

Smallholders' Cost Efficiency in Mozambique: Implications for Improved Maize Seed Adoption

Helder Zavale^{*}, Edward Mabaya^{*}, and Ralph Christy^{*}

^{*} Cornell University

Applied Economics and Management

Ithaca, NY 14853

Contributed Paper Prepared for Presentation at the International Association of Agricultural Economists Conference, Gold Coast, Australia, August 12-18, 2006

Corresponding author: Helder Zavale, Department of Applied Economics and Management, Cornell University, Ithaca, NY 14853. Phone: (607) 319 4088. Email: hz49@cornell.edu.

Copyright 2006 by Helder Zavale, Edward Mabaya, and Ralph Christy. All rights reserved.

Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copy right notice appears on all such copies.

Smallholders' Cost Efficiency in Mozambique: Implications for Improved Maize Seed Adoption

Abstract

Maize is an important staple in Mozambique. It is also a dominant crop produced by smallholder farmers. However, the actual maize yields, currently estimated at 1.4 tons/ha, fall short of potential yields of 5-6.5 tons/ha. With population growth rate increasingly exceeding agricultural (and maize) productivity growth rate, the government of Mozambique faces a serious problem of food insecurity and poverty alleviation. This study examines cost inefficiency among smallholder maize farmers in Mozambique, and the impact of improved maize seed adoption on cost efficiency. A Translog functional form is used to estimate the frontier cost function. A cost-inefficiency function is used to examine the factors that determine cost inefficiency among farmers. Econometric techniques to control for self-selection bias resulting from endogeneity of the adoption variable are used.

JEL Codes: Q12, Q16, D13, O33

Keywords: stochastic frontier, technology adoption, selection bias, Mozambique.

INTRODUCTION

Agriculture is a major sector within Mozambique's economy. However, despite the enormous potential of Mozambique's natural resource available for a higher growth rate of the sector, its performance is relatively low. At the core of the lackluster economic performance is the need to improve crop yields. For example, actual maize yields (generally intercropped) average about 1.4 tons/ha, compared to the potential yield (5-6.5 tons/ha).

Since the 1960s, the maize production in Mozambique has increased rapidly due mainly to expansion in cultivated area. During the same period, yields have stagnated.

With the population growth increasingly outstripping the rate of growth in agricultural output, Mozambique has to improve its agricultural productivity to alleviate poverty.

Agricultural productivity can be decomposed into technical change and efficiency change. It is largely recognized that agricultural output growth is not only influenced by technology enhancements but also by the efficiency with which available technologies are utilized.

The low maize yields in Mozambique suggest that scope exists to increase maize production from the existing technology if resources are efficiently allocated. Hence, the objective of this paper is to estimate the determinants of the cost efficiency of the smallholders using improved and traditional maize seed. A stochastic frontier model is estimated to determine the cost efficiency of smallholders, and a modified version of Heckman's two-step method proposed by Nawata and Ii (2004) is used to correct for the potential bias in the parameter estimates due to self-selection among farmers. The paper is organized as follows. The next section describes the data and methods employed in this study. This is followed by the presentation of estimation results. The final section focuses on the policy implications emanating the research findings.

DATA

The data used in this study were obtained from a national agricultural survey – widely known as *TIA (Trabalho de Inquerito Agrícola)* – conducted by Mozambique's Ministry of Agriculture and Rural Development (MADER) in the agricultural year 2001-2002. The survey collects a wide range of information on the socio-economic and demographic characteristics of households, including income, expenditures, production, capital stock, land use, and demographic characteristics. A total of 4,908 small and medium holdings were

surveyed. Given that this study focuses on maize-growing farm households, the sample entirely used in this study consists of 3,603 small and medium maize-growing farm households. It is worth pointing out that a separate census of all large holdings was also conducted. Large holdings numbered about 400. Table 1 summarizes the sample statistics of the explanatory variables employed in this study.

