Received: 4 June 2020 | Revised: 2 October 2021 | Accepted: 3 December 2021

DOI: 10.1111/agec.12702

AGRICULTURAL ECONOMICS
The Journal of the International Association of Agricultural Economists
IAAE
WILEY

ORIGINAL ARTICLE

Does neighborhood matter? Spatial proximity and farmers' technical efficiency

Wondmagegn Tirkaso[1] | **Atakelty Hailu**[2]

[1] Swedish University of Agricultural Sciences (SLU), Department of Economics, Uppsala, Sweden

[2] University of Western Australia (UWA), School of Agriculture and Environment, Perth, Australia

Correspondence
Wondmagegn Tirkaso, Swedish University of Agricultural Sciences (SLU), Department of Economics, Box 7013, SE-750 07 Uppsala, Sweden.
Email: wondmagegn.tafesse@slu.se, wondewin@gmail.com

Abstract

This article examines the effects of neighborhood on the farmer's technical efficiency (TE) level, adopting a stochastic frontier and spatial Durbin regression models. Our study exploits a three-wave household-level panel data from the Ethiopian Rural Socioeconomic Survey (ERSS) collected between 2011 and 2015. We find that farmers have an average TE score of 53%, implying a substantial potential for improvement in the production level. We further find that there is a positive and statistically significant spatial interdependence in TE levels between farms in neighboring communities. Input use, education, and other demographic characteristics are found to have significant positive direct and indirect effects. The findings suggest that policies and programs targeting productivity improvements in agriculture need to consider spatial spillover effects.

KEYWORDS
Ethiopia, neighborhood effect, panel data, spatial Durbin model, technical efficiency

JEL CLASSIFICATION
Q16, C21, Q18

1 | INTRODUCTION

The evidence that neighborhood influences economic agents' decisions has received considerable attention in recent years (e.g., Bandiera & Rasul, 2006; Bramoullé et al., 2009; Conley & Udry, 2010; Krishnan & Patnam, 2014; Sacerdote, 2001). Economic agents usually have certain, but limited knowledge about their resources and activities, which ultimately shapes their decision-making behavior through available social interaction opportunities. The communication channels could be either through imitation from the neighbor (endogenous effects), due to exogenous characteristics of the neighbor (exogenous effects), or as a result of common unobserved characteristics (e.g., see Anselin, 2013; L. F. Lee, 2004; Manski, 2000).

Early empirical studies have shown that farmer behavior towards new agricultural technology adoption, acquisition of formal and informal education, and information flow is significantly influenced by the existing social interaction with other farmers around the neighborhood (e.g., Abdulai & Huffman, 2014; Aker & Mbiti, 2010; Conley & Udry, 2010; Genius et al., 2014; Krishnan & Patnam, 2014; Maertens & Barrett, 2013; Yamauchi, 2007). This is particularly common in subsistence farming regimes where poor technology adoption, lower educational levels, risk aversion behavior, resource constraints, and higher transaction costs are key features. A study in Mozambique showed that farmers are more likely to adopt a new crop when some farmers in their network also adopt but are less likely to adopt when many others do so (Bandiera & Rasul, 2006). In the absence of well-functioning financial markets, the demand for insurance in the farm sector is influenced by the payout experience of others within the farmer network (Karlan et al., 2014). Another empirical study in rural Ethiopia showed that social networks reduce transaction costs and increase farmers' bargaining power, helping farmers earn higher returns when marketing their products (Kassie et al., 2013).

In this study, we examined the spatial interdependence of technical efficiency score (TE) between farms in neighboring communities using a sample of Ethiopian farmers. Identifying a theoretically consistent parameter estimate for the neighborhood effect usually poses an empirical challenge. Notably the issue of the so called "*reflection problem*" draws much attention in spatial network analysis (e.g., see Brock & Durlauf, 2001; Lin, 2010; Manski, 1993, 2000). This problem is related to uniquely identifying the endogenous, exogenous, and correlated effects associated with social interaction. The endogenous effect reflects how a farmer's TE is shaped by its neighbor's TE, while the exogenous effect relate to how the farmer's TE is affected by the exogenous characteristics of others in his/her neighborhood. The correlated effect is related to separating social interaction from other confounding effects, which are attached to non-random distributions of the group where TE level of the farmer is influenced by unobserved factors, commonly shared by the same group members, rather than social interaction among them. Having similar individual characteristics or institutional environment with a group are examples of correlated effects.

In this article, we seek to identify the effects of the neighborhood on farmer's TE level by controlling households and community-level socioeconomic, demographic, and institutional characteristics. Our empirical strategy is based on the following steps. First, we predicted the individual farmer's TE score using a time-varying stochastic frontier model (SFM) (e.g., Battese & Coelli, 1992; Kumbhakar, 1990; Kumbhakar & Wang, 2005). The standard time-varying SFM's could be limited in handling potential endogeneity problems in the production function.[1] Thus, we addressed this empirical challenge by estimating a copula-based SFM (Amsler et al., 2017; Amsler et al., 2016; Karakaplan & Kutlu, 2017; Tran & Tsionas, 2015). Second, we have estimated the effect of neighborhood and other covariates on the predicted TE score using the spatial Durbin model (SDM) (Anselin, 2013; J. P. LeSage & Pace, 2010; Lin, 2010). Relying on SDM with consideration of endogenous effects, exogenous effects, and individual fixed effects allows for consistent estimation of the parameters (e.g., see Boucher et al., 2014; Hsieh & Lee, 2016; Lin, 2010).

Our study is generally motivated by two fundamental justifications. First, there are spatial externalities due to knowledge spillovers that tend to increase the probability of learning, adopting efficiency improving technologies and farming techniques from the nearby farmers (e.g., Ertur & Koch, 2007; Genius et al., 2014; Manski, 2000). Second, there are spatially interdependent unobserved or latent variables such as work culture, infrastructure, and various amenities that could potentially affect farmers TE level (e.g., LeSage & Pace, 2009). These imply that there are at least two lines of mechanisms through which neighborhood could affect farmers TE level. First, the probability of adopting new farming technologies by a farmer is higher if other farmers initially adopted the same technology in the neighboring communities. This process is expected to improve the TE level of adopters following the integration of new technologies or farming systems. Second, farmers could learn about new farming techniques and optimal input uses from their closest skilled neighbor, such as farmers with better educational level, through formal and informal interaction to improve efficiency. Therefore, analysis of farm efficiency in ways that incorporate neighborhood effects would contribute to policy design and implementation aimed at improving farm performance in Ethiopian agriculture. Particularly, results are expected to provide information useful to integrating spatial dimensions into agricultural education, technology adoption programs and policies aiming at increasing productivity while optimizing the benefits from limited public resources.

We investigate spillover effects by exploiting a recent and unique geo-coded three wave (2011, 2013, and 2015) household-level panel data from the Ethiopia Rural Socioeconomic Survey (ERSS) (World Bank, 2016). This dataset contains information on, among other things, agricultural output, farm size, input use, output and input prices, employment, and income from farm and non-farm sectors.

