African Journal of
Agricultural and Resource Economics
Volume 21, Number 1 (2026), pp 103–133

Impact of cluster farming on wheat commercialisation in Ethiopia

Hassen Nurhussen Ahmed*
Haramaya University, Dire Dawa, Ethiopia. E-mail: hasseno2005@gmail.com

Dawit Alemu
Stichting Wageningen Research Ethiopia (SWR Ethiopia), Addis Ababa, Ethiopia. E-mail: dawit96@gmail.com

Lemma Zemedu
Ethiopian Institute of Agricultural Research, Addis Ababa, Ethiopia. E-mail: zemedul@gmail.com

Kedir Jemal
Agricultural Economics, Haramaya University, Dire Dawa, Ethiopia. E-mail: jemalkedir@gmail.com

* Corresponding author

Received: August 2025
Published: March 2026

DOI: https://doi.org/10.53936/afjare.2026.21(1).5

Abstract

In Ethiopia, wheat cluster farming has gained popularity in recent years due to its potential to transform the agricultural sector and improve the livelihoods of wheat producers. This study aimed to evaluate the impact of cluster farming on the wheat output and input commercialisation level. The study was conducted in the Amhara region of Ethiopia, using household-level survey data collected from 383 wheat-producing farm households selected through multistage sampling during the 2022/2023 production year. Descriptive statistics and the endogenous switching regression (ESR) model, supported by propensity score matching (PSM), were employed for data analysis. The results of the model show that the education status of the farm household head, access to formal credit services, frequency of extension contacts, distance from rural all-weather roads, wheat farm size, land fragmentation (index), livestock size (TLU), cluster farming awareness and distance of the farm to the nearby wheat cluster farms significantly influence farm household participation in wheat cluster farming. In addition, producing wheat in a cluster farming approach has a positive effect on wheat output and input commercialisation levels, and the PSM result ensures that the results are robust. The estimated average treatment effect on the treated (ATT) for wheat output and input commercialisation was 5.0% and 3.1%, respectively, while the average treatment effect on the untreated (ATU) was 1.9% and 0.4%, and the transitional heterogeneity (TH) 3.1% and 2.6%, respectively. All effects were positive and statistically significant, confirming the positive effect of the cluster farming approach on both wheat output and input commercialisation. Therefore, government and non-government policies and interventions should focus on enhancing smallholder farmers' participation in wheat cluster farming to improve their livelihoods through increased wheat output and input commercialisation.

Key words: cluster farming, wheat, commercialisation, ESR, ATT

1. Introduction

Wheat is an important crop that serves as food and source of income in Ethiopia. Nationally, it accounts for 18.68% and 19.97% of the total area coverage of cereal crops and production, respectively. Although the country ranks second in terms of total wheat production area and production in Africa (United States Department of Agriculture (USDA) 2025b), it remains a net importer of wheat products (USDA 2025a). This is due to rapid growth in the demand for wheat products (Senbeta & Worku 2023), the low level of technology adoption (Zenbaba *et al*. 2024) and smallholder producers' inefficiency (Silva *et al*. 2021), resulting in low productivity and commercialisation. For instance, the national average of wheat productivity was 3.11 tons per ha during the 2021/2022 production year (Central Statistics Agency of Ethiopia [CSA] 2022), which is far below the world average of 3.59 tons (USDA 2025a), and the average marketed proportion was only 24.37% (CSA 2021).

Amhara region is one of the largest wheat crop producers in the country. The crop accounts for 18.89% and 19.57% of the total area of cereal crops coverage and production respectively, indicating how important the crop is in the region. However, wheat productivity and the level of commercialisation in the region are low. In the 2021/2022 production year, average productivity was 2.87 tons per hectare, which is far below the national average productivity (CSA 2022). Moreover, 64.2% of wheat output is used for consumption, while a small proportion (16.4%) is sold in the market and the remaining is used for seed and other purposes, indicating that the commercialisation of the crop is low (CSA 2021).

The concept of agricultural commercialisation is complex and multidimensional, and hence there are various definitions with different degrees of emphasis (Jaleta *et al*. 2009; Boka 2017). Agricultural commercialisation can take place in the output side of production through increased marketed surplus, in the input side of production through the increased use of purchased agricultural inputs such as herbicides, fertilisers and mechanised equipment (tractors, harvesters), or on both sides of production (commercialisation that integrates both the output and input side, in which an output in commercialisation drives an input in commercialisation by improving smallholder producers' ability to purchase agricultural inputs. This considers both staple and cash crops (Von Braun 1995; Pingali 1997; Govereh *et al*. (1999), and is defined as the proportion of agricultural production sold in the market. It involves a transition from subsistence-oriented production to highly market-oriented production (both production and input use). Recently, for countries with a high share of agricultural GDP such as Ethiopia, agricultural commercialisation remains a major development strategy. This is because transforming smallholder farming from subsistence to commercial production is critical for sustaining the agriculture sector's vital role in economic growth. The main drivers that speed up such a transformation process are economic growth, urbanisation, and the withdrawal of labour from the sector (Pingali & Rosegrant 1995; Pingali 1997). Economic growth drives the expansion of agro-processing industries, which require a large and sustainable supply of agricultural products, necessitating greater commercialisation. It also influences change in people's dietary behaviour due to rising incomes shifting agriculture toward a more market-oriented approach. Similarly, the expansion of urban areas creates a concentrated market opportunity for agricultural products, encouraging smallholder producers to focus on market-oriented crops. In addition, the withdrawal of labour from the agricultural sector compels smallholder producers to adopt mechanisation and focus on the use of purchased inputs such as herbicides, which may lead to land consolidation as land ownership becomes concentrated in the hands of fewer individuals. However, it should be noted that agricultural commercialisation has its own downsides, such as environmental and health effects arising from dependence on chemical inputs (e.g. chemical fertilisers, herbicides), soil degradation due to specialisation, and increased reliance on chemical input producers (Pingali 2001; Negowetti 2017).

The commercial transformation of smallholder agriculture is an indispensable pathway towards economic growth and development for most developing countries relying on agriculture (Von Braun 1995; Kilimani *et al*. 2022). There are several research reports that reveal the importance of commercialisation in improving farm households' productivity and income. For instance, Adam *et al*. (2010) and Ademe *et al*. (2017) reported that smallholder producers with highly commercial-oriented farming had better crop productivity. In addition, Cazzuffi *et al*. (2020) and Tabe Ojong *et al*. (2022) reported the positive effect of agricultural commercialisation on asset accumulation, livestock ownership and income of farm households. Hence, there is a need to devise alternative policies, strategies and/or programmes that improve commercialisation. The cluster approach (an effective approach in the manufacturing sector) is believed to be a promising key policy alternative that could boost smallholder commercialisation (Guyo *et al*. 2023).

The cluster approach substantially promotes collaborative and cooperative activities that help to address a broad range of economic, social and ecological problems to ensure rural recovery and sustainability (Rosenfeld 2009). Porter (1990) emphasises that geographic and spatial distribution, and the economic sector of the cluster, are the basic attributes to be considered in defining the term cluster. Porter (1990) defines a cluster as a geographically proximate group of interconnected companies and associated institutions in a particular field, linked by commonalities and complementarities. Schmitz (1992) relates the issue of a cluster to agglomeration, and defines a cluster as a geographic and sectoral agglomeration of enterprises. Similarly, an agriculture-based cluster is defined as a concentration of producers, agribusinesses and institutions that are engaged in the same agricultural (crop or livestock) or agro-industrial subsector and is established to address common challenges and pursue opportunities through interconnected and value chain networks (FAO 2010).

The literature supports the importance of the cluster approach for improving productivity, and delivering extension services, training and other services (like financial services), improved technology. These together help in strengthening producers' collective action, creating market linkages, and facilitating the use of mechanisation and commercialisation. The cluster approach reduces barrier in terms of access to the inputs, services, staff and skills required (Schouten & Heijman, 2012). In addition, it makes the delivery of training to cluster members and experience sharing between clusters easy, which makes cluster participants better off in using improved technologies (Axalan *et al*. 2011; Karki *et al*. 2021). It also strengthens producers' collective action and bargaining position (Poulton *et al*. 2010; Kilelu *et al*. 2017) and this is crucial for small-scale producers to remain competitive, overcome the limitations of access to services, and mitigate different risk sources (Poulton *et al*. 2010; Molema *et al*. 2016). The approach is found to be vital in facilitating the delivery of extension services and improved technologies, creating market linkage, and the use of agricultural mechanisation (Te & Milkias 2021).

Farming in clusters facilitates small-scale farmers' engagement in higher productivity practices that are more market orientated and have higher value-added production (FAO 2010), mitigates market risk due to their collective power (Hao *et al*. 2018) and facilitates market reach (Martinidis *et al*. 2021). Cluster farming helps farm households to achieve economies of scale, which lower the costs per unit produced because of lower transport costs, competing input suppliers and specialisation. It provides incentives for entry to a market because of easy access to information about opportunities (Schouten & Heijman 2012). It also plays a significant role in improving the production of major cereal crops (wheat, teff, maize and barley) to improve farm households' level of market participation (Elouhichi *et al*. 2019). All these things contribute to improvements in the level of commercialisation.

The Ethiopian government introduced the agricultural commercialisation cluster (ACC) initiative programme in 2017. This programme is designed to provide institutional support and play a central role in transforming agriculture and improving smallholder producers' livelihoods (ATA 2018). The enhancement of farm productivity, production competitiveness and linkages amongst value chain actors are the main targets of the programme in the process of smallholder agriculture transformation (ATA 2017; Elouhichi *et al*. 2019). Wheat is one of the programme's target commodities, and the Amhara regional state is one of the target areas. To facilitate the diffusion of research outputs, make modern technology affordable and improve bargaining power and market linkage to smallholder producers, the farmer production cluster (FPC) was introduced as a sub-set of the ACC in 2019. To achieve economies of scale, 30 to 200 farmers are clustered together on adjacent land to farm as one (ATA 2019). The Ethiopian government has set an ambitious goal to achieve self-sufficiency in wheat production and improve its commercialisation. To achieve this, cluster farming in wheat production is recognised as the best approach.

Although cluster farming was introduced as a government programme to enhance smallholder producers' livelihoods through the improved adoption of technology, productivity and commercialisation, studies on its effect on wheat output and input commercialisation remain limited. Since its introduction as an alternative farming approach in Ethiopia, research by some scholars indicates a positive association between cluster farming and output commercialisation. For instance, cluster-based teff production significantly enhances teff commercialisation among smallholder producers (Endalew *et al*. 2024). Similarly, farming high-acreage crops (teff, wheat, maize, barley and sesame) in clusters positively affects commercialisation (Guyo *et al*. 2023). However, no studies have examined the impact of cluster farming on the commercialisation of wheat inputs.

This study addresses the questions: What are the key determinants of smallholder producers' participation in wheat cluster farming? Does the newly introduced cluster farming approach significantly improve wheat output and input commercialisation levels? In so doing, this study contributes to the literature on the impact of cluster farming approach on wheat output and input commercialisation. In addition, it provides valuable insights for policymakers and development practitioners, focused on wheat, a cereal crop prioritised by the government, and the scaling up of the wheat cluster farming approach.

2. Research methodology

2.1 Study area

The study was conducted in Amhara regional state, Ethiopia, particularly in the West and East Gojjam administrative zones. The region is the second-largest and most populous national regional state under the federal democratic republic of Ethiopia. It is one of the major wheat-producing regions, accounting for 33.79% and 31.41% of the national wheat area coverage and production, respectively (CSA 2022). Moreover, the region is one of the selected pilot areas for the implementation of cluster farming, and currently the full scale-up of cluster farming has started (Ferede *et al*. 2020). West Gojjam Zone is one of the administrative zones of the region and has a total population of 2.107 million. The overwhelming majority (91%) of the population lives in the rural areas, and their sources of livelihood are primarily agriculture (CSA 2007). Its current administrative centre is Finote Selam, which is in the northwest of Ethiopia – 387 km from Addis Ababa and 176 km from Barhir Dar. The zone is known for the production of cereal and pulse crops. It is among the major wheat-producing zones in Amhara region and accounts for 8.22% and 8.88% of the regional wheat area coverage and production respectively (CSA 2022). Similarly, East Gojam Zone is another administrative zones in the region, with Debre Markos is its administrative centre. The total population in the zone is 2.154 million, of which 90% live in rural areas and whose sources of livelihood depend primarily on agriculture (CSA 2007). The zone has a mainly unimodal type of rainfall pattern, where the average annual rainfall varies from 900 to 1 800 mm, and its average temperature ranges from 7.5°C to 27°C (Ferede *et al*. 2020). It is one of the major producers of grain crops in the region. In particular, the zone contributes the largest share in terms of area coverage, as well as production of wheat crop, at 24.16% and 27.59%, respectively (CSA 2022).

