Fola *et al. BMC Agriculture* (2025) 1:4
https://doi.org/10.1186/s44399-025-00006-w

BMC Agriculture

RESEARCH

Open Access

Effect of maize cluster farming on smallholder farmers' technical efficiency: evidence from Southern Ethiopia

Mulugeta Fola[1*], Genene Tsegaye[1], Samuel Zawde[2] and Mathewos Matsalo[2]

Abstract

Background Agriculture plays a crucial role in Ethiopia's economy in terms of employment and overall output. However, the sector's productivity remains suboptimal due to fragmented production and institutional inefficiencies. As a result, the design of an effective food production policy has emerged as a resilience strategy to enhance productivity. The Ethiopian government implemented a cluster-based crop production strategy in line with this. However, there is inadequate evidence regarding whether this policy has improved production efficiency. This study, therefore, investigated the impact of maize cluster-based production on the technical efficiency of smallholder farmers in southern Ethiopia.

Methods To assess the efficiency and the effects of cluster farming, we employed stochastic frontier and endogenous switching regression (ESR) models, respectively. Data were collected from 421 randomly selected smallholder farmers during the 2021 production season, and the translog production function frontier model was used to analyze efficiency.

Results Of the sample, 49% were in clusters, and 51% were not. The results from the stochastic translog frontier model revealed that maize output is most responsive to land input relative to others. The estimated technical efficiencies were 74% for cluster farmers and 60% for non-cluster farmers, indicating that cluster-based production reduces technical inefficiency and enhances efficiency. The first stage of the ESR model identified several significant factors influencing household participation in cluster farming, including sex, oxen ownership, frequency of extension contacts, market distance, and access to credit. The ESR model also showed that cluster-based farmers would theoretically lose 18% (ATT) in technical efficiency if they did not engage in cluster production, while non-cluster farmers would theoretically gain 33% in technical efficiency if they participated in cluster farming.

Conclusion Consequently, policymakers and development organizations should focus on promoting clustering in crop production while addressing key factors influencing farmers'participation.

Keywords Cluster-based production, Endogenous switching regression, Stochastic frontier model, Technical efficiency

Background

The agricultural sector in sub-Saharan African countries is critical for achieving sustainable development goals. However, it faces production inefficiencies and is highly vulnerable to climate risks [1]. As a result, poverty remains persistent [2–4]. Nevertheless, the use of modern inputs, including knowledge inputs and the

*Correspondence:
Mulugeta Fola
mulugetafola@gmail.com
[1] Sidama Agricultural Research Institute, P.O. Box 06, Hawassa, Ethiopia
[2] South Ethiopia Agricultural Research Institute, Dilla, Ethiopia

adoption of improved technologies, suggests that there is potential to further enhance productivity [5]. Therefore, it is essential to design an appropriate agricultural policy strategy to address productivity bottlenecks.

Early strategic efforts to increase productivity in Ethiopia involved providing agricultural extension services and advice to farmers. However, these efforts became infeasible due to the inadequate supply of inputs, such as seeds and fertilizers. Later, the promotion of crop intensification through the integrated use of agricultural inputs was introduced [4, 6] However, this strategy also failed due to its limited coverage and quality [7]. Furthermore, the establishment of farmer training centers (FTCs) and the increase in the number of development agents (DAs) in each small administrative unit were implemented as means to transfer improved agricultural technologies. However, due to the low retention of DAs within kebeles and the poor functionality of most FTCs, the effective transfer and extension of agricultural knowledge has not been successful in many regions of the country [8]. Additionally, although watershed management and soil conservation practices have been implemented, degradation is expected due to inappropriate layouts, insufficient regular maintenance, and financial constraints for sustainable management [9, 10].

Furthermore, rural land certification has been implemented to enhance land productivity by ensuring tenure security [11, 12]. Despite its positive contributions, fragmented farming practices have become a significant cause of lower productivity [13–15]. This is because smallholder farmers typically have less than one hectare of land, which is often further fragmented into smaller parcels, thereby reducing productivity and exacerbating food insecurity [16].

The agricultural commercialization cluster initiative was introduced in 2010, aimed at scaling up the dissemination of best practices and specialization in high-value crops based on market demand [14, 17]. Clustering is designed around the geography of clusters and the commodity production potential, promoting institutional innovations to help farmers overcome constraints, thereby improving productivity [18, 19]. Specifically, cluster farming supports the transition from subsistence farming to market-oriented production by providing capacity-building through training and demonstrations, fostering cooperation among farmers to share farm inputs, facilitating markets for outputs, and providing access to seeds, fertilizers, and crop protection chemicals. As a result, clusters have the potential to increase efficiency and productivity by applying recommended input rates for specific crops, receiving support from agricultural extension agents, and utilizing shared inputs, thus addressing constraints in access to farm inputs [20].

Maize is one of the staple crops included in this strategy due to its status as the lowest-cost caloric source among all major cereals. Additionally, it dominates household diets and provides twice as many calories per dollar as other cereals in Ethiopia [21]. However, its productivity is low compared to its potential, with farmers only producing 30% of its potential. Contributing factors include inadequate access to advanced technologies, market imperfections, economic constraints, and technical inefficiency [22]. Efficiency studies have revealed that efficiency gaps range from 14% to 28.3% and have identified factors contributing to these gaps, such as land fragmentation, market distance, low adoption of sustainable agricultural intensification practices, and household socioeconomic characteristics [23–26]. These studies indicate that producers lose over one-fourth of their potential production due to technical inefficiencies.

To address this inefficiency, a cluster-based production strategy was implemented to overcome the barriers of land fragmentation and technology adoption [20, 27] reported that cluster-based production increases productivity, as measured by yield per area. However, land is not the only input affecting agricultural production; inputs such as fertilizer, seed types, and other conventional inputs, along with socioeconomic and institutional factors, also play vital roles, leaving areas underexplored. Therefore, efficiency measurement provides a more comprehensive method for evaluating policy strategies [28]. Furthermore, we employed an endogenous switching regression model to account for both observable and unobservable farmer characteristics, thus externalizing the effects of cluster farming on technical efficiency scores.