The rest of this article is organized as follows. Section 2 presents the empirical strategies, starting with the elaboration of a time-varying Stochastic Frontier Analysis (SFA), which is used to predict the average individual-level TE scores and the spatial regression models to examine neighborhood effects. In Section 3, we present a brief description of the data used in this study. Section 4 presents results and discussion. In Section 5, we conclude the article and draw some policy lessons.

2 | EMPIRICAL STRATEGY

Our empirical analysis consists of two main parts. First, we specify a stochastic frontier model to estimate the frontier production function. This specification allows predicting the corresponding farm level TE. Later, we employ a spatial regression model, that is, SDM, for examining the potential spatial interdependence in the predicted TE scores. The subsequent sections provide a detailed description of each model.

[1] A detailed brief on the sources of endogeneity is presented in Section 2.

TABLE 1 Variable definitions

Variable	Definition
Frontier related variables	
Agricultural output	Value of total agricultural output (ET Birr)
Land	Total land (Hectare)
Labor	Planting season household labor (Days)
Farm equipment	The monetary value of agricultural equipment (ET Birr)
Nitrogen use	The total amount of nitrogen use in the farm (Kilograms)
Phosphorus	The total amount of phosphorus use in the farm (Kilograms)
Rainfall	Average 12-month total rainfall (mm) for January to December
Temperature	Annual Mean Temperature (Degree Centigrade × 10)
Soil type	Nutrient availability (No or Slight Constraint = 1, Moderate Constraint = 2, Severe Constraint = 3, Very Severe Constraint = 4)
Inefficiency related variables	
Age	Age of the household head (years)
Gender	Gender of the household head (Male = 1, Female = 0)
Family size	Household size (number)
Education	Highest grade of the household head education (years)
Market distance	Household distance to nearest market (Kilometers)
Extension services	Dummy for extension services in the household (Yes = 1, No = 0)
Improved seed	Dummy for use of improved seed (Yes = 1, No = 0)
Business Share	The share of business-related income out of the total income (%)
Pesticide use	Dummy if the household uses pesticide (Yes = 1, No = 0)

2.1 | Stochastic frontier analysis

Our analysis begins by specifying a log-transformed stochastic frontier model as follows:

$$y_{it} = \alpha_i + f(x_{it}; \beta) + V_{it} - U_{it} \tag{1}$$

where y_{it} is a scalar output of individual farmers i in period t. The parameter α_i is an intercept that varies across individuals, while $f(x_{it}; \beta)$ is a production function that follows a Cobb-Douglas specification.[2] The term x_{it} denotes the $1 \times K$ vector of exogenous inputs such as land, labor, fertilizer, and the value of farm equipment used by the farmer i in period t.[3] Additionally, the vector includes variables representing temperature, precipitation, and soil types. Table 1 presents detailed descriptions of the output and input variables. The term β is a $K \times 1$ vector of technology parameters for the corresponding inputs. The term, V_{it} represents an $N \times 1$ vector of idiosyncratic error components or noise, while U_{it} an $N \times 1$ vector for the one-sided inefficiency term.

The Maximum likelihood technique is a commonly used method to estimate stochastic frontier production function (e.g., Kumbhakar, 1990; Kumbhakar & Wang, 2005; Kumbhakar et al., 2014; Y. H. Lee & Schmidt, 1993). However, this approach is limited in dealing with the potential endogeneity problem in Equation (1). For instance, a farmer may choose the amount of fertilizer used depending on the level of productivity, which could be determined by the soil quality, weather conditions or other influences on production that are not observed by the researcher. For example, a farmer may have an idea on parts of the stochastic noise components, for example, information on the weather forecast that could influence input choice in the production process. In this context, the *iid* assumption on the two error components in the standard SFA framework could be violated and results in biased and inconsistent estimates. Thus, we believe that the frontier in Equation (1) should accommodate potential endogeneity problems.

Instrumental variable (IV) based estimation of SFA is a reliable approach to address the endogeneity problem (e.g., Amsler et al., 2016; Kutlu, 2010; Tran & Tsionas, 2013). However, this approach is highly dependent on the

[2] We preferred the Cobb-Douglas specification since the likelihood-ratio test against the translog functional form shows that the Cobb-Douglas is not too restrictive (not rejected). Table 3 presents the corresponding test statistics.

[3] We used the real monetary value of farm equipment by multiplying the physical quantity of the equipment with corresponding prices at the community level. This allows us to aggregate diverse farm equipment using monetary value.

availability of valid instruments that need to satisfy the exclusion restriction and the independence assumptions. Moreover, the IV approach requires further assumptions regarding the correct specifications of the second-stage model, which is vital in predicting unbiased TE components.

To ease such challenges, we estimated a Copula-based stochastic frontier model following the works by Amsler et al. (2017), Amsler et al. (2016), Tran and Tsionas (2015), and Karakaplan and Kutlu (2017). This approach accounts for the possible dependency among the endogenous input and the two composed errors in the standard stochastic frontier model. The method also relaxes the *idd* assumption on the two error components in the traditional stochastic frontier analysis framework. Therefore, the Copula function allows for constructing the marginal distributions of the endogenous input and the error terms independently. The joint distribution is used to derive the likelihood function and maximized to obtain consistent and unbiased estimates.

To explain details, suppose a nonnegative one-sided error component $U_{it} = u \in \underline{R}_{+}$ in Equation (1) is a random variable with a cumulative density function $F(u) = \Pr(U_{it} \leq u)$. This term is assumed to be continuous and independent of x_{it}. However, the term U_{it} is possibly dependent on unknown vector of distributional parameters (δ_u). Similarly, the idiosyncratic error term $V_{it} = v \in R$ has a cumulative density function of $G(v) = \Pr(V_{it} \leq v)$ which is assumed to be continuous and independent of x_{it}. However, the term is similarly dependent on unknown vector of parameters (δ_v). Applying Sklar's theorem (see Sklar, 1973), the joint cumulative density function for U_{it} and V_{it} can be represented as:

$$H(u, v) = \Pr(U_{it} \leq u, V_{it} \leq v) = C_\theta(F(u), G(v)) \quad (2)$$

where $C_\theta(F(u), G(v))$ denotes the bivariate copula for the joint distributions of U_{it} and V_{it} that can depend on unknown parameters in vector θ. We assumed that the vector θ does not hold a common element within the vectors β, δ_u and δ_v. As mentioned before, Equation (2) relaxes the *iid* assumption on the two error components in the standard SFA framework (e.g., see Amsler et al., 2016; Shi & Zhang, 2011; Smith, 2008; Tran & Tsionas, 2015).

