Agricultural and Food Economics

RESEARCH

Open Access

Correcting for sample selection in stochastic frontier analysis: insights from rice farmers in Northern Ghana

Shaibu Baanni Azumah*, Samuel Arkoh Donkoh and Joseph Agebase Awuni

* Correspondence: raszumah1983@gmail.com; sazumah@uds.edu.gh
Department of Agricultural and Resource Economics, University for Development Studies, P. O. Box TL 1350, Tamale, Ghana

Abstract

This study employs stochastic frontier analysis (SFA) correcting for sample selection bias, to determine technical efficiency (TE) and technology gap using cross-sectional data collected from 543 rice farmers in Northern Ghana. The results showed that corrected sample selection TE estimates were marginally higher. Without the appropriate corrections, inefficiency is overestimated, while the gap in performance between irrigation farmers and their rainfed counterparts is underestimated. We recommend that authorities in Ghana should work with development partners, especially in the implementation of small village-dam projects, and also to expand the existing irrigation schemes. Bunds should also be constructed around rice production valleys across northern Ghana so that farmers could expand their farm sizes to increase production. It is important also that the government's input subsidy programme be structured to cater for experienced and younger farmers who consider agriculture as a business.

Keywords: Rice production, Sample selection, Stochastic frontier, Technical efficiency, Northern Ghana

Introduction

Rice is an important cereal crop, second to maize in terms of consumption in Ghana. The importation of rice continues to surge ahead of production due to increases in domestic consumption. For example, annual per capita consumption of rice in Ghana grew from 17.5 kg in 2001 to 24 kg in 2011 (Ragasa et al. 2014). This has seen a further increase to about 32 kg for 2015 (MoFA 2016). Also, the demand for rice is projected to grow at an annual rate of 11.8%, exceeding that of maize (2.6%) in the medium term (Millennium Development Authority (MiDA) 2010). As only 5% of global production is traded, local production would also protect consumers from price shocks in the world rice market (World Bank 2013).

While substantial investments in national rice production have been made, local production is still not able to keep up with the growing demand for rice. Ghana imported 508,587 MT of rice in 2013 alone, translating into USD$639 million to compensate for domestic shortfall (Ministry of Food and Agriculture (MoFA) 2013). This has increased further by 22% to 620,811 MT in 2016 (MoFA2016). Increasing rice yields through sustainable and efficient production systems is necessary and has therefore become a priority for stakeholders in the rice value chain. Adoption of irrigation technology is one way of improving farmers' efficiency in the production of rice, especially in northern Ghana where the impact of climate change has become more evident. For instance, Azumah et al. (2017) found that rice farmers in northern Ghana who adopted irrigation were not only efficient but had higher yields compared with their non-adopting counterparts.

This present study investigates the output effect of irrigation farming in northern Ghana. By way of methodology, it combines a stochastic production frontier framework correcting for sample selection bias developed by Greene (2010). To the best of the researchers' knowledge and also based on available literature, this study is the first of its kind for efficiency studies in the rice sector of Ghana. Many of the studies conducted in the study area have employed the traditional stochastic frontier approach which did not control for sample selection bias. The rest of the paper is organised as follows: methodology, results and discussion, and conclusions and recommendations.

Methodology

The study location

The study was conducted using data from rice farmers in the Northern and Upper East Regions of Ghana. The area is characterised by poor soil conditions and two climatic seasons (MoFA 2016). The rainy season begins lightly in April and peaks in August/September but gradually declines by October/November. The dry season occurs between November and April each year and is characterised by dry harmattan winds which engulf the whole region. The vegetation of the region is generally the Guinea savannah with its characteristic grass and tree species. The biodiversity in tree vegetation used to be high, but now it is decreasing due to over-exploitation.

The major economic activity of the people is agriculture (combination of food crops and animal husbandry) with most parts of the region being rural (Ghana Statistical Service (GSS) 2014). The agricultural sector employs the largest share of the economically active population in the area with the bulk of production done by smallholder farmers for subsistence purpose (MoFA 2016). The poverty levels in the two regions are about 50% (Ghana Statistical Service (GSS) 2014). Agricultural productivity continues to be low due to a variety of factors including the low uptake of improved agricultural technologies (Ragasa et al. 2013).

Sampling and data collection

This study used cross-sectional data from 543 rice farmers in the Upper East and Northern Region of Ghana in the 2016/2017 cropping season. Multistage sampling method was used to select the respondents from rice-growing communities in the two regions. Primary data was collected from two strata of rice farmers (rainfed and irrigation farmers) in 62 communities located in 10 districts of the two regions.

Analytical framework—stochastic frontier model with sample selection

The stochastic production frontier (SPF) methods have been used extensively in many industries, including agriculture, to model input–output relationships and to measure the technical efficiency of individual producers. These methods have also been used to compare the performance of farmers under different technological regimes. For example, the method has been used to examine the impact of technology adoption on output and TE of rice farmers (Villano et al. 2015).

