that cause errors to be correlated between seed choice and income. At the same time, tests of our major hypotheses concerning the positive impact of growing maize seed on income outcomes appear to be robust. In the years covered by Tegemeo survey data (2000 to 2010), data support the hypotheses that growing hybrid maize seed increases total household income and negatively influences the likelihood that household income falls below the national poverty line. In the combined regression for all areas surveyed, and in the lower potential zones, we found no impacts of growing hybrid maize seed on the relative deprivation of households with respect to income. One reason why is that hybrid maize is not as widely grown in these areas, and households are also more dependent on off-farm income sources.

Several results are of particular interest concerning the high potential maize-growing areas of Kenya. On one hand, our results indicate that growing maize hybrids does not reduce poverty in this region. On the other, growing maize hybrids appears to reduce the relative deprivation of individual households compared to others in their location. This does not surprise us given the long period of maize hybrid use in these areas of Kenya, the higher adoption rates, and the wider dispersion of income among adopters than non-adopters. Finally, female headship has a strong negative effect on income in this area alone, other factors held constant.

Among other factors, consistent with the general development literature and other research conducted in Kenya, we observe the strong, positive effect of education on incomes and poverty reduction, and the significance of household labor constraints in generating income. Distance to tarmac roads and extension do not affect poverty, income, or inequality after controlling for other factors. This probably reflects improvements in marketing infrastructure and alternative information sources over time. The predicted seed-to-grain ratio, and the distance to the nearest certified seed seller are strong instruments and therefore major determinants of hybrid maize seed use, which is consistent with economic theory.

Further research might compare the findings derived from fixed effects estimation to those generated by other estimation procedures, such as propensity score matching, and application of control function approach with correlated random effects to estimate the impacts of the *scale* of hybrid seed use (kgs per hectare) on income and poverty. In addition, we are interested in the impacts of hybrid seed use on the diversification of income sources as Kenyan agriculture develops. Our objective in this body research is to assemble a robust set of econometric results that will be of use to decision-makers in the Kenyan maize seed industry.

References

Angrist, J.D. 2001. Estimation of limited dependent variable models with dummy endogenous regressors: Simple strategies for empirical practice. *Journal of Business and Economic Statistics* 19 (1): 2-16**.**

Angrist, J. D. and A.B. Krueger. 2001. Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic Perspectives* 15 (4): 69-85.

Argwings-Kodhek et al. 1999. "How Can Micro-level Household Information Make a Difference for Agricultural Policy Making? Selected Examples from the KAMPAP Survey of Smallholder Agriculture and Non-Farm Activities for Selected Districts in Kenya.

Byerlee D. and C.K. Eicher. 1997. Africa's Emerging Maize Revolution. Boulder, CO: Lynne Rienner Publishers.

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. Tegemeo Working Paper No. 37. Nairobi, Kenya: Tegemeo Institute.

De Groote, H., G. Owuor, C. Doss, J. Ouma, L. Muhammad, and K. Danda. 2005. The Maize Green Revolution in Kenya Revisited. Electronic Journal of Agricultural and Development Economics 2.1: 32-49. Available online at www.fao.org/es/esa/eJADE

Foster, J., J. Greer and E. Thorbecke. 1984. A class of decomposable poverty measures. Econometrica. 2, 81: 761–766.

Gautam, M. 2000. Agricultural Extension: The Kenya Experience. Washington: World Bank.

Gerhart, J. 1975. The Diffusion of Hybrid Maize in Western Kenya. Mexico City: CIMMYT.

Handa, S. and J.A. Maluccio (2010). Matching the Gold Standard: Comparing experimental and non-experimental evaluation techniques for a geographically targeted program. Economic Development and Cultural Change, 58, 415-447.

Hassan, R. 1998. Maize Technology Development and Transfer. Published jointly by Oxon, UK: Centre for Agricultural Bioscience International (CABI); Mexico, D.F.: CIMMYT; and Nairobi, Kenya: KARI.

Heisey, P.W., M.L. Morris, D. Byerlee, and M. López-Pereira. 1998. Economics of hybrid maize adoption. In M.L. Morris (ed.), *Maize Seed Industries in Developing Countries*, Lynne Rienner and CIMMYT, Boulder and Mexico, DF

Jones, A., T. Dalton, and M. Smale. 2012. A stochastic production function analysis of maize hybrids and yield variability in drought-prone areas of Kenya. WPS 49. Tegemeo Institute of Agricultural Policy and Development, Nairobi.

Marenya, P. and C. Barrett. 2009. “Soil Quality and Fertilizer Use Rates among Smallholder Farmers in Western Kenya.” Agricultural Economics 40: 561–572.

