



Do *Egocentric* information networks influence technical efficiency of farmers? Empirical evidence from Ghana

Sadick Mohammed ¹ · Awudu Abdulai ¹

Accepted: 2 July 2022 / Published online: 16 July 2022
© The Author(s) 2022

Abstract

We investigate the impact of farmers' *egocentric* information network on technical efficiency and its distribution in the network, using observational data of 600 farmers from northern Ghana. We exploit community detection algorithms to endogenously identify homogeneous network communities with known structures to account for spatial heterogeneity, in a spatial stochastic frontier model that controls for social selection bias. The empirical results reveal that at the global network level, farmers' technical efficiency strongly correlate with that of farmers in their *egocentric* networks. Our findings also show that farmers who are technically less efficient tend to depend on the more efficient farmers in their networks to improve efficiency. We further find that estimating spatial dependence of technical efficiency without accounting for spatial heterogeneity can lead to underestimation of technical efficiency of high (efficiency score >0.6) performing farmers, while overestimating that of medium (efficiency scores between 0.36–0.5) and low (efficiency scores between 0.1–0.35) performing farmers. The findings suggest that identifying central farmers in *egocentric* networks and improving their technical knowledge in a farmer-to-farmer extension organization, can contribute to improving the productivity of many farmers.

Keywords Egocentric information network · Stochastic frontier analysis · Spatial heterogeneity · Technical efficiency · Technology adoption

JEL C45 · D83 · D85 · O13 · Q16

1 Introduction

In many industries, the production managers' first point of contact for information on new production technologies is through their personal acquaintances, friends, and networks in the same or similar industry (Chatterjee and Dutta 2016). Personal information networks¹ (or *egocentric* information

networks) serves as an important information channel for diffusion and adoption of new production technologies (Abdul Mumin and Abdulai 2022). The level of knowledge of people who learn from their personal information networks is said to be correlated with the knowledge level of persons from whom they learn (Fafchamps et al. 2021). Yet, the influence of personal information networks on firm productivity, in terms of technical (in)efficiency, appears to be overlooked in the literature. In agricultural production, inadequate information on innovative agricultural technologies continue to be a major constraint and the jinx to low technology adoption among smallholder farmers in developing countries (De Janvry et al. 2017; Foster and Rosenzweig 2010; Suri 2011). The low technology adoption among farmers has been identified as one of the root causes of low productivity and high poverty incidence among smallholder farmers (Takahashi et al. 2020; Abdulai and Huffman 2014). Weak and ineffective extension services due to resource constraints to employ and equip extension agents to serve the needs of the widely dispersed

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1007/s11123-022-00643-2>.

✉ Sadick Mohammed
smohammed@food-econ.uni-kiel.de

¹ Department of Food Economics and Consumption Studies, University of Kiel, Kiel, Germany

¹ An *egocentric* network refers to a network consisting of well-identified focal persons (known as the *egos*) and members within their personal contacts or networks (known as the *alters*) (Yen et al. 2016).

smallholder farmers who constitute majority of the farming population in many developing countries, accounts for the inadequate information (Mohammed and Abdulai 2022; Blum and Szonyi 2011).

The use of farmers' personal information networks is viewed as a potential information channel to leverage the limited number of extension agents to aid in the diffusion of information on improved technologies to farmers (Beaman and Dillon 2018; Valente 1996). The process of using personal information networks to diffuse information about new technologies or products in order to accelerate adoption, or improve organizational performance is described in the literature as the network interventions approach (Valente 2012). One strategy of the network interventions approach that has become popular among development practitioners and organizational managers is the segmentation strategy. This strategy is cost-effective and efficient, as it relies on passing the new information to an identified group of persons who act as change agents rather than trying to reach individual farmers (Fafchamps et al. 2021; Valente 2012). A major area of application of the segmentation strategy of the network intervention approach in agricultural development is the use of the lead farmer concept in peer-to-peer agricultural extension delivery.

The potential of farmers' networks to diffuse information about new technologies, due to social learning have been extensively explored in the literature within the last decade (e.g., Bandeira and Rasul 2006; Conley and Udry 2010; Kondylis et al. 2017; Beaman and Dillon 2018). For instance, Kondylis et al. (2017) found that contact farmers' personal information networks played a significant role in the diffusion and adoption of sustainable land management practices among farmers in central Mozambique. However, evidence suggest that there is strong correlation of knowledge gain on a given technology by farmers who learn from their personal information networks with those from whom they learn (Fafchamps et al. 2021). The observed correlation between farmers' knowledge with their personal information networks suggests that the likelihood of farmers' technical (in)efficiency being correlated with that of the peers from whom they learn could be equally high. This may be due to factors such as informational inadequacies, willingness to share information, common shocks, and differing absorptive capacity among individual farmers (Kondylis et al. 2017; Bochma 2005).