Moreover, there is empirical debate on the effectiveness of maize cluster farming in improving technical efficiency and the existence of regional differences across Ethiopia. Differing socioeconomic and cultural practices across regions in Ethiopia, which affect agricultural production and policy implementation, necessitate decentralized policy design. Region-specific studies have the potential to enhance policy development. Therefore, this study aimed to assess the impact of maize cluster farming on the technical efficiency of smallholder farmers in southern Ethiopia and found sufficient evidence that cluster farming reduces production inefficiencies. Given the relatively more fragmented nature of crop production in southern Ethiopia compared to other regions of the country, this study can serve as an instrument for policy improvement.

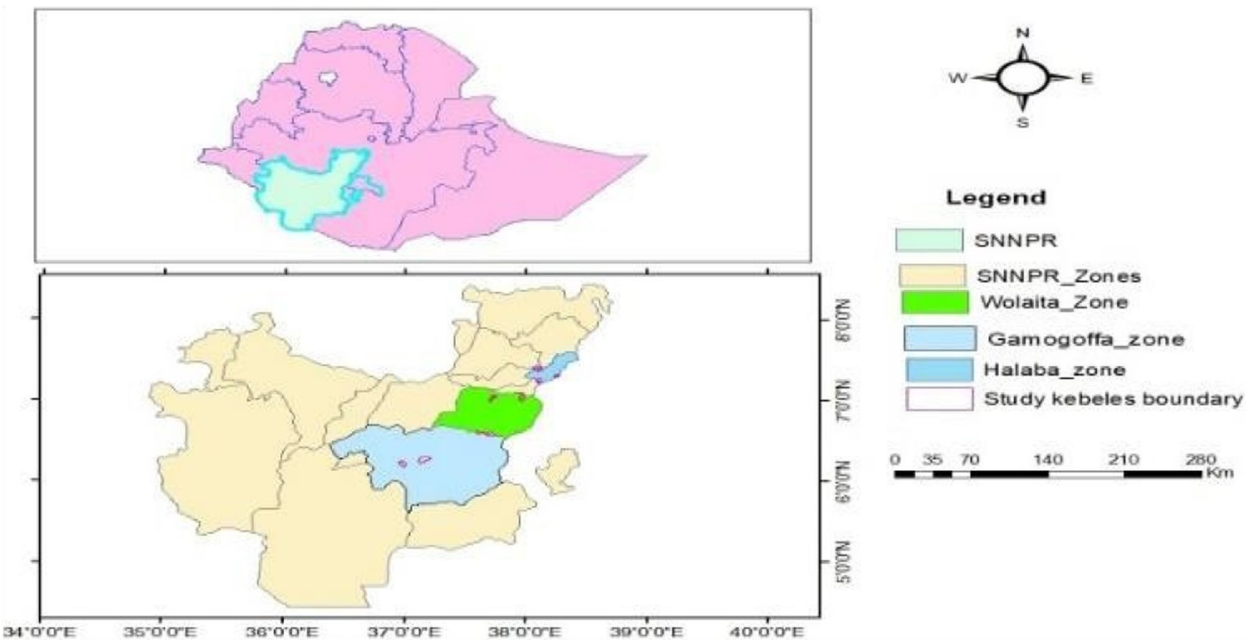

Fig. 1 Study area map. Source: Author (2021)

Methods

Description of the data and sampling techniques

Data were collected from major maize-producing areas in southern Ethiopia by the Hawassa, Areka, and Arbaminch Agricultural Research Centers in 2021. The study areas included the Halaba, Wolaita, Gamo, and Gofa zones, as shown in Fig. 1. One cluster-based producer district and one non-cluster-based district from each zone were selected for data collection based on maize production potential. A two-stage sampling technique was used. First, six districts were chosen based on their maize production potential and the presence of cluster-based production. Accordingly, Atoti Ulo, Damot Gale, and Boreda Woredas were selected as cluster districts, while Woyra, Boloso Sore, and Zala were selected as non-cluster districts. Next, three kebeles were selected from the Atoti Ulo and Woyra districts, and two kebeles were selected from each of the remaining districts based on the availability of clusters and non-clusters. A random sampling method was employed to select households from each kebele, with twenty to thirty-five households selected from each kebele using proportionate sampling. Of the 421 total samples, 206 were from cluster-based producers and 215 from non-cluster producers.

We used a structured questionnaire to collect data from households. The questionnaire included details on household socioeconomic and demographic information, maize production practices, cluster production, and outputs. Additionally, it incorporated institutional variables influencing production and productivity. The questionnaire[1] consists of seven major parts. The preliminary section includes general information about the respondent's geographical location. Part one covers household socioeconomic and demographic characteristics. Parts two and three focus on land use characteristics for maize and other crop types of representative households. The third part includes details on inputs used for maize production, from land preparation to threshing, as well as yields. Part four addresses agronomic practices, soil characteristics, and institutional variables. Parts five and six contain questions regarding cluster farming and barriers to participation. The final part covers livestock ownership and the respondents'source of annual income.

Data analysis

We employed both descriptive and econometric analyses. Means, frequencies, and percentages were used for the descriptive statistics, while the stochastic frontier and endogenous switching regression models was applied.

Technical efficiency model specification

The stochastic nature of agricultural production necessitates the use of the parametric analysis approach of the frontier model. Based on this, we assume that a farmer's production maximization from a given set of direct inputs is influenced by indirect constraints (socioeconomic and institutional factors, including cluster-based

[1] The questionnaire is provided in the supplementary documents.

production) that affect farm inefficiency. The direct inputs generate maize output, while the indirect factors facilitate the production process. Thus, direct inputs (Xi) are incorporated into the deterministic production frontier, and indirect factors (Zi) are included in the inefficiency component. Accordingly, technical efficiency is calculated based on farm household production performance, with Xi's used to produce a single maize output (Yi) for each household.

$$Ti \equiv \{X_i, Y_i)|X_i \text{can produce } Y_i \text{conditional on } Z_i\} \quad (1)$$

Equation (1) assume that all Y_i lie on the frontier, as it represents the upper boundary for production possibilities. We assume that maize production is below the frontier due to technical inefficiency [29]. The stochastic frontier model can be represented as:

$$\ln Y_i = \ln f(X_i; \beta) + v_i - u(z_i) \quad (2)$$

where Y_i*'s* is the logarithms of output, X_i^s are logarithms of the vector of direct inputs, the term $lnf(X_i\,\beta)$ is a deterministic frontier, v_i captures the random noise, and u_i is technical inefficiency which is assumed as half-normally distributed as indicated in Eq. (2) [28].