The limitation of most studies that have used SPFs to compare the TE of adopters versus non-adopters is the failure to account for selectivity bias in a manner that is compatible with the nonlinear nature of the stochastic frontier model. For example, following Heckman's (1979) methodology to account for selection bias, several attempts have been made to address sample selection in a stochastic frontier framework. Sipiläinen and Oude Lansink (2005) added an inverse Mill's ratio (IMR) to the deterministic part of the frontier function to examine possible sample selection bias in the analysis of organic and conventional farms. A similar approach was implemented by Solis et al. (2007) when analysing farmers with different levels of adoption of soil conservation practices in Central America. However, this procedure has proven unsuitable for nonlinear models such as the SPF (Greene 2010).

In recent years, alternative strategies have been proposed to deal with this problem including the one by Kumbhakar et al. (2009) who developed a model where the selection mechanism is assumed to operate through the one-sided error in the frontier, and then used their model to evaluate the performance of organic versus conventional dairy farming in Finland.

Lai et al. (2009) studied wage determination employing a copula function and assumed that selection is correlated with the composed error in the frontier. These two models require computationally demanding log-likelihood functions. This study adopts the framework developed by Greene (2010) who extended Heckman's approach to consider sample selection in a stochastic frontier framework assuming that the unobserved characteristics in the selection equation are correlated with the noise in the stochastic frontier. The model introduced by Greene can be expressed succinctly with the following set of equations[1]: (Eqs. 1 and 2 represent the sample selection and stochastic frontier models, respectively.)

$$d_i = 1\left[\propto^1 z_i + w_i > 0\right], w_i \sim N(0,1) \tag{1}$$

$$y_i = \beta^1 x_i + \varepsilon_i \tag{2}$$

(y_i, x_i) are observed only when $d_i = 1$.

The error structure is specified as follows:

$$\varepsilon_i = v_i - u_i \tag{3}$$

$$u_i = |\sigma_u U_i| = \sigma_u |U_i| \text{where } U_i \sim N(0,1) \tag{4}$$

$$v_i = \sigma_v V_i \text{ where } V_i \sim N(0,1) \tag{5}$$

$$(w_i v_i) \sim N_2\left[(0,0), \left(1, \rho\sigma_v, \sigma_v^2\right)\right]$$

where:

d is a binary variable equal to one for adopters (irrigated farmers), and zero for non-adopters (rainfed farmers);

z is a vector of explanatory variables included in the (binary) sample selection model; and

w_i is the unobservable error term;
y is output for the rice farmers;
x is a vector of inputs in the production frontier; and
ε is the composed error term.

The coefficients $\propto$ and β are parameters estimated, while the elements in the error structure correspond to those typically included in the stochastic frontier formulation. In this model, sample selection arises if the noise in the stochastic frontier, v_i, is correlated with unobserved characteristics in the sample selection equation, w_i (Greene 2010). A statistically significant ρ is evidence that selectivity bias in unobservables is present.

Results and discussion

Definition and descriptive statistics of variables

Table 1 provides a summary of definitions for the variables used in this study. Table 2 presents the summary statistics for the matched sample.[2] The matched sample contains 538 observations, made up of 223 irrigation farmers and 315 for rainfed. t test was performed to compare the mean values of the variables for the irrigated farms to that of the rainfed farms. The pooled results indicate the average age of a rice farmer to be 38.43 years. There were more male respondents (83%) compared to 17% female respondents. This finding does not however suggest that females were least involved in rice production. Focus group discussions conducted with the farmers revealed that the activities in rice production appeared led by the males because they owned the

Table 1 Definition of variables

Variable	Definition/measurement	Sign
Output	Natural log of rice output (measured in 100 kg bags)	+
Age	The total number of years from birth	+
Sex	Dummy: 1 for male, 0 if otherwise	+
HH head	Dummy: 1 for household head, 0 if otherwise	+
Education	Number of years spent in formal schooling	+
Commercial	Dummy: 1 if farmer produces for commercial purpose, 0 if otherwise	+
Experience	The total number of years a farmer has been cultivating rice.	+
Region	Dummy: 1 for a farmer in Northern Region, 0 for a farmer in Upper East Region	+/−
FBO	Dummy: 1 for if the farmer belongs to a farmer group, 0 if otherwise	+
Research/extension	Dummy: 1 for access to research/extension service, 0 if otherwise	+
Credit	Dummy: 1 for access to credit in the last growing season, 0 if otherwise.	+
Training	Dummy: 1 if farmer had access to trainings last season, 0 if otherwise.	−/+
CC perception	Dummy: 1 for farmers who perceived that rainfall was reducing with rising temperatures, 0 if otherwise	+/−
HH size	Total number of people in housing unit that feed from the same source	+/−
Farm size	Natural log of farm size (measured in the acres of land under rice production)	+
Herbicides	Natural log of quantity of herbicides (measured in litres) used	+
Fertilizer	Natural log of total quantity of fertilizer (measured in kg)	+
Seed	Natural log of quantity of improved seed (measured in kg)	+
Labour	Natural log of total number of persons available that worked on the farmers field during the farming season	−