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

Mwabu, G., W. Mwangi and H. Nyangito.2007. Does Adoption of Improved Maize Varieties Reduce Poverty?Evidence from Laikipia and Suba Districts in Kenya. Productive Sector Division, Kenya Institute for Public Policy Research and Analysis. KIPPRA Discussion Paper No. 7, Nairobi, Kenya.

Nambiro, E., De Groote, H. & Oluoch K'Osura, W.,2004. The Hybrid Maize Seed Industry in the Transzoia District of Western Kenya. In: Proceedings of the 7th Eastern and Southern AfricaRegional Maize Conference, Nairobi, Kenya, 11 – 15 February 2002. CIMMYT, Mexico, D. F., pp. 474-479.

Ouma, J., F. Murithi, W. Mwangi, H. Verkuijl, M. Gethi, H. De Groote. 2002. Adoption of seed and fertiliser technologies in Embu District, Kenya. Mexico, D.F.: Kenya Agricultural Research Institute (KARI) and International Maize and Wheat Improvement Center (CIMMYT).

Salasya, B.D.S., W. Mwangi, H. Verkuijl, M.A. Odendo, and J.O. Odenya. 1998. An assessment of adoption of seed and fertilizer packages and the role of credit in smallholder maize production in Kakamega and Vihiga Districts. Mexico, D.F.: Kenya Agricultural Research Institute (KARI) and International Maize and Wheat Improvement Center (CIMMYT).

Schaffer, M.E. 2010. Xtivreg2: Stata module to perform extended IV/2SLS, GMM and AC/HAC/LIML and k-class regression for panel data models. http//ideas.repect.org/c/bod/bocode/s456501.html. Accessed May 5, 2012.

Smale, M., and T.S. Jayne. 2010. “Seeds of Success” in Retrospect: Hybrid Maize in Eastern and Southern Africa. In *Successes in African Agriculture: Lessons for the Future,* ed. S. Haggblade and P.B.R. Hazell. Baltimore: Johns Hopkins University Press for the International Food Policy Research Institute.

Smale, M., J. Olwande, and H. De Groote. 2012. Demand for Maize Hybrids and Hybrid Change on Farms in Kenya. Mimeo. Michigan State University and Tegemeo Institute, East Lansing, MI and Nairobi, Kenya.

Smith, R.J., and R.W. Blundell. 1986. An exogeneity test for a simultaneous equation tobit model with an application to the labor supply. Econometrica 50(3): 679-685.

Stark, O. and E. Taylor. 1989. Relative Deprivation and International Migration. *Demography* 26: pp. 1-14.

Suri, T. 2011. Selection and Comparative Advantage in Technology Adoption. *Econometrica* 79.1: 159-209.

Swanckaert, J.,2012. Mapping the effect of market liberalisation policies on the maize seed systems in Kenya based on micro-evidence from 1992 to 2010. MSc. thesis, University of Ghent, Belgium.

Wangia, C., S. Wangia, and H. De Groote. 2004. Review of maize marketing in Kenya: Implementation and impact of liberalisation, 1989-1999, In D. K. Friesen and A.F. E. Palmer, eds. Integrated Approaches to Higher Maize Productivity in the New Millennium. Proceedings of the 7th Eastern and Southern Africa Regional Maize Conference Nairobi, Kenya, February 2002. CIMMYT,, Mexico, D. F.

Wekesa E., W. Mwangi, H. Verkuijl, K. Danda, H. De Groote. 2002. Adoption of maize production technologies in the Coastal Lowlands of Kenya. Mexico, D.F.: CIMMYT and KARI.

Wooldridge, J.W. 2002. Econometric Analysis of Cross Section and Panel Data. Massachusetts Institute of Technology (MIT) Press, Cambridge.

Yitzhaki, S. 1979. Relative Deprivation and the Gini Coefficient. *The Quarterly Journal of Economics* 93: 321-324.