A number of recent studies have considered the potential correlation of technical efficiencies between contagious production units in the literature. For example, among neighboring electricity and chemical firms (e.g., Orea and Álvarez 2019; Kutlu et al. 2020), provincial and regional administrative units (e.g., Tsionas and Michaelides 2016; Gude et al. 2018; de Graaff 2020), airports and transportation terminals (e.g., Pavlyuk 2019), sport teams (e.g.,

Horrace and Jung 2018) and wine industries (e.g., Fusco and Vidoli 2013; Vidoli et al. 2016). In agricultural production, the influence of geographical and economic proximity on farmers' efficiency have also been considered. Some earlier studies of spatial dependence of technical efficiency in agriculture include, Druska and Horrace (2004) study on rice farmers in Indonesia, Schmidt et al. (2009) on regional farms in Brazil, Areal et al. (2012) on dairy farms in England and Wales. Some recent studies include, Billé et al. (2018) study on olive farms in Italy, Pede et al. (2018), and Areal and Pede (2021) studies on rice farmers in the Philippines. However, the form of contiguity considered by almost all these studies is based on geographical location, position, or distance (i.e., physical contiguity) between the production units or farms. The physical contiguity approach is based on the assumption that farmers embedded in social communities learn from their peers, given the similar environmental and social factors they face and the socio-economic relationships they share, thus creating a local *terroir effect* (Vidoli et al. 2016; Billé et al. 2018).

However, while the physical contiguity approach may account for environmental, climatic, and edaphic factors in the production system, it is insufficient to address the issue of informational inadequacies among interacting farmers that lead to learning. Despite, the voluminous literature on the impact of the information networks on adoption of new technologies and yields, there is paucity of knowledge on how the network contributes to technical (in)efficiency in the production functions of the individual farmers who constitute the information network. This is important because the position of farmers in a network who are first to receive information about a new technology have distributional consequences among members in the network (e.g., Banerjee et al. 2013; Beaman and Dillon 2018).

The present study attempts to fill the knowledge gap on how information networks influence farmers' technical (in) efficiency, by considering farmers' relationship within a social space separated by social distance in a social network, rather than, geographical space separated by physical distance of the farm location. Specifically, we use a unique survey data of 600 soybean farmers to investigate the influence of farmers personal information networks (i.e., the *egocentric* networks) on their technical (in)efficiency and its distributive mechanisms in the network, while controlling for social selection bias. We estimate a spatial stochastic frontier analysis (SSFA) model that accounts for unobserved spatial heterogeneity, which presents a potential source of endogeneity in efficiency analysis and could bias the estimates (Kutlu et al. 2020; Qu and Lee 2015).

Our findings show that *egocentric* information network influence on technical (in)efficiency of farmers is network-specific and heterogeneous according to the nature of the

social ties between farmers in the network. We find that farmers who share farming information with inefficient farmers are more likely to be inefficient, a correlation that suggests farmers learn from efficient peers in order to improve their own efficiency. We observed that failure to account for the influence of farmers' *egocentric* information networks effects in the production process, as a special case of spatial heterogeneity, can lead to negative efficiency bias estimates among high efficient farmers and positive efficiency bias among low to medium efficient farmers.

The present study contributes to the literature by incorporating social network structure into efficiency analysis, using stochastic frontier analysis. To the best of our knowledge, this is the first attempt to consider the impact of social interactions in technical (in)efficiency analysis.

The rest of the paper is organized as follows; Section 2 discusses the conceptual framework and the estimation strategy of the study. We then discuss the data and the empirical results in Sections 3 and 4 respectively, while Section 5 presents the conclusions of the study.

2 Conceptual framework and estimation strategy

In this section, we present a conceptual framework to illustrate the mechanism by which the *egocentric* information network membership could influence the farmers' technical in(efficiency) as well as the empirical strategy employed in this study for identification and estimation.

2.1 Socio-spatial stochastic frontier analysis with social network dependence

We assume that farmers are homogeneous in regards to their production technology. Let Y_i denote an individual farmer's soybean output and X_i be a vector of production factors. The farmer's production function can be specified as follows;

$$Y_i = \ln(f(X_i; \beta_i)) + v_i - u_i, \text{ and } u_i \geq 0, [i = 1, \dots, n] \quad (1)$$

where Y_i is a vector of *log* outputs (yield) of the farmer, X is a vector of production factors, β is a vector of parameters of interest, u represents the inefficiency term, and v the random error term, assumed to be *iid* with; $v \sim iid N(0, \sigma_v^2 I)$ and $u \sim iid N^+(0, \sigma_u^2 I)$, where I is an identity matrix.

The performance of farmers producing under any given technology, without external influence on the farmers' technical abilities can be estimated from Eq. 1. However, when a farmer obtains technical knowledge of a given technology from other farmers through information exchange, the possibility of the farmer's technical ability to

be influenced by informational inadequacies from the farmers with whom they exchange information may increase (Fafchamps et al. 2021; Kondylis et al. 2017). Hence, analysis of the farmer's productivity under any technology that ignores the influence of the informational inadequacies of other farmers in the farmer's production function, could suffer a potential bias, due to the unobserved informational inadequacies (Horrace and jung 2018).