Table 2 Descriptive statistics—matched sample

Variable	Pooled		Irrigated farms		Rainfed farms		Test of means[a]
	Mean	SD	Mean	SD	Mean	SD	
Age	38.43	10.66	40.29	11.92	37.12	9.47	3.425[d]
Sex	0.83	0.38	0.87	0.34	0.81	0.4	1.804[b]
HH head	0.57	0.5	0.60	0.49	0.56	0.5	0.87
Education	4.04	5.13	3.86	5.52	4.17	4.84	− 0.692
Commercial	0.65	0.48	0.64	0.48	0.65	0.48	− 0.304
Experience	11.73	7.68	12.48	8.16	11.19	7.28	1.931[b]
Region	0.68	0.47	0.84	0.37	0.57	0.50	6.821[d]
FBO	0.64	0.48	0.58	0.50	0.68	0.47	− 2.328[d]
Extension	0.54	0.50	0.43	0.50	0.63	0.48	− 4.735[c]
Credit	0.12	0.32	0.07	0.26	0.15	0.36	− 2.767[d]
Training	0.72	0.45	0.91	0.29	0.58	0.49	8.866[d]
CC perception	0.67	0.47	0.86	0.35	0.53	0.50	8.311[d]
HH size	9.33	6.24	9.28	6.65	9.37	5.94	− 0.163
Farm size	2.41	3.64	1.29	1.00	3.20	4.51	− 6.21[d]
Output	31.07	47.19	21.98	16.02	37.5	59.39	− 3.804[d]
Herbicide	3.26	8.66	2.28	1.33	3.96	11.21	− 2.221[d]
Fertilizer	6.39	37.9	3.47	6.26	8.45	49.17	− 1.505
Seed	54.88	108.38	15.93	26.16	82.46	133.28	− 7.353[d]
Labour	16.87	16.37	25.16	17.86	11	12.21	10.92[d]
Obs	538		223		315		

[a]*t* test is used to determine if the sample means are significantly different between the irrigated and rainfed farms
[b], [c] and [d] represent 10%, 5% and 1% level of significance, respectively
Source: Analysis of field data, 2017

lands on which production was carried out. Mostly, the females provided labour for transplanting, weeding and harvesting. Females are also mainly responsible for value addition such as parboiling and processing of rice for onward sale in local markets.

Also, 60% of the irrigation farmers were found to be household heads, as against 56% for their counterpart rainfed farmers. On average, a farmer had up to only 4.04 years of formal education. Irrigation farmers were also found to have more experience in rice production (12.48 years) compared with their rainfed counterparts (11.19 years). Just about 12% of the farmers had access to production credit in the previous season. Credit access among the rainfed farmers was relatively high (15%), compared with 7% for irrigation farmers. The average output of rice was also reported to be 31.07 bags,[3] which translates into an average yield of 12.89 bags/acre (0.52 MT/Ha). While the achievable yield of rice is projected to about 6.0 MT/Ha, the current national average yield of the commodity is about 2.75 MT/Ha, with the Northern and Upper East Regions recording lower yields than the national average (Ministry of Food and Agriculture (MoFA) 2016). The rest of the variables are as presented in Table 2.

Technical efficiency of rice farmers in Northern Ghana

We first discuss the results of the selection equation (adoption of irrigation) before looking also at the frontiers of the different systems of production. The estimates of the probit model were used to obtain a propensity score (the predicted probability of participation in irrigation) for each farmer after which each irrigation farmer was matched to a rainfed counterpart.

Determinants of irrigation technology adoption

Table 3 presents the results of the probit sample selection model (factors that influence the adoption of irrigation) using the matched sample. The McFadden pseudo *R*-squared was low at 0.155, but with a significant chi-squared test statistic (113.5), indicating a joint significance of the parameters for the irrigation adoption variables. Age, sex, location, membership of farmer-based organisation (FBO), farm input subsidy, training, credit, household size and farmers' perception about climate change significantly influenced the adoption of irrigation.

Education was insignificant in explaining the adoption decision of farmers, contradicting the finding of Villano et al. (2015). Age and sex were found to be negatively related to the adoption of irrigation. Younger farmers had a higher tendency of producing rice under irrigated conditions compared to their older counterparts. Again, female farmers in the study area had a greater probability of participating in irrigation than their counterpart male farmers. There was also a positive and significant relationship between location and adoption of irrigation, indicating that farmers in the Northern Region had higher adoption drive for irrigation than those in the Upper East Region of Ghana. Also, we found a significant but negative relationship between FBO and adoption of irrigation, implying that farmers who belonged to FBOs had a lower probability of adopting irrigation. Perhaps, the farmers who belong to groups might have contractual

Table 3 Parameter estimates of probit selection equation for irrigation using matched sample

Variable	Coef.	Std. Err.
Age	− .01711[c]	0.006
Sex	− 0.48832[c]	0.180
HH head	0.14699	0.154
Education	0.01264	0.012
Commercial	− 0.15443	0.125
Experience	0.01211	0.009
Location (region)	0.30812[a]	0.161
FBO	− 0.36539[c]	0.129
Subsidy	0.25765[a]	0.153
Training	0.91004[c]	0.153
Credit	− 0.70239[c]	0.198
CC perception	0.35782[c]	0.138
HH size	− 0.02741[c]	0.010
McFadden pseudo R^2	0.155	
Log-likelihood function	−308.26	
Chi^2 test statistic	113.5[c]	
Number of Obs.	538	

[a], [b] and [c] represent 10%, 5% and 1% level of significance, respectively
Source: Analysis of field data, 2017

obligations that compelled them to increase production by acquiring more land which is not available under irrigation conditions.