2.2 Sampling methods and sample size determination

To select representative sample respondents for interviews, this study employed a multi-stage sampling procedure. First, two zones (West Gojjam and East Gojjam) were selected in consultation with regional Bureau of Agriculture experts and by referring to the ATA farmers' cluster electronic data source. The selection criteria include prevalence of wheat production, area coverage, number of established clusters, and accessibility (in terms of transport and security). Second, in collaboration with the selected zonal and woreda agricultural office experts and by referring to the ATA electronic data source, a list of potential woredas and kebeles was identified. Third, a total of five woredas were randomly selected: three from East Gojjam (Debre Elias, Baso Liben and Gozamin) and two from West Gojjam (Dembecha and Wenberma). In addition, a total of ten kebeles (two per woreda) were randomly selected. Fourth, a list of smallholder wheat-producing households, including their cluster farming participation status, was compiled in consultation with kebele development agents and administrative experts. Finally, sample respondents were randomly selected based on their proportion to the total farm households in the selected kebeles. The lottery method was used for woreda and kebele selection to ensure randomisation. It is the simplest and most commonly used randomisation technique particularly, for small samples. In this study, the list of potential woredas and kebeles, identified in collaboration with zonal and woreda experts, were given identification numbers. The numbers were then written on equally sized pieces of paper, folded, placed on a cup, and the required sample woredas and kebeles were drawn. A table of random numbers was used to select sample farm households. The final sample size was determined using the sample size formula developed by Kothari (2004), as specified below:

$$n = \frac{z^2pqN}{e^2(N-1)+z^2pq}, \tag{1}$$

where p represents the proportion of the target population, z represents the value obtained from the area under the standard normal curve, q is $1-p$, e is the acceptable margin error, N is the total number of households in the study areas (woredas), and n is the required sample size. A challenge in applying this formula is obtaining the value of p, and the authors propose two methods to get the value. The first method is to take an initial estimate of p based on either personal judgement or a pilot survey result. The second is setting $p = 0.5$ to get the most conservative sample size, ensuring at least the desired precision. Given time and resource constraints, the second method ($p = 0.5$) was adopted. With a 5% acceptable margin of error ($e = 0.05$), a 95% confidence level ($z = 1.96$), and a total population of 119 503 farm households ($N = 119\,503$), the required sample size (n) could be calculated as follows:

$$n = \frac{z^2pqN}{e^2(N-1)+z^2pq} = \frac{1.96^2x0.5x(1-0.5)x119{,}503}{0.05^2x(119{,}503\ -1)+(1.96^2x0.5x(1-0.5)} = 383$$

The total sample size for interviews was 383; 49.87% (191 farm households) were participants in wheat cluster farming, and the remaining 50.13% (192 farm households) were non-participants. The details of the sampling process are presented as in Table 1:

Table 1: Details of sample selection procedure

Zones	Selected weredas	Selected kebeles	Total HH number	Proportion	HH number selected
West Gojjam	Dembecha	Enewend	1 534	0.095	36
		Wad	1 362	0.084	32
	Wenberma	Markuma	1 053	0.065	25
		Waz	1 371	0.085	32
East Gojjam	Baso Liben	Limich	1 782	0.110	42
		Dogem	1 675	0.104	40
	Debre Elias	Guay	2 010	0.124	48
		Yekegat	1 222	0.076	29
	Gozamin	Chertekel	2 353	0.146	56
		Libanos	1 807	0.112	43
Total	5 (119 503 households)	10	16 169		383

2.3 Types of data and methods of data collection

The study employed both primary and secondary data. Discussions with relevant experts and reviews of annual reports, policy documents and research papers in relation to agriculture, cluster farming and commercialisation were used to compile the secondary data. On the other hand, primary data were collected in the 2023 production year, mainly through a formal survey using a structured questionnaire and face-to-face interviews with selected wheat cluster farming participant and non-participant farm household heads. The primary data includes farm household characteristics (such as demographic and socioeconomic), village characteristics (such as institutional and infrastructural), wheat production data, and participation decisions relating to cluster farming by the farm household. Well-trained enumerators and supervisors were employed to ensure effective and quality data collection. In addition, data were also collected through focus group discussions and discussions with the respective experts to support the survey data.

2.4 Empirical framework

An evaluation of each given programme, approach and technology was done to assess results and help improve outcomes intended for the intervention (Olmos & Govindasamy 2015). However, the application of quantitative techniques in evaluating the impact of an intervention based on nonexperimental cross-sectional data remains a challenging issue. The challenge is mainly linked to the difficulty of obtaining proper counterfactuals against which the impact is intended to be measured due to self-selection problems (Kassie *et al*. 2011; Shiferaw *et al*. 2014). To effectively measure the influence of an intervention (participation in wheat cluster farming) on an outcome variable (wheat output and input commercialisation), farm households' participation in wheat cluster farming must be random so that the groups' (participants and non-participants) observable and unobservable characteristics have the same effect, and are attributable to the treatment. However, when groups' random assignment to treatments fails, their participation decision is affected by both observable and unobservable heterogeneity, such as skills and motivation (Shiferaw *et al*. 2014).

Different econometric approaches can be employed to deal with selection bias in impact studies using cross-sectional data. Propensity score matching (PSM), which was pioneered by Rubin (1973), is one of the approaches that is used when the problem of selection bias in observable variables is an issue. The instrumental variable (IV) approach is the other well-known approach that can help to deal with unobserved heterogeneity (Shiferaw *et al*. 2014; Wooldridge 2002). However, this approach assumes that the impact would differ only by a constant term between the treated and nontreated groups, although, in the real sense, the difference is more systematic due to the potential interaction between the treatment and outcome determinants (Jaleta *et al*. 2016).

Since PSM does not require linearity, or parametric or distributional assumptions (Heckman & Vytlacil 2007), several scholars have used propensity score matching to deal with structural differences (basically based on observed variables) (Becerril & Abdulai 2010; Kassie *et al*. 2011). However, it is not a consistent estimator when there are unobservable variables that affect the treatment and outcome of interest simultaneously, and it does not take into account the effect of covariates on outcome variables (Jaleta *et al*. 2016). The endogenous switching regression (ESR) model approach, developed by Lee (1982) as a general model of the Heckman's selection correction model (Heckman, 1979), relaxes the assumptions imposed by the PSM and IV approaches (Maddala 1983; Wooldridge 2002; Lokshin & Sajaia, 2004). Therefore, there is a need to select a method that takes into account observable and unobservable endogeneity by simultaneously estimating the participation and the outcome equations for each group, and the effect of covariates on the outcome variable; that is the ESR approach. The model has two main parts, namely the selection and the outcome equations (for participant and non-participant farm households).

2.5 Methods of data analysis

Descriptive statistics, as well as an econometric model, were used to analyse the data. The descriptive statistics included averages, percentages and standard deviations. In addition, a t-test was employed to examine mean differences between groups, while a chi-square test was used to examine frequency differences. In addition to the ESR model, the PSM model was used to check the robustness of the estimated results obtained from the ESR model.

2.5.1 Specification of the ESR model

The econometric model used in this study is derived from utility theory, a measure of relative human satisfaction. Consequently, smallholder farm households decided to participate in cluster farming based on the expected returns from their participation (Adeoti *et al*. 2014). These decisions aim to maximise utility (U), subject to a set of constraints (Otekunrin *et al*. 2019). Thus, smallholder wheat producers are expected to participate in cluster farming (CF) when the expected utility received from participation (U_p) exceeds that from non-participation (U_{np}): $U_p - U_{np} > 0$. Based on this theoretical framework, the mathematical and estimation procedures of the ESR model are specified following Shiferaw *et al*. (2014), Di Falco *et al*. (2011) and Jaleta *et al*. (2016), as follows:

Let A_i^* be the latent variable that captures the benefit from participating in the CF approach by the i^{th} farm household, given as:

$$A_i^* = Z_i\beta + \varepsilon_i \ where \ A_i = \begin{cases} 1, & Z_i\beta + \varepsilon_i > 0 \\ 0, & otherwise \end{cases}, \tag{2}$$

where A_i^* is a binary variable equal to 1 if a wheat-producing farm household participates in the CF approach, and 0 otherwise; Z_i is a vector of sociodemographic factors (age, sex, education, household size, involvement in off-farm activity), institutional factors (access to credit services, extension services, cooperative membership and cluster farming awareness), resource ownership (wheat farm size and livestock) and village-level (distance to rural all-weather roads, distance from main market, land fragmentation and farm proximity to nearby wheat cluster) variables that affect the decision to participate in CF approach or not; β is a vector of parameters to be estimated; and ε_i is an error term normally and independently distributed, with mean 0 and variance σ_ε^2. The variables 'cluster farming awareness' and 'farm distance to the nearby wheat cluster farms' were included as model identification or instrumental variables in the selection equation, but not in the outcome equation. Equation 2 is labelled as participation or selection equation. Since farm households make their cluster farming participation decisions based on the expected utility or benefit of the outcome variable (in this case, wheat output and input commercialisation), the equation for the outcome variable of interest (Y_i) is expressed as:

$$Y_{1i} = X_i\alpha_1 + \mu_{1i}, if \;\; A_i = 1 \tag{3}$$

$$Y_{2i} = X_i\alpha_2 + \mu_{2i}, if \;\; A_i = 0 \tag{4}$$

Estimating the outcome equations (3 and 4) to see if there exists a correlation between the error terms (μ_{1i} and μ_{2i}) and the participation or selection equation (ε_i) (Equation (2)) leads to biased estimates due to the existence of the endogeneity problem. Therefore, to take into account the possible endogeneity problem, the model is specified as:

$$Regime1{:}\; Y_{1i} = X_{1i}\alpha_1 + \sigma_{1\varepsilon}\hat{\lambda}_{1i} + \eta_{1i}, if \;\; A_i = 1, \tag{4}$$

$$Regime1{:}\;\; Y_{2i} = X_{2i}\alpha_2 + \sigma_{2\varepsilon}\hat{\lambda}_{2i} + \eta_{1i}, if \;\; A_i = 0, \tag{5}$$

where Y_{1i} is a dependent variable representing wheat output and the input commercialisation level for CF participants, and Y_{2i} the same for non-participants, X_i is a vector of variables that affect the outcome variable (wheat output and input commercialisation level), α and σ are parameters to be estimated, and $\hat{\lambda}_{1i} = \frac{\phi(Z_i\hat{\alpha}_i)}{\Phi(Z_i\hat{\alpha}_i)}$ and $\hat{\lambda}_{2i} = \frac{\phi(Z_i\hat{\alpha}_i)}{1-\Phi(Z_i\hat{\alpha}_i)}$ are the inverse Mill's ratios (IMRs) computed from the selection Equation (2) to correct for selection bias in the second-stage estimation (outcome equations). η is an independently and identically distributed error term with mean zero and constant variance.