Annexes

Figure A.1. Histograms of total household income, by hybrid seed use, with normal curve

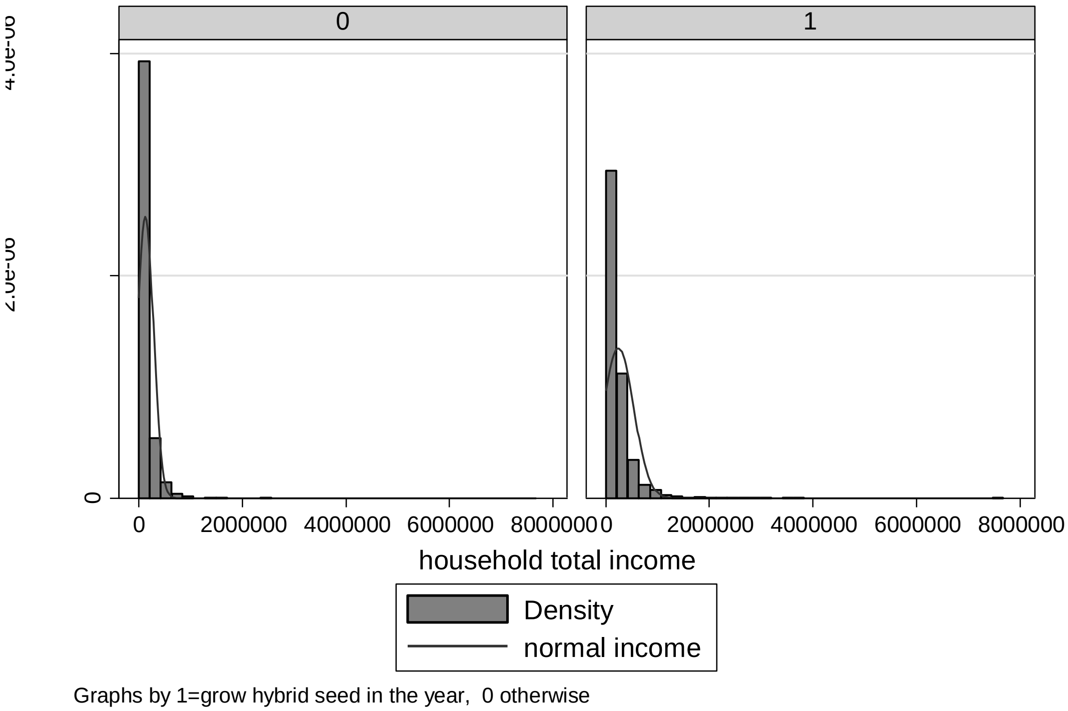

Figure A.2. Histograms of natural logarithm of total household income, by use of hybrid seed, with normal curve