Farmers who had access to subsidised farm inputs participated more in irrigation than those who did not have access. Dorward and Chirwa (2013) found that input subsidies have had a wider impact on economies through increased food crop production, which lead to a reduction in consumer food prices and to the benefit of poor food consumers, and an increase in rural agricultural wages. However, the benefit of agricultural subsidy programmes has varied with the nature of the subsidies and their context in the market, as well as with the weather (Kato and Greeley 2016), justifying the need for irrigation. As expected, farmers who attended trainings had a greater probability of adopting irrigation compared to their untrained counterparts. Access to credit is expected to influence technology adoption decision of farmers positively (Anang et al. 2016). However, our results show otherwise, contrary to our a priori expectation. Our findings in the previous sections suggested that irrigation farmers had smaller land holdings compared to their counterpart rainfed farmers and so may not require any credit to invest in inputs considering that credit acquisition comes with cost associated with interest payments. We found also the perception of farmers about the prevalence of climate change to be significant and positively related to the adoption of irrigation. Also, household size had an inverse relationship with the adoption of irrigation, suggesting that larger households are more averse to adopting irrigation than smaller households.

Production frontiers

We first report the result of hypotheses tests conducted to select the functional form, i.e. the choice between Cobb–Douglas vs. translog functional form (H_0, $\beta_{jk} = 0$). However, given the complexity of our model and the focus on the empirical significance of the framework applied, we concentrated on the choice of an appropriate functional form that is also flexible. The generalised likelihood ratio (LR) test (see Table 4) confirmed that the translog production function is suitable for the production structure in our case. The translog specification presented a smaller AIC (771.7) compared with the 842.7 for the Cobb–Douglas specification, also providing sufficient justification for our choice of the translog production function.

In Table 5, we present the results of the stochastic production frontier model corrected for selectivity bias. In the same table, the results for the conventional frontier without correcting for selectivity bias (with technical inefficiency effects) to allow for comparison are also presented. All variables in the translog models were normalised by their corresponding geometric means so that the first-order coefficients can be interpreted as partial elasticities of output with respect to inputs at mean values (Villano et al. 2015).

The sum of all partial production elasticities, i.e. return to scale (RTS), for the pooled and irrigation farmers in both the conventional and sample selection models are

Table 4 Generalised likelihood ratio test of hypothesis

Null hypothesis	LR statistic (λ)	Critical value[a]	Decision
Production function is Cobb–Douglas	100.96	36.17	Reject H_0. Use Translog PF
AIC	Translog = 771.7	Cobb–Douglas = 842.7	Reject H_0. Use Translog PF

[a]Critical value for the production function is obtained from Kodde and Palm (1986) at 5% two-tail reading
Source: Analysis of field data, 2017

Table 5 Parameter estimates of SPF model

Variable	Convectional SPF						Sample selection SPF					
	Pooled		Irrigation only		Rainfed only		Pooled—sample		Irrigation		Rainfed	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Const.	3.988[c]	0.044	3.944[c]	0.062	4.053[c]	0.063	3.970[c]	0.066	3.942[c]	0.023	4.007[c]	0.114
Farm size	0.920[c]	0.074	1.205[c]	0.011	1.283[c]	0.095	0.923[c]	0.084	1.336[c]	0.034	1.144	0.122
Labour	0.067[a]	0.04	0.044	0.071	0.256[c]	0.093	0.069	0.053	0.023	0.018	0.11	0.102
Seed	− 0.235[c]	0.027	− 0.157[c]	0.045	− 0.235[c]	0.051	− 0.158[c]	0.034	− 0.153[c]	0.016	− 0.107	0.08
Fertilizer	− 0.014	0.047	0.053	0.108	− 0.127[b]	0.057	− 0.018	0.049	0.132[c]	0.028	− 0.071	0.113
Herbicide	0.124[a]	0.064	− 0.266	0.221	− 0.074	0.102	0.062	0.081	− 0.520[c]	0.047	− 0.076	0.138
Farm size^2	0.076[c]	0.026	0.183[c]	0.042	− 0.071[b]	0.035	0.013	0.02	0.224[c]	0.01	− 0.085[b]	0.036
Labour^2	− 0.0198	0.017	− 0.0187	0.017	0.025	0.03	− 0.014	0.02	− 0.0098[a]	0.005	0.0022	0.038
Seed^2	− 0.046[c]	0.006	− 0.014	0.01	− 0.060[c]	0.011	− 0.044[c]	0.007	− 0.012[c]	0.002	− 0.036[b]	0.015
Fertilizer^2	0.004	0.011	0.006	0.014	− 0.025	0.015	− 0.012	0.009	0.051[c]	0.007	− 0.0277	0.035
Herbicide^2	− 0.033	0.03	0.104	0.073	− 0.022	0.034	− 0.054	0.037	0.145[c]	0.022	− 0.04223	0.061
Farm size[a]Labour	0.0756	0.058	− 0.009	0.066	0.207[a]	0.125	0.066	0.067	− 0.052[b]	0.021	0.26375[a]	0.142
Farm size[a]Seed	− 0.036	0.048	− 0.069	0.056	− 0.085	0.064	0.0097	0.049	− 0.071[c]	0.017	− 0.06171	0.089
Farm size[a]Fertilizer	− 0.234[c]	0.072	0.175[a]	0.09	0.009	0.102	0.235[c]	0.072	0.173[c]	0.027	− 0.17683	0.139
Farm size[a]Herbicide	− 0.013	0.088	− 0.379[b]	0.151	0.159	0.11	0.182[c]	0.066	0.473[c]	0.042	0.28069[b]	0.123
Labour[a]Seed	0.049[a]	0.027	0.026	0.034	− 0.145[c]	0.051	− 0.038	0.034	0.0063	0.008	− 0.09272	0.065
Labour[a]Fertilizer	0.037	0.052	− 0.034	0.073	− 0.012	0.068	0.072	0.059	− 0.016	0.021	0.05588	0.105
Labour[a]Herbicide	− 0.065	0.063	− 0.075	0.074	− 0.02	0.125	− 0.087	0.07	− 0.036	0.025	− 0.1116	0.133
Seed[a]Fertilizer	0.091[b]	0.036	− 0.055	0.069	0.214[c]	0.06	0.132[c]	0.04	− 0.093[c]	0.017	0.20680[b]	0.103
Seed[a]Herbicides	0.131[c]	0.046	0.06	0.086	0.178[c]	0.058	0.056	0.054	0.071[c]	0.022	0.08807	0.082
Fertilizer[a]Herbicide	0.042	0.073	− 0.288[b]	0.117	− 0.151	0.093	0.03	0.074	− 0.458[c]	0.035	− 0.01806	0.139
RTS	0.86		0.88		1.10		0.88		0.82		1.00	