To account for the influence of other farmers' informational inadequacies in the production function of the farmer; let g_f represent village level farmer information network, with $g_f(g_f = 1, \dots, G_{FN}) \in G_F$, where G_F is a set of all farmer information networks across N villages. Furthermore, let w_{ij} represent the link that exists between farmer i and j ($i \neq j$), defined as $w_{ij} = 1$, if farmer i shares agricultural information with farmer j , otherwise $w_{ij} = 0$, and $w_{ij} \in W_{ij}$, where W_{ij} is the social contiguity matrix (or adjacency matrix). The social contiguity matrix is assumed to be undirected (i.e., $w_{ij} = w_{ji}$) and i cannot share information with i (i.e., $w_{ii} \neq 1$). In line with Fusco and Vidoli (2013) and Vidoli et al. (2016), we re-specify (suppressing the subscript)² Eq. 1, as a socio-spatial stochastic frontier model² that accounts for the interdependency of the farmer's inefficiency on the information network as follows;

$$\ln Y = \ln(f(X; \beta)) + v - (I - \rho W)^{-1} \tilde{u} \quad (2)$$

where W is the adjacency matrix of the network, ρ is the spatial lag parameter ($\rho \in [0, 1]$), v and \tilde{u} are the random error and latent unknown terms respectively, assumed to be distributed as $v \sim iid N(0, \sigma_v^2 I)$ and $\tilde{u} \sim iid N(0, \sigma_u^2 I)$, respectively. The inefficiency term u in Eq. 1 is expressed as $u = (I - \rho W)^{-1}$, and assumed to be distributed as $u \sim N^+(0, [(I - \rho W)^{-1}(I - \rho W)^{-1}(I - \rho W)^{-1}] \sigma_u^2)$.

This specification was first introduced by Areal et al. (2012), but estimated using the Bayesian approach. However, the specification employed here, is the frequentist approach developed by Fusco and Vidoli (2013) and Vidoli et al. (2016), which is estimated using maximum likelihood approach. This specification is more convenient, because, it adopts a one-stage estimation procedure, which makes it more efficient and easier to compare with the standard stochastic frontier analysis with the spatial stochastic frontier analysis for consistency, since the spatial stochastic frontier model converges to the standard stochastic frontier model in the absence of spatial dependence (i.e., $\rho = 0$). Furthermore, because the specification limits the analysis to only the inefficiency term in the stochastic frontier model,

² Socio-spatial stochastic frontier model as used in this study refers to a spatial stochastic frontier model that uses social distance of nodes within a social network as the contiguity matrix (W), rather than, physical distance between firms or farms within a geographical area. We acknowledge an anonymous reviewer for suggesting this term.

there is substantial reduction in the model's complexity (Vidoli et al. 2016).

2.2 Identification and endogeneity issues of spatial heterogeneity

The spatial stochastic frontier model expressed in Eq. 2 above accounts for the unobserved spatial dependence in the farmer's production function and not the unobserved spatial heterogeneity. This is because spatial heterogeneity arises due to unobserved structural changes within the network structure itself and not in the model parameters³. That is, clusters (or spatial regimes) that are observed in reality varies in structure over geographical or social space, resulting in the inverse problem (Anselin 2010). Spatial heterogeneity, therefore, presents a potential source of endogeneity in efficiency analysis (Kutlu et al. 2020; Qu and Lee 2015), which has proved to be a challenge in empirical analysis, particularly, when employing geographically based proximity measures as the weighting matrix (i.e., the contiguity matrix), since such measures do not easily change in reality. The reason for the empirical challenge is that the identification problem tends to focus on the contiguity matrix, which embodies the network structure rather than the observed covariates. In a recent study, Billé et al. (2018) suggest using a computer-based algorithm that can endogenously identify, in a data-driven approach, spatial homogeneous regimes or clusters from observed real-world spatial data, as a way to account for spatial heterogeneity. Following this approach, let \tilde{W} represent the contiguity matrix of a homogeneous specific network community that can be identified from observed real-world information network data. By substitution, Eq. 2 can be re-specified in terms of the specific network community adjusted contiguity matrix as follows;

$$\ln Y = \ln(f(X; \beta)) + v - (I - \rho(\tilde{W}))^{-1} \tilde{u} \quad (3)$$

2.3 Impact of spatial effects on productivity performance

To assess the benefits that farmers derive from the information network, we employ the structural imbalance distance measure expressed in Vidoli et al. (2016), as well as

Fusco and Vidoli (2013) as follows;

$$d_{i\Delta E} = \frac{\hat{E}_{SFA_i} - \hat{E}_{SSFA_i}}{\hat{E}_{SFA_i}} * 100, \forall_i = 1, \dots, N \quad (4)$$

where \hat{E}_{SFA_i} and \hat{E}_{SSFA_i} are the predicted efficiencies at the standard stochastic frontier model and the spatial stochastic frontier model for individual farmer i , respectively, and $d_{i\Delta E}$ is a distance measure of efficiency difference between the two models. A negative difference indicates improvement in efficiency (the reverse is true for efficiency loss) performance from the network, while the magnitude measures the extent of gains or otherwise from the network (Fusco and Vidoli 2013).