The log-likelihood function can be derived from the composite distributions of the two error components $W = V_{it} - U_{it}$, where the random variable $W = w| -\infty < w < \infty$ is continuous. The probability density function for $W = h_\theta(w)$ can be derived as follows:

$$h(w) = \frac{\partial^2}{\partial u \partial v} H(u, v) = f(u)\, g(v)\, c_\theta(F(u), G(v)) \quad (3)$$

where $f(u) = \partial F(u)/\partial u$, $g(v) = \partial G(v)/\partial v$, and c_θ is the Copula density for C_θ. Importantly, the weighting function given by $c_\theta(F(u), G(v))$ in Equation (3) controls for any potential dependence between U_{it} and V_{it} in the main model. The probability density function for (U, W) is given by:

$$h(u, w) = f(u)\, g(u + w)\, c_\theta(F(u), G(u + w)) \quad (4)$$

Therefore, the probability density function for $W$is expressed as follows:

$$h_\theta(w) = \int_{\underline{R}_{+}} h(u, w) du = E_U\left[g(U + w)\, c_\theta(F(U), G(U + w))\right] \quad (5)$$

where the right-hand side of Equation (5) represents expectation with respect to the distributions of inefficiency component U. Subsequently, the joint density function in Equation (4) allows for the construction of the likelihood function as follows:

$$L(\beta, \delta_u, \delta_v, \theta) = \sum_i \ln c_c(F(U), G(U + w)) + \sum_i \ln g(U + w) \quad (6)$$

The Copula term c_θ reflects the potential dependence between the two error components U_{it} and V_{it}. Therefore, we used the likelihood function given by Equation (6) to derive the parameter of interests. For robustness check, we also estimated and reported three additional time-varying inefficiency models. These are the Kumbhakar model (Kumbhakar, 1990), the Battese and Coelli model (Battese & Coelli, 1992) and the Kumbhakar and Wang model (Kumbhakar & Wang, 2005). Following Battese and Coelli (1998), the mean technical efficiency score (TE_θ) for individual farmers can be predicted using the following relationship:

$$\begin{aligned} TE_\theta &= E\left[\exp^{(-U)|W=w}\right] \\ &= \frac{1}{h_\theta(w)} \int_{\underline{R}_{+}} \exp(-u) h(u, w)\, du \\ &= \frac{E_U\left[\exp(-U)\, g(U + w)\, c_\theta(F(U), G(U + w))\right]}{E_U\left[g(U + w)\, c_\theta(F(U), G(U + w))\right]} \end{aligned} \quad (7)$$

The TE_θ in Equation (7) will be used as a dependent variable in the spatial model where the identification of neighborhood effect is examined. For notational purposes, we use u for TE score in the remaining sections.

2.2 | Spatial regression analysis

We begin our spatial analysis by specifying the standard non-spatial pooled ordinary least square regression model (OLS). This specification provides information on whether the model should be further extended by a spatial component. Thus, the baseline non-spatial specification is given as:

$$u = \iota_N c + Z\alpha + \xi \tag{8}$$

where u is an $N \times 1$ vector denoting the predicted TE scores for the N observations in the sample from Equation (7), ι is an $N \times 1$ identity matrix for the constant term c, the vector Z represents an $N \times K$ explanatory variables with α as the corresponding $K \times 1$ parameter vector and ξ is a vector of disturbances. It is assumed that the disturbance terms are independently and identically distributed with zero mean and constant variance, σ_ξ^2, that is, $\xi_i \sim N(0, \sigma_\xi^2)$, $i = 1, \ldots, N$.

Implementing the Lagrangian Multiplier (LM) and robust-LM (RLM) tests on the OLS residuals allows one to decide whether to use the specification in Equation (8) or alternative spatial models such as Spatial Autoregressive Model (SAR), Spatial Error Model (SEM) or SDM (Anselin, 1988; Anselin et al., 1996; LeSage & Pace, 2009). Accordingly, rejecting the null hypothesis in the OLS residuals which is would be a confirmation for the presence of spatial correlation and suggests the need to specify an alternative general spatial model. The specification in Equation (8) could be extended into a full spatial specification as follows:

$$\begin{aligned} u &= \iota_N c + \lambda W u + Z\alpha + WZ\pi + \varepsilon \\ &= \rho W\varepsilon + \xi, \xi \sim N(0, \sigma_\xi^2) \end{aligned} \tag{9}$$

where λ represents the endogenous interaction effect among the farmers TE level, π denotes the exogenous interaction effect among the explanatory variables, and ρ stands for the interaction effect among the spatially differentiated stochastic disturbance terms. The parameters, λ and ρ are the spatial autoregressive and the spatial autocorrelation coefficients, respectively, while α is a $K \times 1$ vector of parameters to be estimated.

The vector W is an $N \times N$ matrix showing the extent of spatial interdependence between farms in neighboring communities from our sample. It should be noted that only the off-diagonal elements of the matrix can take non-zero values ($w_{ij} > 0$) indicating that farmers i and j are neighbors. Constructing spatial weight or defining neighborhood is a crucial step prior to the estimation of any spatial models (e.g., Anselin, 2013; Elhorst, 2010; J. P. LeSage & Pace, 2010). In our study, we have used the K-nearest neighbor criterion in defining the spatial weight matrix (see J. P. LeSage & Pace, 2010). Assigning a relative weight based on the number of neighbors at a certain threshold is done by row standardization of the spatial weight matrix where the sum of the row elements equals one (see Anselin, 2001; Pace & Barry, 1997). Such standardization assigns weights that measure the degree of interconnectedness among farmers in the sample.

Recalling the elaboration in Section 1, the specification in Equation (9) distinguishes three potential spatial interaction effects, namely the endogenous, exogenous and contextual effects (Manski, 1993). The endogenous interaction effects denoted by λ captures whether a farmer's TE level is dependent on its neighbor(s) TE level, while the exogenous interaction effect represented by π considers whether a farmers TE level is dependent on the exogenous socioeconomic characteristics of its neighbors. The contextual effect indicted by ρ identifies the effects of common unobserved characteristics that affect the TE level of all farmers.

Estimating the parameters associated with the three interaction effects in Equation (9) usually poses an empirical challenge due to the reflection problem (Brock & Durlauf, 2001; L. Lee, 2007; Manski, 1993). However, it is indicated that identification of these parameters is possible with the consideration of group fixed effects in Equation (9) (e.g., see Elhorst, 2010; Lin, 2010). Therefore, our study follows the LeSage and Pace (2009) SDM specification with the exclusion of spatially autocorrelated error terms ($\rho = 0$).[4] Our SDM is given as follows:

$$u = \iota_N c + \lambda W u + Z\alpha + WZ\pi + \varepsilon, \varepsilon \sim N\left(0, \sigma^2\right) \tag{10}$$

For robustness check, we have implemented a restriction test for SAR and SEM models. Accordingly, accepting the null hypothesis, $H_0: \pi = 0$, would indicate that the SAR model is adequate to describe our data. On the other hand, accepting the null hypothesis, $H_0: \pi + \lambda\alpha = 0$, suggests the SEM model is adequate for describing our data. Rejecting both restrictions implies the SDM given by Equation (10) to fit our data in a better way. Our final model is estimated by employing the ML estimator given its flexibility on the distributional assumption of the disturbance terms (L. F. Lee, 2004; L. Lee, 2007; Pace, 2009).