Table 5 Parameter estimates of SPF model *(Continued)*

Variable	Convectional SPF						Sample selection SPF					
	Pooled		Irrigation only		Rainfed only		Pooled—sample		Irrigation		Rainfed	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
L. Likelihood	− 362.855		− 35.986		− 235.691		− 671.929		− 187.078		− 408.302	
gamma(γ)	0.947		0.969		0.972							
lambda(λ)							2.056		9.093		1.324	
$\sigma(u)$	4.212[c]	0.525	5.618[c]	1.593	5.866[c]	1.336	0.689[c]	0.047	0.644[c]	0.012	0.627[c]	0.098
σ (v)	0.807[c]	0.001	0.503[c]	0.002	0.907[c]	0.002	0.335[c]	0.031	0.071[c]	0.01	0.474[c]	0.06
Selectivity bias ($\rho_{w,v}$)							0.727[c]	0.085	0.99998[c]	0.001	0.823[c]	0.092
N	538		223		315		538		223		315	

[a], [b] and [c] represent 10%, 5% and 1% level of significance, respectively
Source: Analysis of field data, 2017

consistently less than one, showing decreasing returns to scale. The RTS for rainfed farmers in the conventional SPF was estimated to be more than one, indicating increasing return to scale. For the same group of rainfed farmers in the sample selection model, the RTS was estimated to be exactly one, implying a constant RTS. These results indicate that the rainfed farmers in the study area are able to increase their output with increases in input usage. On the contrary, irrigation farmers are not able to achieve proportionate increase in output with upward adjustments in their input usage. Both the estimates of σ_u and σ_v in the conventional and sample selection models are significantly different from zero at the 1% level, indicating goodness of fit of the model. The coefficients of the selectivity bias variables ($\rho_{w,v}$) were significantly different from zero at the 1% level for all the sample selection frontiers, which confirm that selection bias existed, providing justification for the use of a sample selection framework in the SPF model. In other words, estimation using observations from only a single system of production (either rainfed or irrigation) will provide biassed estimates of the frontier, which will then be carried on to the biassed estimates of efficiency scores as well (Villano et al. 2015).

Results from the SPF controlling for selectivity bias revealed that output of rice increased with farm size, differing from the finding of Donkoh et al. (2013), but reduced with quantity of seed used in the pooled frontier. Four out of five estimated linear coefficients in the selectivity-corrected SPF for irrigation were significant in explaining the output of rice farmers, with all of them being insignificant in explaining output in the frontier of rainfed farmers. As expected, farm size and fertiliser had a positive relationship with output of irrigation rice farmers corroborating with Addison et al. (2016).

Output of irrigation rice farmers was however found to be inversely related to the quantity of seed and herbicide used, diverging from the findings of Anang et al. (2016). Continuous increases in the amount of fertiliser and herbicides were also found to increase output marginally. Most of the interaction variables were only significant in explaining the output of rice farmers in the irrigation frontier. The interactions of farm size and fertiliser, farm size and herbicide, and seed and herbicides were necessary for increased production of rice under irrigation ecology.

In both the conventional and sample selection models, farm size had the highest elasticity value, corroborating with Rahman and Barmon (2015). The elasticity of farm size in the sample selection frontiers was 0.92 and 1.34 for the pooled and irrigation frontiers, respectively, implying that a 100% increase in land allocated for rice production will increase output by 92% and 134% for the pooled and irrigation frontiers, respectively. Ragasa et al. (2013), noted that the increases in the output of rice in the study area have largely been due to expansion in farm sizes and not necessarily due to the use of farm inputs and improved production techniques. This phenomenon should be of serious concern to stakeholders of the agricultural sector in Ghana as the present population statistics do not support the theory of a positive relationship between farm output and farm size.