To estimate the empirical model, heteroscedastic stochastic frontier analysis (SFA) is employed to examine the effect of cluster-based production and other Zi on the variance of the efficiency distribution [29]. The assumption of constant (homoscedastic) inefficiency variance is impractical, as smallholder farmers operate under varying socioeconomic and institutional conditions. Moreover, neglecting heteroscedasticity in the variance of inefficiency may significantly bias the model estimates [28].

The model has the following form

$$\ln Y_i = \alpha + \ln f(X_i; \pi, \theta) + \varepsilon_i$$

$$\varepsilon_i = v_i - u_i$$

$$v_i \sim \mathrm{N}(0, \sigma_v{}^2) \quad (3)$$

$$ui \sim \mathrm{N}^+(0, \sigma_\mathrm{v}{}^2(\delta, Z_i))$$

where δ is a parameter vector associated with socioeconomic and institutional factors including cluster farming in the variance of inefficiency.

We used the translog production functional form for production function estimation due to its flexibility compared to the restricted Cobb–Douglas production function. Additionally, our preliminary analysis shows that the Cobb–Douglas production function was rejected in favor of the more flexible translog production function, with LR chi2(15) =31.11 and Prob >chi2 =0.0085. To maintain the homogeneity property of production, we scaled the inputs and outputs by their means. Therefore, the first-order coefficients of the estimated deterministic function are interpreted as elasticities of output evaluated at sample mean values. The equation for the technical efficiency translog model is as follows:

$$\ln Y_i = \ln A_{\pi i \theta ij} + \sum_{i=1}^{n} \pi_i \ln X_i + (\frac{1}{2}) \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \theta_{ij} \ln X_i . \ln X_j + v_i - u_i \quad (4)$$

where n is the number of inputs.

The parameters of Eq. (4) are estimated using the maximum likelihood (ML) method. The parameters of the deterministic production function, as well as the socioeconomic and institutional factors in the inefficiency effect function, are estimated simultaneously using the standard one-step modeling approach. This approach fits the production and inefficiency functions simultaneously, rather than separately in two stages [30]. The scale parameters for the ML estimation are expressed using variance parameters. The estimated parameters are then used to calculate farmers' specific technical efficiency scores (TE) using Battese and Coelli's (1988) estimator as explained below.

$$E[\exp(-ui|\varepsilon_i)] = \exp(-\mu_{*i} + \frac{1}{2}\sigma_*^2) \frac{\Phi(\frac{\mu_{*i}}{\sigma_*} - \sigma_*)}{\Phi(\frac{\mu_{*i}}{\sigma_*})} \quad (5)$$

where $\mu_{*i} = \frac{\sigma_u^2 \varepsilon_i}{\sigma_v^2 + \sigma_u^2}$, $\sigma_*^2 = \frac{\sigma_v^2 \sigma_u^2}{\sigma_v^2 + \sigma_u^2}$ and σ_v^2 and σ_u^2 are variance parameters of error and inefficiency terms [28].

Endogenous switching regression model specification

To accurately assess the impact of cluster-based maize production on the efficiency of smallholder farmers, both observable and unobservable characteristics of cluster producers (treatment group) and non-cluster producers (control group) must be accounted for. However, many impact analysis approaches using non-experimental data fail to capture the observable and unobservable factors affecting treatment and outcome variables. For example, instrumental variables can only address unobserved heterogeneity but rely on the assumption that the parallel shift of outcome variables represents the treatment effect [2]. Similarly, using regression models to analyze the impact of an intervention with pooled samples of participants and non-participants may be inappropriate, as it assumes similar effects on both groups [31]. In contrast, the endogenous switching regression model (ESR) is a methodological approach that addresses these issues [2, 32].

For this study, the parametric approach of endogenous switching regression (ESR) was employed to reduce selection bias, compared to the non-parametric approach (propensity score matching method). The impact of cluster-based maize production on technical efficiency under ESR is analyzed in two stages. In the first stage, participation in cluster-based farming is estimated using the binary probit model as the selection equation. In the second stage, both linear regression and the binary probit model are used to assess the association between the outcome variable and cluster farming. Specifically, the model adopts the expected utility maximization theory to explain farmers' participation in cluster farming (CF). Individual i participates in CF if the expected utility from production within CF exceeds the expected utility from non-participation.

$$C_i^* = \theta X_i + v_i \text{ Where } C_i^* = \begin{cases} 1 \\ 0 \end{cases} \text{if } C_i^* > 0 \tag{6}$$

where C_i^* is the latent variable capturing the unobserved preferences associated with the participation of CF determined by observed farm and socioeconomics characteristics of X_i and the error term (v_i). C_i is an observed binary indicator variable that equals one if farmers participated on CF and zero otherwise, while θ is a vector of parameters to be estimated.