The error terms of the selection equation (ε_i), and the two output equations (μ_{1i} and μ_{2i}), are assumed to have a trivariate normal distribution with mean vector zero and the following covariance matrix:

$$cov(\varepsilon_i, \mu_{1i}, \mu_{2i}) = \begin{bmatrix} \sigma_{\mu_{2i}}^2 & \sigma_{\mu_{1i}\mu_{2i}} & \sigma_{\mu_{1i}\varepsilon_i} \\ \sigma_{\mu_{1i}\mu_{2i}} & \sigma_{\mu_{1i}}^2 & \sigma_{\mu_{1i}\varepsilon_i} \\ \sigma_{\mu_{2i}\varepsilon_i} & \sigma_{\mu_{1i}\varepsilon_i} & \sigma_{\varepsilon_i}^2 \end{bmatrix}, \tag{6}$$

where $\sigma_{\mu_{2i}}^2$ and $\sigma_{\mu_{1i}}^2$are the variances of the outcome equations for non-participant and participant farm households, while $\sigma_{\mu_{2i}\varepsilon_i}$ and $\sigma_{\mu_{1i}\varepsilon_i}$ represent the covariance between ε_i , μ_{2i} and ε_i, μ_{1i} respectively. If the selection equation error term (ε_i) is correlated with the participant and non-participant error terms (μ_{1i}, and μ_{2i} respectively), the expected values of μ_{1i}, and μ_{2i} conditional on the sample selection are non-zero. That is:

$$E(\mu_{1i}|A_i = 1) = \sigma_{\mu_{1i}\varepsilon_i}\frac{\phi(Z_i\hat{\alpha}_i)}{\Phi(Z_i\hat{\alpha}_i)} = \sigma_{\mu_{1i}\varepsilon_i}\hat{\lambda}_{1i}, \tag{7}$$

$$E(\mu_{2i}|A_i = 0) = \sigma_{\mu_{2i}\varepsilon_i}\frac{-\phi(Z_i\widehat{\alpha}_i)}{1-\Phi(Z_i\widehat{\alpha}_i)} = \sigma_{\mu_{2i}\varepsilon_i}\hat{\lambda}_{2i}, \quad (8)$$

where ϕ and Φ are the probability density and the cumulative distribution function of the standard normal distribution, respectively. Hence, if $\sigma_{\mu_{1i}\varepsilon_i}$ and $\sigma_{\mu_{2i}\varepsilon_i}$ are statistically significant, this would be evidence that farm households' decision to participate in wheat cluster farming and their wheat output commercialisation level are correlated, suggesting evidence of sample selection bias. This implies that the use of ordinary least squares (OLS) to estimate the outcome equation would lead to biased and inconsistent results. Therefore, in such a scenario, the full information maximum likelihood (FILM) estimator can be used to fit an endogenous switching regression that simultaneously estimates the selection and outcome equations to generate consistent estimates.

Hence, the actual expected outcomes of the cluster farming (CF) participant (Equation 10) and non-participant (Equation 11) farm households, and the counterfactual cases that the participants had not participated (Equation 12) and the non-participants had they participated (Equation 13), can be compared using ESR and are specified as:

$$E[Y_{1i}|X_{1i}, A_i = 1] = X_{1i}\alpha_1 + \sigma_{1\varepsilon}\hat{\lambda}_{1i} \quad (9)$$

$$E[Y_{2i}|X_{2i}, A_i = 0] = X_{2i}\alpha_2 + \sigma_{2\varepsilon}\hat{\lambda}_{2i} \quad (10)$$

$$E[Y_{2i}|X_{1i}, A_i = 1] = X_{1i}\alpha_2 + \sigma_{2\varepsilon}\hat{\lambda}_{1i} \quad (11)$$

$$E[Y_{1i}|X_{2i}, A_i = 0] = X_{2i}\alpha_1 + \sigma_{1\varepsilon}\hat{\lambda}_{2i} \quad (12)$$

Among the above equations, Equations (10) and (11) are observed from the survey data, but Equations (12) and (13) are the counterfactual outcomes. On this basis, following Heckman *et al*. (2001) and Di Falco *et al*. (2011), the average treatment effect on the treated (ATT), the average treatment effect on the untreated (ATU), the base heterogeneity effect for participant (BH_1) and non-participant (BH_2) farm households, and the transitional heterogeneity (TH) effect were computed as follows (see also Table 2):

$$ATT = (10) - (12) = \left(X_{1i}\alpha_1 + \sigma_{1\varepsilon}\hat{\lambda}_{1i}\right) - \left(X_{1i}\alpha_2 + \sigma_{2\varepsilon}\hat{\lambda}_{1i}\right) = X_{1i}(\alpha_1 - \alpha_2) + \hat{\lambda}_{1i}(\sigma_{1\varepsilon} - \sigma_{2\varepsilon}) \quad (13)$$

$$ATU = (13) - (11) = \left(X_{2i}\alpha_1 + \sigma_{1\varepsilon}\hat{\lambda}_{2i}\right) - \left(X_{2i}\alpha_2 + \sigma_{2\varepsilon}\hat{\lambda}_{2i}\right) = X_{2i}(\alpha_1 - \alpha_2) + \hat{\lambda}_{2i}(\sigma_{1\varepsilon} - \sigma_{2\varepsilon}) \quad (14)$$

$$BH_1 = (10) - (13) = \left(X_{1i}\alpha_1 + \sigma_{1\varepsilon}\hat{\lambda}_{1i}\right) - \left(X_{2i}\alpha_1 + \sigma_{1\varepsilon}\hat{\lambda}_{2i}\right) = \alpha_1(X_{1i} - X_{2i}) + \sigma_{1\varepsilon}(\hat{\lambda}_{1i} - \hat{\lambda}_{2i}) \quad (15)$$

$$BH_2 = (12) - (11) = (X_{1i}\alpha_2 + \sigma_{2\varepsilon}\hat{\lambda}_{1i}) - (X_{2i}\alpha_2 + \sigma_{2\varepsilon}\hat{\lambda}_{2i}) = \alpha_2(X_{1i} - X_{2i}) + \sigma_{2\varepsilon}(\hat{\lambda}_{1i} - \hat{\lambda}_{2i}) \quad (16)$$

$$TH = ATT - ATU \quad (17)$$

Table 2: Summary of conditional expectations, treatment and heterogeneity effects

Subsample	Participation decision		Participation effect
	To participate	Not to participate	
Wheat cluster farming participant	$(a)\ E[Y_{1i} \mid X_{1i}, A_i = 1]$	$(c)\ E[Y_{2i} \mid X_{1i}, A_i = 1]$	ATT= $(a) - (c)$
Wheat cluster farming non-participant	$(d)\ E[Y_{1i} \mid X_{2i}, A_i = 0]$	$(b)\ E[Y_{2i} \mid X_{2i}, A_i = 0]$	ATU= $(d) - (b)$
Heterogeneity effects	$BH_1 = (a) - (d)$	$BH_2 = (c) - (b)$	TH = ATT – ATU

2.5.2 Specification of the PSM model

On the other hand, following Caliendo and Kopeinig (2005), the PSM model was specified as:

$$\tau_i = (Y_i \mid D_i = 1) - (Y_i \mid D_i = 0), \tag{18}$$

where τ_i is the treatment effect (effect due to participation in wheat cluster farming), Y_i is the outcome variable (wheat output and input commercialisation level), D_i is a binary variable equal to 1 if a wheat-producing farm household participates in the CF approach, and 0 otherwise. The most frequently used treatment effects in various empirical studies are the population average treatment effect (ATE) and average treatment effect on the treated (ATT). However, ATT is applied for most impact studies, as it targets only the effects on those for whom the treatment is intended (Caliendo & Kopeinig 2005). Therefore, ATT was estimated in this study, and its mathematical expression is:

$$\tau_{ATT} = E(\tau \mid D_i = 1) = E(Y_1 \mid D_i = 1) - E(Y_0 \mid D_i = 1) \tag{19}$$

However, it is difficult to get data on the counterfactuals, $E(Y_0 \mid D_i = 1)$, because they are unobserved. This leads to difficulty in estimating ATT. Due to this problem, one must choose a proper substitute for it to estimate ATT, and the possible solution is to use the mean outcome of untreated individuals, $E(Y_0 \mid D_i = 0)$, given that the conditional independence and common support or overlap assumptions are fulfilled. Hence, using the logit model and covariates used in the ESR model above, the propensity scores (P) were predicted and the ATT is estimated as:

$$\tau_{ATT} = \{[E(Y_0 \mid D_i = 1, P_i(X)] - [E(Y_0 \mid D_i = 0, P_i(X)]\}\ ) \tag{20}$$

Once the propensity score has been obtained, it is necessary to identify the best matching algorithm that performs well in matching the treated with the untreated observations using pseoudo-R^2 and mean standardised bias performance criteria. The most frequently used matching algorithms in PSM are nearest neighbour (NN), radius calliper and kernel matching (Caliendo & Kopeinig 2005). The best matching estimator is the one that results in better low R^2 and mean standardised bias values, among others (Dehejia & Wahba 2002).

2.6 Variables used in the ESR and PSM models

The dependent, outcome and explanatory variables included in the ESR and PSM models are the same. This is because the PSM model is estimated just to check the robustness of the results.

Dependent variables: The farm household participation decision is the dependent variable used in the selection equation in the ESR and the logit model in the PSM models. It is measured as a dummy variable by being given a value of one if the farm household is a participant in wheat cluster farming, and zero otherwise. A farm household is considered as a participant in wheat cluster farming if it has a plot of land adjacent to other participants, has agreed to perform agricultural operations in a coordinated manner, and is willing and able to apply the recommended farm inputs and practices. On the other hand, a farm household is considered as non-participant in wheat cluster farming if it is producing wheat independently as usual, and uses farm inputs and practices based on personal experience, resource availability and extension officer advice.

Outcome variables: Since the main interest of this paper is to evaluate the impact of smallholder producers' wheat cluster farming participation on wheat output and input commercialisation, the wheat output and input commercialisation levels measured as an index are the outcome variables and are represented as WOCI and WICI.

Explanatory variables: A review of related studies indicates that sociodemographic characteristics (e.g., age, sex, education, household size, involvement in off-farm activity) play a significant role in smallholder producers' participation in cluster farming (Endalew *et al*. 2024; Degefu *et al*. 2024; Gidelew *et al*. 2025). Similarly, institutional factors, such as access to extension services, credit services, cooperative membership and awareness training significantly influence the participation of smallholder producers in cluster farming (Degefu *et al*. 2024; Endalew *et al*. 2024; Degefu *et al*. 2025; Gidelew *et al*. 2025). Likewise, variables related to resource ownership (e.g., wheat land size and livestock) and village-level variables (e.g. all-weather rural roads, distance to market, proximity of farm to nearby wheat cluster and land fragmentation) also play a critical role (Guyo & Tabe-Ojong 2024; Gidelew *et al*. 2025). Therefore, these are the explanatory variables that were included in the regression models to identify the determinants of participation in cluster farming and to evaluate the impact of participation on wheat output and input commercialisation. Similarly, except for the instrumental variables (cluster farming awareness and farm proximity to nearby wheat cluster), all the variables included in the selection equation are used in the outcome equation to estimate the impact of participation in wheat cluster farming on the commercialisation level of wheat output and input.

The variables sex of the head of the farm household, education status, credit access, cooperative membership, involvement in off-farm activity and awareness of cluster farming were measured as dummy variables. In this study, a farm household head was considered literate if he or she had attended formal education and was able to read and write; otherwise they were considered as illiterate. Similarly, the variable cluster farming awareness was included in the participation model, because it is one of the most important factors influencing a farm household's participation in wheat cluster farming. Before the introduction of a new or improved farming approach such as cluster farming, it is primarily the office of agriculture that creates awareness among farm households on its importance, establishment and implementation. Hence, those who have information about cluster farming (either through participating in the training or by other means) are considered as being aware of cluster farming, while those who do not have such information are considered as not being aware. On the other hand, household size, extension services, distance to rural all-weather roads, distance from the main market, wheat farm size, land fragmentation, livestock size and farm proximity to nearby wheat clusters are measured as continuous variables. Land fragmentation represents farm households' degree of land fragmentation and is measured in index form. Its value is computed by using the Simpson index (SI), as follows:

$$SI_i = 1 - \frac{\sum_{j=1}^{J} a_{ij}^2}{A_i^2}, 0 \leq SI \leq 1,$$

where a_{ij} represents the area of the j^{th} plot and A_i is the total area of annual crop land operated by a farm household. A zero value for SI indicates complete land consolidation (one parcel only), while values closer to one indicates numerous parcels and that the farm is fragmented.