3 | DATA AND DESCRIPTIVE STATISTICS

This study uses three-period balanced panel data (2011, 2013, and 2015) from the ERSS collected at the household

[4] Setting the parameter $\rho = 0$ offer unbiased coefficient estimates in relation to the spatial lag or spatial error model. In addition, it produces correct standard errors for the estimated coefficients. A more detailed explanation can be found in Elhorst (2010) and LeSage and Pace (2009).

level in Ethiopia (World Bank, 2016). This data set constitutes information on household socioeconomic characteristics, crop and livestock production, farm input use, access to services and resources, employment and farm and non-farm income. The ERSS holds a total number of 3996, 5469, and 5262 households in three waves, that is, year 2011, 2013, and 2015, respectively. The first wave of the survey was mainly conducted in rural and small-town households across 333 unique enumeration areas (EA), while the second and third waves incorporated urban areas and 433 EA's across all regions of the country. After consideration of missing values, outliers, household locations and a balanced data requirement for spatial regression models, the study used data from 5490 rural households, that is, 1800 households in each wave.[5] Table 1 presents the definition of the variables used in the analysis.

There are several advantages in using the ERSS data in Ethiopia. First, the sample is representative as it constitutes the entire regions of the country. These are 10 politically classified regional states in Ethiopia, that is, Tigray, Afar, Amhara, Oromia, Somalia, Benshagul Gumuz, Southern Nations Nationalities and Peoples Regional State (SNNP), Gambella, Harari, and Dire Dawa. This allows us to explore wider and richer information collected from households in different enumeration areas. Second, it allows us to consider regional heterogeneity. Third, households are referenced with a geographical location of their community. This information allows us to exploit the spatial dimension in our analysis. Accordingly, households in the data are distributed across 290 geographically distinguished rural communities, that is, enumeration areas.

Descriptive statistics on the sample used are given in Table 2. The total monetary value of agricultural output represents the revenue from crop and livestock production in the household.[6] The corresponding mean values for 2011, 2013, and 2015 are, respectively, 6407, 7630, and 12,209 ET Birr in real terms.[7] The average values show an increasing pattern over the study period which is in line with the observed growth in the agricultural sector (e.g., see Bezawagaw et al., 2018). Summary of statistics on input use in households are also presented. For instance, the average household landholding is 1.2 hectare, while planting season labor use in a year is 120 days. The average real value of farm equipment is about 392.7 ET Birr which shows the low level of capital use in the sector. On average, farmers also use 24 and 13 kilograms of nitrogen and phosphorous, respectively. This figure is relatively smaller than the average fertilizer use for the Sub-Saharan African region (Holden, 2018). Regarding demographic characteristics, the average household head is 46 years old and has 3.7 years of schooling, while family size is 5.4 persons per household.

The overall farming sector in Ethiopia is characterized by a smallholder farming practices with low levels of productivity. The literature identifies several factors as the root causes of low agricultural productivity (e.g., Dorosh & Minten, 2020; Mekonnen et al., 2018). Some of them include inadequate input usage (such as improved seed varieties and fertiliser), poor linkages between the market and the farming sector, backward technological setups coupled with diminishing cultivated land size, and backward institutional setup. Our data also complement this situation where farmers are influenced by input and market constraints.

4 | RESULTS

4.1 | Frontier production model

Table 3 presents estimates of the stochastic frontier model specified in Equation (1). As baseline specifications, three time-varying stochastic frontier models are included in the table for comparison, namely, the Kumbhakar (1990) model in (1), the Battese and Coelli (1992) model in (2), and Kumbhakar and Wang (2005) model in (3). These models will be compared and will provide insight into the validity of our preferred specification. Our estimates based on the copula-based stochastic frontier model are presented in column (4).

To verify whether there is an endogeneity problem, we have implemented the Durbin–Wu–Hausman test for endogeneity in each of the three baseline models. The corresponding test statistics confirmed the prevalence of endogeneity in the baseline models, that is, the coefficient for the predicted residuals becomes statistically significant once it is considered in the respective frontier models. Therefore, we extend our analysis based on the Copula-based frontier model, which is robust to potential endogeneity problems. A likelihood-ratio (LR) test was used to compare the model without technical inefficiency (OLS) against the Copula-based stochastic frontier model. The corresponding LR test rejects the null hypothesis of no technical inefficiency at the 99% level.

[5] Our analysis entirely focused only on those households living in rural areas. Also, the spatial regression models need balanced panel data to construct an invertible weighting matrix (Elhorst, 2010).

[6] We use the Total Livestock Unit (LTU) coefficients for Sub-Saharan Africa region to calculate the monetary value of livestock production (for details see FAO, 2011) as the dependent variable in the production frontier model. This approach allows aggregation of livestock value using specific coefficients that are established based on the types of animal, age and nutritional requirement.

[7] 1 ET Birr ≈ 0.04 USD based on the average exchange rate for the year 2015. The nominal values are adjusted to real term using the consumer price index where the 2011 price is taken as a base year, that is, 2011=100.

TABLE 2 Summary of the descriptive statistics

Variable	Whole panel			Year = 2011			Year = 2013			Year = 2015		
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Agricultural output	8748.6	16.4	14 1565	6 407	30.6	14 1565	7630	19.4	89686	12208.6	16.4	75837.3
Land	1.21	.019	7.227	1.068	.019	3.499	1.255	.019	7.227	1.308	.019	7.23
Labor	120	0	2 304	133.36	0	2 273	108.12	0	2 304	118.6	0	1 802
Farm equipment	392.7	0	3 447.5	742.01	10	3 447.48	126.37	0	924.96	126.72	0	883.9
Nitrogen use	24.2	0	1 334	11.64	0	416.7	17.52	0	1 334	43.52	0	1 334
Phosphorus	12.59	0	180	11.63	0	180	14.14	0	180	12	0	180
Family size	5.415	1	17	5.363	1	14	5.454	1	15	5.429	1	17
Gender	.826	0	1	.841	0	1	.825	0	1	.812	0	1
Education	3.715	1	13	3.507	1	13	3.944	1	13	3.72	1	13
Age	46.06	0	97	44.46	18	97	45.9	0	97	47.82	0	97
Market distance	65.88	2.9	222	65.88	2.9	221.8	65.9	2.9	221.8	65.85	3	222
Business Share	.133	0	1	.296	0	1	.225	0	1	.063	0	1
Extension services	.371	0	1	.302	0	1	.398	0	1	.412	0	1
Pesticide use	.088	0	1	.094	0	1	.078	0	1	.093	0	1
Improved seed	.211	0	1	.186	0	1	.223	0	1	.225	0	1
Rainfall	939.65	290	1692	908.93	290	1600	964.39	345	1636	945.62	294	1692
Temperature	188.87	102	289	188.94	106	289	188.84	102	289	188.83	109	289
Soil type	1.38	1	4	1.38	1	4	1.38	1	4	1.37	1	4
Enumeration area (EA)	147.9	1	281	147.9	1	281	147.9	1	281	147.9	1	281
Observation	5 490			1 830			1 830			1 830		

The input variables, including land, labor, and nitrogen are positive and statistically significant at the 99% level. These results imply that labor use, besides land and nitrogen, has a positive effect on agricultural production. The results also indicate that the mean annual rainfall and temperature have statistically significant effects, having positive and negative coefficients, respectively. The estimated coefficient for the effect of time trend represented by the Year variable is positive and statistically significant at 99% level and suggests that the rate of technical progress amounted to about 3% every second year (around 1.5% annually).