Efficiency estimates of rice farmers in Northern Ghana

Summaries of technical efficiency (TE) scores of the matched samples are presented in Table 6. The first sets of TE estimates are from the conventional stochastic production frontier (SPF). The second set of TE estimates is from the selectivity-corrected SPFs.

Table 6 Frequency distribution of technical efficiency of rice farmers

Eff. Score	Conventional SPF						Sample selection SPF					
	Pool		Irrigation		Rainfed		Pool		Irrigation		Rainfed	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%
10–20	0	0	0	0	0	0	10	1.9	1	0.4	2	0.6
21–30	0	0	0	0	0	0	15	2.8	7	3.1	8	2.5
31–40	118	21.9	0	0	0	0	35	6.5	20	9	13	4.1
41–50	53	9.9	0	0	127	40.3	68	12.6	22	9.9	32	10.2
51–60	80	14.9	57	25.6	45	14.3	74	13.8	27	12.1	60	19
61–70	105	19.5	30	13.5	36	11.4	128	23.8	26	11.7	70	22.2
71–80	73	13.6	64	28.7	43	13.7	132	24.5	40	17.9	105	33.3
81–90	99	18.4	42	18.8	56	17.8	71	13.2	53	23.8	25	7.9
91–100	10	1.9	30	13.5	8	2.5	5	0.9	27	12.1	0.0	0.0
Total	538	100	223	100	315	100	538	100	223	100	315	100
Min	33.7		59.4		40.6		13.3		18		18.4	
Max	95.2		96.6		85.9		93.3		97.5		89.2	
TE-Mean	60.6		74.4		60		62.2		68		63.4	

Source: Field data, 2017

The results reveal that TE estimates improved upon implementing the sample selection framework.

Using the pooled estimates, the mean TE increased from 60.6 to 62.2%, comparing the conventional and sample selection SPF respectively. The mean technical efficiency of irrigation farmers, corrected for selectivity bias, was estimated to be 68%, implying that 47% [(100–68)/68] of the production is lost due to technical inefficiency alone. This implies that the average farmer producing under irrigation could increase production by about 47% by improving their technical efficiency.

The mean technical efficiency of rainfed farmers, corrected for selectivity bias, was estimated to be 63.4%, implying that 57.7% [(100–63.4)/63.4] of the production was lost due to technical inefficiency alone. This implies that the average farmer producing under rainfed condition could increase production by about 57.7% by improving their technical efficiency.

Overall, the efficiency scores for irrigation farmers were relatively high for both the conventional and the corrected selectivity bias SPFs, implying that the farmers who produced under irrigation were more technically efficient than those who produced under rainfed condition. For example, while 13.5% of the irrigation farmer operated at efficiency level of 91% and above, only 2.5% of the rainfed farmers operated at this efficiency level for the conventional frontiers. In the corrected selectivity bias frontiers, about 12% of irrigation farmers operated at efficiency level of 91% and above as against 0% for their counterpart rainfed farmers.

Previous studies have estimated the technical efficiency of rice farmers in the study area to be high. This obviously could be due to the estimation processes adopted by those authors which did not account for selectivity bias. For example, Donkoh et al. (2013), estimated the technical efficiency of rice farmers in the Tono irrigation scheme to be 81%.

The empirical results show that without the appropriate corrections, inefficiency was overestimated, while the gap in performance between irrigation farmers and their rainfed counterparts was underestimated, corroborating with Villano et al. (2015). However, Mayen et al. (2010) reported the evidence of bias to be the opposite, where the differentials between treated and control units decreased as the correction for bias was implemented.

Determinants of technical efficiency among rice farmers in Northern Ghana

The determinants of efficiency (or inefficiency) indicate the potential sources of efficiency that could be relevant for policy formulation. In Table 7, the translog maximum likelihood estimates of the determinants of technical inefficiency are presented. The translog maximum likelihood frontier estimates are from a two-stage selectivity-corrected pooled sample SPF and inefficiency models. For comparison, we present separate estimates for the group (irrigation and rainfed) as well as that of the pooled corrected selectivity bias data. The coefficients for the technical inefficiency results are interpreted by their signs, such that a positive (negative) coefficient indicates a positive (negative) effect on inefficiency. To be simplistic, we only discuss the determinants focusing on variables that are statistically significant at conventional levels.

In the inefficiency model of the pooled results, sex, location, household size, credit and perception of farmers about climate change were found to be significant at conventional levels and positively related to technical inefficiency (negatively related to technical efficiency). The coefficients of subsidy, experience, commercialisation and household head are also found to be significant at conventional levels but negatively related to technical inefficiency (positively related to technical efficiency). In the 'irrigation' group, technical inefficacy was influenced by age, sex, education, farmers' commercialisation drive, location, membership of FBO and household size.