If the selection equation (first stage) is endogenous in the outcome equation (second stage), the results will be biased and inefficient [32]. Therefore, using an instrumental variable to identify the second-stage equation from the first stage is crucial. The instrumental variable should influence participation in cluster farming (CF) but not affect the outcome (technical efficiency). Empirically selecting appropriate instrumental variables can be challenging; however, in this study, the frequency of extension contact (measured by the number of days) during maize production is used as the selection instrument. This is because agricultural extension services encourage farmers to participate in government initiatives such as CF [33]. Thus, this variable is likely to be correlated with participation in CF directly but not with technical efficiency. The outcome variable for both clusters (Regime 1) and non-clusters (Regime 2) can be expressed as the ESR model.

$$\begin{aligned} &\text{Regimes 1}: \ \mathrm{Y_{1i}} = \theta_1 \mathrm{Z_{1i}} + \varepsilon_{1i} \text{ if i} = 1 \\ &\text{Regimes 2}: \ \mathrm{Y_{2i}} = \theta_2 \mathrm{Z_{2i}} + \varepsilon_{2\mathrm{i}} \text{ if i} = 0 \end{aligned} \tag{7}$$

where Y_i represents the outcome variable (technical efficiency score) of smallholder farmers i for each regime (1 = CF participant and 0 = non-CF participant), Z_i is a vector of farmer farm and socioeconomic characteristics of households that affect technical efficiency of maize production, and θ_i is a vector of parameters to be estimated. The error terms inEquations 1 and 2 are distributed to be trivariate normal, with mean zero and the variance of the selection equation assumed to be equal to one since the coefficients are estimable only up to a scale factor.

Accordingly, by comparing the real and counterfactual scenarios of the expected outcomes for CF participants, the average treatment effect on the treated (ATT) is obtained. Similarly, the average treatment effect on the untreated (ATU) can be calculated by comparing the expected values for non-participants in real and counterfactual scenarios. Thus, the expected value of the outcomes for both participants and non-participants in both the real and counterfactual scenarios is given as:

Participants with participation in CF (real)

$$E[Y_{1i}|X = 1,] = \theta_1 X_{1i} + \sigma_{1v}\lambda_{1i} \tag{8}$$

Non-participants without participation in CF (real)

$$E[Y_{2i}|X = 0,] = \theta_2 X_{2i} + \sigma_{2v}\lambda_{2i} \tag{9}$$

If the participant had not participated in CF(counterfactual)

$$\begin{aligned} ATT &= E[Y_{2i}|X = 1,] - E[Y_{2i}|X = 1,] \\ &= (\theta_1 - \theta_2)X_{1i} + (\sigma_{1v} - \sigma_{2v})\lambda_{1i} \end{aligned} \tag{10}$$

If non-participants had participated in CF (counterfactual)

$$E[Y_{1i}|X = 0,] = \theta_1 X_{2i} + \sigma_{1v}\lambda_{2i} \tag{11}$$

Hence ATT of the participant is the difference Eq. 7 and 9

$$\begin{aligned} ATT &= E[Y_{1i}|X = 1,] - E[Y_{2i}|X = 1,] \\ &= (\theta_1 - \theta_2)X_{1i} + (\sigma_{1v} - \sigma_{2v})\lambda_{1i} \end{aligned} \tag{12}$$

Likewise, the ATU of non-participants is computed as the difference Eqs. 8 and 10

$$\begin{aligned} ATU &= E[Y_{1i}|X = 0,] - E[Y_{2i}|X = 0,] \\ &= (\theta_1 - \theta_2)X_{2i} + (\sigma_{1v} - \sigma_{2v})\lambda_{2i} \end{aligned} \tag{13}$$

Definitions of the variables and their measurements

For data analysis, we used two main types of variables. First, direct inputs—land, seeds, oxen, labor, and fertilizers—along with their respective measurement units, are explained in Table 1. The fertilizers used for maize production were NPS and UREA; thus, we combined them and referred to them as"fertilizer."Second, socioeconomic and institutional variables (including cluster-based

Table 1 Variable definitions and their respective measurements

Variables	Variable definitions and measurement units
Output (*Yi*)	Quantity of maize output produced in kilograms (kg)
Direct inputs (X_i)	
Land	Land area in hectares (ha)
Seed	Quantity of seed used in kilograms (kg)
Oxen	Oxen draught power used (oxen-days)
Labor	Quantity of labor used for production (person-days)
Fertilizer	Quantity of chemical fertilizers (NPS and Urea) in kilograms (kg)
Socioeconomic and institutional factors (Z_i)	
Sex	1 = if a respondent household head is male, 0 otherwise
Cluster	1 = if household produced in the cluster, 0 otherwise
Age	Age of the household head in years
Education	Education of the household head in years of schooling
Ox	Number of oxen owned by farm household
Credit	1 = if a household had access to credit, 0 otherwise
Income	The total annual income of a household measured in thousands of Birr
Extension contacts	The frequency of extension contacts during maize production

Source: Author (2021)

production) that affect farmers'technical (in)efficiency were considered, with their respective measurement units. Cluster farming is treated as a binary variable indicating whether farmers participate in clusters. Participation in cluster-based production is hypothesized to positively influence technical efficiency, thereby reducing inefficiency. Additionally, differences among farmers in terms of age, education, oxen ownership, access to credit, income, and frequency of extension contact are also expected to influence households'participation in maize cluster production and efficiency performance.

Table 2 Descriptive statistics of the variables

Variables		Mean	St. Dev
Output		1686.7	1897.7
Direct inputs (X_i)			
Land	Land amount	0.91	0.696
Seed	Seed amount	17.78	14.167
Oxen	Oxen number	17.54	27.86
Labor	Labor amount	55.24	69.52
Fertilizer	Fertilizer amount	126.82	144.52
Socioeconomic and institutional factors (Z_i)			
Sex		0.95	0.22
Cluster		0.49	0.50
Age		39.97	11.13
Education		4.32	3.37
Ox		1.22	1.03
Income		39.33	50.81
Extension contacts		4.21	6.33

Results

We used a structured questionnaire, validated through internal review by our office experts and external review by agricultural extension experts to ensure its content aligned with the theoretical foundation of the study. Both groups confirmed that the questions were appropriately structured. Additionally, we conducted a pilot survey with a small group of respondents, who provided positive feedback indicating that they clearly understood the questions. Furthermore, we observed consistency in responses across households, supporting the reliability and validity of our questionnaire.