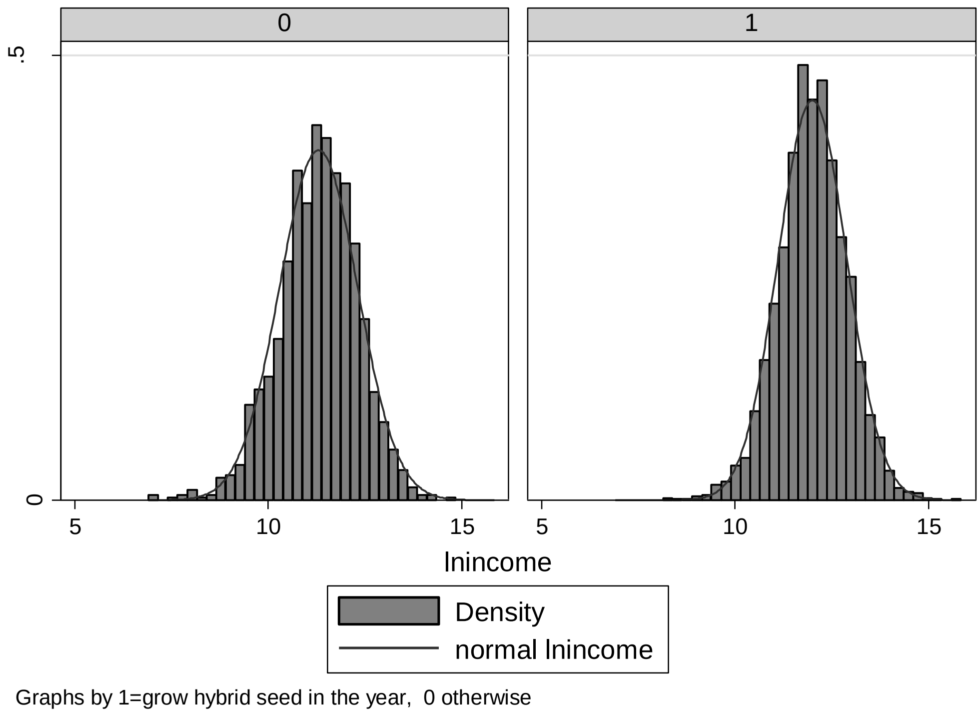

Figure A.3. Histogram of relative deprivation (income), by use of maize hybrids

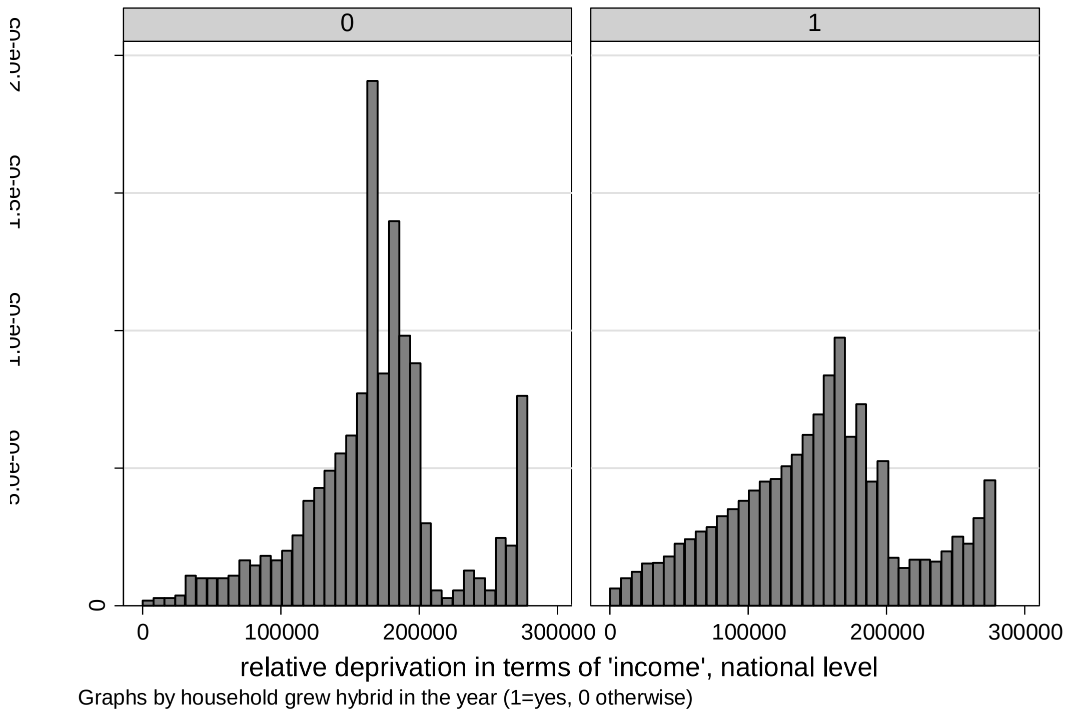

Table A.1. Preliminary pooled, random effects, fixed effects regressions

	(1) Pooled OLS	(2) Random Effects	(3) Fixed Effects	(4) PooledIV	(5) FEIV2
Grow hybrid	0.47**	0.30**	0.07*	0.95**	2.06*
	(0.026)	(0.028)	(0.032)	(0.222)	(0.872)
Female head	-0.32**	-0.28**	-0.12*	-0.26**	-0.02
	(0.029)	(0.035)	(0.053)	(0.038)	(0.082)
Education	0.09**	0.08**	0.02**	0.08**	0.02*
	(0.004)	(0.005)	(0.007)	(0.007)	(0.010)
Young adults	0.05**	0.05**	0.04**	0.05**	0.03*
	(0.007)	(0.007)	(0.008)	(0.007)	(0.013)
Mature adults	0.13**	0.12**	0.10**	0.14**	0.10**
	(0.009)	(0.009)	(0.010)	(0.010)	(0.015)
Extension	-0.01*	-0.00	0.00+	-0.00	0.00
	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
Tarmac	0.00	-0.00	-0.00	0.00*	-0.00
	(0.002)	(0.002)	(0.004)	(0.002)	(0.005)
Land	0.02**	0.02**	0.00+	0.02**	0.01*
	(0.001)	(0.002)	(0.003)	(0.002)	(0.003)
Rain	-0.00**	-0.00**	-0.00	-0.00**	-0.00+
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
2004	0.20**	0.17**	0.11**	0.24**	0.26**
	(0.032)	(0.026)	(0.027)	(0.039)	(0.075)
2007	0.28**	0.28**	0.27**	0.26**	0.16**
	(0.030)	(0.025)	(0.025)	(0.033)	(0.061)
2010	0.39**	0.44**	0.51**	0.30**	0.15
	(0.032)	(0.028)	(0.030)	(0.053)	(0.163)
Constant	10.39**	10.58**	11.08**	10.19**	
	(0.053)	(0.062)	(0.079)	(0.105)	
Observations	4,895	4,895	4,895	4,878	4,874
R-squared	0.414		0.154	0.371	-0.823
Number of hhid		1,243	1,243		1,239