The mean of estimated TE scores is 53% implying the there is a great potential to increase agricultural productivity in Ethiopia by eliminating inefficiency. Additionally, the production technology exhibits a decreasing return to scale given the sum of the input coefficient estimates is less than unity, that is, $\sum \beta_i < 1$. The evidence of decreasing returns to scale could be related to overutilization of the prominent inputs such as labor or other constraints to production. Low returns to inputs inflates the cost of output expansion. This can trigger farmers to invest more in inputs compared to what they received. Thus, farmers will have to decrease the amount of input used to reach the point where the cost per unit of inputs is equal to returns per unit.

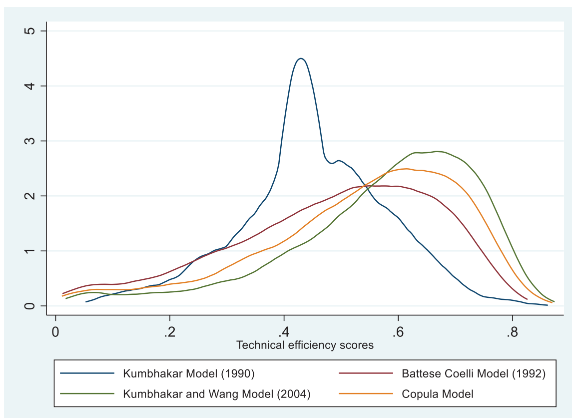

FIGURE 1 Density plots of estimated technical efficiency (TE) scores from alternative models

The distributions of TE scores from the four models are presented in Figure 1. The TE scores from the Battese and Coelli (1992), Kumbhakar and Wang (2005), and the Copula-based models exhibit similar patterns with a slightly right-skewed smooth distribution. In contrast, the

TABLE 3 Estimates for the frontier production models

Frontier estimates	The dependent variable is log of total agricultural output (1)	(2)	(3)	(4)
Land	.240*** (.066)	.191*** (.065)	.201*** (.064)	.230*** (.065)
Labor	.104*** (.023)	.084*** (.023)	.083*** (.023)	.094*** (.023)
Nitrogen	.087*** (.024)	.069*** (.023)	.071*** (.023)	.064*** (.024)
Phosphorus	.014 (.024)	−.010 (.023)	−.010 (.023)	.003 (.023)
Farm equipment	.056*** (.019)	.040** (.019)	.041** (.019)	.028 (.019)
Rainfall	.625*** (.098)	.531*** (.091)	.538*** (.091)	.534*** (.097)
Temperature	−.783*** (.173)	−.502*** (.164)	−.520*** (.163)	−.525*** (.172)
Soil type[11]	−.021 (.053)	−.042 (.050)	−.041 (.049)	−.020 (.053)
Year	.034*** (.003)	.030*** (.003)	.030*** (.003)	.030*** (.003)
Constant	7.110*** (1.088)	6.902*** (1.016)	6.853*** (1.013)	6.825*** (1.065)
Inefficiency determinants				
Family size	−4.665*** (1.589)	−2.988* (1.739)	−.442*** (.095)	−.818*** (.181)
Gender	−.542 (.852)	−.684 (.672)	−.082 (.085)	−.111 (.175)
Education	−1.287** (.632)	−.931 (.644)	−.113** (.047)	−.170* (.095)
Age	.647 (.564)	−1.414 (1.031)	−.215** (.103)	−.349* (.207)
Market distance	2.784*** (.797)	.205 (.262)	.033 (.037)	.050 (.075)
Business share	2.973** (1.252)	2.722 (1.762)	.339*** (.130)	.627** (.254)
Extension services	3.060** (1.356)	.027 (.577)	.008 (.075)	−.005 (.147)
Pesticide use	−2.800*** (1.021)	−6.988 (5.303)	−.743*** (.187)	−1.366*** (.367)
Improved seed	−3.821** (1.535)	−2.140 (1.451)	−.263*** (.090)	−.493*** (.175)
Mean Efficiency score	.47	.48	.57	.53
Endogeneity test (*F-test*)	6394.71***	3037.96***	2868.23***	
Likelihood-ratio (LR) test	−4407.79***	−4405.27***	−4404.15***	−4375.44***
Observation	5490	5490	5490	5490

Note: Standard errors in parentheses. Results in (1) denotes the Kumbhakar (1990) estimates, results in (2) represent the Battese and Coelli (1992) estimates, results in (3) represent Kumbhakar and Wang (2005) specification, and results in (4) represent Copula based specification. The output and input variables are in log form.
$^{***}p < .01$.
$^{**}p < .05$.
$^{*}p < .1$.

Kumbhakar (1990) model reveals a potential misspecification error as its TE distribution is neither right-skewed nor smooth compared to other distributions. The estimates of the inefficiency determinants, such as family size, educational level, distance to the market, and extension services, also have relatively larger magnitudes in for the Kumbhakar (1990) model.[8] This also highlights the potential misspecification error in the Kumbhakar (1990) model.

Figure 2 illustrates the spatial distributions of TE score across administrative zones in Ethiopia. Accordingly, there is a substantial inter-regional variation. Such spatial discrepancies in farmers' TE level across regions could be linked to farmer's socioeconomic heterogeneity between

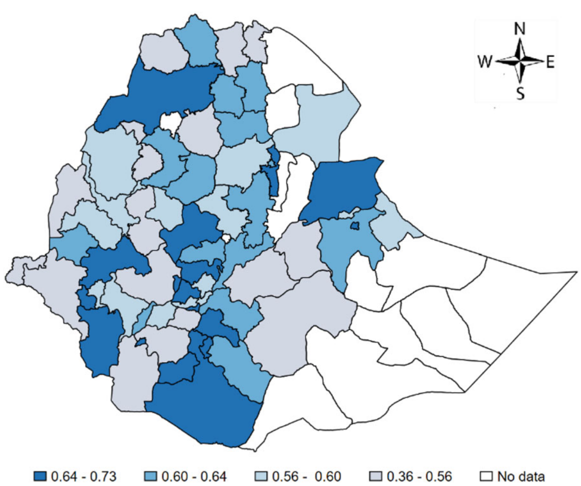

FIGURE 2 Regional distributions of the average technical efficiency (TE) scores in Ethiopia

[8] The coefficient estimates for the inefficiency determinants in Table 3 cannot be considered as marginal effects. This is due to the non-linear relationship between the predicted TE and its determinants (see Kumbhakar & Lovell, 2000). For a meaningful comparison, we have computed the marginal effects and presented them under Table A1 in the appendix.

TABLE 4 Spatial autocorrelation tests

K – nearest neighbor	Moran's I statistics			Geary's C statistics		
	Statistics	Std. Dev.	*p*-Value	Statistics	Std. Dev.	*p*-Value
K = 1	.329	21.772	.000	.655	12.488	.000
K = 3	.319	33.515	.000	.645	23.938	.000
K = 5	.330	42.215	.000	.644	33.400	.000
K = 7	.319	46.804	.000	.658	40.235	.000
K = 10	.285	49.583	.000	.703	41.522	.000

and within regions. There has been evidence that cash crop farmers are more likely to be technically efficient than subsistence farmers (Tirkaso & Hess, 2018). In this regard, the higher TE score in the north western, south western and eastern parts of the country could be linked to the dominance of cash crop production in these regions (e.g., oilseeds and coffee).