Summary statistics of the variables of sample households

This section summarizes the variables used for frontier model estimation: direct inputs (Xi) and socioeconomic and institutional variables (Zi). As described in Table 2, the average land size used for maize production is 0.91 hectares, with an average of 55.24 labor days used. The average inorganic fertilizer applied was 126.82 kg, and the average oxen power used for plowing was 17.54 days. The average amount of seeds applied was 17.78 kg. Among the total sample, 95 percent were male-headed households, 49 percent produced maize through cluster farming, and 28 percent used manure for maize production. The average household age is approximately 40 years, the average educational level is 4.32 grades completed, oxen ownership is 1.22, the annual income (measured in thousands of Birr) is 39.33, and the frequency of extension contact is 4.21 days.

Results of stochastic frontier model estimation

The results of the translog efficiency model estimation in Table 3 show the first-order terms, second-order terms, and inefficiency factors. To estimate these, we applied a standard one-step modeling approach and incorporated socioeconomic and institutional variables (including cluster-based production) into the variance of the frontier models.

Table 3 Technical efficiency frontier model results

Direct inputs (Xi)	Coefficient	St. Error
Constant	0.342***	0.07
Land	0.469***	0.097
Seed	0.0272	0.09
Fertilizer	0.220***	0.06
Oxen	– 0.00844	0.05
Labor	0.248***	0.05
0.5 land^2	0.332	0.28
0.5 seed^2	0.228	0.31
0.5fertilizer^2	0.0907**	0.05
0.5oxen^2	– 0.0827	0.09
0.5 labor^2	– 0.0634	0.05
Land x seed	– 0.365	0.26
Land x *fertilizer*	– 0.00929	0.15
Land x *oxen*	– 0.114	0.16
Land x labor	0.170	0.14
Seed x fertilizer	– 0.0543	0.12
Seed x oxen	0.235	0.14
Seed x labor	– 0.0400	0.14
Fertilizer x oxen	– 0.0982	0.06
Fertilizer x labor	0.0855	0.05
Oxen x labor	– 0.0772	0.05
Inefficiency effects component of the frontier model		
Constant	0.175	1.29
Socioeconomic and institutional factors (z_i)		
Cluster	– 0.484**	0.24
Sex	– 0.660	0.42
Age	0.0236	0.06
Age2	– 0.000106	0.0006
Education	0.0507	0.03
Extension contacts	– 0.0424	0.03
Oxen ownership	– 0.427**	0.15
Income	– 0.028***	0.006
Vsigma	– 1.788***	0.15
Estimated parameters		
σu	0.61	
σv	0.41***	
337.0535	(0.03) Log likelihood = -	

SE standard error

*, **, and *** are significant at the 10%, 5%, and 1% probability levels, respectively

The estimated input elasticities are presented in Table 4. Translog production function elasticities are observation-specific and depend on the inputs used. Additionally, Table 4 provides the estimated efficiency scores from the model parameters for both cluster and non-cluster farmer categories. We also computed the returns to scale based on the estimated elasticities. Furthermore, we used the kernel density distribution to illustrate the differences in technical efficiency distributions between farmers in cluster and non-cluster production zones. As shown in Fig. 2, the efficiency of clusters is skewed to the right, indicating that their efficiency scores are higher than those of non-clusters.

Factors affecting smallholder farmers' participation in cluster farming

The first-stage probit results of the ESR model are presented in Table 5. The model fits the data reasonably well (LR-chi2 = 113.91, P = 0.000). The results indicate that smallholder farmers' socioeconomic and institutional factors significantly affected their participation in cluster farming. Factors such as sex, ownership of oxen, frequency of extension contacts, market distance, credit access, and household income positively and significantly influenced their participation.

Effect of cluster farming on technical efficiencies

The effect of cluster farming participation on the technical efficiencies of smallholders was the primary focus of this study. Table 6 presents the results of the ESR model-based ATT and ATU for the key outcome variable (technical efficiency scores) related to participation in cluster farming. As mentioned earlier, the technical efficiency scores of smallholders were analyzed using the estimated parameters of the stochastic frontier model with a translog production function and used as outcome variables.

Table 4 Estimated elasticities of a translog production function

Direct inputs	Output elasticities	
	Mean	St. Dev
Land	0.36	0.34
Seed	– 0.06	0.24
Fertilizer	0.16	0.18
Oxen	0.08	0.22
Labor	0.27	0.15
Total technical efficiency	0.67	0.19
Clusters technical efficiency	0.74	0.17
Nonclusters technical efficiency	0.60	0.19
Returns to scale	**0.81**	

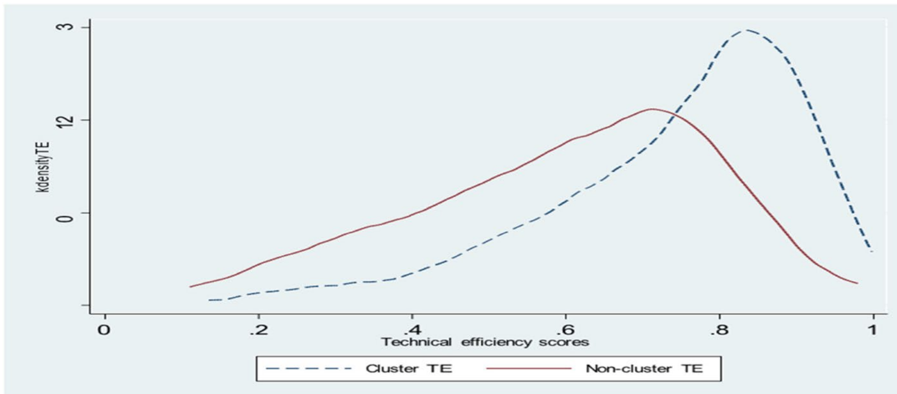

Fig. 2 Technical efficiency score distribution. Source: Author (2021)

Table 5 First stage ESR model probit model estimation

Observations = 421			
LR chi2(11) = 113.91			
Prob > chi2 = 0.0000			
Pseudo R2 = 0.1952			
Log-likelihood = – 234.76378			
Variables	**Coefficient**	**Standard error**	**Marginal effects**
Constant	– 1.904*	0.85	
Age	– 0.00937	0.04	– 0.004
Age2	0.0000386	0.0004	0.00002
Education	0.0303	0.022	0.01
Sex	0.753*	0.34	0.28
Family size	0.00231	0.03	0.0009
Land holdings	0.0987	0.07	0.04
Ownership of the oxen	0.325***	0.08	0.13
Extension contact	0.0758***	0.02	0.03
Nearest market distance	0.0913***	0.02	0.04
Credit	0.428*	0.21	0.17
Income	0.00274	0.002	0.001

Source: Author (2022) *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$ statistically significance levels

The ESR model output indicates that maize production in clusters enhances technical efficiency.