2.7 Measurement of output and input commercialisation

To evaluate the influence of participation in cluster farming on wheat output and input commercialisation among smallholder producers, the wheat output commercialisation index (WOCI) and input commercialisation index (WICI) – the dependent variables in the outcome equations – need to be computed. Scholars have conceptualised smallholders' commercialisation in various ways. However, it can generally be viewed from three perspectives: input and output, sales and purchases, and type of crops grown (Alemu *et al*. 2006; Boka 2017). The output side of smallholder commercialisation is the ratio of the value of agricultural outputs sold to the total value of agricultural outputs produced by a farm household. Similarly, the input side of smallholder commercialisation is the ratio of the value of agricultural inputs purchased to the total value of agricultural outputs produced by a farm household (Von Braun 1995; Strasberg *et al*. 1999). Smallholder commercialisation, based on the type of crops grown, is primarily measured by the production of cash crops, and is calculated as the ratio of the gross value of cash crops produced to the total gross value of production (Govereh *et al*. 1999). Smallholder producers may also participate in the purchase of agricultural produce in addition to selling. Their purchases can involve their own produce (such as storage, depressed sale, diversification) or market purchases, which may lead to specialisation (Alemu *et al*. 2006). In developing countries like Ethiopia, smallholder producers are characterised by diversified farming and the use of both purchased and own farm inputs. It is well known that, due to the diverse dietary uses of wheat crops and increases in urbanisation, smallholder wheat producers supply a significant portion of their produce to the market, indicating that staple crops can also be commercialised.

Therefore, considering commercialisation, which encompasses both output and input sides, the household level of WOCI and WICI was computed as follows, following Von Braun (1995) and Strasberg *et al*. (1999).

$$WOCI_{HH_i} = \frac{S_i}{P_i}\, x\, 100, \tag{21}$$

where, HH_i is household $i, (i = 1,2, \ldots n)$, $WOCI_{HH_i}$ represents the wheat output commercialisation index of the i^{th} household, S_i is the total value of wheat output obtained from sales in the market by the i^{th} household, and P_i is the gross value of wheat output produced by the i^{th} farm household.

Following the commercialisation framework used for the WOCI, the wheat input commercialisation index (WICI) measures the extent of market-purchased inputs for wheat production by smallholder farmers. As inorganic fertilisers, seeds and pesticides are the primary market-purchased inputs, the WICI is computed as follows:

$$WICI_{HH_i} = \frac{\sum_{m=1}^{M} IP_{im}}{P_i}\, x\, 100, \tag{23}$$

where HH_i is household $i, (i = 1,2, \ldots N)$, $WICI_{HH_i}$ represents the wheat input commercialisation index of the i^{th} household, IP_{im} is the value of inputs $m, (m = 1,2, \ldots M)$ purchased from the market for wheat production (inorganic fertiliser, seed and pesticide) by the i^{th} household, and P_i is the gross value of wheat output produced by the i^{th} farm household.

3. Results and discussion

3.1 Results

3.1.1 Descriptive summary of variables

This section provides an overview of the wheat yield, input use and other explanatory variables used in the econometric model for impact estimation. As presented in Table 3, smallholder producers participating in cluster farming have higher size of land than non-participants. The average wheat yield in the study area was 2 891.22 kg per hectare, with cluster farming participants achieving significantly higher yields than non-participants, as confirmed by t-tests. Similarly, the average wheat area was 0.56 ha, with significant differences between participants and non-participants at the 1% level of significance. Key inputs in wheat production include DAP fertiliser, urea, seeds, labour and oxen power. The t-test result indicates significant differences in DAP and urea fertiliser application for wheat production between participant and non-participant farm households at the 1% level of significance. Labor usage averages 41.90 man-days for participants and 47.19 for non-participants, with significant differences at the 1% level of significance.

Table 3: Summary of wheat yield and production input use by treatment

Variables	Full sample (N = 383)		Participant (n1 = 191)		Non-participant (n2 = 192)		Mean difference	t-value
	Mean	SD	Mean	SD	Mean	SD		
Total land size (ha)	1.09	0.32	1.16	0.30	1.02	0.33	0.14	4.32***
Wheat yield (kg)	2 891.22	486.41	3 078.70	402.13	2704.72	492.16	373.98	8.14***
Wheat area (ha)	0.56	0.21	0.64	0.20	0.48	0.18	0.16	8.02***
DAP fertiliser (kg)	114.28	31.60	118.91	31.66	109.68	30.95	9.24	2.89***
Urea fertiliser (kg)	124.01	39.85	136.03	39.25	112.06	36.83	23.96	6.16***
Seed rate (kg)	103.73	17.14	102.46	17.17	105.00	17.05	-2.54	1.46
Labour (man-day)	44.55	17.80	41.90	14.92	47.19	19.95	-5.29	2.94***
Use of oxen power (oxen-day)	22.40	4.26	22.50	4.05	22.29	4.47	0.21	0.49

Note: *** represents statistically significant at the 1% level of significance.
Source: Own computation from field survey data (2023)

The Ethiopian government has prioritised cluster farming in the wheat sector to enhance commercialisation. As shown in Table 4, participant smallholder farmers exhibit higher than average levels of wheat output and input commercialisation than non-participants, as supported by a t-test. Cluster participants also have a higher frequency of extension contact, larger wheat land size, and their wheat farm is close to the nearby wheat cluster farms, all statistically significant at the 1% level. In contrast, non-participants are located significantly further from rural all-weather roads and main markets, as confirmed by the t-test at the 1% level of significance. Furthermore, farm households owned multiple parcels of land, as indicated by the Simpson index (SI) of land fragmentation, where values close to zero reflect greater consolidation and one indicates higher fragmentation. In terms of livestock ownership, measured in tropical livestock units (TLU), this averaged 4.97 for participants and 4.34 for non-participants, with a significant mean difference at the 1% level of significance.

Table 4: Summary of descriptive statistics for outcome and continuous variables by treatment

Variables	Full sample (N = 383)		Participant (n1 = 191)		Non-participant (n2 = 192)		Mean difference	t-value
	Mean	SD	Mean	SD	Mean	SD		
Outcome variable								
Wheat output commercialisation (index)	0.49	0.15	0.55	0.12	0.43	0.16	0.12	8.40***
Wheat input commercialisation (index)	0.11	0.03	0.12	0.03	0.10	0.03	0.02	5.64***
Continuous explanatory variables								
Age (years)	43.97	0.15	44.32	7.53	43.62	7.63	0.70	0.90
Household size (AEU)	3.82	1.37	3.92	1.26	3.72	1.47	0.20	1.42
Extension contacts (frequency)	5.04	2.37	6.06	2.19	4.03	2.09	2.03	9.29***
Wheat land size (ha)	0.56	0.21	0.64	0.20	0.48	0.18	0.16	8.02***
Distance to rural all-weather roads (km)	1.89	1.52	1.47	1.07	2.31	1.76	-0.84	5.63***
Land fragmentation (index)	0.57	0.38	0.54	0.52	0.60	0.12	-0.07	1.68*
Livestock size (TLU)	4.66	2.21	4.97	2.27	4.34	2.11	0.63	2.81***
Farm distance to the nearby wheat cluster farms (km)	1.32	1.12	0.68	0.57	1.97	1.16	-1.29	13.78***
Distance from main market (km)	9.79	5.83	8.88	5.69	10.69	5.84	-1.81	3.08***

Note: * and *** represents statistical significance at the 10 and 1% level respectively
Source: Own computation from field survey data (2023)

Table 5 summarises the descriptive results of the dummy explanatory variables used in the econometric model evaluating the impact of wheat cluster farming participation on wheat output and input commercialisation among smallholder farmers. Of the smallholder farmers sampled, 94.8% were male-headed households, compared to 5.2% female-headed ones. Among the literate smallholder farmers, 55.6% were cluster farming participants, compared to 45.7% of illiterate ones, with the chi-square test confirming significant differences, highlighting the role of education in participation. Access to formal credit services was limited, with only 32.1% of smallholder farmers having access, and with significant differences in participation based on access to formal credit, as supported by the chi-square test at the 5% level of significance. In terms of farmers' cooperatives, a notable 88.3% of households were members. This plays a crucial role in accessing agricultural inputs. Awareness of cluster farming was also high, with 85.9% of households being aware of it. There also was a significant difference in participation among the aware and non-aware farm households. Finally, only 25.1% of smallholder farmers engaged in off-farm activities, while the majority (74.9%) did not participate in such activities.

3.1.2 Results of endogenous switching regression model

The endogenous switching regression (ESR) model was used to identify the cluster farming participation determinants and evaluate the impact of cluster farming participation on wheat output and input commercialisation among smallholder farmers. Thus, before interpreting the model outputs, diagnostic tests were conducted, including tests for heteroscedasticity, normality of residuals, instrument validity, correlation between participation and outcome equations, and overall model significance.

Table 5: Summary of descriptive statistics for dummy explanatory variables by treatment

Variables		Full sample (N = 383)		Cluster farming				χ2-test
				Participant (n1 = 191)		Non-participant (n2 = 192)		
		Count	%	Count	%	Count	%	
Sex	Male	363	94.8	182	50.1	181	49.9	0.20
	Female	20	5.2	9	45.0	11	55.0	
Education status	Literate	162	42.3	90	55.6	72	44.4	3.63*
	Illiterate	221	57.7	101	45.7	110	54.3	
Access to formal credit	Yes	123	32.1	71	57.7	52	42.3	4.47**
	No	260	67.9	120	46.2	140	53.8	
Cooperative membership	Yes	338	88.3	167	49.4	171	50.6	0.25
	No	45	11.7	24	53.3	21	46.7	
Cluster farming awareness	Yes	329	85.9	174	52.9	155	47.1	8.5***
	No	54	12.5	17	68.5	37	31.5	
Involvement in off-farm activity	Yes	96	25.1	45	46.9	51	53.1	0.46
	No	287	74.9	146	50.9	141	49.1	

Note: *, ** and *** represent statistical significance at 10%, 5% and 1% level respectively
Source: Own computation from field survey data (2023)

As presented in the Appendix (Tables 1 and 2), the Breusch and Pagan (1979) heteroscedasticity test indicates that the variance of the residuals for the WOCI equation is constant for both participants and non-participants (χ^2 = 2.01; p = 0.156 and χ^2 = 0.05; p = 0.823, respectively). Similarly, the WICI equation residuals exhibit constant variance for both groups (χ^2 = 0.57; p = 0.451 and χ^2 = 1.460; p = 0.227, respectively). Likewise, Tables 3 and 4 in the Appendix shows that the Jarque and Bera (1987) test confirms the normality of residuals in WOCI equations for both participants and non-participants (χ^2 = 3.252; p = 0.197 and χ^2 = 1.814; p = 0.404, respectively) and WICI equation (χ^2= 3.13; p= 0.209 and χ^2 = 2.42; p = 0.298, respectively).

The ESR model's participation and outcome equations were identified through the nonlinearity of the selection correction term (λ) (Lokshin & Glinskaya 2008), or by including at least one instrumental variable in the participation equation but not in the outcome equation (Jaleta *et al*. 2016). Thus, this study uses smallholder farmers' cluster farming awareness and farm proximity to wheat cluster farms as an exclusion or instrumental variable. Farm household heads' awareness of the benefits of cluster farming and those with farms closer to wheat cluster farms were expected to be more likely to participate than their counterparts. These variables are key in the process of cluster establishment and participant selection for wheat cluster farming, and are expected to have a direct effect on participation, but not on WOCI or WICI. For instance, Shiferaw *et al*. (2014) used distance to seed market and sources of a variety of information as instrumental variables in their study of the adoption of improved wheat varieties and impacts on household food security. To examine the validity of the selected instruments in estimating the impact of participation in wheat cluster farming on output and input commercialisation, a falsification test was conducted following the statistical procedure of Di Falco *et al*. (2011). The results of the test confirm that the included instrumental variables are jointly statistically significant at the 1% level in the participation decision (selection equation: χ^2 = 75.25; p = 0.00) but not in the non-participant wheat output and input commercialisation equations (outcome equation: F= 0.71; p = 0.495 and F = 0.85; p = 0.429, respectively) (Appendix Table 5). In addition, as shown in Appendix Table 6 and 7, the likelihood ratio test of independence for wheat output and input commercialisation indicates that the participation and outcome equations are dependent (χ^2 = 5.42; p = 0.02 and χ^2 = 4.70; p = 0.03, respectively), leading to the rejection of the null hypothesis that the selection and outcome equations are independent. This result supports the use of the ESR model. Furthermore, the Wald chi-square test confirms the overall model is significant at the 1% level for both outcome equations (Wald χ^2 = 133.65; p = 0.00 and Wald χ^2 = 75.52; p = 0.00, respectively).

