Hybrid Seed, Income, and Inequality among Smallholder Maize Farmers in Zambia

by

Melinda Smale and Nicole Mason

Working Paper 72

March 2013

Indaba Agricultural Policy Research Institute (IAPRI)

Lusaka, Zambia

Downloadable at: http://www.aec.msu.edu/fs2/zambia/index.htm

Hybrid Seed, Income, and Inequality among Smallholder Maize Farmers in Zambia

by

Melinda Smale and Nicole Mason

Working Paper 72

March 2013

Smale is professor and Mason is assistant professor, Department of Agricultural, Food, and Resource Economics, Michigan State University, East Lansing, MI. Mason is currently based at the Indaba Agricultural Policy and Research Institute, Lusaka.

ACKNOWLEDGEMENTS

The Indaba Agricultural Policy Research Institute (IAPRI) is a non-profit company limited by guarantee and collaboratively works with public and private stakeholders. IAPRI exists to carry out agricultural policy research and outreach, serving the agricultural sector in Zambia so as to contribute to sustainable pro-poor agricultural development.

The authors gratefully acknowledge funding from the USAID Zambia Mission (USAID/Zambia) and the International Maize and Wheat Improvement Center (CIMMYT) through a grant from the Standing Panel on Impact Assessment (SPIA) of the CGIAR.

The authors would like to thank Patricia Johannes for her helpful formatting and editing assistance.

Any views expressed or remaining errors are solely the responsibility of the authors.

Comments and questions should be directed to:

The Executive Director
Post Net 99
Indaba Agricultural Policy Research Institute
26A Middleway, Kabulonga,
Lusaka.
Landline: +260 211 261194;
Fax: +260 211 261199;
Email: kabaghec@iconnect.zm

INDABA AGRICULTURAL POLICY RESEARCH INSTITUTE TEAM MEMBERS

The Zambia-based Indaba Agricultural Policy Research Institute (IAPRI) research team is comprised of Chance Kabaghe, Thomas. S. Jayne, Stephen Kabwe, Rhoda Mofya Mukuka, Munguzwe Hichaambwa, Auckland Kuteya, Mary Lubungu, Nicole Mason, Brian Mulenga, Arthur Shipekesa, Nicholas Sitko, and Solomon Tembo. Michigan State University-based researchers associated with IAPRI are Margaret Beaver, Eric Crawford, Steven Haggblade, Chewe Nkonde, Melinda Smale, and David Tschirley.

EXECUTIVE SUMMARY

The successful development and diffusion of improved maize seed in Zambia is a major achievement in African agriculture. Since independence in 1964, and most notably from the 1970s, the commitment of the Government of Zambia (GoZ) to ensuring food security through the involvement of smallholder farmers was an important aspect of this process. Most recently, the commitment of the GoZ has entailed the re-introduction of input subsidies.

Over the decades, a number of in-depth case studies have explored the patterns and determinants of hybrid seed adoption among smallholder farmers in Zambia, but to our knowledge, none of these had the opportunity to utilize a large panel data set to test the impacts of adoption on smallholder farmers. Here, we are able to assess the impacts of hybrid seed use on the income and inequality of maize-growing, smallholder farmers in Zambia using survey data collected in 2002/3 and 2006/7 from a nationally-representative, statistical sample of roughly 3,500 households.

We test the effects of hybrid seed use on the total household income (including farm and nonfarm sources) of smallholder maize growers and on two measures in income inequality. The first, Stark's index of relative deprivation, compares the relative rank of households with respect to total household income. This measure has the advantage that it is specific to each household, and is thus well-suited to econometric models based on household data. We also explore two more well-known indexes that are derived from general entropy theory, the Theil index, and the Mean Logarithmic Deviation (MLD). The disadvantage of these is that they are designed for use in summary statistics comparing one population group to another. To use them in our model, we must employ only the portion of the index that varies over households.

Panel data enables us to apply Fixed Effects (FE) and Correlated Random Effects (CRE) models to control for time-invariant, unobserved heterogeneity among farmers. Use of a two-year panel renders this estimate analogous to a difference-in-difference analysis. Conveniently, the first survey season also marks the inaugural year of the Fertilizer Support Program. Since hybrid seed may be endogenous in the equations explaining income outcomes, we apply instrumental variables regression. Applied with a binary dependent variable (adoption of maize hybrids) in the first stage, the two-stage, instrumental variables approach generates estimates of the local average treatment effect (LATE), or the impact of hybrid seed use on users. Similarly, seed subsidy receipt may be endogenous in the decision to plant hybrid seed, and we test this hypothesis using a Control Function Approach (CFA), also controlling for unobserved heterogeneity with CRE model.

We find that growing hybrids increased gross nominal income of smallholder maize growers by an average of 29% between 2002/3 and 2006/7, other factors held constant. Relative deprivation with respect to income, or income inequality as measured by Stark's index, was reduced when maize growers plant hybrids. Considering that the dependent variable was scaled by 1,000, the average partial effect as a percentage of mean of relative deprivation was 5%. As we have modified the entropy indices of inequality for use in a household model, the Mean Logarithmic Deviation also declines significantly with use of hybrid seed. Thus, by all three of the income outcomes we used, the impact of hybrid seed use on maize-growing smallholders in the early years of the subsidy program was positive.

These findings are consistent with the widespread perception that smallholder maize growers who do not grow hybrids in Zambia are strongly disadvantaged relative to those who do. The impact of not growing hybrids is pervasive is not only associated with lower overall household income but worsening relative status with respect to income. Thus, despite other

evidence that farmers who receive subsidies may not be among the worse off in terms of land and other endowments, including political capital, our results indicate that the subsidy program may be successfully addressing the goals of raising smallholder income and flattening inequality with respect to income. Further research should examine other welfare indicators, and test these hypotheses during the more recent subsidy period.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii
INDABA AGRICULTURAL POLICY RESEARCH INSTITUTE TEAM MEMBERS iv
EXECUTIVE SUMMARY v
LIST OF TABLES viii
LIST OF FIGURES viii
ACRONYMS ix
1. INTRODUCTION 1
2. HISTORICAL CONTEXT 3
3. METHODS 5
3.1. Conceptual Approach 5
3.2. Data 6
3.3. Outcome Variables 7
3.4. Econometric Models 10
3.5. Explanatory and Instrumental Variables 11
4. FINDINGS 13
4.1. Descriptive Statistics 13
4.2. Econometric Analysis 15
5. CONCLUSIONS 19
REFERENCES 21

LIST OF TABLES

TABLE **PAGE**

1. Variable Definitions 7
2. Percent Distribution of Households Planting Hybrid Maize Seed, by Subsidy Receipt.....13
3. Comparison of Means of Outcome Variables, Users and Non-users of Maize Hybrids.....14
4. Comparison of Means of Explanatory Variables, Users and Non-users of Maize Hybrids 14
5. Impact of Seed Subsidy on Likelihood of Planting Hybrid Seed.....16
6. Impact of Hybrid Seed Use on Income and Income Inequality 18

LIST OF FIGURES

FIGURE **PAGE**

1. Histogram of Theil Index for Individual Households, by Hybrid Seed Use 9
2. Histogram of MLD Index for Individual Households, by Hybrid Seed Use.....9

ACRONYMS

CFA	Control Function Approach
CGIAR	Consultative Group on International Agricultural Research
CIMMYT	International Center for Maize and Wheat Improvement
CRE	Correlated Random Effects
CSO	Central Statistical Office
CV	Coefficient of Variation
DTMA	Drought Tolerant Maize for Africa
FAO	Food and Agricultural Organization of the United Nations
FE	Fixed Effects
FISP	Farmer Input Support Program
FSP	Fertilizer Support Program
FSRP	Food Security Research Project
GoZ	Government of Zambia
ha	Hectare
hh	Households
IAPRI	Indaba Agricultural Policy Research Institute
kg	Kilogram
LATE	Local Average Treatment Effect
MACO	Ministry of Agriculture and Cooperatives, Zambia
MAL	Ministry of Agriculture and Livestock
MLD	Mean Logarithmic Deviation
MSU	Michigan State University
RD	Relative Deprivation
SPIA	Standing Panel on Impact Assessment
U.S.	United States
USD	U.S. Dollar
USAID	United States Agency for International Development
ZARI	Zambian Agricultural Research Institute
ZMK	Zambia Kwacha

1. INTRODUCTION

The successful development and diffusion of improved maize in Zambia represents a major scientific and policy achievement in African agriculture (Smale and Jayne 2010). The commitment of the Government of Zambia (GoZ) to ensuring food security has played a pivotal role in realizing maize productivity gains on smallholder farms. During the boom period of 1970-89, the GoZ incurred large fiscal costs on parastatal grain and seed marketing, provision of services to maize growers, and a pan-territorial pricing scheme (Howard 1994; Howard and Mungoma 1997). Under pressure from the International Monetary Fund, the World Bank, and other donors, this burdensome, integrated system was gradually dismantled during the 1990s.