Table 7 Maximum likelihood estimates of determinants of technical inefficiency

Variable	Irrigation only		Rainfed only		Pooled	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Const.	2.542	0.792	− 5.15[a]	0.981	− 1.552[a]	0.479
Age	− 0.038[b]	0.016	0.057	0.016	0.014	0.009
Sex	− 1.005[a]	0.366	1.916	0.459	0.656[a]	0.237
HH head	− 0.218	0.315	− 0.734	0.299	− 0.54[a]	0.194
Education	0.055[a]	0.021	− 0.018[b]	0.026	− 0.002	0.015
Commercial	− 1.467[a]	0.283	− 0.221	0.24	− 0.579[a]	0.163
Experience	− 0.002	0.021	− 0.078[a]	0.021	− 0.042[a]	0.012
Location	−0.722[b]	0.368	1.886[a]	0.436	0.631[a]	0.222
FBO	− 0.565[c]	0.311	− 0.432	0.287	0.052	0.177
Subsidy	− 0.31	0.375	− 0.869[a]	0.285	− 0.713[a]	0.179
Training	− 0.273	0.376	0.731[a]	0.278	0.112	0.184
Credit	0.553	0.508	0.465	0.332	0.575[b]	0.246
CC perception	− 0.064	0.378	0.83[a]	0.264	0.344[c]	0.187
HH size	0.006[b]	0.022	0.035	0.021	0.046[a]	0.015
Log-likelihood	5.77		− 160.88		− 318.61	
N	223		315		538	

[a], [b] and [c] represent 1%, 5% and 10% level of significance respectively
Source: Field data, 2017

The coefficient of age was negative and significant, implying that younger farmers were more inefficient compared to older farmers. The coefficient of sex was negative and significantly different from zero at 1% level, indicating that female farmers were more inefficient compared to their male counterparts. The positive and significant sign of the coefficient of the education variable indicates that farmers who received more formal education were more inefficient, contrary to our a priori expectation. The coefficient of commercialisation was also positive and significant, implying that irrigation farmers who produced for subsistent purpose were less inefficient compared with their counterparts who had commercialisation drive, against our a priori expectation that commercial farmers were most likely going to commit more resources to production and will invest in improved practices to increase output and their incomes, to compensate for their investments. Location was also found to bear a significant and negative relationship with technical inefficiency. This means that rice farmers in the Upper East Region were found to be more technically inefficient compared to their colleague farmers in the Northern Region of Ghana. FBO was positive and significant, indicating that farmers who belonged to farmer-associations were less efficient as compared to those who did not belong to any farmer group. The significant and positive association between household size and technical inefficiency imply that households with larger membership were more technically inefficient, while smaller size households exhibited better levels of technical efficiency.

Technical inefficiency among the 'rain-fed' group was influenced by education, experience, subsidy, location, training and the perception of the farmers about climate change in the study area. The negative and significant sign of the coefficient of the education indicates that farmers who received more formal education were less inefficient, agreeing with our a priori expectation. This was rather the reverse for the 'irrigation only' group. The reasons for this diverging situation was not sufficiently explored by this present study and so need further investigations. Experience was found to be negative and significantly related to inefficiency, implying that farmers with relatively long years of experience of rice production under rainfed conditions were more efficient.

Location was also found to bear a significant and positive relationship with technical inefficiency, meaning that farmers in the Northern Region who produced under rainfed conditions were found to be more technically inefficient compared to their colleague farmers in the Upper East Region. An important policy variable, subsidy, was also found to be positive and significantly related to technical inefficiency of rice farmers producing under rainfed conditions. The negative sign of subsidy implied that farmers who received and used subsidised farm inputs were more technically efficient than those who did not receive subsidy. Also, rice farmers producing under rainfed conditions who received training, and those who perceived climate change to be present and dominant in the study area were found to be less technically efficient, as the covariates of training and climate change perception have positive and significant relationship with technical inefficiency.

Conclusions and Recommendations

This study employs a sample selection-corrected stochastic production frontier model to determine the TE of rice farmers in Northern Ghana. We conclude that TE estimates improve upon implementing the model, as the mean TE increased from 60.6 to 62.2% for the pooled results. The empirical results show that without the appropriate corrections, inefficiency is overestimated, while the gap in performance between irrigation farmers and their rainfed counterparts is underestimated. In terms of policy, we recommend that the government of Ghana should work with development partners to develop new and existing irrigation schemes and also construct bunds around the rice production valleys in northern Ghana so that rice farmers could expand their farm sizes to increase production. It is important that the input subsidy programme by the government be structured to cater for experienced and younger farmers who consider agriculture as a business. Farmers are also advised to form groups to be able to learn new techniques of production from one another. Forming or joining groups could also offer them the opportunity to contract loans and apply technologies which could increase efficiency and output. Agricultural policies of Ghana should emphasise intensification and the adoption of productivity-improving practices by farmers, as the per capita land area continues to reduce due to high population growth.

Endnotes

[1]The model was estimated directly using of LIMDEP 11 Software.

[2]Propensity score matching was used to select both rainfed and irrigation farmers to control for biasses stemming from observed variables. The result is available upon request.