After implementing these statistical tests, the determinants of smallholder wheat producers' participation in cluster farming and the impact of the cluster farming approach on wheat commercialisation (output and input) was estimated. The selection equation result of the ESR model presented in Table 6 shows the participation determinants of smallholder producers in wheat cluster farming. The result indicates that, among the included explanatory variables, farm household education status, access to formal credit, frequency of extension contacts, cluster farming awareness, wheat farm size and number of livestock owned are the key variables that have a positive and significant influence. On the other hand, distance of the farm household's residence to rural all-weather roads, the land fragmentation index and farm proximity to a nearby cluster have a significant and negative influence on farm households' participation in wheat cluster farming.

Table 6: Determinants of farm households' participation in wheat cluster farming

Variables	Participation (participant = 1)	
	Coef.	Std. error
Age (years)	0.035	0.024
Sex (male = 1)	-0.449	0.473
Education status (literate = 1)	0.780***	0.211
Household size (AEU)	-0.220	0.143
Formal credit access (yes = 1)	0.348*	0.204
Extension contacts (frequency)	0.183***	0.055
Cooperative membership (yes = 1)	0.088	0.319
Distance to rural all-weather roads (km)	-0.209**	0.091
Distance from main market (km)	-0.026	0.020
Wheat farm size (ha)	1.200**	0.597
Land fragmentation (index)	-1.983**	0.867
Livestock size (TLU)	0.096**	0.048
Involvement in off-farm activity (yes = 1)	-0.147	0.230
Cluster farming awareness (aware = 1)	1.118**	0.457
Farm proximity to nearby wheat cluster (km)	-0.929***	0.118
Constant	-0.897	1.099
Number of observations	383	
Pseudo/R-squared	0.536	

Note: *, ** and *** represent statistical significance at the 10%, 5% and 1% level respectively
Source: Own computation from field survey data (2023)

In addition to this, the ESR model shows that smallholder producers' participation in wheat cluster farming has a positive and significant impact on wheat output and input commercialisation levels (Table 7).

As shown in Table 7, the wheat output commercialisation level is, on average, 5% higher for smallholder producers participating in cluster farming compared to their counterfactuals (i.e., if these smallholder producers had used a conventional farming approach), representing the average treatment effect on the treated (ATT). Similarly, the wheat output commercialisation level for smallholder producers using the conventional farming approach would, on average, have increased by 1.9% if they had produced in a cluster, which is termed the average treatment effect on the untreated (ATU). In addition, wheat production under the cluster farming approach yielded a positive transitional heterogeneity (TH) effect. Furthermore, the model results show that the wheat input commercialisation level was higher for farm households using the cluster farming approach compared to their counterfactuals (ATT). Likewise, on average, the wheat input commercialisation level for farm households producing wheat under conventional farming would have been significantly higher if they had produced wheat in a cluster (ATU). Moreover, wheat production under cluster farming resulted in a significant and positive transitional heterogeneity (TH) effect.

Table 7: Expected average wheat output commercialisation, treatment and heterogeneous effects

Outcome variables	Subsample	Participation decision		Participation effect
		To participate	Not to participate	
wheat output commercialisation index	wheat cluster farming participant	(a) 0.549 (0.005)	(c) 0.499 (0.008)	ATT = 0.050*** (0.005)
	wheat cluster farming non-participant	(d) 0.445 (0.005)	(b) 0.426 (0.007)	ATU = 0.019*** (0.004)
	heterogeneity effects	BH1 = 0.104 (0.007)	BH2 = 0.073 (0.011)	TH = 0.031*** (0.006)
wheat input commercialisation index	wheat cluster farming participant	(a) 0.119 (0.020)	(c) 0.088 (0.017)	ATT = 0.031*** (0.002)
	wheat cluster farming non-participant	(d) 0.103 (0.018)	(b) 0.099 (0.014)	ATU = 0.004** (0.002)
	heterogeneity effects	BH1 = 0.015 (0.002)	BH2 = -0.011 (0.001)	TH = 0.026*** (0.001)

Notes: *** represents statistically significance at the 1% level; numbers in parenthesis represent standard errors
Source: Own computation from field survey data (2023)

3.1.3 Analysis of propensity score matching model

The impact of smallholder producers' participation in cluster farming on wheat output and input commercialisation levels was estimated using the PSM model to check the robustness of the results. Before the final model estimation, the performance of various matching algorithms – namely nearest neighbour (NN) with no replacement and up to three replacements, radius calliper with 0.01 and 0.05 callipers, and kernel matching with 0.03, 0.06, and 0.1 bandwidths – was evaluated using pseudo-R^2 and mean standardised bias criteria (Table 8). As shown in Table 8, the test results indicate that nearest neighbour (NN) with one replacement, radius calliper with a 0.01 calliper, and kernel with a 0.03 bandwidth were the best-performing matching estimators, achieving low pseudo-R^2 and mean standardised bias after matching. However, kernel matching with a 0.03 bandwidth was selected as the optimal estimator, as it balanced nearly all covariates included in the model after matching (Appendix Table 9), while maintaining low pseudo-R^2 and mean standardised bias. The final PSM model results, including the logistic regression (Appendix Table 8) and ATT were estimated based on this estimator.

Table 8: Performance of different matching estimators

Matching estimator	Performance criteria					
	Pseudo-R^2				Mean standardised bias	
	Before matching	LR χ^2 (p-value)	After matching	LR χ^2 (p-value)	Before matching	After matching
NN	0.536	284.39 (0.00)	0.416	175.52 (0.00)	49.0	31.70
NN (1)	0.536	284.39 (0.00)	0.048	19.50 (0.108)	49.0	8.50
NN (2)	0.536	284.39 (0.00)	0.068	27.88 (0.015)	49.0	15.20
NN (3)	0.536	284.39 (0.00)	0.078	32.77 (0.005)	49.0	12.20
Radius calliper						
0.01	0.536	284.39 (0.00)	0.059	16.96 (0.201)	49.0	7.80
0.05	0.536	284.39 (0.00)	0.096	39.00 (0.00)	49.0	17.0
Kernel						
bandwidth 0.03	0.536	284.39 (0.00)	0.042	17.67 (0.171)	49.0	7.34
bandwidth 0.06	0.536	284.39 (0.00)	0.046	19.47 (0.010)	49.0	10.5
bandwidth 0.1	0.536	284.39 (0.00)	0.082	34.38 (0.003)	49.0	14.8

As reported in Table 9 below, farm household participation in cluster farming had a positive and significant effect on wheat output and input commercialisation under the nearest neighbour (NN) with single replacement, radius calliper with a 0.01 calliper, and kernel with 0.03 bandwidth matching estimators. The participation effect on wheat output and input commercialisation under these estimators ranged from 4.8% to 12.5% and 0.9% to 1.4%, respectively. The results from ESR and PSM confirm that farm household participation in cluster farming contributed significantly to improvements in wheat output and input commercialisation, and ensured the robustness of the data.

Table 9: Average treatment effect on the treated (ATT): PSM results

Outcome variables	Matching algorithm	Mean of outcome variables based on matched observations for		Participation effect (ATT)
		Participant	Non-participant	
wheat output commercialisation index	NN (1)	0.531	0.426	0.105 (0.014) ***
	Calliper (0.01)	0.549	0.426	0.123 (0.013) ***
	Kernel (0.03)	0.531	0.483	0.048 (0.020) **
wheat input commercialisation index	NN (1)	0.110	0.100	0.01 (0.004) **
	Calliper (0.01)	0.114	0.101	0.009 (0.004) **
	Kernel (0.03)	0.109	0.100	0.014 (0.003) ***

Note: ** and *** indicate significance at the 5% and 1% level respectively
Source: Survey result (2023).

3.2 Discussion

3.2.1 Determinants of smallholder producers' participation in wheat cluster farming

Education is an important tool that helps to facilitate farm households' access to information on agricultural production, production technologies, marketing and prices. Accordingly, the findings of the study reveal that education status of farm household head plays a positive role in improving the likelihood of farm households' participation in wheat cluster farming and is significant at the 1% level of significance (Table 6). This result is consistent with Endalew *et al*. (2024) and Guyo and Tabe-Ojong (2024), who found similar results in related studies.

The arrangement of formal credit, provision of extension services and creating awareness of new or improved production technologies and farming approaches are among the common institutional support services delivered to farm households to facilitate the implementation of improved or new farming approaches. As shown in Table 6, the results of the model indicate that access to a formal credit service has a positive and significant effect at the 10% level of significance, implying that, those farm households that have access to formal credit services have a better chance of participation in wheat cluster farming than those who do not have access. Participation in wheat cluster farming requires the ability and willingness to apply comprehensive production packages, and such requirements need financial support to implement. In view of this, and the attention the government pays to wheat production, the positive effect of credit is not surprising and the result is in consistent with a related study conducted by Degefu *et al*. (2025).

Similarly, the results of the study indicate that the frequency of extension contacts plays a positive role in improving farm household heads' participation in wheat cluster farming and is significant at the 1% level. This finding is in consistent with the studies of Hussen and Geleta (2021), Guyo and Tabe-Ojong (2024) and Degefu *et al*. (2025). Likewise, farm households that are aware of wheat cluster farming are more likely to participate compared to those who are not aware, and the effect of this is significant at the 1% level (Table 6). The cluster farming approach needs a series of awareness creation training for different stakeholders, particularly farm households, from its establishment to its implementation. Such a priori awareness creation training plays an important role in improving farm households' participation in cluster farming, and this may be one of the possible reasons for the positive effect of the awareness of farm households of cluster farming on their participation decision in this study. Guyo and Tabe-Ojong (2024) found similar results, and state that farmers who have awareness of the importance of cluster crop production are more likely to participate in cluster crop production.

One of the most important rural infrastructures that facilitate farm households' membership in group activities like cluster farming is rural roads. In this study, distance to rural all-weather roads (km) had a negative influence on farm households' participation in wheat cluster farming, and this was significant at the 5% level (Table 6). This means that farm households residing close to rural all-weather roads tended to have a greater chance of participating in wheat cluster farming. The result is in line with studies conducted by Mussema *et al*. (2013) and Mekonnen and Alamirew (2017) whose findings show that easy access to rural all-weather roads makes an important positive contribution to smallholder producers' decision to participate in the market.

Land size allocated to wheat cultivation is one of the key criteria to select participant smallholder producers and establish their participation in a wheat production cluster. In line with this criterion, the findings of the study reveal that farm size allocated to wheat production has a positive effect on smallholder producers' decision to participate in wheat cluster farming and is significant at the 5% level. This positive result may be linked to the minimum land size requirement to participate in wheat cluster farming. This result is consistent with Guyo and Tabe-Ojong (2024), who found that, in general, the participation rate of smallholder producers in cluster farming has a positive relationship with farm size. Land fragmentation (measured as an index) has a negative effect on participation in wheat cluster farming and is significant at the 5% level (Table 6). High and fragmentation gives rise to multiple parcels of smaller land and this hinders farm households from formally participating in wheat cluster farming.

Farm proximity to nearby clusters (km) is one of the most important attributes that influence smallholder producers' participation in wheat cluster farming. In this study, the variable was found to influence their participation negatively and is significant at the 1% level (Table 6). The result implies those farm households with farmland far from the surrounding established wheat clusters tends to have less chance of participating in wheat cluster farming than their counterparts. This may be related to the concept of cluster farming (the cultivation of a given crop by grouping neighbouring farms). The result is inconsistent with related studies conducted by Ahmed and Mesfin (2017) and Gemechu *et al*. (2024), whose studies reveal the important role played by distance to cooperatives in farm households' membership.

Livestock (measured in terms of TLU) had a positive role in smallholder producers' participation in wheat cluster farming and was significant at the 5% level (Table 6). This implies that smallholder producers with a large number of livestock have a better chance of participating in wheat cluster farming. This may be associated with the key role livestock play as a source of wealth, income, food and power for crop farming activities like ploughing and threshing for Ethiopian farm households. However, there is no consistent research report on the effect of livestock size (TLU) on smallholder producers' participation in collective farming activities such as wheat cluster farming. For instance, Wossen *et al*. (2017) reported that livestock have a positive effect, whereas Hussen and Geleta (2021) reported that it has a negative effect. Such inconsistent research results could be associated with the competitive nature of the two farming activities (crop and livestock), particularly when the size of livestock owned by a farm household is large.