The seed industry was among the first sectors to be liberalized during structural adjustment. By 2010, 203 maize varieties had been released to farmers, over 100 of which were grown by farmers in the 2010/11 grow season (De Groote et al. 2012). The Zambian maize seed industry can boast a larger number of seed companies compared to other maize-reliant nations in the region (Kassie et al. forthcoming); five of these are major companies that finance their own seed research (SeedCo, MRI, Pannar, Zamseed, and Monsanto).

Nonetheless, the consequences of structural adjustment during the 1990s, which were aggravated by poor rainfall conditions, appeared to be unfavorable for food security–which has long been equated with national maize production in Zambia. Use of fertilizer and improved seed plummeted, area planted to maize contracted as farmers shifted to other crops, and maize yields declined. In 2002, in response to this situation, the Government of Zambia re-established input subsidies for maize production. GoZ's justification for the establishment of the Fertilizer Support Program (FSP) in 2002 was that a public role was needed to manage the transition toward full market liberalization, rebuild the resource base of smallholders, and help instill a sense of self-reliance (MACO 2008). In general, the subsidies have emphasized the fertilizer component of the input package, given the country's landlocked status and the fact that progress in seed industry liberalization was already evident. However, hybrid maize seed has been included in the subsidized input pack since the inception of FSP.

Although a number of insightful case studies about the adoption and impacts of maize hybrids on smallholder farmers have been conducted, to our knowledge, no nationally representative analysis has been conducted since the re-introduction of input subsidies. Previous studies on maize hybrids in Zambia have generally focused on factors affecting adoption and the impacts of adoption on maize production. To our knowledge, none of these studies has examined the income and inequality effects of adoption on smallholders.

For example, during the late 1980s, survey-based research conducted by the International Food Policy Research Institute (IFPRI) in the Eastern Province of the country explored the fertilizer use and gender aspects of hybrid seed use (Jha and Hojjati 1993; Kumar 1994). Jha and Hojjati found high rates of fertilizer use on both hybrids and local maize, well supported by the favorable price environment of that time. Kumar concluded that although virtually all farmers with over five hectares adopted hybrids, farmers with smaller landholdings also adopted hybrid maize, and grew it more profitably. Only a fifth of female-headed, maize-growing households planted hybrids, as compared to a third of male-headed households. Kumar also explored the effects of hybrid seed use on consumption and nutrition with an instrumental variables approach.

More recently, case study, survey-based research has been conducted by the Zambian Agricultural Research Institute (ZARI) and the International Center for Maize and Wheat

Improvement (CIMMYT). The ZARI/CIMMYT study was conducted with a sample of 300 households selected in ten randomly selected villages of three randomly selected districts (Mkushi, Katete, and Sinazongwe) in three provinces (Eastern, Central, and Southern) representing moderate, optimal, and marginal maize-growing conditions. After stratifying households according to their access to productive assets, Langyintuo and Mungoma (2008) estimated a double-hurdle model adoption model, concluding that the factors influencing both the adoption and use intensity of improved maize seed differed by wealth group. In the CIMMYT study conducted in two drought-prone districts of southern Zambia, Kalinda et al. (2010) also found that better-off households grew more improved maize varieties. Furthermore, the authors signaled a marked difference in technology use between better-off and poorer households, as well as between male- and female-headed households.

Our contribution in this analysis is to rigorously assess the impacts of hybrid seed use on the income and income inequality of maize-growing, smallholder farmers in Zambia using panel survey data collected in 2002/3 and 2006/7 from a nationally-representative, statistical sample of roughly 3,500 households. We measure income outcomes in terms of three indicators: total household income, Stark's index of relative deprivation, and the mean logarithmic deviation. Panel data enables us to apply FE and CRE models to control for time-invariant, unobserved heterogeneity among farmers. Use of a two-year panel renders this estimate analogous to a difference-in-difference analysis. Conveniently, the first survey season also marks the inaugural year of the Fertilizer Support Program.

Hybrid seed use may be endogenous in the equations explaining income outcomes. To test and control for this potential endogeneity in the equations, we apply instrumental variables regressions. Applied with a binary dependent variable (adoption of maize hybrids) in the first stage, the two-stage, instrumental variables approach generates estimates of the local average treatment effect (LATE), or the impact of hybrid seed use on users (Angrist and Krueger 2001). Similarly, seed subsidy receipt may be endogenous in the decision to plant hybrid seed. We test this hypothesis using a combination a Control Function Approach (CFA) (Wooldridge 2010), controlling for unobserved heterogeneity with Correlated Random Effects (CRE) model (Chamberlain 1984; Mundlak 1978).

The next section provides some useful historical context regarding maize research, seed diffusion, and input subsidies in Zambia. Section 3 presents the conceptual approach, data, and econometric methods. Findings are presented in the fourth section, including descriptive characterization of users and non-users of hybrid seed, and regression estimates. In the fifth section, we draw conclusions and policy implications.

2. HISTORICAL CONTEXT

Maize became the dominant food crop in Zambia during the first half of the 20th century, yielding higher returns than the previous staple cereals, sorghum, and millet. Easier to process and market, particularly as an export to the British starch market, maize was also an easy way to pay workers on the large-scale farms and in the mines of the Federation of Rhodesia and Nyasaland (Smale and Jayne 2010).

Howard (1993) described the structure and evolution of Zambia's national agricultural research system and maize research from the colonial period. The orientation of maize research, as in Zimbabwe (then Southern Rhodesia), Kenya, and Malawi (then Nyasaland), was originally toward the needs of expatriate commercial farmers. The research network in Zambia, distributed across agro-ecological zones, dates back to the 1920s. The first permanent station was established at Mount Makulu in 1950.

Before independence, Zambia (Northern Rhodesia) relied on its Federation partner Southern Rhodesia (now Zimbabwe) for maize seed suitable for its commercial farmers. Southern Rhodesia had a hybrid maize breeding program as early as 1932, which developed the SR52 and SR11 hybrids. Except for some local variations in grain color, all maize varieties bred for commercial farmers and grown by small-scale farmers for subsistence were white in color. Small-scale farmers who could not afford hybrids generally planted unimproved maize varieties, which were open-pollinated, long-season, and typically flinty in grain texture. Hybrids introduced from Southern Rhodesia before Zambian independence are reported to have more than doubled yields on commercial farms (Howard 1994).

After independence in 1964, the focus of the research remained the same because of the economic pre-eminence of mining and the need to provide cheap food to a relatively large urban population in Zambia. A major turning point occurred during the mid-1970s, when the collapse of copper prices led the government to explore the potential for small-scale farming. A number of maize breeders worked for the Zambian national research system during those years, supported by British, Yugoslav, Swedish, U.S., and FAO funding. In the late 1970s, the first Zambian professionals joined the maize breeding program. Zambia's maize breeders released an impressive array of double and three-way cross hybrids with yield advantages, even without fertilizer, in all but the most difficult growing environments (Howard 1994).