4.2 | Technical inefficiency model

Our examination of the neighborhood effects started by implementing the Global Moran's I (Moran, 1948) and Geary's C local spatial autocorrelation tests on the TE model (Geary, 1954). These tests use data on efficiency scores and household locations to determine whether observed efficiency distribution patterns expressed are clustered, dispersed, or random. Both measures require a spatial weight matrix and the K-nearest neighbor algorithm (*K-NN*) was used to define Euclidean distance-based spatial interdependence.[9] The null hypothesis for these tests is that TE scores are randomly distributed between farms in neighboring communities, that is, there is no spatial interdependence.

The test results are presented in Table 4. Accordingly, the null hypothesis is rejected at different K- level indicating the existence of spatial interdependence in TE score between farms in neighboring communities. For robustness purposes, the tests are implemented at five alternative K-nearest neighbor levels and the results are consistent across the K levels. There is significant spatial autocorrelation in farmer efficiency scores.

Additionally, it is possible to trace and reaffirm the presence of spatial interdependence using residual from the pooled OLS estimate (Elhorst, 2010; Ertur & Koch, 2007). The procedure also provides a baseline estimate for the determinants of TE scores. Therefore, we have implemented OLS residuals-based spatial autocorrelation test which is based on the Lagrange Multiplier (LM) test statistics as suggested by Anselin et al. (1988) and robust Lagrange Multiplier (RLM) test statistics as suggested by Anselin et al. (1996). The test results are presented under Table A2 in the online appendix. The null hypotheses of no spatial autocorrelation are rejected at the 99% significance level in all cases. The result endorses the presence of spatial interdependence on the TE level. Moreover, all variables have the expected sign with key variables including family size, education, age, market distance, pesticide use, and improved seed use being statistically significant at the 99% level. It should be noted that these estimates are potentially inconsistent due to omitted variable bias and the exclusion of spatial dimension.

Although the SDM model presented in Equation (10) is our preferred specification, it is recommended that estimates are compared with those from the SAR and SEM models for robustness purposes (Elhorst, 2010; LeSage & Pace, 2009). Therefore, we have estimated the two models using the spatial weight matrix for the 10-nearest neighbor. The corresponding results are presented in the online appendix (Table A3). The spatial autoregressive coefficient, ρ, is positive and statistically significant in the SAR model case confirming the endogenous lag variable effect. Meanwhile, the spatial autocorrelation coefficient, ρ is statistically insignificant under the SEM model with the household fixed effects are controlled. These results generally suggest estimation of the spatial model.

The spatial autoregressive coefficients have a similar sign across the three specifications (SAR, SEM, or SDM models) with strong evidence of spatial interdependence. The reported LM and RLM tests also provide additional decision guidelines to identify the preferred model among the three.[10] The test statistics favors the SDM where the null hypotheses for both SAR and SEM models are rejected

[9] In the context of this study, the social network or social structure of farmers is represented by the Euclidean distance between farms in neighboring communities that allows the exchange of ideas, norms and resources in closest proximity.

[10] If the OLS model is rejected in favor of SAR, SEM, or in favor of both models, the SDM model should be estimated. Subsequently, the likelihood ratio (LR) test is applied to test the hypotheses $H_0 : \pi = 0$ (to check whether the SDM can be simplified to the SAR model) and $H_0 : \pi + \lambda\alpha = 0$ (to check whether Equation (8) can be simplified to the spatial error model) in Equation (8). Note that both tests are based on the OLS residuals and follow a χ^2 distribution with one degree of freedom.

[11] Soil type category with a "Very Severe Constraint" is used as a reference.

AGRICULTURAL ECONOMICS
The Journal of the International Association of Agricultural Economists

TABLE 5 Spatial Durbin model estimates (SDM)

	The dependent variable is TE			Marginal effects		
Variable	**(1)**	**(2)**	**(3)**	**Direct**	**Indirect**	**Total**
Family size	.075*** (.021)	.257*** (.019)	.074*** (.021)	.267***	.263***	.530***
Gender	−.035 (.026)	.075*** (.019)	−.034 (.026)	.078***	.077***	.155***
Education	.001 (.009)	.018* (.009)	−.001 (.009)	.019*	.019*	.038*
Age	.006 (.009)	.010 (.011)	.006 (.009)	.010	.010	.020
Market distance	−.312* (.185)	−.039* (.020)	−.308* (.186)	−.040*	−.040*	−.080*
Business share	−.143*** (.030)	−.275*** (.039)	−.143*** (.030)	−.286***	−.282***	−.568***
Extension services	.012 (.012)	.044*** (.016)	.012 (.012)	.046***	.045***	.091***
Pesticide use	.056*** (.017)	.170*** (.025)	.056*** (.017)	.176***	.174***	.350***
Improved seed	.035** (.014)	.052* (.019)	.035** (.014)	.055*	.054*	.109*
Rainfall	.007 (.092)	−.109 (.072)	−.077 (.137)	−.113	−.112	−.225
Temperature	.284 (.207)	−.057 (.083)	.290 (.207)	−.059	−.058	−.117
Soil type	.092** (.043)	−.069** (.027)	.090** (.043)	−.071**	−.070**	−.142**
$W\times$ Family size	−.079 (.060)	.001 (.040)	−.079 (.060)			
$W\times$ Gender	−.098* (.056)	−.120*** (.039)	−.098* (.056)			
$W\times$ Education	.055* (.031)	.035* (.019)	.055* (.031)			
$W\times$ Age	.066 (.052)	.020 (.030)	.067 (.052)			
$W\times$ Market distance	.018 (.028)	.040* (.022)	.018 (.028)			
$W\times$ Business share	.155 (.125)	.134* (.071)	.154 (.125)			
$W\times$ Extension services	−.013 (.063)	−.017 (.029)	−.013 (.063)			
$W\times$ Pesticide use	.067 (.092)	.137*** (.047)	.065 (.092)			
$W\times$ Improved seed	.098 (.072)	.040 (.035)	.097 (.072)			
$W\times$ Rainfall	−.061 (.107)	.196** (.079)	−.061 (.107)			
$W\times$ Temperature	−.186 (.135)	−.056 (.098)	−.186 (.135)			
$W\times$ Soil type	.106** (.046)	.088*** (.032)	.107** (.046)			
Spatial lagged dependent variable (λ)	.405*** (.021)	.515*** (.017)	.404*** (.021)			
Year fixed effects	No	YES	YES			
Household fixed effects	YES	YES	NO			
Diagnostic tests (Wald test)						
SAR vs. SDM	25.549**	37.989***	25.680**			
SEM vs. SDM	25.549**	37.989***	25.680**			
AIC	7271.307	2941.527	2944.435			
Observation	5490	5490	5490			

Note: Values in the parenthesis are standard errors. Estimates are based on the spatial weight matrix constructed at the 10-nearest neighbor level.