3.2.2 Impact of cluster farming on wheat output and input commercialisation

The result of the ESR model reveals that producing wheat in the cluster approach has a positive and significant impact on wheat output and input commercialisation compared with conventional farming (Table 7). Related studies on different crops, such as teff (Endalew *et al*. 2024), major cereal crops (wheat, teff and maize) (Gidelew *et al*. 2025), and high-acreage crops (*teff*, wheat, maize, barley, and sesame) (Guyo *et al*. 2023), show that cluster farming has a positive and significant impact on output commercialisation. Similarly, the results of the model show that farming in the cluster approach results in a positive and significant impact on wheat input commercialisation. Such a positive impact of cluster farming on wheat output and input commercialisation may be associated with the role of the farming approach in terms of improving farm households' bargaining power and information sharing, access to improved agricultural technologies, market participation, access to different awareness training and access to rural institutional services like credit and extension services. In addition, the Ethiopian government's ambition of scaling up wheat cluster farming and wheat self-sufficiency in the country is expected to play a positive role in making cluster farming result in a positive impact on wheat output commercialisation.

4. Conclusion and recommendations

Recently, the issue of cluster farming in general, and wheat cluster farming in particular, has gained popularity in Ethiopia. By grouping wheat-producing smallholder farmers together and providing them with resources and training, this farming model is expected to improve yields and access to new markets and help them negotiate better prices for their output. This not only benefits individual farmers, but also contributes to the overall economic growth of the country. This study tried to analyse the wheat cluster farming approach in relation to the level of wheat output and input (measured as an index). The study was conducted in Ethiopia's Amhara region using household-level survey data collected from 383 randomly selected wheat-producing rural farm households. Their selection took place through a multistage sampling method during the 2022/2023 production year. The study employed descriptive statistics and the ESR econometric model in the data analyses.

Based on the results of the study, the following conclusions are drawn. The overall average wheat output and input commercialisation levels were 49% and 11%, respectively. Specifically, smallholder producers using the cluster farming approach exhibited higher output and input commercialisation levels compared to those using the conventional way. The results of both the ESR and PSM models demonstrate that participation in cluster farming has a positive and significant impact on wheat output and input commercialisation levels. The model results further confirm the robustness of these findings.

The results of the ESR model indicate that the education status of the farm household head, access to formal credit services, frequency of extension contacts, distance from rural all-weather roads, wheat farm size, land fragmentation (index), livestock size (TLU), cluster farming awareness, and farm distance to the nearby wheat cluster farms significantly influenced farm household participation in wheat cluster farming. In addition, the average treatment effect on the treated (ATT) for the cluster farming approach was 5% for wheat output commercialisation and 3.1% for input commercialisation. The average treatment effect on the untreated (ATU) was 1.9% for output and 0.4% for input commercialisation, indicating that smallholder producers using conventional farming would increase their wheat output and input commercialisation levels by 1.9% and 0.4%, respectively, if they adopted cluster farming. In addition, the transitional heterogeneity (TH) effect was 3.1% for output and 2.6% for input commercialisation, highlighting the significant role of the cluster farming approach in improving both wheat output and input commercialisation levels.

Given the positive and significant impact of cluster farming on wheat output and input commercialisation levels, there is a need to create a favourable environment to make more farmers join clusters for wheat production. By fostering collaboration through the cluster approach, stakeholders in the wheat sector can promote a more prosperous future for all involved. Therefore, government and non-government interventions aimed at strengthening institutions' roles in supporting rural farmers are key for enhancing farm households' participation in wheat cluster farming, thereby boosting wheat output and input commercialisation.

References blues

Adam B, Kassa B, Belaineh L & Tesfaye L, 2010. Effects of crop commercial orientation on productivity of smallholder farmers in drought-prone areas in the Central Rift Valley of Ethiopia. Journal of Rural Development 33(4): 105–28.

Adeoti AI, Oluwatayo IB & Soliu RO, 2014. Determinants of market participation among maize producers in Oyo State, Nigeria. British Journal of Economics, Management & Trade 4(7): 1115–27.

Ademe A, Legesse B, Mohammed JH & Goshu D, 2017. Smallholder farmers' crop commercialization in the Highlands of Eastern Ethiopia. Review of Agricultural and Applied Economics 20(2): 30–7. https://doi.org/10.15414/raae.2017.20.02.30-37

Agricultural Transformation Agency (ATA) 2017. Annual report 2016-17. Addis Ababa, Ethiopia: ATA.

Agricultural Transformation Agency (ATA), 2018. Annual report 2017-18. Addis Ababa, Ethiopia: ATA.

Agricultural Transformation Agency ATA, 2019. Annual report 2018-19. Addis Ababa, Ethiopia: ATA.

Ahmed MH & Mesfin HM, 2017. The impact of agricultural cooperatives membership on the wellbeing of smallholder farmers: Empirical evidence from eastern Ethiopia. Agricultural and Food Economics 5: 6. https://doi.org/10.1186/s40100-017-0075-z

Alemu D, Gabre-Madhin E & Aredo SD, 2006. From farmer to market and market to farmer: Characterizing smallholder commercialization in Ethiopia. Paper presented at the ESSP 2006 Policy Conference on "Bridging, balancing, and scaling up: Advancing the rural growth agenda in Ethiopia", 6–8 June, Addis Ababa, Ethiopia.

Axalan JT, Israel FT, Batt PJ, Murray-Prior R & Loma AL, 2011. Socio-economic impact of cluster marketing: The case of Ned Landcare Association Sweet Pepper Cluster. Acta Horticulturae 895: 37–44.

Becerril J & Abdulai A, 2010. The impact of improved maize varieties on poverty in Mexico: A propensity score-matching approach. World Development 38(7): 1024–35. https://doi.org/10.1016/j.worlddev.2009.11.017

Boka GT, 2017. Climate change challenges, smallholder commercialization and progress out poverty in Ethiopia. Working Paper No. 253, African Development Bank Group, Abidjan, Côte d'Ivoire. https://www.afdb.org/ar/documents/document/working-paper-253-climate-change-challenges-smallholder-commercialization-and-progress-out-poverty-in-ethiopia-94540

Breusch TS & Pagan AR, 1979. A simple test for heteroscedasticity and random coefficient variation. Econometrica 47(5): 1287–94.

Caliendo M & Kopeinig S, 2005. Some practical guidance for the implementation of propensity score matching. IZA Discussion Paper No. 1588. Bonn, Germany: IZA. https://docs.iza.org/dp1588.pdf

Cazzuffi C, McKay A & Perge E, 2020. The impact of agricultural commercialisation on household welfare in rural Vietnam. Food Policy 94: 101811. https://doi.org/10.1016/J.FOODPOL.2019.101811

Central Statistics Agency of Ethiopia (CSA), 2007. Population and housing census: Statistical report for Amhara Region. CSA: Addis Ababa, Ethiopia.
Central Statistics Agency of Ethiopia (CSA), 2021. Agricultural sample survey report on crop and livestock product utilization. Statistical Bulletin Volume VII, Addis Ababa, Ethiopia.
Central Statistics Agency of Ethiopia (CSA), 2022. Agricultural sample survey report on area and production of major crops. Statistical Bulletin volume I, Addis Ababa, Ethiopia.
Degefu S, Andualem HM, Sileshi M, Ogeto MA & Bacsi Z, 2025. Impact of cluster farming on technical efficiency and food security among smallholder wheat producers in Southeastern Ethiopia. Discover Food 5: 129. https://doi.org/10.1007/s44187-025-00399-2
Degefu S, Demelash M & Tsegae T, 2024. Determinants of smallholder wheat farmers' participation in cluster farming in Lemu-Bilbilo and Hetosa Districts of Arsi Zone, Ethiopia. Discover Agriculture 2: 93. https://doi.org/10.1007/s44279-024-00111-5
Dehejia RH & Wahba S, 2002. Propensity score-matching methods for nonexperimental causal studies. Review of Economics and Statistics 84(1): 151–61. https://doi.org/10.1162/003465302317331982
Di Falco S, Veronesi M & Yesuf M, 2011. Does adaptation to climate change provide food security? A micro-perspective from Ethiopia. American Journal of Agricultural Economics 93(3): 825–42. https://doi.org/10.1093/ajae/aar006
Elouhichi K, Temursho U, Colen L & Gomez Y Paloma S, 2019. Upscaling the productivity performance of the agricultural commercialization cluster initiative in Ethiopia. Luxembourg: Publications Office of the European Union. https://doi.org/10.2760/57450
Endalew B, Elias A & Yasunobu K, 2024. Impact of cluster farming on smallholder farmers teff commercialization in Ethiopia. CABI Agriculture and Bioscience 5: 18. https://doi.org/10.1186/s43170-024-00220-7
FAO, 2010. Agro-based clusters in developing countries: Staying competitive in a globalized economy. Agricultural Management, Marketing and Finance Occasional Paper. Rome: Food and Agriculture Organization (FAO) of the United Nations.
Ferede S, Agegnehu G, Ashagrie AK, Alemu T & Yirga C, 2020. Farming systems characterization and analysis in East Gojjam Zone: Implications for research and development (R&D) interventions. Research Report No. 127, EIAR, Addis Ababa, Ethiopia. https://www.researchgate.net/publication/344848973
Gemechu A, Jaleta M, Zemedu L & Beyene F, 2024. Impact of membership of seed-producer cooperatives on commercialisation among smallholder farmers in the central highlands of Ethiopia. African Journal of Agricultural and Resource Economics 19(1): 54–84. https://doi.org/10.53936/afjare.2024.19(1).4
Gidelew EG, Alemu BA & Kassie KE, 2025. The crop commercialization impact of cluster farming initiative in Northwestern Ethiopia. Frontiers in Sustainable Food Systems 9: 1548304. https://doi.org/10.3389/fsufs.2025.1548304
Govereh J, Jayne TS & Nyoro J, 1999. Smallholder commercialization, interlinked markets and food crop productivity: Cross-country evidence in Eastern and Southern Africa. Department of Agricultural Economics and Department of Economics, Michigan State University, East Landing, Michigan, USA. https://www.researchgate.net/publication/238075353
Guyo GD & Tabe-Ojong MP, 2024. Land as a binding constraint to cluster-based development in Ethiopia: To cluster or not to cluster? PLoS ONE 19(4): e0298784. https://doi.org/10.1371/JOURNAL.PONE.0298784
Guyo GD, Tabe-Ojong Jr MP & Owusu-Sekyere E, 2023. The new normal? Cluster farming and smallholder commercialization in Ethiopia. Agricultural Economics 54(6): 900–20. https://doi.org/10.1111/agec.12790
Hao J, Bijman J, Gardebroek C, Heerink N, Heijman W & Huo X, 2018. Cooperative membership and farmers' choice of marketing channels – Evidence from apple farmers in Shaanxi and