Discussion

The input amounts determine the respective output levels, as shown in Table 2. The larger standard deviations of most direct inputs and outputs from their means are attributed to the significant variations in farm sizes. Additionally, the results indicate that there are variations in the socioeconomic and institutional conditions of maize-producing farm households, which may lead to differences in production efficiency.

Based on the results in Table 3, the estimated translog model shows that, among the direct inputs, land, fertilizer, and labor have statistically significant and positive effects on maize output, with land being the most responsive input. The second-order estimates reveal that the coefficients for oxen and labor are negative, which is intuitive and indicates that farmers experience diminishing returns from the use of oxen and labor. The second-order effect of fertilizer significantly increases returns to output. This is primarily due to the low application of inorganic fertilizers, which results in stagnant crop yields [34] due to liquidity and institutional constraints that affect farmers'ability to apply inorganic fertilizers.

Table 6 Endogenous switching regression model results (Average Treatment Effects)

Outcome variables	Obs	Type of treatment	Cluster	Non-cluster	Treatment effect
Technical Efficiency	206	ATT	0.74 (0.11)	0.56 (0.16)	0.18***
	215	ATU	0.93 (0.09)	0.60 (0.13)	0.33***

*** Significant at the 1percent probability level. Statistics within parenthesis imply standard deviations

In addition, Table 3 shows that socioeconomic and institutional variables influence technical (in)efficiency. The results indicate that maize cluster-based production reduces technical inefficiency and increases technical efficiency, which is expected. This suggests that farmers involved in cluster farming receive specialized support from both governmental and non-governmental organizations that promote agriculture through cluster farming. Oxen ownership also negatively affects the variability of technical inefficiency, meaning that having more oxen allows farmers to begin plowing at the optimal time in the cropping season. Delayed sowing due to a shortage of oxen results in lower crop performance, thereby reducing yields [35]. Oxen ownership differs from the efficient use of oxen power. Household income is another factor that negatively and statistically significantly reduces technical inefficiency, as higher income alleviates liquidity constraints.

To determine the elastic nature of the inputs, the estimated elasticities are presented in Table 4, which reveals that all elasticities are positive except for the seed in the output elasticities. The estimated output elasticities for land, seed, fertilizer, oxen, and labor were 36%, 6%, 16%, 8%, and 27%, respectively. This shows that maize output is more elastic to land. The estimated returns to scale of the technical efficiency function show that the rate of return is 0.81 for output elasticity, indicating that maize production exhibits decreasing returns to scale. We also expected an efficiency difference between clusters and non-clusters. The results showed that households producing maize in clusters had higher technical efficiency than those producing maize in non-cluster areas. Specifically, the technical efficiency scores for clusters and non-clusters were 74% and 60%, respectively.

Table 5 shows that, compared to female household heads, male household heads are more likely to participate in cluster farming. This is due to gaps in access to agricultural extension advice between male and female heads. This finding is consistent with prior studies; [36, 37] argued that being a male household head positively influences participation in cooperative initiatives like agro-clusters. Ownership of oxen also increases the likelihood of participation, as having oxen motivates farmers to begin plowing their fields with fewer days of variation compared to other cluster members. This aligns with the cluster strategy's principle, which allows a maximum of five days'variation for both land preparation and sowing. The frequency of extension contacts also significantly affects participation in clusters, implying that government agricultural extension services are the primary pathway for farmers, and the more frequent the contact, the higher the probability of receiving information about cluster farming. This finding is consistent with [36, 38].

In addition, the distance to the market also significantly affects the probability of participation in cluster farming. This implies that an increase in market distance raises the likelihood of participating in cooperative initiatives like agricultural clusters. Farmers who are closer to the market tend to be less dependent on group activities, whereas farmers farther from the market expect higher returns and reduced transaction costs from cooperation. This finding aligns with studies by [39, 40] Ahmed et al. (2017), who found that cluster or group-based crop production facilitates the provision of inputs to members and reduces transaction costs. Access to credit also significantly increases the likelihood of smallholders participating in cluster-based production. This suggests that having access to credit reduces liquidity constraints and encourages the timely purchase of improved inputs, which is consistent with the findings of [41].

Finally, Table 6 shows that farmers who produced maize in the cluster but theoretically would not have done so outside the cluster experienced an 18% decrease in their technical efficiency level. This suggests that, on average, the ATT of technical efficiency would decline by 18% if they did not produce in the cluster. Similarly, if non-cluster producers had participated in cluster farming, their average technical efficiency level would increase by 33%. Moreover, maize production within clusters enhanced the technical efficiency of households. This finding is consistent with prior studies that examined whether group-based initiatives such as agro-clusters and cooperative-based crop production have a positive and significant effect on farmers' technical efficiency [42–45].

Conclusion

This study finds that maize cluster-based production is negatively associated with technical inefficiency, implying that it enhances efficiency. Additionally, the levels of technical efficiency are higher in clusters compared to non-clusters. We find that sex, oxen ownership, extension contact, market distance, and credit access were key factors influencing cluster participation. Additionally, our results suggest that non-participants would gain efficiency if they participated, while participants would experience efficiency loss if they did not participate.

The overall findings indicate a persistent efficiency gap in maize production, which significantly hinders the livelihood of agrarian communities. Given that maize is a staple crop in the area, targeted support is essential to enhance productivity and efficiency. This can be achieved by strengthening the improved seed production system, modernizing the agricultural extension system through advanced communication technologies, and promoting climate-smart agricultural practices. Furthermore, expanding cluster-based production has the potential to reduce production inefficiency. Therefore, we conclude that cluster-based production significantly contributes to the social and economic development of smallholder farmers by improving average technical efficiency scores and addressing extension support barriers, such as the provision of information on comprehensive agricultural technology packages. Thus, we suggest that policymakers and development organizations should focus on supporting farmers'clustering in crop production while considering significant socioeconomic and institutional factors. Moreover, collective action among farmers, governments, the private sector, and donors in the region is essential.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s44399-025-00006-w.