2.4 Distributive mechanisms of gains in egocentric networks

In this section, we specify the determinants of efficiency gains and its distribution among farmers within an information network and across different networks. This is important for an informed policy on agriculture extension service delivery that employs network structures for technology information dissemination. In the spatial stochastic frontier analysis literature, contextual environmental factors are normally regressed on the efficiency distance measure (i.e., $d_{i\Delta E}$) and the coefficients interpreted as determinants. The information network is highly hierarchical because it is composed of individual farmers sampled from a cross-section of smaller units of personal information networks and pooled together to form the village network (see Figure A1 in the Appendix for a sample village network). As such, we employ the spatial effect Cox proportional hazard model with individual level covariate adjustment, which is more appropriate (Bai et al. 2020; Banerjee and Dey 2005).

Given that the distribution of benefits within a social network is assumed to be nonlinear within the framework of social proximity and social embeddedness theory (Bochma 2005). We follow the approach of Bai et al. (2020), which estimates a generalized additive spatial effect Cox model by employing a spatial smoothing function to adjust for individual farmer and network characteristics. Specifically, we estimate a spatial survival time-event Cox model, which is more appropriate for smaller number of units (Banerjee and Dey 2005). In addition, the interpretation of the sign of the distance measure of benefit $d_{i\Delta E}$, makes it amenable to survival analysis. We convert $d_{i\Delta E}$ to a binary event occurrence variable in which the negative sign indicating positive gains on efficiency performance is equal to 1 (implies the farmer benefits from efficiency gain due to the network) and 0, otherwise. Next, the individual farmer predicted mean efficiency score from the SSFA model (i.e., \hat{E}_{SSFA_i}) representing the efficiency level then becomes the

³ The spatial heterogeneity being refer to in this study arises due to the differences in the social network structures of the individual egocentric networks in which farmers belong. Since our data is not panel, we are unable to estimate a spatial panel model to account for individual farmer heterogeneity, in addition to the network structural heterogeneity, a limitation we acknowledge in this study.

survival time variable in the estimation. That is, the level of technical efficiency at which the individual farmer is said to have benefited as a result of being a member of the information network. In line with Bai et al. (2020), the generalized additive spatial effect Cox model for individual farmer i in information network g_f is specified as follows;

$$\lambda_i(g_f) = \lambda_0(g_f) \exp\{X_i\beta_i + s_i\} = \eta_i = X_i\beta_i + s_i \quad (5)$$

where $\lambda_i(\cdot)$ is the benefit hazard function of farmer i in network g_f . X is a vector of observed factors that determine the farmer's spatial efficiency gains and its distribution across individual networks, s is a network-specific structural property. However, because the network-specific structural property has been accounted for in the adjusted weighting matrix (i.e., \tilde{W}), in order to ensure identification, we assume s_i in Eq. (5) to be equal to zero in the estimation.

2.5 Estimation strategy

We estimate both Eqs. 2 and 3 using maximum likelihood estimation procedure implemented in the R software (R Core Team 2017), by combining the packages offered by Fusco and Vidoli (2013) and Pavlyuk (2019). Equation 5 is estimated using partial likelihood estimation approach also in the R package offered by Bai et al. (2020). To ensure identification within the framework of the social network analysis, we account for social selection bias in all the models estimated by controlling for correlated peer effects and contextual effects (Manski 1993)⁴. A parsimonious empirical model we estimate can be specified as follows;

$$y_{i, sf} = x_i\beta_i + \gamma_p + \theta_c + \tau_d + \varepsilon_i \quad (6)$$

where y is the outcome variable (in this case, log yields and spatial efficiency performance gains) of farmer i in network g_f . x is a vector of observed farm characteristics, p , c and d denote farmer's peers, community/village and district level indicators, respectively, β is a parameter of interest, γ , θ , τ are vector of peer, community, as well as district level network fixed-effects, respectively that may correlate with the observed characteristics of the farmer and ε is a composite error term, defined as $(\varepsilon = v - (I - \rho W)^{-1}\tilde{u})$. The efficiency calculation for each farmer follows the

approach of Jondrow's et al. (1982) as expressed in Fusco and Vidoli (2013)⁵.

3 Context and data

3.1 Study context

The study context is northern Ghana, where over the last decade scientific research organizations such as the International Institute of Tropical Agriculture (IITA) and the Council for Scientific and Industrial Research-Savannah Agricultural Research Institute (CSIR-SARI) and their partner organizations employed Farmer Based Organizations (FBOs) concept to disseminate a new agricultural technology (known as the *Rhizobia* inoculant) to small-holder grain-legume farmers. The organizations used conventional extension approaches (e.g., field visits, on-farm and off-farm demonstrations, etc) as well as innovative communication channels such as Radio Listening Clubs (RLCs) and Video Documentaries (VDs) to disseminate and offer technical training to farmers through the FBOs in three regions (Northern, Upper East and Upper West) of northern Ghana. After the training, members of the FBOs became the initial farmers to disseminate or share their knowledge with other farmers in their communities to facilitate the adoption of the new inoculant technology. Thus, the dissemination program sought to use the farmers' personal information networks (i.e., the *egocentric* social network) to diffuse and promote adoption of the *Rhizobia* inoculant technology. This dissemination approach is likely to result in unobserved spatial heterogeneity in the performance of the technology across the population, due to differences in individual disseminating farmers' cognitive proximity (Bochma 2005) and willingness to share knowledge on the new technology (Di Falco et al. 2018).