Shandong Provinces, China. Food Policy 74: 53–64. https://doi.org/10.1016/J.FOODPOL.2017.11.004
Heckman JJ, 1979. Sample selection bias as a specification error. Econometrica 47(1): 153–61.
Heckman JJ & Vytlacil EJ, 2007. Econometric evaluation of social programs, Part I: Causal models, structural models and econometric policy evaluation. In Heckman JJ & Learner E (eds), Handbook of econometrics, Vol. 6, Part B, (pp. 4779–4874). https://doi.org/10.1016/S1573-4412(07)06070-9
Heckman J, Tobias JL & Vytlacilt E, 2001. Four parameters of interest in the evaluation of social programs. Southern Economic Journal 68(2): 210–23. https://doi.org/10.2307/1061591
Hussen CH & Geleta FT, 2021. Factors affecting smallholder farmers participations in cluster crop production: Evidence from selected districts of West Shewa Zone, Oromia National Regional State, Ethiopia. Sarhad Journal of Agriculture 37(3): 818–29. https://doi.org/10.17582/journal.sja/2021/37.3.818.829
Jaleta M, Gebremedhin B & Hoekstra D, 2009. Smallholder commercialization: Processes, determinants and impact. Discussion Paper No. 18, Improving Productivity and Market Success (IPMS) of Ethiopian Farmers Project, ILRI (International Livestock Research Institute), Nairobi, Kenya.
Jaleta M, Kassie M, Tesfaye K, Teklewold T, Jena PR, Marenya P & Erenstein O, 2016. Resource saving and productivity enhancing impacts of crop management innovation packages in Ethiopia. Agricultural Economics 47(5): 513–22. https://doi.org/10.1111/agec.12251
Jarque CM & Bera AK, 1987. A test for normality of observations and regression residuals. International Statistical Review 55(2): 163–72.
Karki L, Shange R, Hill WA, Bonsi C, English H, Agbodjan E, Stewart M, Chukwuma F, Patin DM & Idassi J, 2021. Assessing the impact of cluster farming initiatives on small and socially disadvantaged farmers, ranchers, and forestland owners: A case of the Southeastern region of the United States. Professional Agricultural Workers Journal 7(2): 26–41.
Kassie M, Shiferaw B & Muricho G, 2011. Agricultural technology, crop income, and poverty alleviation in Uganda. World Development 39(10): 1784–95. https://doi.org/10.1016/j.worlddev.2011.04.023
Kilelu C, Klerkx L, Omore A, Baltenweck I, Leeuwis C & Githinji J, 2017. Value chain upgrading and the inclusion of smallholders in markets: Reflections on contributions of multi-stakeholder processes in dairy development in Tanzania. European Journal of Development Research 29(5): 1102–21. https://doi.org/10.1057/S41287-016-0074-Z/
Kilimani N, Buyinza F & Guloba M, 2022. Crop commercialization and nutrient intake among farming households in Uganda. Food Policy 113: 102328. https://doi.org/10.1016/J.FOODPOL.2022.102328
Kothari CR, 2004. Research methodology: Methods & techniques. New Delhi: New Age International Publishers.
Lee L-F, 1982. Some approaches to the correction of selectivity bias. The Review of Economic Studies 49(3): 355–72.
Lokshin M & Glinskaya E, 2008. The effect of male migration for work on employment patterns of females in Nepal. Policy Research Working Paper 4757, Development Research Group, The World Bank.
Lokshin M & Sajaia Z, 2004. Maximum likelihood estimation of endogenous switching regression models. The Stata Journal 4(3): 282–9.
Maddala GS, 1983. Limited-dependent and qualitative variables in econometrics. Cambridge UK: Cambridge University Press.
Martinidis G, Adamseged ME, Dyjakon A, Fallas Y, Foutri A, Grundmann P, Hamann K, Minta S, Ntavos N, Råberg T, Russo S & Viaggi D, 2021. How clusters create shared value in rural areas:

An examination of six case studies. Sustainability 13(8): 4578. https://doi.org/10.3390/SU13084578
Mekonnen N & Alamirew B, 2017. The effect of rural road transport on crop commercialization: Evidence from the Ethiopian Living Standard Measurement Survey. Ethiopian Journal of the Social Sciences and Humanities 13(2): 81–108. https://doi.org/10.4314/ejossah.v13i2.4
Molema M, Segers Y & Karel E, 2016. Introduction: Agribusiness clusters in Europe, 19th and 20th centuries. Tijdschrift Voor Sociale en Economische Geschiedenis/The Low Countries Journal of Social and Economic History 13(4), 1–16. https://doi.org/10.18352/tseg.894
Mussema R, Kassa B, Alemu D & Rashid S, 2013. Analysis of the determinants of small-scale farmers' grain market participations in Ethiopia: The contribution of transaction costs. Ethiopian Journal of Agricultural Science 23: 75–94.
Negowetti NE, 2017. Exposing the invisible costs of commercial agriculture: Shaping policies with true costs accounting to create a sustainable food future. Valparaiso University Law Review 51(2): 447–82. https://scholar.valpo.edu/vulr/vol51/iss2/5
Olmos A & Govindasamy P, 2015. Propensity scores: A practical introduction using R. Journal of MultiDisciplinary Evaluation 11(25): 68–88. https://doi.org/10.56645/jmde.v11i25.431
Otekunrin OA, Momoh S & Ayinde IA, 2019. Smallholder farmers' market participation: Concepts and methodological approach from Sub-Saharan Africa. Current Agriculture Research Journal 7(2): 139–57. https://doi.org/10.12944/carj.7.2.02
Pingali PL, 1997. From subsistence to commercial production systems: The transformation of Asian agriculture. American Journal of Agricultural Economics 79(2): 829–34. https://doi.org/10.2307/1244162
Pingali PL, 2001. Environmental consequences of agricultural commercialization in Asia. Environment and Development Economics 6(4): 483–502. https://doi.org/10.1017/S1355770X01000274
Pingali PL & Rosegrant MW, 1995. Agricultural commercialization and diversification: Processes and policies. Food Policy 20(3): 171–85.
Porter ME, 1990. The competitive advantage of nations. New York, NY: Free Press.
Poulton C, Dorward A & Kydd J, 2010. The future of small farms: New directions for services, institutions, and intermediation. World Development, 38(10), 1413–28. https://doi.org/10.1016/J.WORLDDEV.2009.06.009
Rosenfeld S, 2009. Generating wealth, opportunity, and sustainability through rural clusters, Volume I. Carrboro, NC: Regional Technology Strategies. https://doi.org/10.13140/RG.2.1.5151.3048
Rubin DB. 1973. Matching to remove bias in observational studies. Biometrics 29(1): 159–83.
Schmitz H, 1992. On the clustering of small firms. IDS Bulletin 23(3): 64–9. https://doi.org/10.1111/j.1759-5436.1992.mp23003012.x
Schouten M & Heijman WJM, 2012. Agricultural clusters in the Netherlands. Visegrad Journal on Bioeconomy and Sustainable Development 1(1): 20–6. https://vua.uniag.sk/sites/default/files/20-26_0.pdf
Senbeta AF & Worku WW, 2023. Ethiopia's wheat production pathways to self-sufficiency through land area expansion, irrigation advance, and yield gap closure. Heliyon 9(10): e20720. https://doi.org/10.1016/J.HELIYON.2023.E20720
Shiferaw B, Kassie M, Jaleta M & Yirga C, 2014. Adoption of improved wheat varieties and impacts on household food security in Ethiopia. Food Policy 44: 272–84. https://doi.org/10.1016/j.foodpol.2013.09.012
Silva JV, Reidsma P, Baudron F, Moti J, Kindie T & Van Ittersum MK, 2021. Wheat yield gaps across smallholder farming systems in Ethiopia. Agronomy for Sustainable Development, 41: 12. https://doi.org/10.1007/S13593-020-00654-Z/
Strasberg PJ, Jayne TS, Yamano T, Nyoro J, Karanja D & Strauss J, 1999. Effects of agricultural commercialization on food crop input use and productivity in Kenya. International Development

Working Paper No. 71, Michigan State University. East Lansing, Michigan. http://www.aec.msu.edu/agecon/fs2/index.htm

Tabe Ojong Jr MP, Hauser M & Mausch K, 2022. Does agricultural commercialisation increase asset and livestock accumulation on smallholder farms in Ethiopia? Journal of Development Studies 58(3): 524–44. https://doi.org/10.1080/00220388.2021.1983170

Te AG & Milkias D, 2021. Review on agricultural extension systems in Ethiopia: A cluster farming approaches. Journal of Biology, Agriculture and Healthcare 11(14): 1–6. https://doi.org/10.7176/jbah/11-14-01

United States Department of Agriculture (USDA), 2025b. World agricultural production. Circular Series WAP 05-25, United States Department of Agriculture Foreign Agricultural Service, Washington, DC.

United States Department of Agriculture (USDA), 2025a. Grain: World markets and trade. Washinton, DC: United States Department of Agriculture Foreign Agricultural Service Global Market Analysis.

Von Braun J, 1995. Agricultural commercialization: Impacts on income and nutrition and implications for policy. Food Policy 20(3): 187–202. https://doi.org/10.1016/0306-9192(95)00013-5

Wooldridge JM, 2002. Econometric analysis of cross section and panel data. Second edition. Cambridge, MA: The MIT Press.

Wossen T, Abdoulaye T, Alene A, Haile MG, Feleke S, Olanrewaju A & Manyong V, 2017. Impacts of extension access and cooperative membership on technology adoption and household welfare. Journal of Rural Studies 54: 223–33. https://doi.org/10.1016/j.jrurstud.2017.06.022

Zenbaba OS, Ketema M, Jaleta M & Jemal K, 2024. Adoption determinants of wheat production technology packages by smallholder farmers in Horo Guduru Wollega zone, Ethiopia. Discover Food 4: 134. https://doi.org/10.1007/S44187-024-00193-6/

Appendix

Appendix Table 1: Heteroscedasticity test for wheat output commercialisation equation

Variables	Wheat output commercialisation level			
	Participant		Non participant	
	Coef.	Std. error	Coef.	Std. error
Age (years)	0.001	0.001	-0.001	0.002
Sex (male = 1)	0.037	0.030	0.069**	0.034
Education status (literate = 1)	0.031**	0.014	0.008	0.018
Household size (AEU)	-0.005	0.008	-0.004	0.010
Formal credit access (yes = 1)	0.037***	0.014	0.039**	0.019
Extension contacts (frequency)	0.004	0.004	0.008*	0.004
Cooperative membership (yes = 1)	0.066**	0.026	0.047**	0.022
Distance to rural all-weather roads (km)	-0.011***	0.006	0.007	0.005
Distance from main market (km)	-0.004***	0.001	-0.002	0.002
Wheat farm size (ha)	0.204***	0.038	0.443***	0.048
Land fragmentation (index)	-0.015	0.012	0.025	0.079
Livestock size (TLU)	0.002	0.003	0.006	0.004
Involvement in off-farm activity (yes = 1)	-0.006	0.015	-0.001	0.019
Constant	0.277	0.058	0.087	0.079
Breusch-Pagan/Cook-Weisberg test of heteroskedasticity				
Number of observations	191		192	
R-squared	0.435		0.498	
H_o: Constant variance				
chi^2 (1)	2.01		0.05	
Prob > chi^2	0.156		0.823	
Decision	Accept H_o		Accept H_o	

Appendix Table 2: Heteroscedasticity test for wheat input commercialisation equation

Variables	Wheat output commercialisation level			
	Participant		Non participant	
	Coef.	Std. error	Coef.	Std. error
Age (years)	001	0.001	0.002**	0.001
Sex (male = 1)	0.015	0.011	0.008	0.010
Education status (literate = 1)	0.018***	0.005	0.011*	0.005
Household size (AEU)	-0.003	0.003	-0.006**	0.003
Formal credit access (yes = 1)	0.009*	0.005	0.007	0.005
Extension contacts (frequency)	0.006***	0.001	0.003**	0.001
Cooperative membership (yes = 1)	-0.010	0.009	-0.014**	0.006
Distance to rural all-weather roads (km)	-0.003	0.002	-0.004	0.002
Distance from main market (km)	-0.001	0.001	0.001	0.001
Wheat farm size (ha)	0.042***	0.014	0.043***	0.014
Land fragmentation (index)	0.004	0.004	0.006	0.023
Livestock size (TLU)	-0.001	0.001	-0.002	0.001
Involvement in off-farm activity (yes = 1)	0.010*	0.005	0.006	0.006
Constant	0.083***	0.021	0.063***	0.023
Breusch-Pagan/Cook-Weisberg test of heteroskedasticity				
Number of observations	191		191	
R-squared	0.322		0.170	
H_o: Constant variance				
chi^2 (1)	0.57		1.46	
Prob > chi^2	0.451		0.227	
Decision	Accept H_o		Accept H_o	