With abundant land, favorable weather, and the establishment of a seed industry (including the parastatal seed company, Zamseed), the period from 1970 to 1989 was characterized by adoption rates that climbed to 65% and a 4.9% annual rate of growth in maize yields (Smale and Jayne 2010). Fertilizer subsidies, pan-seasonal and pan-territorial pricing, and geographically dispersed market depots bolstered rates of return to maize production even in remote areas. Considering the full cost of seed development, extension, and marketing, however, the rate of return to maize research investment was negative (Howard and Mungoma 1997). GoZ's integrated system of parastatal grain and seed marketing, provision of services to maize growers, and pan-territorial pricing for maize proved fiscally unsustainable, and pressure from the International Monetary Fund, World Bank and other donor agencies led to a series of market reforms to remove restrictions on trade and encourage withdrawal of the state from market operations. Given the importance of maize as a wage good in Zambia, riots engendered by rising maize meal prices, and drought, reform implementation was gradual during the 1990s.

Combined with a decade of droughts and food crises in 1991–1992 and 2002–2003, policy shifts induced a retraction of maize production areas, a drop in hybrid adoption from an estimated 72% in 1990 to under 22% in 1996 (López-Pereira and Morris 1994; Hassan, Mekuria, and Mwangi 2001), and a decline in fertilizer use on maize. In response to this situation, seed and fertilizer subsidies were re-established in 2002 under the Fertilizer Support Program. In each cropping season since 2002, fertilizer and maize seed have been distributed through the program. Seed and fertilizer were subsidized at the same rate (50 to 60%) from 2002/03 through 2007/08. Since 2008/09, GoZ has set different subsidy rates for seed and fertilizer: 75 to 79% for fertilizer and 50 to 53% for seed (MAL 2012).

The stated goal of the current Farmer Input Support Program (FISP), formally established in 2009/10, has been to improve access of small-scale farmers to inputs while enhancing the participation and competitiveness of the private sector. During that season, the size of the package was halved from 400 kilograms (kg) of fertilizer and 20 kg of hybrid seed (the standard package under FSP) to facilitate diffusion to a larger number of smallholders. Through the time of this writing, the FISP has operated by selecting private suppliers through a tender process. Local transporters distribute inputs to designated collection points, and selected cooperatives and other farmer organizations issue inputs to approved farmers.

3. METHODS

3.1. Conceptual Approach

We view impact pathways among smallholder farmers in Zambia from the viewpoint of the non-separable model of the household farm, in which family members organize their labor to maximize utility over consumption goods and leisure in an economic environment with market failures, such as overt input price distortions caused by subsidies (de Janvry, Fafchamps, and Sadoulet 1991). Households produce goods for consumption or sale, and cash constraints are relaxed primarily through farm sales and off-farm employment of family members.

In the Zambian setting, household characteristics affect commercial input demand and also access to subsidies via registered farmer organizations. Growing higher-yielding hybrid maize can contribute to augmenting household income by releasing land to produce other crops or by generating cash from maize sales. Thus, households growing hybrids may be ranked more highly in their communities with respect to income relative to the period before adoption. 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 as a component of total household income. In that case, we would hypothesize very little impact on income inequality.

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 (e.g., de Janvry, Dustan, and Sadoulet 2010). The motivation for these approaches is understood as essentially one of missing data (Ravallion 2005). That is, we observe the values of outcome variables, such as indicators of income or income inequality, 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. In the case of hybrid seed use among smallholders in Zambia, we observe values of outcome variables for adopters and non-adopters, but not for adopters had they not adopted. In Zambia, there is potential for both self-selection into the treatment group and targeting through the subsidy program.

From a large body of previous empirical research on the adoption of agricultural technologies, we hypothesize that adopters are generally those who are wealthier in terms of various types of capital (human, natural, physical) 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 as well as to explicit targeting via registered farmers' groups, reflects the fact that adopters may be 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, since, as described above, Zambian farmers have a long exposure to maize hybrids. 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 Zambian farmers.

Panel data methods (e.g., fixed effects, first-differencing, and CRE methods) are also designed to control for unobservable factors that do not vary over time but could be correlated with observed variables (de Janvry, Dustan, and Sadoulet 2010; Ravallion 2005; Wooldridge 2010). Often, this unobserved heterogeneity is considered to capture intrinsic features that are not easily measured (such as farm management capability, unmeasurable soil quality, etc.). Difference-in-difference models compare 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 those obtained through first-differencing.

We consider instrumental variables combined with panel data methods as the best suited to our data-generating process and hypothesized relationship of adoption to outcome variables. 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 income and income inequality through hybrid seed use.

In summary, we consider that either because of selection on observables (e.g., farm size, education, labor supply) or unobservables (e.g., intrinsic management ability, unmeasurable soil quality), simultaneity in decision-making or feedback processes, the decision by Zambian smallholder farmers to grow hybrid maize seed is potentially endogenous in the income outcomes. In the presence of endogeneity, estimators generated by ordinary least squares are biased. Our approach estimates the local average treatment effect (LATE): the effect of self-selection or targeted selection into the group of farmers that uses maize hybrids on outcome variables identified through instrumentation.

3.2. Data

The data were collected using a nationally-representative sample of smallholder farm households in 70 districts of Zambia. The sampling frame was designed by the Zambian Central Statistical Office (CSO) and the survey was conducted as a supplement to the Post-Harvest Survey by the CSO, Ministry of Agriculture and Cooperatives (MACO) and Food Security Research Project (FSRP) of Michigan State University. In the first wave of the survey (1999/2000), which is not used here because maize seed details were not recorded, a total of 6,922 smallholder farmers were interviewed. During the second wave, covering the 2002/03 agricultural year, a total of 5,358 households were re-interviewed. Of these, 4,286 were re-interviewed in the 2006/07 agricultural year. Only maize-growing households were used in the current study, numbering 3,553, and 3,542 in the latter two survey waves, respectively.

Although attrition bias may be a source of concern, there is no regression-based test to test for attrition bias when FE or CRE modes are employed with two time periods only. Earlier research (Mason 2011) based on the same data set and all three waves of data found no evidence of bias when the test specified by Wooldridge (2010) was applied.

3.3. Outcome Variables

We consider three indicators of smallholder farmer income: total household income, Stark's index of relative deprivation (RD), and the Mean Logarithmic Deviation (MLD) (Table 1).

Table 1. Variable Definitions

Variable	Definition
Dependent variables	
plant hybrid seed	1=planted F1 hybrid seed in survey year, 0 else
income	natural logarithm of total household income from farm and off-farm sources
RD	Stark's index of relative deprivation with respect to income, scaled by 1000 ZMK (see text)
MLD	Mean logarithmic deviation in income (see text)
Explanatory variables	
distance to district town	km from center of SEA to nearest district town
distance to tarmac road	km from center of SEA to nearest tarred/main road
distance to feeder road	km from center of SEA to nearest feeder road
distance to transport	km from homestead to nearest point to get vehicular transport
female household head	1=household headed by female, 0 else
education of head	education of household head (years)
maximum adult education	maximum years of education of adults 15-59 years
adults 15-59 years	number of adults in prime age group (15-59 years)
farm size (ha)	total farm size in hectares
farm size, squared	square of total farm size
crop richness	count of crops grown
crop evenness	Simpson index of crop grown (1 - sum of squared area shares)
own livestock	1= own livestock, 0 else
off-farm employment	count of off-farm employment categories (0-4) (business/informal; salaried; remittance; work on other farm)
rainfall	Growing season rainfall (mm, Nov.-Mar.)
rainfall stress	Rainfall stress (# of 20-day periods, Nov.-Mar., with <40 mm total rainfall)
expected rainfall	5 yr. average growing season rainfall (Oct-May, '00 mm)
expected rainfall, squared	5 yr. average growing season rainfall (Oct-May, '00 mm), squared
rainfall cv	CV of 5 year growing season rainfall (Oct-May, %)
maize price	Log of lagged real price of maize (2006=100)
groundnut price	Log of lagged real price of ground nuts (2006=100)
mixed bean price	Log of lagged real price of mixed beans (2006=100)
sweet potato price	Log of lagged real price of sweet potatoes (2006=100)
weeding wage	Log of real wage to weed 0.25 ha field (2006=100)
fertilizer price	Log of real commercial fertilizer price (2006=100)
2006	1=agricultural year is 2006/7, 0 else
Instrumental variables	
maize seed subsidy	1=received subsidized seed, 0 else
fertilizer allocation	government fertilizer distributed in district under FSP (kg/agricultural household)
point spread	Percentage-point spread between ruling party & lead opposition

Source: Authors. N=7075. ZMK=Zambian Kwacha. Prime-age refers to ages 15-59. CV=coefficient of variation.