3.2 Survey of farm households

Our data is from a recent survey of farm households in the northern region of Ghana. The survey was conducted from June to August, 2018. The sample was drawn using a multistage sampling technique. Based on the proportion of beneficiary communities (78%) in the inoculant dissemination program and intensity of soybean production in Ghana, northern region was purposively selected. Cluster sampling technique was employed to zone the region into two clusters, consisting of eastern corridor zone (ECZ) and western corridor zone (WCZ). Based on participation status of districts in the dissemination program and intensity of

⁴ Peer effect (or correlated effect) refers to a situation whereby farmers with similar characteristics tend to behave the same. Contextual effect (or exogenous effect) refers to the situation whereby farmers' behavior varies due to exogenous characteristics or factors. Endogenous effect arises due to joint occurrence of the individual farmer's own outcome and that of the peers' or neighbors' outcomes in the same model specification resulting in the reflection problem of Manski (1993). The endogenous effect does not arise in the context of this study, since the study is not estimating a spatially lagged (or spatial autocorrelation – SAR) model.

⁵ We refer readers interested in the likelihood specifications to the relevant references cited in this paper for details.

soybean production at the district level within the clusters, eight (8) districts, comprising of four (4) from each cluster were purposively sampled. From the ECZ, Yendi, Saboba, Chereponi and Karaga districts were selected, while in the WCZ, East Mamprusi, East Gonja, Savelugu and Kumbungu districts were selected. In consultation with the field officers and agriculture extension agents (AEAs) in the selected districts, 5–7 communities were proportionately sampled, based on dissemination program participation and the extension channel employed, as well as farmer population. One farmer-based organization (FBO) was randomly selected from a list of FBOs that participated in the dissemination program and another randomly selected from a list of FBOs that did not participate in the program, to compose the observed intervention network pool. We then employed a two-stage random sampling technique, which is detailed in the next section below, to sample 600 farm households used for this analysis.

3.3 Data on egocentric networks

An *egocentric* network sampling technique is employed to sample members of *Ego-Alter* networks (see Krivitsky and Morris 2017; Schweinberger et al. 2020)⁶. Similar approaches have been employed in the literature (e.g. Badham et al. 2021; Yen et al. 2016; Cai et al. 2015) to sample real world networks in empirical studies. A two-stage random sampling technique is employed to sample members of the *egocentric* networks. In stage one, using a lottery approach, we randomly drew five farmers (as the *Egos*—seed or focal farmers) from each FBO in the observed intervention network pool. Following an initial interview with the *Egos*, using computer assisted personal interview (CAPI), a list of each farmer's information network members (INMs) was compiled as the *Alters*. In the second stage, the CAPI random number generator used farmers' unique identification numbers to randomly sample three *Alters* from each *Ego's* *Alter* list for interview. The total number of *Egos* and *Alters* for each village is 20, resulting in 20×20 undirected social contiguity matrix (i.e. $W_{ij} = W_{ji}$) for each sampled village or community. In order to avoid missing links due to missing information, a major problem confronting studies that employ sampled social networks, the data used for this analysis is restricted to five villages per district, totaling 30 villages across six districts (East Mamprusi, East Gonja, Savelugu, Kumbungu, Yendi and Karaga), where we have

full data on both the *Egos* and the *Alters*. A total undirected social contiguity matrix size for this analysis is 600×600 block matrix, representing the aggregate village networks for the sampled farm households.

3.4 Network community detection (\tilde{W})

Based on network ecology theory and in line with Billé et al. (2018), we employ three computer software algorithms (i.e., Clauset et al. 2004; Newman and Girvan 2004; Pons and Latapy 2006) to identify three homogeneous virtual network communities with known structural properties, from observed real-world *egocentric* network data. The virtual network community approach is often employed as a pseudo experimental design in the network intervention literature to overcome data challenges that threaten identification and valid statistical inferences. Recent applications of this approach in the social network literature include; Simpson (2020), who use observed real-world *egocentric* network data of Cai et al. (2015;) in Stochastic Actor-Oriented Models (SAOM—a simulation based algorithm), to study the relationship between farm size and social ties formation among rice farmers in China. In the technology adoption literature, Valente and Yon (2020) used a similar approach as in the current paper, to study diffusion of health practices in social networks, while others employed purely simulation studies based on observed real-world network data to study network structure on adoption behavior, knowledge transfer and productivity (see e.g., Badham et al. 2021; Beaman et al. 2021). Though not in the context of social network but in productivity analysis, Billé et al. (2018), employed geographical weighting and adaptive weight smoothing algorithms (Cleveland and Devlin 1988; Polzehl and Spokoiny 2000) to study spatial regimes in olive farm technologies in Italy.

In this study, we denote the three virtual network communities constructed as $T0$, $T1$ and $T2$, each using a specific algorithm. The algorithms employed are based on modularity⁷ maximization procedure, which optimizes a global criterion over all possible clustering in the network for community detection (Li et al. 2021; Geng et al. 2019). Clauset's et al. (2004) algorithm is employed to detect $T1$ communities. The algorithm identifies virtual network communities around farmers (or edges) with high *eigenvector* centrality measure (i.e., a measure of social importance in the network community) from the observed real-world network data. Intuitively, the $T1$ is assumed to correspond to choosing a lead farmer in a community based on the farmer's social importance. Newman and Girvan (2004) algorithm is employed to detect $T2$ communities.