$^{***}p < .01$.

$^{**}p < .05$.

$^{*}p < .1$.

at a 99% significant level. The Akaike information criterion (AIC) suggests Model 2 as a better specification. To check whether the estimated spatial dependence parameter is sensitive to the selected weight matrix, we estimated Model 2 using different K-nearest neighbor classifications. Results from these sensitivity tests are presented under Table A4 in the online appendix.

Table 5 presents the final SDM model estimates which are based on Equation (10). This specification contains the spatial lag of all the explanatory variables and thus the coefficients need to be interpreted in the right way. Accordingly, the effects are decomposed into a direct effect, indirect effect and total effect. The direct effect reflects a change in farmer's TE score, u, due to change in the farmer's own exogenous variables, while the indirect effect shows a change in u as a result of a change in the farmer's neighbor exogenous characteristics. The total effects represent the sum of direct and indirect effects.

We further estimated the SDM model in three different specifications. In Model 1, we have controlled only for household fixed effects, while both household and time effects are controlled in Model 2. The year effects are included in Model 3. Consideration of household level fixed effects would accommodate the effects of any potential household level omitted variables. Similarly, the time fixed effects capture those unobserved time-variant factors that are common to all households in our sample. Having these specifications allows us to choose the most complete specification that satisfies both the theoretical and statistical properties.

The marginal effects associated with the direct, indirect and total effects of the explanatory variables are computed and presented in Table 5. Accordingly, two important interaction effects, that is, the endogenous and contextual interaction effects elaborated by LeSage and Pace (2009); Elhorst (2010) can be explained. The spatial autoregressive coefficient, λ that represents the endogenous effect is statistically significant at the 99% level with a magnitude of .52. This suggests that a farmer's TE level depends on its neighbors' relative TE levels. Under contextual interaction effects, it is found that family size, gender, age and educational levels of the household head and pesticide use are statistically significant at least at 99% level. The same variables have indirect impacts (spillover effects). As elaborated in sub Section 2.2, the SDM model provides sufficient information for the identification of endogenous and exogenous effects by controlling the household level fixed effects that could confound the neighborhood effect.

5 | CONCLUSIONS AND POLICY IMPLICATIONS

This study estimated farmers TE level in Ethiopia, and investigated influences on efficiency including spatial interdependence or neighborhood effects. Our empirical strategy started by estimating a time-varying stochastic frontier model that allows to predict farmers efficiency scores (TE) and then estimating a SDM to investigate the spatial interdependence of TE level between farms in a neighboring communities. The application of this approach provides empirical evidence on the existence of spatial interdependence and an indication for identifying better policies towards efficiency improvement.

The result indicates that there is evidence of significant neighborhood effects on farmer's TE scores. Specifically, we found that neighborhood plays a key role to influence farmer's TE levels through peer effects coming from social interaction. Further, the contextual interaction effects represented by neighbor's specific socioeconomic characteristics such as peer educational level are found to be vital in shaping the farmer's managerial skill set.

The main contributions of our study can be seen from the following standpoints. First, our results suggest that disregarding the spatial dimension in our efficiency analysis would result in biased estimates. Second, we addressed the reflection problem by distinguishing both the endogenous and contextual social interaction effects. To the best of our knowledge, previously related studies on this matter did not reconcile this problem and that could make their estimates disputable. Third, our study shows that efficiency clusters exist at sub-regional levels reflecting local farmer interaction effects as well as local resource use patterns (e.g., market orientation) which is indicated by some of the parameter estimates in our final model. For instance, it is shown that own and neighbor farmer pesticide use and education levels have positive direct and indirect effects on efficiency. Efficiency is also found to be positively related to the household size. Further, the level of efficiency is found to be low over all, averaging about 53%, and varies greatly both within and across administrative regions.

The key results proved that policymakers should develop spatially targeted policies and programs that improve farmers' TE level. Notably, access to education, market, extension service, pesticide, and improved seed use are not only imperative to the individual farmer TE level but are beneficial to the neighboring farmers. This demonstrates the role of spatial spillover effects in improving farmers' TE level through the exchange of skills, knowledge, and other related resources. This study also has vital policy implications for optimal public investment in TE-enhancing projects. In this aspect, the investment decisions may be prioritized based on locations that could maximize the spatial spillover effects. Overall, our study points out that peer effects need to be considered carefully acknowledging that their effects might be more complex than normally expected (e.g., peer effects on technology adoption). Specifically, it suggests the need to evaluate the benefit (effectiveness) of efficiency targeting agricultural programs with respect to spatial dimensions and the usefulness of local empirical evidence to improve such programs.

ACKNOWLEDGEMENTS

We are grateful to Prof. Abdulai (the editor in charge) and the two anonymous referees for their valuable critics, comments, and suggestions, which helped to improve the manuscript. We also want to thank Torben Mideksa (Uppsala University), Abenezer Zeleke (Södertorn University), and George Marbuah (Stockholm Environment Institute) for their general feedbacks in the earlier version of the manuscript.

REFERENCES

Abdulai, A., & Huffman, W. (2014). The adoption and impact of soil and water conservation technology: An endogenous switching regression application. *Land Economics*, *90*(1), 26–43.

Aker, J. C., & Mbiti, I. M. (2010). Mobile phones and economic development in Africa. *Journal of Economic Perspectives*, *24*(3), 207–232.

Amsler, C., Prokhorov, A., & Schmidt, P. (2016). Endogeneity in stochastic frontier models. *Journal of Econometrics*, *190*(2), 280–288.

Amsler, C., Prokhorov, A., & Schmidt, P. (2017). Endogenous environmental variables in stochastic frontier models. *Journal of Econometrics*, *199*(2), 131–140.

Anselin, L. (1988). Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity. *Geographical Analysis*, *20*(1), 1–17.

Anselin, L., Bera, A. K., Florax, R., & Yoon, M. J. (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics*, *26*(1), 77–104.

Anselin, L. (2001). Spatial econometrics. in: (B. Baltagi Ed.), *A companion to theoretical econometrics*, Basil Blackwell, Oxford, *pp*. 310–330

Anselin, L. (2013). Spatial econometrics: Methods and models. *Studies in operational regional science*, Springer. https://doi.org/10.1007/978-94-015-7799-1

Bandiera, O., & Rasul, I. (2006). Social networks and technology adoption in northern Mozambique. *The Economic Journal*, *116*(514), 869–902.

Battese, G. E., & Coelli, T. J. (1992). Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India. *Journal of Productivity Analysis*, *3*(1), 153–169.

Battese, G. E., & Coelli, T. J. (1995). A model for technical inefficiency effects in a stochastic frontierproduction function for panel data. *Empirical economics*, *20*(2), 325–332.

Bezawagaw, G., Dihel, N. C., Geiger, M. T., & Getachew, K. Z. (2018). *Ethiopia economic update: The inescapable manufacturing services nexus*., World Bank Group. https://doi.org/10.1596/29919

Boucher, V., Bramoullet́, Y., Djebbari, H., & Fortin, B. (2014). Do peers affect student achievement? Evidence from Canada using group size variation. *Journal of Applied Econometrics*, *29*(1), 91–109.