Household income is comprised of crop income (gross value of crop production); livestock income (gross value of livestock and livestock product sales); salaries of household members; business income of household members; income from informal labor employment of household members; and remittances, pension and share dividends received by all household members. We take the natural logarithm of nominal income, smoothing the skewness in the distribution and enabling us to interpret the regression coefficient on hybrid seed use as the percentage change. According to Wooldridge (2010), the use of natural logarithms in combination with year dummies achieves the same results as using real incomes.

We measure inequality with Stark and Taylor's (1989) index of relative deprivation, calculated as:

$$(1)\quad RD(y_i) = AD(y_i) * P(y_i)$$

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 household *i* and $P(y_i)$ is the proportion of households in the sample that are richer than household *i* (Stark and Taylor 1989). To construct the index, sample households were ranked by income from lowest to highest. 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 with respect to gross nominal household income. We scale the index by ZMK 1,000 for easier interpretation. An important advantage of Stark's index is that it has been developed specifically for use in household modeling. Stark and Taylor (1989) applied the index to analyze the impacts of migration on rural households in Mexico.

For validation of results, we considered the Theil index and the MLD as second outcome indicators. Both are derived from general entropy theory, and thus satisfy important criteria related to symmetry, mean independence, and testability, among others (Haughton and Khandker 2009). Values of these summary measures, when computed over the full distribution, vary between zero and infinity, with zero representing an equal distribution and higher values representing higher levels of inequality. The Theil index (1967) is computed as:

$$(2)\quad Theil = \frac{1}{N}\sum_{i-1}^{N}\left(\frac{y_i}{\mu} * \ln\frac{y_i}{\mu}\right)$$

Where y_i is the income of the *i*th household among *N* households, and μ is mean income. The Theil index is sensitive to changes in the higher tail of the income distribution. When this expression is inverted, the index can be recomputed as the mean logarithm deviation (MLD). By construction, the MLD emphasizes lower income values:

$$(3)\quad MLD = \frac{1}{N}\sum_{i-1}^{N}\left(\ln\frac{\mu}{y_i}\right)$$

Typically, as shown in equations (2) and (3), inequality measures of this type are used to summarize income distributions for population subgroups. In our application, we use only the expression within the summation sign, which reflects the relative position of the individual household. This is important because our indicator is not a summary index but is distributed across all households in the sample.

Figures 1 and 2 show the distributions of the Theil and MLD indexes as calculated from our data. The distribution is relatively smooth for the MLD but not for the Theil index. Given these shapes, we use only the MLD index.

Figure 1. Histogram of Theil Index for Individual Households, by Hybrid Seed Use

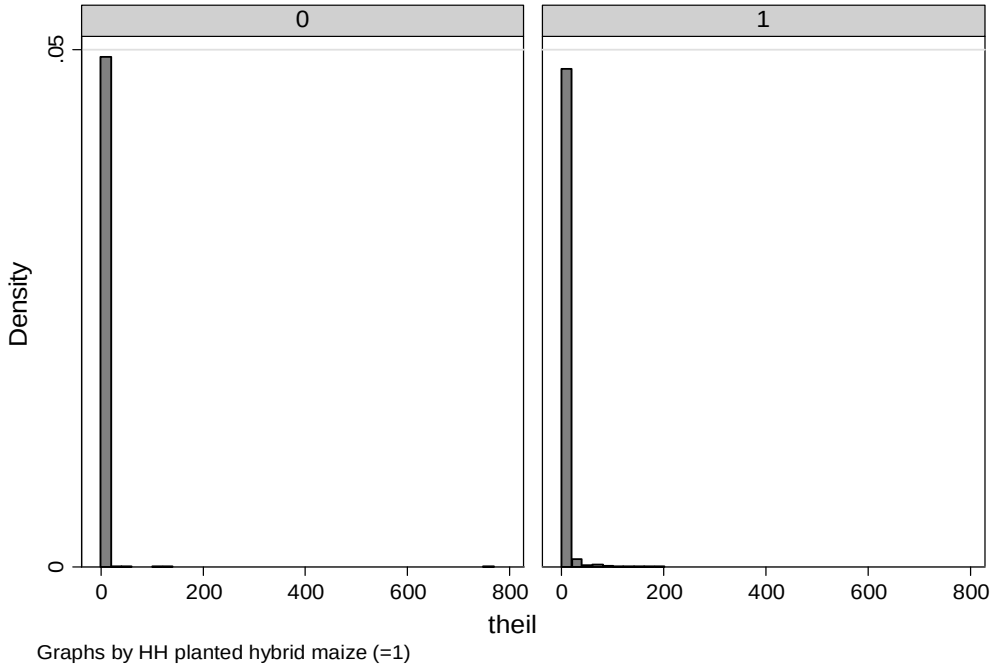
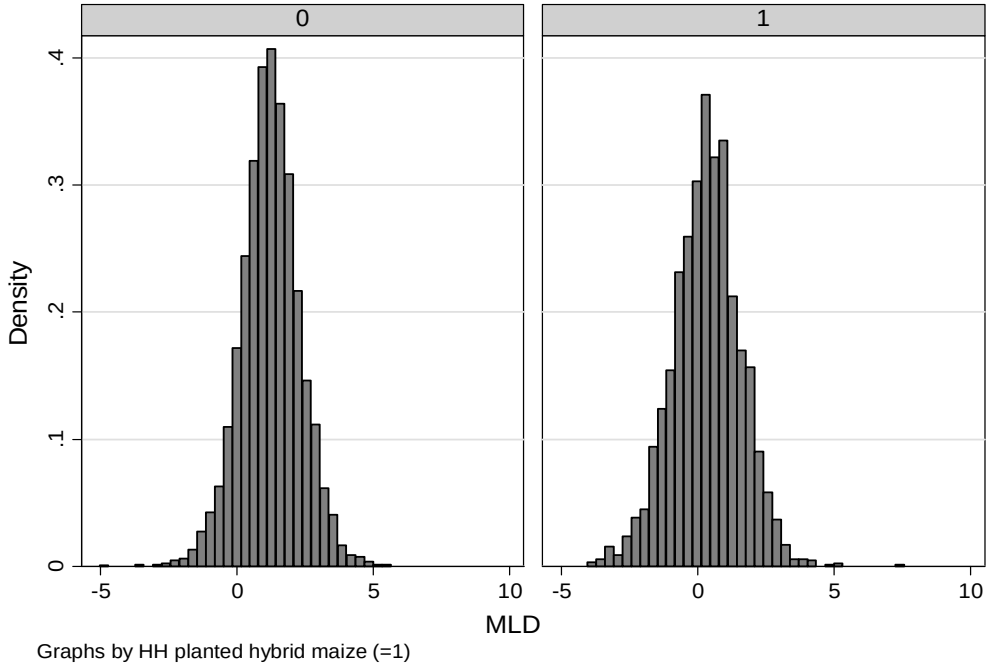


Figure 2. Histogram of MLD Index for Individual Households, by Hybrid Seed Use



3.4. Econometric Models

We apply panel data methods to estimate separate regressions for each outcome variable of interest according to the following general model:

$$(4)\quad Y_{it} = \beta_0 + \boldsymbol{\beta_1 X}_{it} + \boldsymbol{\gamma} Z_{it} + \alpha_i + \varepsilon_{it}, \text{ where } i = 1,..,N \; t=1,..,T.$$

The dependent variable Y is the outcome indicator of interest, $\boldsymbol{X}$ is a vector of exogenous explanatory variables, Z is a binary variable for hybrid seed use, and α_i represents time-invariant unobserved factors that affect the outcome variable. The random error term, $\boldsymbol{\varepsilon}_{it}$, is also unobserved but varies over time.