⁶ The *egocentric* sampling design probabilistically generates a sample of nodes along with their edges who constitute the network. The network data generated consist of the *egos* and their *alters*. The *egos* are the focal persons (or nodes) of the network. The *alters* are the personal acquaintances, friends, etc, who share specific relationship or tie (in this study *farming information*) with the *egos* (see Yen et al. 2016; Krivitsky and Morris 2017; Schweinberger et al. 2020).

⁷ *Modularity*, is defined as a natural division of network nodes into densely connected subgroups (Newman and Girvan 2004).

This algorithm identifies virtual network communities around farmers with high *betweenness* centrality (i.e., a measure of power based on being a bridge for other farmers to pass through for information in the network) measure in the observed real-world network data. Intuitively, the *T2* is assumed to correspond to choosing a lead farmer in a community based on the farmer's power derived from being a bridge to access information. The *edge-eigenvector* and *edge-betweenness* community structures have received wide empirical application in the literature (e.g., Beaman and Dillon 2018; Beaman et al. 2021; Fafchamps et al. 2021), due to their importance in information diffusion required for technology adoption. In order to identify the effects of network community structure on the economic outcomes of interest, we employed the algorithm of Pons and Latapy (2006) to construct a third network community *T0*, which assumes a randomly distributed centrality measure in the network, as the virtual control community for comparison. This algorithm provides an *iid* situation for comparison, since it identifies virtual network communities based on the assumption that, the virtual communities observed in the network are randomly formed, and do not necessarily form around any influential farmer (or node) within the network. Intuitively, the *T0* is assumed to correspond to randomly choosing any farmer in the community to be a lead farmer for the community. After identifying homogeneous network communities with known network structural properties around influential farmers in the network, we then construct a network-specific contiguity matrices (\tilde{W}) for each network community. The influence of three network characteristics or properties namely; *transitivity*, *degree-centrality* and *eccentricity*, are analyzed for each detected network community⁸. These social ties are chosen based on their importance and wider application in technology adoption studies using social networks in the literature (e.g., Beaman et al. 2021; Fafchamps et al. 2021; Simpson 2020; Beaman and Dillon 2018). Table 1 presents the layout of the adjusted matrices and the social ties.

By iterative substitution, each adjusted matrix (\tilde{W}) is then employed in the estimation of Eq. 3 of the empirical specifications to account for spatial heterogeneity effect, while the global contiguity matrix (W) from the observed real-world network data is used to account

Table 1 Adjusted weighting matrices

Centrality Measure	Network Community Structure		
	<i>T0 (Random)</i>	<i>T1 (Edge-eigenvector)</i>	<i>T2 (Edge-betweenness)</i>
<i>Transitivity</i>	\tilde{W}_1	\tilde{W}_2	\tilde{W}_3
<i>Degree-Centrality</i>	\tilde{W}_4	\tilde{W}_5	\tilde{W}_6
<i>Eccentricity</i>	\tilde{W}_7	\tilde{W}_8	\tilde{W}_9

\tilde{W} denotes the adjusted weighting matrix for the respective community

for spatial dependence at the global network level, as expressed in Eq. 2.

3.5 Descriptive statistics

Table 2 presents descriptive statistics of the data⁹. Average soybean yield of a farmer is 830 kg/ha, cultivating on average 5 ha of land to soybean and using an average labor of 8 worker days/ha. About 51% of the farmers used inoculant, averaging 14 g/ha, of which 70% of the farmers also used improved soybean seed variety. Average age of farmers in the sample is 42 years, who are predominantly male farmers (71%), with average years of schooling of 3 years, living in an average of 6 member households.

Table 2 also presents the average network structural properties. Note that because the algorithms employed to construct the virtual network communities are based on the modularity maximization procedure, they are interpreted as modularity measures of the respective network communities. The three virtual communities are therefore described in terms of their modularity measures. The table shows that, average modularity of *T1* communities is 0.324, indicating that at least 32% of links in the information network is formed around an agriculturally important (i.e., successful farmer or past award winning farmer) farmer in the network. Average modularity of *T2* communities, is 0.332, indicating that at least 33% of links in the information networks is formed around powerful farmers (i.e., farmers serving as bridges for others to pass through for information). Average modularity of *T0* communities is 0.318, suggesting that about 32% of links in the network may be formed around any randomly chosen farmer within the networks. Although, the average modularity across the treatment arms look similar, graphical visualization of their full distribution in Fig. 1 shows that the treatment arms are heterogenous.

Table 2 further shows that, average transitivity (which measures the structural strength of ties or links in the network or cohesion) of the global network is 0.471, suggesting that at least 47% of farmers (or adjacent vertices) are

⁸ *Transitivity* refers to ties among connected triple nodes in social network graph. It measures cohesiveness in the structure and social ties among farmers (or nodes) within a network (Wasserman and Faust 1994). *Degree-centrality* is a measure of the number of direct ties that a given farmer (or node) has in a given network (Marsden 2002). The degree centrality is an important index measuring the communicative activity of a node in a network (Freeman 1979). *Eccentricity* of a node refers to the largest geodesic distance between that node and any other node. It measures how far a node is from the node most distant from it in a network graph (Wasserman and Faust 1994).