METHODS

Considerable literature has been devoted to the estimation of efficiency since the pioneering work of Farrell (1957). Farrell showed how to define cost efficiency and decompose cost efficiency into its technical and allocative components. The stochastic frontier approach, based on a specific functional form and introduced by Aigner, Lovell and Schmidt (1977) and Meeusen and Van den Broeck (1977), is motivated by the idea that deviations from the frontier may not be entirely attributed to inefficiency, because random shocks outside the control of farmers can also affect output.

This study uses a cost-efficiency approach and combines the concepts of technical and allocative efficiency in the cost relationship. The most commonly used functional forms for cost functions are Cobb-Douglas and Translog. This study estimates both the Cobb-Douglas and the Translog stochastic frontier cost functions, and uses the likelihood ratio (LR) test to choose which between them fit the data well. The results from the LR test came out in strong support of the Translog model (at 1% significance level); hence, the Translog model was chosen. Consider the Translog stochastic cost function (equation 1) based on the composed error model (Aigner et al., 1977).

$$\ln C = \beta_0 + \beta_A \ln A + \sum_{i=1}^2 \beta_i \ln P_i + \frac{1}{2} \sum_{i=1}^2 \sum_{j=1}^2 \gamma_{ij} \ln P_i \ln P_j + \sum_{i=1}^2 \gamma_{Ai} \ln A \ln P_i + v + \mu \quad (1)$$

$$v \sim \text{iidN}(0, \sigma_v^2) \text{ and } \mu \sim \text{iidN}(0, \sigma_\mu^2)$$

Where, C represents household's observed total variable costs, A denotes the area devoted to maize, P_i is the price of variable inputs (seed and labor in this case), $\varepsilon = v + \mu$ is the disturbance term consisting of two independently distributed elements. The two-sided random disturbance (v_i) reflects the effect of random factors such as weather and the one-sided nonnegative disturbance (μ_i) represents the cost inefficiency component. Since there is no record on family labor costs, the market wage for hired labor is approximated. Since data on maize seed prices is unavailable, the study assumes that price for maize grain is the same as that for maize seed. Economic theory requires that cost functions be homogenous of degree one. To satisfy this homogeneity property, the following restrictions on parameters were imposed prior to estimation: $\sum_i \gamma_{ij} = \sum_j \gamma_{ij} = \sum_i \gamma_{Ai} = 0$ and $\sum_i \beta_i = 1$.

The stochastic cost frontier may be estimated by maximum likelihood (ML). The measurement of the farm level inefficiency, $\exp(-\mu_i)$, requires first the estimation of the nonnegative disturbance, μ_i . That is, it requires decomposing ε_i into its two individual components. Jondrow et al. (1982) showed that, in the case of normal distribution of v_i and half-normal distribution of μ_i , the conditional mean of μ_i given ε_i is given by equation 2.

$$E(\mu_i | \varepsilon_i) = \frac{\sigma_\mu \sigma_v}{\sigma^2} \left[\frac{\phi(\varepsilon_i \lambda / \sigma)}{1 - \Phi(-\varepsilon_i \lambda / \sigma)} + \frac{\varepsilon_i \lambda}{\sigma} \right] \quad (2)$$

Where $\lambda = \frac{\sigma_\mu}{\sigma_v}$, $\sigma^2 = \sigma_\mu^2 + \sigma_v^2$, ϕ is the probability density function, and Φ is the cumulative distribution function. Once point estimates of μ_i are obtained, a measure of the cost inefficiency of each farmer can be provided by equation 3 below.

$$CE_i = E(\exp\{-\mu_i\} | \varepsilon_i) = \left[\frac{1 - \Phi(\sigma_* - \mu_{*i} / \sigma_*)}{1 - \Phi(-\mu_{*i} / \sigma_*)} \right] \exp\left\{ -\mu_{*i} + \frac{1}{2} \sigma_*^2 \right\} \quad (3)$$

A farmer may not reach the cost frontier due to socioeconomic, demographic, and environmental factors. In order to examine the effect of the potential determinants (h_{ji}) of cost inefficiency, equation 4 presented below was estimated.