As noted above, panel data methods help to control for unobservable factors that do not vary for individual farm households (α_i) but could be correlated with the included explanatory variables ($\boldsymbol{X}$ and Z). In addition, we hypothesize that hybrid seed use (Z) may be endogenous in the outcome equations as a consequence of self-selection, subsidy selection or simultaneity in decision-making. To control for potential endogeneity of hybrid seed use when outcome variables are continuous (maize production, income, income inequality), we apply the Fixed Effects Instrumental Variables method developed for estimation with panel data using two-stage least squares (FEIV2SLS, Schaffer 2010). Based on Kelejian (1971), Angrist and Krueger (2001) have argued 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.

Standard model diagnostics of the FEIV2SLS procedure (Schaffer 2010) 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 i) the evaluation of the joint F-test for excluded instruments in the first stage regression; ii) Hansen's J test for over-identifying restrictions; and iii) the Kleibergen-Paap statistic, which provides a test for the weakness (under-identification) of instruments. Rejection of the null hypothesis in the F-test 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 variables are exogenous in the structural equation, thereby supporting the validity of the instruments. We estimated all FEIV2SLS and FE (when we failed to reject exogeneity of hybrid seed use) models with fully robust variance-covariance matrices.

Given the significance of maize seed subsidy receipt as a selection mechanism for the use of hybrid seed, subsidy receipt is a potential instrumental variable for hybrid seed use. However, subsidy receipt may also be endogenous to seed use. Therefore, before estimating the general model (the effects of hybrid seed adoption on household welfare), we test the potential endogeneity of subsidy receipt in the decision to plant hybrid seed using a Control Function Approach. The CFA enables us to test and control for endogeneity bias in nonlinear models. As in a 2SLS model, the 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 both included 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 CFA is described in early work by Blundell and Smith (1989), and is now widely used. If the test leads to a failure to reject exogeneity, the structural regression simply includes the binary variable for subsidy receipt as one of the exogenous instruments; if the null hypothesis of exogeneity is rejected, Wooldridge (2010) recommends

estimating the two regressions with binary dependent variables (plant hybrid seed, subsidy receipt) with bivariate probit to control for endogeneity.

Following Mundlak (1978) and Chamberlain (1984), we apply the CRE model along with the CFA to control for unobserved heterogeneity. The CRE model requires that the means of time-varying explanatory variables be included in the regression, generating a fixed-effects interpretation.

3.5. Explanatory and Instrumental Variables

Explanatory variables, including instrumental variables, are shown in Table 1. Following the underlying conceptual framework of decision-making on the household farm, determinants of seed demand include household and market characteristics that influence shadow prices, as well as observed prices, and agro-environmental parameters that influence choices through crop production conditions. Farmgate seed prices were not collected in the survey, and only national seed prices are available for analysis. These are perfectly collinear with the year dummy included in the model. We use a year dummy to capture between year variation, and market access variables to exploit cross-sectional variation in seed prices.

Variables that capture household characteristics include proxies for human capital, or labor quality (education of the household head, maximum education of prime-age adults in the household) and labor quantity (the number of adults in the prime-age category, 15-59 years). A dummy variable tests the importance of female as compared to male headship, which past research in Zambia (Section 1) suggests is an important parameter for access to goods and services, resources, and information. Asset endowments are measured by farm size and its square and by livestock ownership. Since farm size is highly correlated with the total value of farm assets, only farm size, along with its square (to capture second-order effects), is included in the regression. To express the role of diversification of household income sources using exogenous variables, we include counts of crop and off-farm income sources, and a Simpson index of the evenness of crop area shares. Market characteristics are measured by a vector of price variables (maize, competing crops, fertilizer and wage), each in logarithmic form, and distances to district town, feeder road, tarmac road, and vehicular transport. The first three distance variables are constant across the two time periods and drop out in the fixed effects estimation. Prices are lagged, corresponding to a naïve price expectations model of input choices.

Agroenvironmental factors are included as actual rainfall and rainfall stress in the agricultural year of the survey, since this will affect the outcome variables, as well as expected rainfall based on a five-year period and the coefficient of variation over the same period, which affect input decisions via expectations.

Two sets of instrumental variables are used: those that identify receipt of the maize seed subsidy and those that identify hybrid maize adoption. For the former, we hypothesize that the amount of fertilizer allocated to the district by the FSP program, per agricultural household, is pre-determined and although it increases the probability that households in a district will receive a fertilizer subsidy and a seed subsidy through input complementarity, it is uncorrelated with unobserved factors affecting individual households' receipt of subsidized inputs. Use of the fertilizer allocation, instead of the seed allocation, as the instrumental variable enables us to exploit the complementarity between fertilizer and seed. A second instrumental variable in the maize seed subsidy equation is the percentage point spread between the ruling party and lead opposition in the constituency of the household. Variables

of this type, which express the political capital of an area, have been significant in a number of empirical analysis of input subsidies (Banful 2011; Mason and Ricker-Gilbert forthcoming; Pan and Christiaensen 2012).

The maize seed subsidy is the instrumental variable in the hybrid seed use equation, since we hypothesize that it affects outcome variables only through seed and previous empirical studies have found seed subsidy receipt to be exogenous to input choices (Ricker-Gilbert, Jayne, and Chirwa 2010; Mason and Ricker-Gilbert forthcoming; Smale and Birol 2013).

4. FINDINGS

4.1. Descriptive Statistics

The percentages of households receiving subsidized seed and planting hybrid maize in the two study years are summarized in Table 2. Sample data suggest that 11% and 10% of maize-growing households received a seed subsidy in 2002 and 2006, respectively. Of farmers who planted hybrid seed, over a quarter in 2002 (28%) and slightly under a quarter in 2006 (23%) received subsidized seed. Considering both years, only 0.63% of maize-growing households (numbering 27) reported that they received subsidized seed but did not plant hybrids. These farmers appear to have planted seed of improved open-pollinated varieties that was received from the Food Security Pack Programme, which distributed both types of improved seed.

Average values of outcome indicators are presented in Table 3 by hybrid seed use, along with several other key descriptors. At the mean, hybrid growers produced over three times as much per farm—although these data do not take into count the scale of maize area planted. In fact, hybrid seed users had more than twice the mean maize area (1.6 ha) of maize growers who did not plant hybrids (0.8 ha). Although differences are highly significant, average maize yields per farm are not as high as would be hoped for hybrid growers (1.9 t/ha) compared to farmers who did not plant hybrids (1.3 t/ha), likely due to the great variability in management, fertilizer application, and biophysical features of the production environments.

Mean differences in gross total household income are even more pronounced. The average RD with respect to income is substantially greater among non-users of maize hybrids. The MLD is significantly higher among farmers who do not grow hybrids as a group. Differences in means are statistically significant across all outcome indicators at less than 1%. Kolmogorov-Smirnov tests also indicate that each of the outcome variables is distributed differently between farmers who did and did not plant hybrid seed.