⁹ See Table A1 in the Appendix for the full descriptive statistics of the data.

Table 2 Definition and summary statistics

Variable	Definition	Mean	SD	Min	Max
<i>Panel A: Farmer and Farm Level Factors</i>					
Yield	Soybean yield per hectare (Kg/ha)	829.64	888.24	32.41	5703.87
Age	Age of farmer (years)	41.56	13.32	18	87
Gender	1 If farmer is male, 0 for female	0.71	0.46	0	1
Edu	Years of schooling	2.79	4.69	0	21
Hhsize	Number of people in a household	5.79	3.05	1	27
Land	Total area of land planted with soybean (ha)	5.05	4.37	1	22
Labor	Total labor used in soy cultivation (Worker-days/ha)	7.81	24.23	0.20	274.73
Agrochem	Total amount of active ingredient in chemical used (kg/ha)	4.00	7.19	0	87.22
Chemdummy	1 If farmer uses agrochemical, Otherwise = 0	0.03	0.16	0	1
Amtinouse	Total amount of inoculant used (kg/ha)	13.91	18.35	0	118.93
Inodummy	1 If farmer uses agrochemical, Otherwise = 0	0.51	0.50	0	1
Improvar	1 If farmer uses improve seed variety, Otherwise = 0	0.70	0.46	0	1
Creditconst	1 If farmer is not credit constrained, Otherwise = 0	0.83	0.38	0	1
Extcont	Number of extension contacts	1.37	1.22	0	5
Distmkt	Distance to nearest market (km)	2.36	4.14	0.10	50.10
Soil	1 If soil quality is good, Poor soil quality = 0	0.51	0.50	0	1
Rain	Amount of rainfall in (%)	61.63	16.24	20	100
Elgrid	1 If community is connected to the national grid for electricity supply, Otherwise = 0	0.51	0.50	0	1
<i>Panel B: Network Structural Characteristics</i>					
Random Structure ($T0$)	Average modularity of intervention communities	0.324	0.059	0.197	0.404
Edge-eigenvector Centrality ($T1$)	Average modularity of intervention communities	0.318	0.053	0.160	0.397
Edge-betweenness Centrality ($T2$)	Average modularity of intervention communities	0.332	0.067	0.151	0.424
Transitivity	Average transitivity of a network	0.471	0.034	0.391	0.530
Eccentricity	Average shortest path distance from the farthest nodes in the network	2.718	0.063	2.5	2.75
Degree-Centrality	Average centrality of the network based on degree connections	0.242	0.059	0.137	0.421

SD is standard deviation; Min and Max are minimum and maximum values respectively (See Appendix Table A5).

connected together. Average eccentricity, which measures the shortest path distance (or geodesic) from the farthest node to any other node within a network is 2.7, meaning on average a farmer in the network need to take 3 steps to reach the farthest farmer within the network, which is very short and easier for information flow within the network. Average degree-centrality of a network is 0.242, implying that at least a randomly chosen farmer in the network is connected to 24% of the farmers within the network.

4 Empirical Results

This section presents estimates based on Eqs. 2–5 in the empirical specifications, representing spatial dependence, spatial heterogeneity, efficiency gains from spatial heterogeneity, as well as determinants of efficiency gains and the distribution of the efficiency gains among farmers within the social space. For brevity, we focus the discussion on the

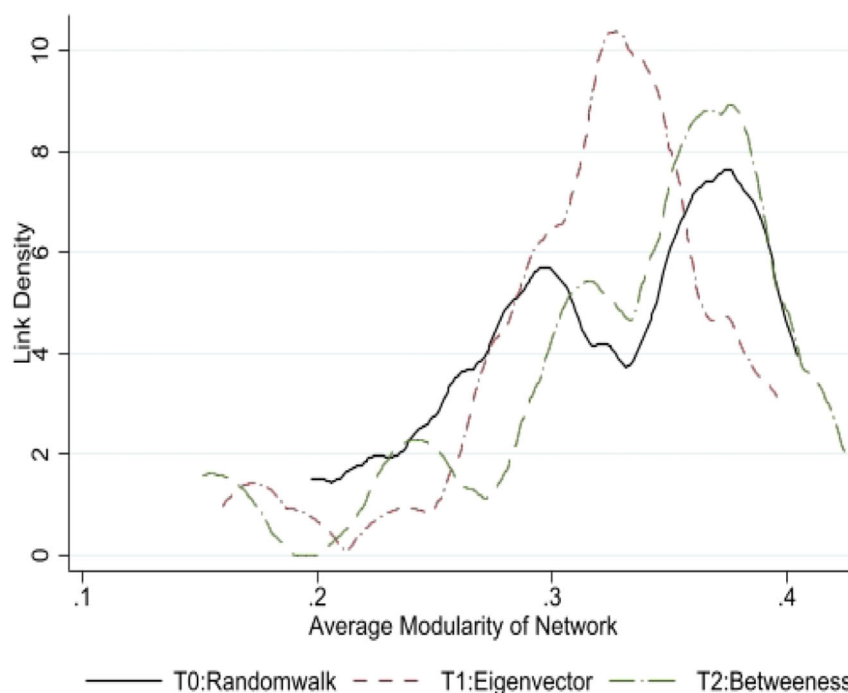
parameter estimates that are germane to the objective set out in this study¹⁰. Thus, to save space, we present the rest of the estimates in the appendix (see the supplementary online materials). However, we will refer to these estimates in the discussion.