$$CE_i = \delta_0 + \sum_{j=1}^n \delta_j h_{ji} + \alpha_i Y_i + \tau_i \quad (4)$$

Where Y is the adoption variable taking the value one if the farmer adopted improved maize seed and 0 otherwise. The farmers' decision to adopt improved maize seed is dependent on both the farm and the farmer's characteristics. Therefore, the adoption decision of a farmer is based on each farmer's self-selection instead of random assignment. The farmers that make the adoption decision may possess certain characteristics that are unobservable to the researcher but known to the farmers. These unobserved effects may cause a systematic correlation between the adoption variable, Y_i , and the error term, which makes the adoption variable to be endogenous. OLS estimation of the cost efficiency equation ignoring this endogeneity results in biased and inconsistent parameter estimates. Hence, sample selection bias model (Heckman, 1978; 1979) was used to control for this endogeneity. We use an ML Probit adoption function (equation 5) to correct the error term for potential self-selection bias.

$$P(Y_i = 1 | X) = \Phi(X'\beta) = \int_{-\infty}^{X'\beta} \frac{1}{\sqrt{2\pi}} \exp(-z^2/2) dz \quad (5)$$

Where, P is the probability that the i th household used improved seed; X is the $K \times 1$ vector of the explanatory variables; z is the standard normal variable, i.e., $Z \sim N(0, \sigma^2)$; and β is the $K \times 1$ vector of the coefficients to be estimated. To correct for self-selection bias, the cost-inefficiency function (equation 6) was estimated.

$$CE_i = \delta_0 + \sum_{j=1}^n \delta_j z_{ji} + \alpha_i Y_i + \rho \sigma_\tau \sigma_\mu \lambda_i + \tau_i \quad (6)$$

Where the terms ρ , σ_τ , and σ_μ represent the covariance of the adoption equation and the cost equation, respectively. It is assumed that τ and μ have a bivariate normal distribution with zero means and correlation ρ . The covariances can be broken down into the standard deviations, σ_τ and σ_μ , and the correlation ρ . However, given the structure of the model and the nature of the derived data, σ_μ can not be estimated so it is normalized to 1.0. The term λ_i , given by equation 7, is the Inverse Mills Ratio.

$$\lambda_i = \frac{\phi(\gamma'Z_i)}{\Phi(\gamma'Z_i)} \quad (7)$$

The cost-inefficiency function and the Probit model can be estimated by the Heckman's two-step estimator. Although this estimator is consistent, Nawata and Ii (2004) pointed out that it is not asymptotically efficient. Thus, the Maximum Likelihood (ML) estimator is employed to jointly estimate the cost-inefficiency function and Probit model. The above two-stage method, consisting of ML estimation of a stochastic cost frontier followed by the regression of the predicted cost inefficiency on the determinants of cost inefficiency, has been criticized. While the merits of the two-stage estimation procedure have been established, Coelli (1996) shows that its assumption that the inefficiency effects are independent and identically distributed may lead to inconsistent parameter estimates. An alternative to the two-step procedure is proposed by Battese and Coelli (1995), which combines the two-stages into a single step. However, Liu and Zhuang (2000) argue that both approaches have a common drawback. In particular, Liu and Zhuang point that unless the efficiency variables are independent of the input variables, the production function estimates will be biased and inconsistent.

In this study, the two-stage approach was used. First, the stochastic cost frontier was estimated. Second, the cost-inefficiency function and the Probit model are jointly estimated using ML. Given that the cost inefficiency is censored between 0 and 1, OLS procedure may

result in biased estimates usually toward zero. To correct for this, we estimate the cost inefficiency function using the Tobit model developed by Tobin 1958 (Greene, 2003).

RESULTS

LIMDEP 8.0 software was used to derive estimates for the ML function of the frontier cost and cost-inefficiency functions. Estimates of both λ and σ are statistically different from zero, suggesting that one-side error component dominates the random error term in the determination of $\varepsilon = \mu + v$. Results from this estimation are presented in Table 2. Thus, the deviation of observed variable cost from the frontier cost is due to both technical and allocative inefficiency. This deviation can be avoided without any lost in output.