Table 2. Percent Distribution of Households Planting Hybrid Maize Seed, by Subsidy Receipt

Plant hybrid maize seed	Received subsidy, 2002		All maize growers	Receive subsidy, 2006		All maize growers
	no	yes		no	yes	
no	2,210	22	2,232	2,061	5	2,066
	(99.01%)	(0.99%)	(100%)	(99.76%)	(0.24%)	(100%)
yes	932	368	1,300	1,128	343	1,471
	(71.69%)	(28.31%)	(100%)	(76.68%)	(23.32%)	(100%)
All maize growers	3,142	390	3,532	3,189	348	3,537
	(88.96%)	(11.04%)	(100%)	(90.16%)	(9.84%)	(100%)

Source: Authors. Pearson chi-squared test shows that subsidy distributions differ by seed use at less than 1% statistical significance.

Table 3. Comparison of Means of Outcome Variables, Users and Non-users of Maize Hybrids

	Plant hybrid			All maize growers
	no	yes		
Maize production (kgs)	990	3,340	***	1,911
Maize area planted (ha)	0.777	1.632	***	1.112
Maize yields (kgs/ha)	1346.122	1971.727	***	1590.944
Income (ZMK)	2,860,202	8,086,364	***	4,908,822
RD	4315.046	3706.693	***	4077.172
MLD	1.231561	.3308254	***	.8784174

Source: Authors. Significant differences with two-tailed difference of means tests (***)=1%. The exchange rate was 4785.62 ZMK/USD during the 2002/03 agricultural season (Oct-Sept) and 4048.72 ZMK/USD during the 2006/07 season (Bank of Zambia).

Table 4. Comparison of Means of Explanatory Variables, Users and Non-users of Maize Hybrids

	Plant hybrid maize seed					All maize growers	
	no		yes				
	mean	st. dev.	mean	st. dev.		mean	st. dev.
distance to district town	35.45	22.91	30.23	19.85	***	33.41	21.91
distance to tarmac road	30.23	41.48	17.11	25.42	***	25.09	36.61
distance to feeder road	3.41	3.23	2.93	2.66	***	3.22	3.03
distance to transport	8.51	15.06	5.44	10.06	***	7.31	13.41
female household head	0.23	0.42	0.16	0.37	***	0.20	0.40
education of head	4.50	3.49	6.47	3.97	***	5.27	3.81
maximum adult education	6.76	3.18	8.80	3.16	***	7.57	3.32
adults 15-59 years	3.02	1.78	3.79	2.23	***	3.32	2.00
farm size (ha)	2.04	1.92	3.37	4.61	***	2.56	3.31
crop richness	2.65	1.25	2.53	1.27	***	2.61	1.25
crop evenness	0.39	0.23	0.31	0.24	***	0.36	0.24
own livestock	0.79	0.41	0.87	0.34	***	0.82	0.39
off-farm employment	0.56	0.69	0.75	0.76	***	0.64	0.72
rainfall	1.57	1.56	2.10	1.73	***	1.78	1.64
rainfall stress	9.40	1.92	8.67	2.12	***	9.11	2.03
expected rainfall	92.03	36.66	79.61	39.32	***	87.16	38.21
rainfall cv	6.68	0.38	6.66	0.39	**	6.67	0.39
maize price	7.49	0.19	7.54	0.22	**	7.51	0.20
groundnut price	7.46	0.12	7.43	0.11	***	7.45	0.12
mixed bean price	5.85	0.34	5.83	0.35	***	5.84	0.35
sweet potato price	3.45	0.32	3.47	0.35	**	3.46	0.33
weeding wage	7.86	0.23	7.85	0.20	*	7.86	0.22
fertilizer price	7.86	0.23	7.85	0.20	***	7.86	0.22

Source: Authors. Significant differences with two-tailed difference of means tests (***)=1%; (**)5%; (*)=10%.

Means of explanatory variables are compared for users and non-users of maize hybrids in Table 4. Strong statistical differences are also evident for all hypothesized determinants of outcome indicators. Not only are hybrid users closer to any means of transport and market

infrastructure, but rainfall was higher and rainfall stress was lower in their areas during the survey year. Interestingly, expected rainfall was higher among farmers who did not plant hybrids. Hybrid users are less rich in crop numbers and evenness of crop value shares than non-users, suggesting greater specialization in maize or a cash crop, but they have larger overall farm sizes. Hybrid users have more adult labor, their household heads have more education, and household heads are less likely to be female. They are more likely to own livestock. Average prices of maize, other major crops, weeding labor and fertilizer also appear to be statistically different between users and non-users of hybrid seed in this large sample; however, differences do not appear to be meaningful in magnitude.

Bivariate statistics confirm expected differences between means of maize-growing households who planted and did not plant hybrid seed in the survey years, not only for outcome variables but for hypothesized determinants of impact outcomes. This lends support to our proposed regression model. Econometric findings are presented next.

4.2. Econometric Analysis

In this sub-section, we first present the preliminary regression estimating the impact of the maize seed subsidy on the likelihood that a smallholder farmer plants hybrid maize. Next, we present the results of models estimating the impacts of hybrid seed use on outcome variables of interest.

Table 5 shows the structural (second-stage) regression results predicting the impact of the maize seed subsidy on growing hybrid maize with the CRE probit model. Both hypothesized instruments, fertilizer allocation and political capital, are statistically significant in the first-stage regression at less than 5%. The residual from the reduced-form regression in the first stage (predicting seed subsidy receipt) is not, however, significant in the structural model. Thus, we fail to reject the exogeneity of the subsidy in the maize hybrid adoption equation. This finding reflects the fact that the program has been nationally-managed, with seed distributed through registered farmers associations that meet specified criteria. The marginal effect of the seed subsidy is a 58 percentage point increase in the likelihood that a farmer grows hybrid seed, all other factors held constant. Looking back at the descriptive statistics, this makes sense given that the overall coverage of the subsidy is roughly 10% of maize growers, while about 25% of farmers who planted hybrid seed over the two years combined received subsidized seed. Similar findings with respect to the exogeneity of the subsidy in input choices were reported for Malawi by Ricker-Gilbert, Jayne, and E. Chirwa (2010) and for Zambia by Mason and Ricker-Gilbert (forthcoming) and Smale and Birol (2013).

Other findings of interest in this preliminary regression include the strong positive effects of closer proximity to roads and transport on hybrid seed use, as well as higher education among the adults in the household and farm size, which remains positive even after including the effect of the squared term. That is, throughout the range of the sample data, farm size has an increasingly positive effect on the probability that maize-growing households plant hybrids. According to Ai and Norton (2003), compared to a linear regression model, in logit and probit regressions such as this one, the coefficient of the squared term is estimated manually and its effect is included in the coefficient of the linear term.

The greater the evenness in the cropping portfolio, however, the less likely hybrid seed is planted. Higher rainfall in the survey year, and higher expected rainfall based on previous seasons, also positively influences the chances that maize growers plant hybrid seed.

Table 5. Impact of Seed Subsidy on Likelihood of Planting Hybrid Seed

	Plant hybrid seed		
	Average partial effect		P-value
Received seed subsidy	0.583	***	0.000
distance to district town	-0.002	***	0.000
distance to tarmac road	0.0000611		0.708
distance to feeder road	-0.00707	***	0.001
distance to transport	-0.00148	**	0.024
female household head	-0.0196		0.489
education of head	0.00405		0.174
maximum adult education	0.00608	*	0.076
adults 15-29 years	-0.000425		0.931
farm size (ha)	0.00805	**	0.012
farm size, squared			
crop richness	0.00415		0.639
crop evenness	-0.124	***	0.005
own livestock	0.0444	***	0.000
off-farm employment	-0.0000000		0.843
rainfall	0.000105	***	0.008
rainfall stress	-0.000695		0.931
expected rainfall	-0.0154	*	0.088
expected rainfall, squared			
rainfall cv	-0.00109		0.317
maize price	-0.0375		0.366
groundnut price	-0.00428		0.936
mixed bean price	0.198	***	0.004
sweet potato price	0.00413		0.841
weeding wage	-0.0391		0.101
fertilizer price	0.111	**	0.012
2006	0.0682	**	0.048
Number of observations=6755			
Wald chi2(52) = 230.19			
Prob > chi2 = 0.0000			
Log pseudo likelihood = -296408.03			
Std. error adjusted for 3751 clusters in hhid			

Source: Authors. Model estimated with probit CRE. Both instruments (fertilizer allocation and point spread) significant at less than 5% in stage 1; residual p-value=0.169 in stage 2. Effects of means of time-varying explanatory variables are not reported.