Supplementary Material 1.

Acknowledgements
We sincerely thank Mr. Tizazu Toma, Coordinator of Agricultural Economics Research at the Institute, as well as the work process coordinators of the Areka Agricultural Research Center (Mr. Yidnekachew Alemayehu) and the Arbaminch Agricultural Research Center (Mr. Endrias Oyka) for their follow-ups and administrative supports. We also appreciate the facilitators, including both woreda and kebele experts, in each sample area. Finally, we extend our gratitude to all colleagues who participated in the data collection.

Authors' contributions
MF initiated the study and contributed to all parts of the research. GT facilitated and arranged the training necessary for manuscript preparation and participated in the data analysis. SZ and MM participated in the data collection and analysis.

Funding
The authors did not receive direct funding for this research.

Data availability
The authors declare that the data will be available upon request from the corresponding author through the indicated email address.

Declarations

Ethics approval and consent to participate
The Sidama Agricultural Research Institute's ethical committee, the Director of Agricultural Economics Research, and the Work Process Coordinator approved the study with the number SARI/2021/017. Additionally, informed oral consent was obtained from all respondents during the survey.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 11 September 2024 Accepted: 27 March 2025
Published online: 08 April 2025

References

1. Zilberman D, Goetz R, Garrido A. Climate smart agriculture building resilience to climate change. Available: http://www.springer.com/series/6360.
2. Ahmed SA, Diffenbaugh NS, Hertel TW. Climate volatility deepens poverty vulnerability in developing countries. Environ Res Lett. 2009;4(3). https://doi.org/10.1088/1748-9326/4/3/034004.
3. Arias-Hidalgo M, et al. A decision framework for wetland management in a river basin context: the "Abras de Mantequilla" case study in the Guayas River Basin, Ecuador. Environ Sci Policy. 2013;34:103–14. https://doi.org/10.1016/j.envsci.2012.10.009.
4. Louhichi K, Temursho U, Colen L, Paloma SG. Upscaling the productivity performance of the agricultural commercialization cluster initiative in Ethiopia : an assessment using a farm household level model. Publications Office of the European Union, 2019.
5. Dercon S. Growth and shocks: evidence from rural Ethiopia. J Dev Econ. 2004;74(2):309–29. https://doi.org/10.1016/j.jdeveco.2004.01.001.
6. Degife A, Worku H, Gizaw S. Environmental implications of soil erosion and sediment yield in Lake Hawassa watershed, south-central Ethiopia. Environ Syst Res. 2021;10(1). https://doi.org/10.1186/s40068-021-00232-6.
7. Di Falco S. Adaptation to climate change in Sub-Saharan agriculture: Assessing the evidence and rethinking the drivers. In European Review of Agricultural Economics. Oxford University Press;2014. pp. 405–430. https://doi.org/10.1093/erae/jbu014.
8. Sc M, Mengistu FT. Problems and prospects of farmers training centers: the case of ada'a woreda, East Shewa, Oromia Eegion. 2009.
9. Yericho BM. Technical viability of physical soil and water conservation structures implemented in Lake Hawassa watershed, southern Ethiopia. J Soil Sci Environ Manage. 2019;10(4):68–74. https://doi.org/10.5897/jssem2019.0740.
10. Fola M, Ketema M, Alamerie K. Households' willingness to pay for the services of watershed management in lake Hawassa watershed, southern Ethiopia. J Forest Natural Resour. 2022;1(1):9–19.
11. Gedefaw AA, Atzberger C, Seher W, Agegnehu SK, Mansberger R. Effects of land certification for rural farm households in ethiopia: Evidence from Gozamin District, Ethiopia. Land (Basel). 2020;9(11):1–23. https://doi.org/10.3390/land9110421.
12. Melesse MB, Bulte E. Does land registration and certification boost farm productivity? Evidence from Ethiopia. Agric Econ (United Kingdom). 2015;46(6):757–68. https://doi.org/10.1111/agec.12191.
13. Alemu A. ISSN(e): 24086851; ISSN(Print); 1119944X Food and Agricultural Organization (FAO). Electronic Journals Service (EJS). 2021;25(2). https://doi.org/10.11226/v25i2.
14. Ethiopian Agricultural Transformation Agency (ATA). Action agenda for a new food and land use economy in Ethiopia. Policy document. Addis Ababa; 2020.
15. Rahman S, Rahman M. Impact of land fragmentation and resource ownership on productivity and efficiency: the case of rice producers in Bangladesh. Land Use Policy. 2009;26(1):95–103. https://doi.org/10.1016/j.landusepol.2008.01.003.
16. Knippenberg E, Jolliffe D, Hoddinott J. Land Fragmentation and Food Insecurity in Ethiopia. Am J Agric Econ. 2020;102(5):1557–77. https://doi.org/10.1002/ajae.12081.
17. National Panning Commision. Growth and Transformation Plan II (GTP II) 2015/16 - 2019/20. Addis Ababa; 2015.
18. Agro-based clusters in developing countries: staying competitive in a globalized economy. Food And Agriculture Organization Of The United Nations;2010.
19. Burger K, Kameo D, Sandee H. Clustering of small agro-processing firms in Indonesia Daniel Kameo. 2001.
20. Degefu S, Sileshi M, Ogeto MA. Impact of cluster farming on wheat productivity and net benefit among smallholder farmers in Lemu-Bilbilo and Hetosa districts of Arsi Zone, Ethiopia. Discover Food. 2024;4(1). https://doi.org/10.1007/s44187-024-00128-1.
21. Ababa A . The Federal Democratic Republic Of Ethiopia Central Statistical Agency Statistical Report On The 2012 Urban Employment Unemployment Survey. 2012. Available: http://www.alientools.com/