As expected, the estimates suggest that the relationship between the total variable cost and input prices (seed and labor) is positive and significant. Also, the total variable cost of producing maize statistically increase in all the explanatory variable included in the model with the exception of the interactions between seed price and labor price, and labor price and cropped land. The interaction between labor price and cropped land is not statistically significant.

Table 3 shows results from Probit adoption function and corrected cost-inefficiency function estimations. Fifteen of the twenty five parameter estimates of the Probit model were statistically significant. Household size; age; education; off-farm employment; location (southern, central, and northern agro-ecological zone); access to extension service, credit, seed stores, and electricity; use of pesticide, fertilizer, and irrigation; and farming of traditional cash crops (cotton and tobacco) are the determining factors influencing the probability of adopting improved maize seed. For a detailed discussion of the factors influencing the likelihood of adopting improved maize seed in Mozambique, see Zavale (2005). This study focuses on the determinants of cost inefficiency of producing maize.

After correcting for self selection bias, the results presented in Table 3 show that twelve out of twenty explanatory variables are statistically related to cost inefficiency. Household size, gender, age of household head, years of schooling, distance, maize cropped area, fragmentation of land, use of pesticide, location of household in terms of macro agro ecological zone, access to electricity, and access to credit have a significant impact on cost inefficiency of the farm households surveyed.

The findings suggest that the larger the household size, the more cost efficient the household is. On average, a unit increase in household size drops off cost inefficiency by nearly 2 percent. A possible reason for this result might be that a larger household size guarantees availability of family labor for farm operations to be accomplished in time. Also, a large household size ensures availability of a broad variety of family workforce (children, adults, and elderly), which suggest that household heads can rationally assign farm operations to the right person. This finding is consistent with a previous study by Parikh, Ali, and Shah (1995).

Education increases the ability to perceive, interpret, and respond to new events, enhancing farmers' managerial skills including efficient use of agricultural inputs. The negative and highly significant impact of education on cost inefficiency indicates that farmers with higher education levels are more cost efficient, supporting Schultz hypothesis. This result is similar to the findings of Binam et al (2004).

Further, the variable "distance to county seat" was found to be negatively associated with cost inefficiency. Surprisingly, the further the county seat is away from farm location, the less cost inefficient the farm household is. This result is inconsistent with the findings of Binam et al (2004) that found technical inefficiency increases with the distance of the plot from the main access road, underscoring the importance of better infrastructure in agricultural development.

The link between efficiency and farm size measured as cropped area has been widely investigated using stochastic frontier methodology. The findings of this study do not support the notion of “efficiency economy of scale” that states that larger farms have efficiency advantage over smaller ones. The relationship between cost inefficiency and maize cropped area is positive and statistically significant, suggesting that smaller maize-growing farms are more cost efficient than their counterparts. The results on land fragmentation suggest that it has a negative and significant effect on cost inefficiency. This result does not support the prior expectation that a fragmented farm will cost more in terms of time wasted in moving from one plot to another.

With regard to location of the farm household, households located in the northern and central macro agro-ecological zones were found to be less cost efficient than the ones in the southern, suggesting that location has an impact on farm efficiency. The location variable can be understood as an interaction amongst agro-ecological conditions, infrastructure, and agricultural policies. The differences in cost efficiency due to location may be attributed to distortions introduced by maize policies that subsidize maize production in the southern and tax it in the northern and central. Also, the southern macro agro-ecological zone is characterized by better infrastructure compared to the northern and central.

Access to electricity was found to enhance cost efficiency of the maize-growing farm households. The positive effect of credit availability on cost efficiency is not surprising. Similar results have been reported by Ali, Parikh, and Shah (1996); and Binam et al (2004). Credit availability shifts the cash constraints outward, enabling the farmers to timely purchase agricultural inputs that they can not provide from their own resources.