Interestingly, the mixed bean price, which is a complementary output to maize because it is often intercropped in portions of the field, also has a positive association with planting hybrid seed. In 2006, as shown in the descriptive statistics, more maize growers planted hybrids.

Following this first step of the analysis, we used receipt of the maize seed subsidy as an instrumental variable to test the endogeneity of hybrid seed use in the three outcome equations. We fail to reject the exogeneity of planting hybrid seed in all three outcomes. With respect to the income model, this finding is consistent with the evidence that despite their reliance on maize for consumption and sale, Zambian smallholders have diversified income sources. They also have a long history of growing maize hybrids and relatively high adoption rates with broad diffusion, due in part to a strong maize seed industry backed by decades of seed-fertilizer subsidies. Failure to reject exogeneity of planting hybrid seed also attests to a

limited role of self-selection in adoption. With respect to our inequality indexes, which are constructed as metrics from household income data and are not choice variables, failure to reject exogeneity also makes sense.

The fixed effects models (Table 6) confirm the strong, positive effect of planting maize hybrids on total household income. The log-linear form of the income equation permits us to interpret the estimated coefficient as evidence of a 29% increase associated with planting hybrid seed among users, other factors held constant. Female headship offsets this effect almost entirely! Educational attainment and prime-age labor supply influence income positively. Gross household nominal income increases in the scale of the farm operation at a decreasing rate, and increases in 2006 relative to 2002 by 82%. The richness of cropping activities and off-farm employment are positive factors in income generation. Expected rainfall, but not actual rainfall, exerts a positive influence on gross household income through affecting decision-making; but so does the coefficient of yield variation, perhaps because it leads households to modify their labor allocation decisions toward income-earning activities that buffer their income streams because there are uncorrelated.

Signs and significance for relative deprivation, which is based on gross nominal income, mirror the income regression. That is, planting hybrid seed reduces deprivation of the household with respect to income, relative to other households in the sample. In addition, increases in human capital reduce deprivation, as do land endowments and the richness of the cropping portfolio and off-farm employment. Evenness among crop value shares worsens the income status of households, attesting to the advantages of crop specialization among smallholder farmers in Zambia. Although nominal income grew in 2006 relative to 2002, so did relative deprivation, on average.

Similarly, growing hybrid seed reduces inequality among maize growers by 28% (given the logarithmic structure of the dependent variable, and binary explanatory variable). The behavior of the regression is very similar to the income regression, reflecting the construction of the independent variable. Again, female headship increases inequality, while human capital reduces it, as does landholding size, at a decreasing rate. The numbers of crops grown and income-earning activities, and ownership of livestock, improve equality. Mean logarithmic deviations declined by a large percentage in 2006 relative to 2002.

Table 6. Impact of Hybrid Seed Use on Income and Income Inequality

	Income			Relative Deprivation			Mean Logarithmic Deviation		
	Coeff		p-value	Coeff		p-value	Coeff		p-value
Plant hybrid seed	0.287	***	0.000	-216.0	***	0.000	-0.287	***	0.000
distance to transport	0.000419		0.802	-0.0112		0.993	-0.00042		0.802
female household head	-0.240	***	0.006	162.467	***	0.008	0.241	***	0.006
education of head	0.0210	**	0.019	-13.774	**	0.050	-0.021	**	0.019
maximum adult education	0.0242	**	0.010	-13.175	*	0.062	-0.0242	**	0.010
adults 15-59 years	0.0474	***	0.002	-39.914	***	0.001	-0.0474	***	0.002
farm size (ha)	0.0816	***	0.000	-67.588	***	0.000	-0.0816	***	0.000
farm size, squared	-0.000883	***	0.000	0.565	***	0.000	0.000883	***	0.000
crop richness	0.231	***	0.000	-159.562	***	0.000	-0.231	***	0.000
crop evenness	-0.190		0.150	484.594	***	0.000	0.19		0.150
own livestock	0.146	***	0.009	-52.959		0.182	-0.146	***	0.009
off-farm employment	0.0000003	***	0.000	-0.000297	***	0.000	-0.0000003	***	0.000
rainfall	-0.000127		0.190	0.0772		0.300	0.000127		0.190
rainfall stress	0.0728	***	0.005	-47.240	**	0.015	-0.0728	***	0.005
expected rainfall	0.269	**	0.025	-198.539	**	0.047	-0.269	**	0.025
expected rainfall, squared	-0.0122	**	0.032	9.217	**	0.049	0.0123	**	0.032
rainfall cv	0.0102	***	0.004	-3.329		0.193	-0.0102	***	0.004
maize price	0.280	**	0.017	-30.253		0.742	-0.279	**	0.017
groundnut price	-0.227		0.140	419.610	***	0.001	0.227		0.140
mixed bean price	0.445	**	0.048	-536.223	***	0.004	-0.445	**	0.048
sweet potato price	-0.106		0.104	123.0475	**	0.017	0.106		0.104
weeding wage	0.0283		0.714	14.118		0.814	-0.0283		0.714
fertilizer price	0.103		0.455	-309.286	***	0.003	-0.103		0.455
2006	0.819	***	0.000	1598.765	***	0.000	-0.819	***	0.000
constant	7.61	***	0.008	7828.792	***	0.000	8.09	***	0.005
Observations	6,751			6,751			6,751		
R squared (within)	0.2630			0.797			0.2701		
F (24, 3749)	41.69			469.81			41.69		
Number of hh	3,750			3,750			3,750		

Source: Authors. Model estimated with FE and fully robust standard errors.

5. CONCLUSIONS

Zambian smallholder farmers have been exposed to improved maize seed since before independence, when some of the early blockbuster successes of maize breeding in southern Africa (e.g., SR52) were diffused among commercial farmers. Many of these early materials may have also diffused through the smallholder farmers and farm laborers who worked on and alongside these farms. It was not until the 1970s, however, when the strong commitment of the Government of Zambia to food security led to maize research focused on smallholder needs, and to the elaboration of an integrated input and output marketing scheme to support it. Though this scheme was gradually removed during the 1990s by structural adjustment programs, the threat of food (read *maize*) insecurity motivated the reinstatement of input subsidies in 2002.

Though a number of insightful, in-depth case studies have analyzed the impacts of hybrid seed use on smallholder farmers in Zambia, a contribution of this study is to exploit Zambia's statistically representative, national database in order to assess impacts using the standard econometric approach of the treatment model. The study also goes beyond previous work on maize hybrids in Zambia by estimating the effects of adoption on income and income inequality. We use a two-year panel of about 3,500 maize-growing households, and apply instrumental variables regressions to measure the local average treatment effect of hybrid seed use out these outcomes.

We consider first the role of maize seed subsidies, and failing to reject the exogeneity of the subsidy in the decision to use maize hybrids, we use subsidy receipt as an exogenous explanatory variable in subsequent regressions. The exogeneity of the subsidy is not surprising given the manner in which the subsidy has been centrally administered via registered farmers' associations.