[3]A bag of rice is standardised at 100 kg (0.1 MT). Yield is calculated from Table 6.3 as output ÷ farm size

Abbreviations

GSS: Ghana Statistical Service; Ha: Hectare; MiDA: Millennium Development Authority; MoFA: Ministry of Food and Agriculture; MT: Metric tonne; SFA: Stochastic frontier analysis; SPF: Stochastic production frontier; TE: Technical efficiency

Acknowledgements

We acknowledge all the rice farmers in the Upper East and Northern Regions of Ghana who provided the necessary data for this study. We also acknowledge the staff of the Ministry of Food and Agriculture in the two regions availing themselves for expert interviews.

Funding

This study was entirely sponsored by the authors as part of PhD programme of the corresponding author.

Availability of data and materials

The dataset supporting this paper will be provided upon request.

Authors' contributions

SBA designed the data collection instruments, gathered the data, analysed the data and wrote the first draft of the manuscript. SAD and JAA provided guidance, corrections, inputs and supervision to the entire study. All authors read and approved the final manuscript. We also confirm that the content of the manuscript has not been published or submitted for publication elsewhere except this journal.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Received: 23 August 2018 Accepted: 7 May 2019
Published online: 20 May 2019

References

Addison M, Ohene-Yankyera K, Fredua-Antoh E (2016) Gender role, input use and technical efficiency among rice farmers at Ahafo Ano North District in Ashanti Region of Ghana. J Food Secur 4(2):27–35

Anang BT, Bäckman S, Sipiläinen T (2016) Agricultural microcredit and technical efficiency: The case of smallholder rice farmers in Northern Ghana. J Agric Rural Dev Trop Subtrop 117(2):189–202

Azumah SB, Donkoh SA, Ansah IGK (2017) Contract farming and the adoption of climate change coping and adaptation strategies in the Northern Region of Ghana. Environ Dev Sustain 19(6):2275–2295
Donkoh SA, Ayambila S, Abdulai S (2013) Technical efficiency of rice production at the Tono irrigation scheme in northern Ghana. Am J Exp Agric 3(1):25–42
Dorward A, Chirwa E (2013) Impacts of the Far m Input Subsidy Programme in Malawi: informal rural economy modelling. Future Agricultures Consortium, Brighton
Ghana Statistical Service (GSS) (2014) National accounts statistics. Final 2012 gross domestic product & revised 2013 gross domestic product. http://www.statsghana.gov.gh
Greene W (2010) A stochastic frontier model with correction for sample selection. J Prod Anal 34(1):15–24
Heckman JJ (1979) Sample selection bias as a specification error. Econometrica 47:153–162
Kato T, Greeley M (2016) Agricultural input subsidies in Sub-Saharan Africa. Institute of Development Studies (IDS Bull). Library Road, Brighton BN1 9RE, UK, 47(2)
Kodde DA, Palm F (1986) Wald criteria for jointly testing equality and inequality restrictions. Econometrica 54(5):1243–1248
Kumbhakar SC, Tsionas EG, Sipiläinen T (2009) Joint estimation of technology choice and technical efficiency: an application to organic and conventional dairy farming. J Prod Anal 31(3):151–161
Lai H, Polachek S, Wang H (2009) Estimation of a stochastic frontier model with a sample selection problem. Working Paper, Department of Economics, National Chung Cheng University, Taiwan
Mayen CD, Balagtas JV, Alexander CE (2010) Technology adoption and technical efficiency: organic and conventional dairy farms in the United States. Am J Agric Econ 92:181–195
Millennium Development Authority (MiDA) (2010) Investment opportunity in Ghana: maize, rice, and soybean. MiDA, Accra
Ministry of Food and agriculture (MoFA) (2013) Agriculture in Ghana: facts and figures (2012). Statistics, research and information directorate (SRID), Accra
Ministry of Food and Agriculture (MoFA) (2016) Agriculture in Ghana. Facts and Figures 2015. Statistics, Research and Information Directorate (SRID), October 2016. Accra, Ghana
Ragasa C, Dankyi A, Acheampong P, Nimo-Wiredu A, Chapoto A, Asamoah M, Tripp R (2013) Patterns of adoption of improved rice technologies in ghana, GSSP-IPFRI Working Paper, vol 35
Ragasa C, Takeshima H, Chapoto A, Kolavalli S (2014) Substituting for rice imports in Ghana. GSSP-IFPRI. Policy Note #6. August 2014
Rahman S, Barmon BK (2015) Productivity and efficiency impacts of urea deep placement technology in modern rice production: an empirical analysis from Bangladesh. J Dev Areas 49(3):119–134
Sipiläinen T, Oude Lansink A (2005) Learning in switching to organic farming. NJF-Seminar 369, NJF Report Vol 1 No 1, Nordic Association of Agricultural Scientists
Solis D, Bravo-Ureta BE, Quiroga R (2007) Soil conservation and technical efficiency among hillside farmers in Central America: a switching regression model. Aust J Agric Resourc Econ 51:491–510
Villano R, Bravo-Ureta B, Solís D, Fleming E (2015) Modern rice technologies and productivity in the Philippines: disentangling technology from managerial gaps. J Agric Econ 66(1):129–154
World Bank (2013) Growing Africa: unlocking the potential of agribusiness. World Bank, Washington, DC