Figure 1 illustrates the wide variation in levels of cost inefficiency across maize-growing farm households. The maximum and minimum cost inefficiency was 0.896 and 0.127 respectively. Table 4 summarizes the cost inefficiency index. The average cost

inefficiency was 0.70 percent, suggesting that on average 70 percent of the cost observed in the production of maize is due to inefficiency that can be avoided without any loss in total output from a given mix of production inputs. Hence, in the short run, cost efficiency can be enhanced by 70 percent by adopting technology and management practices used by the best maize-growing farm households.

SUMMARY AND CONCLUSION

The results indicate that one-sided error component dominates the random error term in the determination of $\varepsilon = \mu + v$, suggesting that the conventional cost function is not an adequate representation of the data. The findings suggest that with the current technology, in the short run, scope exists for fostering cost efficiency by 70 percent without any loss in total output from a given mix of production inputs. The results show that larger household size, male-headed households, older household head, better education, use of pesticides, and access to credit can bridge the gap between the efficient and inefficient maize-growing farm households. Furthermore, Geographic location is associated with lesser cost efficient maize-growing farm. Surprisingly, the further away from the county seat, the more land fragmented, and bigger maize cropped area, the less cost efficient the farm household is. The cost efficiency of maize-growing farm households and adoption rate of improved maize seed could considerably be improved by: i) improving rural infrastructures, ii) providing better access to education, iii) providing better access to credit, and iv) providing better extension services.

References

- Aigner, Dennis; C. A. Knox Lovell; and Peter Schmidt. 1977. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics* 6 (1):21-37.
- Ali, Farman; Ashok Parikh; and Mir Kalan Shah. 1996. Measurement of economic efficiency using the behavioral and stochastic cost frontier approach. *Journal of Policy Modeling* 18 (3):271-87.
- Battese, George E. and Tim J. Coelli. 1995. A Model for Technical Efficiency Effects in a Stochastic Frontier Production Function for Panel Data. *Empirical Economics* 20 (2):325-32.
- Binam, Joachim Nyemeck; Jean Tonye; Njankoua wandji; et al. 2004. Factors affecting the technical efficiency among smallholder farmers in the slash and burn agriculture zone of Cameroon. *Food Policy* 29 (5):531-45.
- Farrell, M. J. 1957. The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society Series A* 120 (3):253-90.
- Greene, William H. 2003. *Econometrics Analysis*. Fifth ed. Upper Saddle River, New Jersey: Prentice Hall.
- Heckman, James J. 1978. Dummy endogenous variables in a simultaneous equations system. *Econometrica* 46 (6):931-60.
- . 1979. Sample selection bias as a specification error. *Econometrica* 47 (1):153-62.
- Jondrow, James; C. A. Knox Lovell; Ivan S. Materov; et al. 1982. On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics* 19 (2-3):233-38.
- Kumbhakar, Subal C. and C. A. Knox Lovell. 2000. *Stochastic Frontier Analysis*. Cambridge: Cambridge University Press.

- Liu, Zinan and Juzhong Zhuang. 2000. Determinants of technical efficiency in post-collective chinese agriculture: Evidence from farm-level data. *Journal of Comparative Economics* 28 (3):545-64.
- Meeusen, Wim and Julien Van Den Broeck. 1977. Efficiency Estimation From Cobb-Douglas Production Functions With Composed Error. *International Economic Review* 18 (2):435-44.
- Nawata, Kazumitsu and Masako Ii. 2004. Estimation of the labor participation and wage equation model of Japanese married women by the simultaneous maximum likelihood method. *Journal of the Japanese and International Economies* 18 (3):301-15.
- Parikh, Ashok; Farman Ali; and Mir Kalan Shah. 1995. Measurement of economic efficiency in Pakistani agriculture. *American Journal of Agricultural Economics* 77 (3):675-85.
- Zavale, Helder. 2005. Analysis of the Mozambique's Maize Seed Industry: Factors Influencing the Adoption Rates of Improved Seed and Determinants of Smallholders' Cost Efficiency. Masters thesis, Department of Applied Economics and Management, Cornell University, Ithaca.