Appendix Table 3: Normality test for wheat output commercialisation equation

Variables	Wheat output commercialisation level			
	Participant		Non-participant	
	Coef.	Std. error	Coef.	Std. error
Age (years)	0.001	0.001	-0.001	0.002
Sex (male = 1)	0.037	0.030	0.069**	0.034
Education status (literate = 1)	0.031**	0.014	0.008	0.018
Household size (AEU)	-0.005	0.008	-0.004	0.010
Formal credit access (yes = 1)	0.037***	0.014	0.039**	0.019
Extension contacts (frequency)	0.004	0.004	0.008*	0.004
Cooperative membership (yes = 1)	0.066**	0.026	0.047**	0.022
Distance to rural all-weather roads (km)	-0.011***	0.006	0.007	0.005
Distance from main market (km)	-0.004***	0.001	-0.002	0.002
Wheat farm size (ha)	0.204***	0.038	0.443***	0.048
Land fragmentation (index)	-0.015	0.012	0.025	0.079
Livestock size (TLU)	0.002	0.003	0.006	0.004
Involvement in off-farm activity (yes = 1)	-0.006	0.015	-0.001	0.019
Constant	0.277	0.058	0.087	0.079
Jarque-Bera normality test				
Number of observations	191		192	
R-squared	0.435		0.498	
H_o: The residuals are normally distributed				
chi^2 (2)	3.252		1.814	
Prob > chi^2	0.197		0.404	
Decision	Accept H_o		Accept H_o	

Appendix Table 4: Normality test for wheat input commercialisation equation

Variables	Wheat output commercialisation level			
	Participant		Non-participant	
	Coef.	Std. error	Coef.	Std. error
Age (years)	001	0.001	0.002**	0.001
Sex (male = 1)	0.015	0.011	0.008	0.010
Education status (literate = 1)	0.018***	0.005	0.011*	0.005
Household size (AEU)	-0.003	0.003	-0.006**	0.003
Formal credit access (yes = 1)	0.009*	0.005	0.007	0.005
Extension contacts (frequency)	0.006***	0.001	0.003**	0.001
Cooperative membership (yes = 1)	-0.010	0.009	-0.014**	0.006
Distance to rural all-weather roads (km)	-0.003	0.002	-0.004	0.002
Distance from main market (km)	-0.001	0.001	0.001	0.001
Wheat farm size (ha)	0.042***	0.014	0.043***	0.014
Land fragmentation (index)	0.004	0.004	0.006	0.023
Livestock size (TLU)	-0.001	0.001	-0.002	0.001
Involvement in off-farm activity (yes = 1)	0.010*	0.005	0.006	0.006
Constant	0.083***	0.021	0.063***	0.023
Jarque-Bera normality test				
Number of observations	191		191	
R-squared	0.322		0.170	
H_o: The residuals are normally distributed				
chi^2 (1)	3.13		2.42	
Prob > chi^2	0.209		0.298	
Decision	Accept H_o		Accept H_o	

Appendix Table 5: IV test for wheat output and input commercialisation equation (non-participant)

Variables	Participation (participant = 1)		Wheat commercialisation			
			Output		Input	
	Coef.	Std. error	Coef.	Std. error	Coef.	Std. error
Age (years)	0.035	0.024	0.001	0.002	0.002	0.001
Sex (male = 1)	-0.449	0.473	0.062*	0.034	0.007	0.010
Education status (literate = 1)	0.780***	0.211	0.011	0.019	0.012**	0.005
Household size (AEU)	-0.220	0.143	-0.004	0.010	-0.006**	0.003
Formal credit access (yes = 1)	0.348*	0.204	0.040**	0.019	0.007	0.005
Extension contacts (frequency)	0.183***	0.055	0.008*	0.004	0.003**	0.001
Cooperative membership (yes = 1)	0.088	0.319	0.044*	0.023	0.014**	0.007
Distance to rural all-weather roads (km)	-0.209**	0.091	0.007	0.005	-0.004**	0.002
Distance from main market (km)	-0.026	0.020	-0.002	0.002	0.001	0.001
Wheat farm size (ha)	1.200**	0.597	0.439***	0.049	0.046***	0.014
Land fragmentation (index)	-1.983**	0.867	0.025	0.079	0.006	0.023
Livestock size (TLU)	0.096**	0.048	0.007	0.004	-0.002	0.001
Involvement in off-farm activity (yes = 1)	-0.147	0.230	-0.001	0.019	0.006	0.006
Cluster farming awareness (aware = 1)	1.118**	0.457	-0.001	0.020	-0.004	0.006
Farm proximity to nearby cluster (km)	-0.929***	0.118	-0.009	0.007	-0.003	0.002
Constant	-0.897	1.099	0.098	0.081	0.069***	0.023
Instrumental variable (IV) test						
Number of observations	383		192		192	
Pseudo/R-squared	0.536		0.502		0.178	
H_o: The instrumental variables (cluster farming awareness and farm proximity to nearby wheat cluster) are valid						
Wald chi^2/F-stat	75.25***		0.71		0.85	
Prob > chi^2/F-stat	0.000		0.495		0.429	
Decision	Accept H_o		Accept H_o		Accept H_o	

Note: For a given instrumental variable to be valid, its effect must be significant in the participation equation but not in the outcome equation.

Appendix Table 6: ESR result for wheat output commercialisation

Variables	Participation (participant = 1)		Wheat output commercialisation level			
			Participant		Non-participant	
	Coef.	Std. error	Coef.	Std. error	Coef.	Std. error
Age (years)	0.035	0.024	0.000	0.001	-0.001	0.002
Sex (male = 1)	-0.449	0.473	0.027	0.029	0.063*	0.033
Education status (literate = 1)	0.780***	0.211	0.040***	0.015	0.004	0.018
Household size (AEU)	-0.220	0.143	-0.003	0.007	-0.003	0.010
Formal credit access (yes = 1)	0.348*	0.204	0.041***	0.013	0.037**	0.018
Extension contacts (frequency)	0.183***	0.055	0.006**	0.003	0.007*	0.004
Cooperative membership (yes = 1)	0.088	0.319	0.072***	0.025	0.044**	0.022
Distance to rural all-weather roads (km)	-0.209**	0.091	0.008	0.006	0.007	0.005
Distance from main market (km)	-0.026	0.020	-0.004***	0.001	-0.002	0.002
Wheat farm size (ha)	1.200**	0.597	0.216***	0.037	0.430***	0.047
Land fragmentation (index)	-1.983**	0.867	-0.021*	0.012	0.040	0.076
Livestock size (TLU)	0.096**	0.048	0.003	0.003	0.006**	0.003
Involvement in off-farm activity (yes = 1)	-0.147	0.230	-0.009	0.015	0.002	0.018
Cluster farming awareness (aware = 1)	1.118**	0.457				
Farm proximity to nearby cluster (km)	-0.929***	0.118				
Constant	-0.897	1.099	0.254***	0.058	0.079	0.076
Model diagnosis parameters						
Number of observations	383					
rho1(ρ_1)	0.394**	0.193				
rho2(ρ_2)	-0.264	0.170				
Log likelihood	259.66					
Wald chi^2	133.65***	p = 0.00				
LR test of independent equations: chi^2	5.42**	P = 0.02				

Appendix Table 7: ESR result for wheat input commercialisation

Variables	Participation (participant = 1)		Wheat output commercialisation level			
			Participant		Non-participant	
	Coef.	Std. error	Coef.	Std. error	Coef.	Std. error
Age (years)	0.027	0.023	0.001	0.001	0.003***	0.001
Sex (male = 1)	-0.506	0.457	0.016	0.010	0.005	0.009
Education status (literate = 1)	0.843***	0.210	0.018***	0.005	0.009*	0.005
Household size (AEU)	-0.159	0.119	-0.003	0.003	-0.006*	0.003
Formal credit access (yes = 1)	0.374*	0.201	0.009*	0.005	0.006	0.005
Extension contacts (frequency)	0.180 ***	0.053	0.006***	0.001	0.003***	0.001
Cooperative membership (yes = 1)	0.089	0.315	-0.010	0.009	-0.016***	0.006
Distance to rural all-weather roads (km)	-0.180**	0.088	-0.003	0.002	-0.004**	0.002
Distance from main market (km)	-0.026	0.020	-0.001	0.002	0.001	0.001
Wheat farm size (ha)	1.471**	0.590	0.043***	0.013	0.050***	0.014
Land fragmentation (index)	-1.853**	0.886	0.004	0.004	0.013	0.022
Livestock size (TLU)	0.082*	0.048	-0.001	0.001	-0.002*	0.001
Involvement in off-farm activity (yes = 1)	-0.180	0.229	0.010**	0.005	0.008	0.005
Cluster farming awareness (aware = 1)	1.111**	0.448				
Farm proximity to nearby cluster (km)	-0.872***	0.116				
Constant	-1.157	1.066	0.084***	0.021	0.057**	0.022
Model diagnosis parameters						
Number of observations	383					
rho1(ρ_1)	-0.021	0.228				
rho2(ρ_2)	-0.481**	0.197				
Log likelihood	694.98					
Wald chi^2	75.52***	P = 0.00				
LR test of independent equations: chi^2	4.70**	P = 0.03				

Appendix Table 8: Logistic regression results for PSM model estimation

Variables	Participation decision (participant = 1)	
	Coef.	**Std. error**
Age (years)	0.031	0.045
Sex (male = 1)	-0.892	0.808
Education status (literate = 1)	1.444***	0.385
Household size (AEU)	-0.219	0.229
Formal credit access (yes = 1)	0.701*	0.379
Extension contacts (frequency)	0.300 ***	0.095
Cooperative membership (yes = 1)	0.357	0.570
Distance to rural all-weather roads (km)	-0.263*	0.156
Distance from main market (km)	-0.048	0.035
Wheat farm size (ha)	2.565**	1.053
Land fragmentation (index)	-2.913*	1.570
Livestock size (TLU)	0.147*	0.087
Involvement in off-farm activity (yes = 1)	-0.162	0.413
Cluster farming awareness (aware = 1)	2.579**	0.950
Farm proximity to nearby cluster (km)	-1.619***	0.222
Constant	-2.206	1.988
Model diagnosis parameters		
Number of observations	383	
Log likelihood	-122.20	
LR chi^2	286.54***	
Pseudo R^2	0.540	

Appendix Table 9: Covariate balancing test

Variables	Mean of covariates before matching		t	p > t	% bias	Mean of covariates after matching		t	p > t	% bias
	Part.	Non-part.				Part.	Non-part.			
Age (years)	44.32	43.62	0.90	0.37	9.2	44.25	43.05	1.56	0.12	15.80
Sex (male = 1)	0.99	0.95	2.30**	0.02	20.1	0.95	0.94	0.45	0.66	4.60
Education status (literate = 1)	0.62	0.27	7.65***	0.00	78.1	0.55	0.50	0.85	0.40	10.5
Household size (AEU)	4.04	3.68	2.57**	0.01	26.2	3.92	3.72	1.42	0.14	14.50
Formal credit access (yes = 1)	0.42	0.27	3.18***	0.00	32.5	0.50	0.36	2.50**	0.013	10.3
Extension contacts (frequency)	6.15	3.94***	10.32	0.00	105.4	5.84	5.71	0.50	0.62	6.0
Cooperative membership (yes = 1)	0.93	0.83	3.02***	0.00	30.9	0.93	0.92	0.27	0.79	2.60
Distance to rural all-weather roads (km)	1.47	2.31	-5.63***	0.00	-57.6	1.53	1.84	1.81*	0.07	-10.2
Distance from main market (km)	8.41	11.16	-4.73***	0.00	-48.4	8.72	8.70	0.02	0.98	0.2
Wheat farm size (ha)	0.64	0.48	8.02***	0.00	82.0	0.62	0.59	1.55	0.12	7.1
Land fragmenta-tion (index)	0.54	0.60	-1.68*	0.09	-17.1	0.58	0.61	-1.35	0.18	-7.5
Livestock size (TLU)	4.97	4.34	2.81***	0.00	28.7	4.76	4.37	1.45	0.15	12.4
Involvement in off-farm activity (yes = 1)	0.26	0.24	0.38	0.70	3.9	0.25	0.26	-0.11	0.91	-1.3
Cluster farming awareness (aware = 1)	0.99	0.74	7.87***	0.00	80.6	0.99	0.98	0.09	0.92	0.40
Farm proximity to nearby cluster (km)	0.68	1.96	-13.78***	0.00	-141.0	0.77	0.70	0.84	0.4	7.1

Note: part. and non-part. indicate participant and non-participant, respectively