We also fail to reject exogeneity of the decision to plant hybrid seed in total household income, which highlights the diversification of income sources among maize growers in this relatively land-abundant agricultural economy. Failure to reject exogeneity of hybrid seed use in the inequality regressions reflects that these are metrics calculated from income data rather than choice variables. We find that growing hybrids increased gross nominal income by 29% during the 2002 and 2006 study period, other factors held constant. Relative deprivation with respect to income, or income inequality as measured by Stark's index, is reduced when maize growers plant hybrids. Considering that the dependent variable was scaled by 1,000, the average partial effect as a percentage of mean of relative deprivation is 5%. As we have modified the indices for use in a household model, the MLD also declines significantly with use of hybrid seed.

Our findings underscore two fundamental points for Zambian policy. The first is that maize growers who do not grow hybrids are strongly disadvantaged relative to those who do. Our regression results confirm what our descriptive statistics suggested–users and non-users of hybrids are distinct populations. The impact of not growing hybrids is pervasive, since it is not only associated with lower maize yields, but also lower overall household income and a lower relative status with respect to income. Thus, despite evidence of enduring clientelism (Mason and Ricker-Gilbert forthcoming), and that farmers who receive subsidies may not be among the worse off in terms of land and other endowments, our results provide evidence that the subsidy program may be successfully addressing the goals of raising smallholder income and flattening inequality with respect to income.

Further research is needed to assesses the cost-effectiveness of the program and identify ways to enhance its inequality-reducing impacts. At the time these data were collected, adoption rates remained relatively low (36.8%), although more recent estimates attain 58.4% for F1 hybrid seed alone, with another nearly 20% of farmer growing improved open-pollinated varieties and unidentified or recycled improved seed. Overall, both the revealed demand for improved maize seed and the estimates presented here attest to the positive effects of rising adoption rates among smallholder farmers in Zambia.

REFERENCES

Ai, C. and E.C. Norton. 2003. Interaction Terms in Logit and Probit Models. *Economics Letters* 80.1:123-29.

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.

Banful, A.B. 2011. Old Problems in the New Solutions? Politically-Motivated Allocation of Program Benefits and the "New" Fertilizer Subsidies. *World Development* 39.7: 1166-76.

Bank of Zambia. http://www.boz.zm/

Blundell, R.W. and R.J. Smith. 1989. Estimation in a Class of Simultaneous Equation Limited Dependent Variable Models. *Review of Economic Studies* 56: 37-58.

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

De Groote, H., Z. Gitonga, M. Smale, E. Kasuta, D. Asare-Marfo, K. Sonder, E. Birol. 2012. Adoption and Diversification of Modern Maize Varieties in Zambia in 2011: A Baseline Report Prepared for HarvestPlus. Washington, D.C.: HarvestPlus. Mimeo.

de Janvry, A., A. Dustan, and E. Sadoulet. 2010. Recent Advances in Impact Analysis Methods for Ex-post Impact Assessments of Agricultural Technology: Options for the CGIAR. Increasing the Rigor of Ex-post Impact Assessment of Agricultural Research: A Discussion on Estimating Treatment Effects. Berkeley, CA: CGIAR. Mimeo.

de Janvry, A., M. Fafchamps, and E. Sadoulet. 1991. Peasant Household Behavior with Missing Markets: Some Paradoxes Explained. *The Economic Journal* 101.409: 1400-17.

Hassan, R., M. Mekuria, and W. Mwangi. 2001. *Maize Breeding Research in Eastern and Southern Africa: Current Status and Impacts of Past Investments Made by Public and Private Sectors, 1966-1997*. Mexico, D.F.: CIMMYT.

Haughton, J. and S.R. Khandker. 2009. Inequality Measures. In *Handbook on Poverty and Inequality, Chapter 6*. Washington, D.C.: The Work Bank.

Howard, J. 1993. *The Impact of Investments in Maize Research and Dissemination in Zambia: Part One, Main Report*. East Lansing: Michigan State University.

Howard, J.A. 1994. The Economic Impact of Improved Maize Varieties in Zambia. Ph.D. dissertation, Michigan State University.

Howard, J. and C. Mungoma. 1997. Zambia's Stop-and-Go Maize Revolution. In *Africa's Emerging Maize Revolution*, ed. D. Byerlee and C.K. Eicher. Boulder, CO: Lynne Rienner Publishers.

Jha, D. and B. Hojjati. 1993. *Fertilizer Use on Smallholder Farms in Eastern Province, Zambia*. IFPRI Research Report No. 94. Washington, D.C.: International Food Policy Research Institute.

Kalinda, T., G. Tembo, E. Kuntashula, A. Langyintuo, W. Mwangi, and R. La Rovere. 2010. *Characterization of Maize Producing Households in Monze and Kalomo Districts in Zambia*. Drought Tolerant Maize for Africa (DTMA) Project Country Report. Harare: International Maize and Wheat Improvement Center (CIMMYT) and Lusaka: University of Zambia.

Kassie, G.T., O. Erenstein, W. Mwangi, P. Setimela, and J. MacRobert. Forthcoming. *Analysis of the Maize Seed Industry in Southern Africa: Trends and Implications.* CIMMYT Socioeconomics Program Working Paper. Mexico, D.F.: CIMMYT.

Kelejian, H.H. 1971. Two-stage Least Squares and Econometric Systems Linear in Parameters but Nonlinear in the Endogenous Variables. *Journal of the American Statistical Association* 66.334: 373-74.

Kumar, Shubh K. 1994. *Adoption of Hybrid Maize in Zambia: Effects on Gender Roles, Food Consumption, and Nutrition.* IFPRI Research Report No. 100. Washington, D.C.: International Food Policy Research Institute.

Langyintuo, A. and C. Mungoma. 2008. The Effect of Household Wealth on the Adoption of Improved Maize Varieties in Zambia. In *Food Policy* 33: 550-59.

López-Pereira, M.A. and M.L. Morris. 1994. *Impacts of International Maize Breeding Research in the Developing World, 1966-1990*. Mexico, D.F.: CIMMYT.

Mason, N. 2011. Marketing Boards, Fertilizer Subsidies, Prices, and Smallholder Behavior: Modeling and Policy Implications for Zambia. Ph.D. dissertation, Michigan State University.

Mason, N. and J. Ricker-Gilbert. Forthcoming. Disrupting Demand for Commercial Seed: Input Subsidies in Malawi and Zambia. Forthcoming in *World Development.*

Ministry of Agriculture and Cooperatives (MACO). 2008. *Fertilizer Support Programme Internal Evaluation*. Lusaka, Zambia: MACO. Accessed 16 November 2012. Available at http://www.aec.msu.edu/fs2/zambia/tour/FSP_Internal_Evauation_2008.pdf

Ministry of Agriculture and Livestock (MAL). 2012. *Farmer Input Support Programme Implementation Manual: 2012/2013 Agricultural Season*. Lusaka, Zambia: MAL.

Mundlak, Y. 1978. On the Pooling of Time Series and Cross Section Data. *Econometrica* 46: 69-85.

Pan, L. and L. Christiaensen. 2012. Who is Vouching for the Input Voucher? Decentralized Targeting and Elite Capture in Tanzania. *World Development* 40.8: 1619-33.

Ravallion, M. 2005. Evaluating Anti-poverty Programs. In *Handbook of Development Economics, Volume 4,* ed. J. Strauss and T.P. Schultz. Amsterdam, The Netherlands: North-Holland.

Ricker-Gilbert, J., T.S. Jayne, and E. Chirwa. 2011. Subsidies and Crowding Out: A Double-Hurdle Model of Fertilizer Demand in Malawi. *American Journal of Agricultural Economics* 93.1: 26-42.

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. Accessed 5 May, 2012. http//ideas.repect.org/c/bod/bocode/s456501.html.

Smale, M. and E. Birol. 2013. *Smallholder Demand for Maize Hybrids and Selective Seed Subsidies in Zambia*. HarvestPlus Working Paper No. 9. Washington, D.C.: HarvestPlus/IFPRI. Can be accessed at http://www.harvestplus.org/

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 IFPRI.

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

Theil, H. 1967. *Economics and Information Theory*. Chicago: Rand McNally and Company.

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