Table 1 Descriptive statistics of the explanatory variables of the adoption model

Variable	Definition	Mean	Standard Deviation
COST	Variable cost (US \$)	590.03	678.34
PRLABOR	Wage rate of labor (US \$ per hectare)	0.71	0.38
PRISEED	Price of maize seed (US \$/Kg)	0.08	0.04
MAIZE	Maize production (Kg)	609.07	1,627.3
AREA	Cultivated area under maize (hectares)	0.94	1.51
HHSIZE	Household size	5.60	3.33
SEX	Gender of the household head (1 = male; otherwise = 0)	0.761	
AGE	Age of the household head (years)	43.88	14.89
EDUC	Highest formal schooling completed by household head (years)	2.80	4.02
JOB	Household head had off-farm employment = 1; otherwise = 0)	0.326	
DISTANCE	Distance to seat county (Km)	27.00	16.61
COTTON	Farm household grew cotton = 1; otherwise = 0	0.067	
TOBACCO	Farm household grew tobacco = 1; otherwise = 0	0.047	
FRAGMEN	Number of plots farming by household	2.55	1.39
EXTENS	Household had contact with extension service = 1; otherwise = 0	0.155	
FERTIL	Household used fertilizer = 1; otherwise = 0	0.053	
PESTIC	Household used pesticide = 1; otherwise = 0	0.071	
IRRIG	Household used irrigation = 1; otherwise = 0	0.155	
NORTH	Household located in northern macro agro-ecologic zone = 1; otherwise = 0	0.442	
CENTRAL	Household located in central macro agro-ecologic zone = 1; otherwise = 0	0.305	
ELECTRIC	Household had access to electricity = 1; otherwise = 0	0.080	
CREDIT	Household had access to credit = 1; otherwise = 0	0.117	
MARKET	Household had access to market = 1; otherwise = 0	0.269	
ROAD	Household had access to paved road = 1; otherwise = 0	0.192	

Table 2 Maximum likelihood estimates of the frontier Translog cost function

Variable	Coefficient	Standard Error	
Constant	6.605693	0.0613	***
Land	0.311460	0.0270	***
Land x land	0.046331	0.0038	***
Seed price	0.235211	0.0564	***
Labor price	0.764789	0.0564	***
Seed price x seed price	0.037591	0.0134	***
Seed price x labor price	-0.091540	0.0225	***
Seed price x land	0.015171	0.0113	
Labor price x labor price	0.053948	0.0114	***
Labor price x land	-0.015171	0.0113	
Variance			
λ	1.322580	0.0627	***
σ	0.591693	0.0137	***
Log likelihood	-2,267.95		
observations	3,603		

*** Statistically significant at the 1% level.

Table 3 Estimates of the determinants of cost inefficiency corrected for self-selectivity