Robust standard errors in parentheses** p<0.01, * p<0.05, + p<0.1

Table A.2. Impact of hybrid maize on income, areas of high and low agricultural productivity potential

	(1) Income High potential	(1) Income Low potential
Grow hybrid	4.13*	1.06+
	(2.003)	(0.601)
Female head	-0.21+	0.15
	(0.122)	(0.122)
Education	0.01	0.03+
	(0.018)	(0.016)
Young adults	0.03	0.04*
	(0.020)	(0.019)
Mature adults	0.09**	0.12**
	(0.027)	(0.022)
Extension	-0.00	0.00
	(0.006)	(0.005)
Tarmac	0.00	-0.00
	(0.009)	(0.007)
Farm size	0.00	0.01+
	(0.004)	(0.006)
Rainfall	0.00+	-0.00
	(0.000)	(0.000)
2004	0.01	0.39**
	(0.067)	(0.099)
2007	0.05	0.37**
	(0.090)	(0.090)
2010	0.22	0.34
	(0.137)	(0.209)
Observations	3,336	1,538
Number of hhid	845	394

Robust standard errors in parentheses ** $p<0.01$, * $p<0.05$, + $p<0.1$

Table A.3. Impact of hybrid maize on relative deprivation, areas of high maize potential and other maize-growing areas

	(1) Relative Deprivation High agricultural productivity potential	(1) Relative Deprivation Low agricultural productivity potential
Grow hybrid	-157,046.19	-84,001.47*
	(96,763.629)	(34,680.969)
Female head	14,775.81**	-5,382.34
	(5,197.120)	(6,647.332)
Education	-663.13	-1,045.10
	(778.998)	(839.857)
Young adults	-1,092.73	-2,108.64+
	(903.705)	(1,196.204)
Mature adults	-5,353.59**	-7,142.07**
	(1,164.300)	(1,386.842)
Extension	91.76	4.15
	(285.444)	(307.921)
Tarmac	17.00	267.46
	(398.386)	(410.576)
Farm size	-319.19	-916.63*
	(215.854)	(384.281)
Rainfall	-30.79*	7.05
	(15.031)	(12.331)
2004	15,775.69**	-5,360.76
	(2,933.883)	(5,398.881)
2007	25,291.81**	24,267.25**
	(4,122.949)	(5,049.916)
2010	86,600.69**	106,747.00**
	(6,502.433)	(12,040.385)
Observations	3,336	1,539
Number of hhid	845	394

Robust standard errors in parentheses ** p<0.01, * p<0.05, + p<0.1

Table A.4. Impact of hybrid maize on poverty status, areas of high and low agricultural productivity potential

	Controlling for potential endogeneity		Exogeneity	
	High potential margins	Low potential margins	High potential margins	Low potential margins
Grow hybrid	-0.185**	-0.207**	-0.148***	-0.0815***
	(0.0775)	(0.102)	(0.0161)	(0.0253)
Residual	0.0401	0.132		
	(0.0745)	(0.0964)		
Female head	0.0462	-0.0209	0.0473	-0.0165
	(0.0357)	(0.0639)	(0.0326)	(0.0522)
Education	-0.0125**	-0.0208**	-0.0127**	-0.0208**
	(0.00498)	(0.00819)	(0.00591)	(0.0105)
Young adults	0.0292***	0.0249*	0.0290***	0.0258***
	(0.00744)	(0.0139)	(0.00597)	(0.00989)
Mature adults	0.0136*	0.0162	0.0140*	0.0156
	(0.00736)	(0.0124)	(0.00842)	(0.0146)
Extension	-0.00132	-0.00267	-0.00125	-0.00238
	(0.00225)	(0.00266)	(0.00243)	(0.00296)
Tarmac	-0.00347	-0.0117	-0.00309	-0.0107
	(0.00625)	(0.00776)	(0.00609)	(0.0104)
Farm size	-0.00319	0.00216	-0.00309	0.00255
	(0.00338)	(0.00429)	(0.00310)	(0.00469)
Rainfall	-5.65e-05	-2.41e-05	-6.09e-05	-4.87e-05
	(8.94e-05)	(0.000116)	(9.34e-05)	(0.000104)
2004	0.0635***	-0.0497	0.0684***	-0.0367
	(0.0206)	(0.0351)	(0.0241)	(0.0355)
2007	0.0588***	0.0482	0.0580***	0.0360
	(0.0199)	(0.0437)	(0.0198)	(0.0391)
2010	0.0753***	0.0642	0.0704***	0.0339
	(0.0279)	(0.0468)	(0.0258)	(0.0481)
Observations	3,315	1,529	3,330	1,540