22. Van Dijk M, Morley T, van Loon M, Reidsma P, Tesfaye K, van Ittersum MK. Reducing the maize yield gap in Ethiopia: Decomposition and policy simulation. Agric Syst. 2020;183. https://doi.org/10.1016/j.agsy.2020.102828.
23. Yimer F, Alemayehu M, Taffesse S. Bringing Rigour and Evidence to Economic Policy Making in Africa The Short-run Impact of the COVID-19 Crisis on Poverty in Ethiopia. 2020.
24. Alemu M, Berihun D, Lokossou JC, Yismaw B. Productivity and efficiency heterogeneity among maize smallholder farmers in Ethiopia. Cogent Food Agric. 2024;10(1). https://doi.org/10.1080/23311932.2023.2300191.
25. Oumer AM, Mugera A, Burton M, Hailu A. Technical efficiency and firm heterogeneity in stochastic frontier models: application to smallholder maize farms in Ethiopia. J Prod Anal. 2022;57(2):213–41. https://doi.org/10.1007/s11123-022-00627-2.
26. Motbaynor Workneh W, Kumar R. The technical efficiency of large-scale agricultural investment in Northwest Ethiopia: a stochastic frontier approach. Heliyon. 2023l9(9). https://doi.org/10.1016/j.heliyon.2023.e19572.
27. Cheffo A, Ketema M, Mehare A, Shumeta Z, Habte E. Impact of malt-barley commercialization clusters on productivity at household level: the case of selected Districts of Oromia Region, Ethiopia. 2023.
28. Kumbhakar SC, Wang HJ, Horncastle AP. A practitioner's guide to stochastic frontier analysis using stata ebook converter DEMO Watermarks.
29. Jondrow J, Lovell CK, Materov IS, Schmidt P. On the estimation of technical inefficiency in the stochastic frontier production function model*.
30. Wang HJ, Schmidt P. One-Step and two-step estimation of the effects of exogenous variables on technical efficiency levels. 2002.
31. Kassie M, Zikhali P, Pender J, Köhlin G. Sustainable agricultural practices and agricultural productivity in ethiopia: does agroecology matter? 2009. Available: www.handels.gu.se.
32. Jaleta M, Kassie M, Marenya P, Yirga C, Erenstein O. Impact of improved maize adoption on household food security of maize producing smallholder farmers in Ethiopia. Food Secur. 2018;10(1):81–93. https://doi.org/10.1007/s12571-017-0759-y.
33. Biswas B, Mallick B, Roy A, Sultana Z. Impact of agriculture extension services on technical efficiency of rural paddy farmers in southwest Bangladesh. Environ Challenges. 2021;5. https://doi.org/10.1016/j.envc.2021.100261.
34. Zhang L. et al., 'Rainfall patterns influence poverty levels throughout Sub-Saharan Africa. 2024. https://doi.org/10.21203/rs.3.rs-4917281/v1.
35. Geren H, Ozdogan T, Sımıc A, Dzeletovıc ZS. Effect of different sowing dates on the grain yield and some yield characteristics of teff [eragrostis teff (Zucc.) trotter]. Turkish J Field Crops. 2020;25(2):107–13. https://doi.org/10.17557/tjfc.831853.
36. Hussen CH, Geleta FT. Factors affecting smallholder farmers participations in cluster crop production: evidence from selected Districts of West Shewa Zone, Oromia National Regional State, Ethiopia. Sarhad J Agric. 2021;37(3):818–29. https://doi.org/10.17582/journal.sja/2021/37.3.818.829.
37. Endalew B, Elias A, Yasunobu K. Impact of cluster farming on smallholder farmers teff commercialization in Ethiopia', CABI Agriculture and Bioscience. 2004;5(1). https://doi.org/10.1186/s43170-024-00220-7.
38. Wiredu AN et al. Factors Influencing Farmer's Participation in Agricultural Projects: The case of the Agricultural Value Chain Mentorship Project in the Northern Region of Ghana. 2013. Available: www.iiste.org.
39. Ahmed MH, Geleta KM, Tazeze A, Andualem E. The impact of improved maize varieties on farm productivity and wellbeing: evidence from the east hararghe zone of Ethiopia. Dev Stud Res. 2017;4(1):9–21. https://doi.org/10.1080/21665095.2017.1400393.
40. Verhofstadt E, Maertens M. Can agricultural cooperatives reduce poverty? Heterogeneous impact of cooperative membership on farmers' welfare in Rwanda. Appl Econ Perspect Policy. 2015;37(1):86–106. https://doi.org/10.1093/aepp/ppu021.
41. Mathewos T, Temesgen D, Hamza D, Fesseha H. Determinants of smallholder farmers' participation in improved sheep production: the case of Doyogena District, Kembata Tembaro Zone, Southern Ethiopia. Advances in Agriculture. 2021;2021, https://doi.org/10.1155/2021/5514315.
42. Olagunju KO, Ogunniyi AI, Oyetunde-Usman Z, Omotayo AO, Awotide BA. Does agricultural cooperative membership impact technical efficiency of maize production in Nigeria: An analysis correcting for biases from observed and unobserved attributes. PLoS One. 2021;16(1). https://doi.org/10.1371/journal.pone.0245426.
43. Mwalupaso GE, Wang S, Rahman S, Alavo EJP, Tian X. Agricultural informatization and technical efficiency in maize production in Zambia. Sustainability (Switzerland). 2019;11(8):2451. https://doi.org/10.3390/su11082451.
44. Abate GT, Francesconi GN, Getnet K. Impact of agricultural cooperatives on smallholders' technical efficiency: evidence from Ethiopia. SSRN Electron J. 2013,https://doi.org/10.2139/ssrn.2225791.
45. Adetoyinbo A, Otter V. Can producer groups improve technical efficiency among artisanal shrimpers in Nigeria? A study accounting for observed and unobserved selectivity. Agric Food Econ. 2022;10(1). https://doi.org/10.1186/s40100-022-00214-x.

Publisher's Note