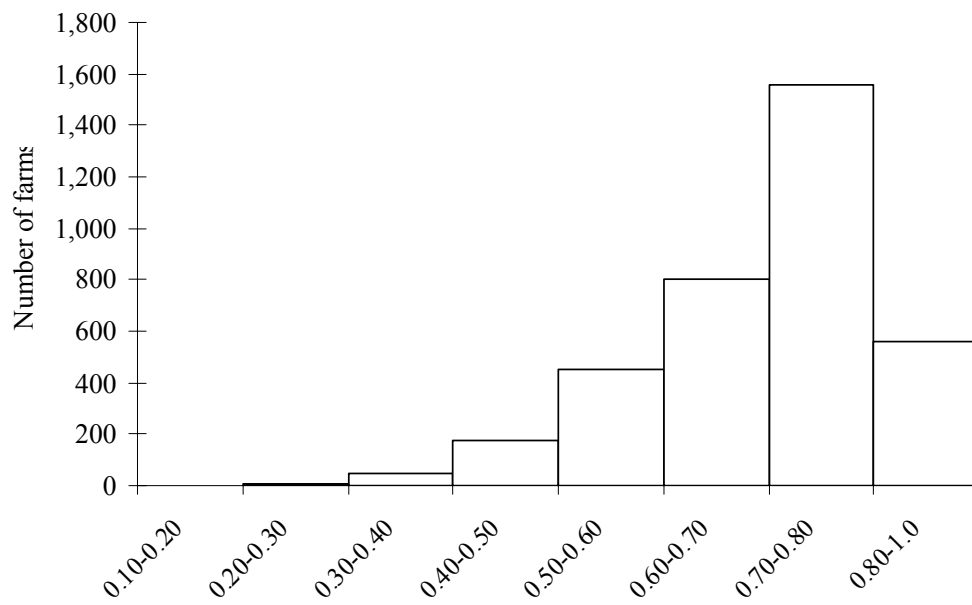
Probit function			Corrected cost-inefficiency function				
Variable	Coefficient		Variable	Coefficient			
Constant	0.231377	(0.2244)	Constant	0.856696	(0.0072)	***	
Distance to seat county	-0.001402	(0.0015)	Distance to seat county	-0.000324	(0.0001)	***	
Household size	0.019516	(0.0073)	***	Household size	-0.022046	(0.0004)	***
Gender	0.000495	(0.0573)		Gender	-0.030531	(0.0033)	***
Age of the household head	-0.014771	(0.0087)	*	Age of the household head	-0.000792	(0.0001)	***
Age of the household head squared	0.000057	(0.0001)		Years of schooling	-0.000604	(0.0003)	*
Years of schooling	0.011257	(0.0058)	**	Off-farm employment	0.003954	(0.0030)	
Off-farm employment	0.162429	(0.0493)	***	North	0.033153	(0.0036)	***
North	-0.678464	(0.0779)	***	Central	0.044983	(0.0040)	***
Central	-0.454732	(0.0698)	***	Extension service	0.003962	(0.0039)	
Extension service	-0.128939	(0.0677)	**	Use of fertilizer	-0.005208	(0.0074)	
Association membership	-0.030954	(0.1008)		Use of pesticide	-0.015640	(0.0066)	***
Access to price information	-0.025184	(0.0530)		Use of irrigation	-0.004281	(0.0039)	
Use of fertilizer	0.243128	(0.1168)	**	Electricity access	-0.011977	(0.0051)	***
Use of pesticide	0.188518	(0.1145)	*	Credit access	-0.012662	(0.0040)	***
Use of irrigation	0.139375	(0.0654)	**	Market access	-0.004471	(0.0035)	
Use of animal traction	0.014907	(0.0632)		Paved road access	-0.005123	(0.0036)	
Electricity access	0.343897	(0.0930)	***	Cotton farming	0.008785	(0.0070)	
Credit access	-0.266283	(0.0782)	***	Tobacco farming	0.004564	(0.0077)	
Market access	-0.035982	(0.0589)		Fragmentation of land	-0.003536	(0.0010)	***
Access to seed shop	0.102922	(0.0584)	*	Maize cropped area	0.018506	(0.0004)	***
Paved road access	-0.001531	(0.0605)		Sigma	0.080725	(0.0009)	***
Cotton farming	-0.211723	(0.1244)	*	Rho	-0.018131	(0.0226)	
Tobacco farming	-0.288330	(0.1234)	***				
Drought last 2 years	0.140419	(0.0931)					
Flood last 2 years	-0.092701	(0.0773)					
Log likelihood	1,869.79						
observations	3,603						

Standard error in parentheses

* Statistically significant at the 10% level; ** Statistically significant at the 5% level; and *** Statistically significant at the 1% level

Table 4 Summary statistics of the cost inefficiency indexes

	Cost inefficiency index
Mean	0.6977
Standard deviation	0.1140
Minimum	0.1268
Maximum	0.8962
Observations	3,603

**Figure 1 Frequency distribution of cost inefficiency index**