Bramoullet́, Y., Djebbari, H., & Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics*, *150*(1), 41–55.

Brock, W. A., & Durlauf, S. N. (2001). Interactions-based models. In *Handbook of econometrics*, Elsevier Sciences, volume 5, pages 3297–3380.

Conley, T. G., & Udry, C. R. (2010). Learning about a New Technology: Pineapple in Ghana. *American Economic Review*, *100*(1), 35–69.

Dorosh, P. A., & Minten, B. (2020). *Ethiopia's agrifood system: Past trends, present challenges, and future scenarios*. International Food Policy Research Institute.

Elhorst, J. P. (2010). Applied spatial econometrics: Raising the bar. *Spatial Economic Analysis*, *5*(1), 9–28.

Ertur, C., & Koch, W. (2007). Growth, technological interdependence and spatial externalities: Theory and evidence. *Journal of Applied Econometrics*, *22*(6), 1033–1062.

FAO, Food and Agricultural Organization. (2011). Guidelines for the preparation of livestock sector reviews. Animal Production and Health Guidelines. No. 5. Rome.

Geary, R. C. (1954). The contiguity ratio and statistical mapping. *The Incorporated Statistician*, *5*(3), 115–146.

Genius, M., Koundouri, P., Nauges, C., & Tzouvelekas, V. (2014). Information transmission in irrigation technology adoption and diffusion: Social learning, extension services, and spatial effects. *American Journal of Agricultural Economics*, *96*(1), 328–344.

Greene, W. (2005). Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of econometrics*, *126*(2), 269–303.

Holden, S. T. (2018). Fertilizer and sustainable intensification in sub-Saharan Africa. *Global Food Security*, *18* (2018), 20–26.

Hsieh, C., & Lee, L. F. (2016). A social interactions model with endogenous friendship formation and selectivity. *Journal of Applied Econometrics*, *31*(2), 301–319.

Karakaplan, M. U., & Kutlu, L. (2017). Endogeneity in panel stochastic frontier models: An application to the Japanese cotton spinning industry. *Applied Economics*, *49*(59), 5935–5939.

Karlan, D., Osei, R., Osei-Akoto, I., & Udry, C. (2014). Agricultural decisions after relaxing credit and risk constraints. *The Quarterly Journal of Economics*, *129*(2), 597–652.

Kassie, M., Jaleta, M., Shiferaw, B., Mmbando, F., & Mekuria, M. (2013). Adoption of interrelated sustainable agricultural practices in smallholder systems: Evidence from rural Tanzania. *Technological Forecasting and Social Change*, *80*(3), 525–540.

Krishnan, P., & Patnam, M. (2014). Neighbors and extension agents in Ethiopia: Who matters more for technology adoption? *American Journal of Agricultural Economics*, *96*(1), 308–327.

Kumbhakar, S. C. (1990). Production frontiers, panel data, and time-varying technical inefficiency. *Journal of Econometrics*, *46*(1), 201–211.

Kumbhakar, S. C., Lien, G., & Hardaker, J. B. (2014). Technical efficiency in competing panel data models: A study of Norwegian grain farming. *Journal of Productivity Analysis*, *41*(2), 321–337.

Kumbhakar, S. C., & Lovell, C. K. (2000). *Stochastic frontier analysis*. Cambridge University Press.

Kumbhakar, S. C., & Lovell, C. K. (2003). *Stochastic frontier analysis*. Cambridge University Press.

Kumbhakar, S. C., & Wang, H. J. (2005). Estimation of growth convergence using a stochastic production frontier approach. *Economics Letters*, *88*(3), 300–305.

Kutlu, L. (2010). Battese-Coelli estimator with endogenous regressors. *Economics Letters*, *109*(2), 79–81.

Lee, L. (2007). Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. *Journal of Econometrics*, *140*(2), 333–374.

Lee, L. F. (2004). Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica*, *72*(6), 1899–1925.

Lee, Y. H., & Schmidt, P. (1993). *A production frontier model with flexible temporal variation in technical efficiency*. The measurement of productive efficiency: Techniques and Applications, 237–255.

LeSage, J., & le, R. K. (2009). Introduction to spatial econometrics. *International Statistical Review*, *77*(3), 472–472.

LeSage, J. P., & Pace, R. K. (2010). Spatial econometric models. In *Handbook of applied spatial analysis*, pages 355–376. Springer.

Lin, X. (2010). Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservable. *Journal of Labor Economics*, *28*(4), 825–860.

Maertens, A., & Barrett, C. B. (2013). Measuring social networks' effects on agricultural technology adoption. *American Journal of Agricultural Economics*, *95*(2), 353–359.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, *60*(3), 531–542.

Manski, C. F. (2000). Economic analysis of social interactions. *Journal of Economic Perspectives*, *14*(3), 115–136.

Mekonnen, D. A., Gerber, N., & Matz, J. A. (2018). Gendered social networks, agricultural innovations, and farm productivity in Ethiopia. *World Development*, *105*(2018), 321–335.

Moran, P. A. P. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society: Series B (Methodological)*, *10*(2), 243–251.

Pace, R. K., & Barry, R. (1997). Sparse spatial autoregressions. *Statistics & Probability Letters*, *33*(3), 291–297.

Shi, P. & Zhang, W. (2011). A copula regression model for estimating firm efficiency in the insurance industry. *Journal of Applied Statistics*, *38*(10), 2271–2287, https://doi.org/10.1080/02664763.2010.545376

Sacerdote, B. (2001). Peer effects with random assignment: Results for Dartmouth roommates. *The Quarterly Journal of Economics*, *116*(2), 681–704.

Sklar, A. (1973). Random variables, joint distribution functions, and copulas. *Kybernetika*, *9*(6), 449–460.

Smith, M. D. (2008). Stochastic frontier models with dependent error components. *The Econometrics Journal*, *11*(1), 172–192.

Tirkaso, W., & Hess, S. (2018). Does commercialisation drive technical efficiency improvements in Ethiopian subsistence agriculture? *African Journal of Agricultural and Resource Economics*, *13*(311-2018-2942), 44–57.

Tran, K. C., & Tsionas, E. G. (2013). GMM estimation of stochastic frontier models with endogenous regressors. *Economics Letters*, *118*(1), 233–236.

Tran, K. C., & Tsionas, E. G. (2015). Endogeneity in stochastic frontier models: Copula approach without external instruments. *Economics Letters*, *133*(2018), 85–88.

World Bank, W. B. (2016). Ethiopia - Living Standards Measurement Study (LSMS). The World Bank Group. Available at http://www.nber.org/papers/w14434

Yamauchi, F. (2007). Social learning, neighborhood effects, and investment in human capital: Evidence from green-revolution India. *Journal of Development Economics*, *83*(1), 37–62.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Tirkaso W., & Hailu, A. (2022). Does neighborhood matter? Spatial proximity and farmers' technical efficiency. *Agricultural Economics*, *53*,374–386. https://doi.org/10.1111/agec.12702