4.1 Spatial dependence of efficiency

In Fig. 2, Panels (a) and (b) present the global Moran's I plot of the residuals in the SFA and SSFA models, respectively. This is a correlation test for spatial dependence between the individual farmer's inefficiency and the inefficiency of the farmers in the information network, as expressed in Eq. 2 of the empirical specifications. This test

¹⁰ Since the estimates are based on a relatively small sample size (600) of soybean farmers from one region of the country (i.e., the northern region of Ghana), the external validity cannot be guaranteed. As such, we interpret the estimates with caution and do not seek to generalize the results to all crops, settings or countries.

Fig. 1 Comparing the heterogeneity of average modularity across the virtual treatment arms. T0, T1, and T2 are the virtual network communities detected base on random walk, network eigenvector centrality, and network betweenness centrality algorithms, respectively that forms the treatment arms. The distribution shows all the treatments are heterogenous notwithstanding the closeness of the means



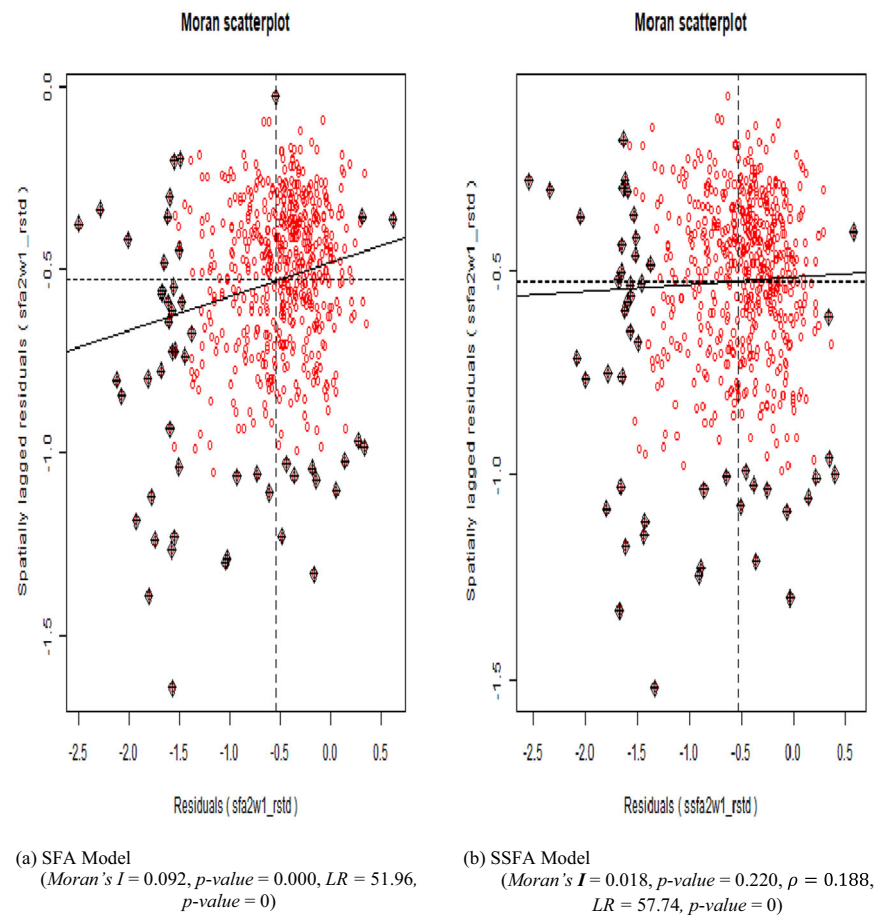
is performed, using the residuals obtained from estimating Eq. 2 with the adjacency matrix (W_{ij}) that defines the social contiguity or proximity of a farmer to other farmers in the observed real-world network. Panel (a) of Fig. 1 shows a Moran's plot of the residuals from the SFA model, assuming no spatial dependence between the farmer's inefficiency and the inefficiency of farmers in their information network (i.e., $\rho = 0$). As reported below in Panel (a), the Moran's statistic (Moran's $I = 0.092$, $p\text{-value} = 0.0001$) is positive and statistically significant at the 1% level, indicating that the assumption of no spatial dependence as implied by the SFA model is rejected, in favor of SSFA. Figure 1 also shows a dense distribution of the residuals in the first quadrant of the Moran's plot in Panel (a), suggesting that highly inefficient farmers are more likely to be connected to highly inefficient farmers. The intuition is that inefficient farmers seek farming advice from peers who themselves are inefficient, thus, low quality advice leading to low performance. This finding is in line with Fafchamps et al. (2021), who found that knowledge of farmers tend to correlate with their peers from whom they seek farming advice. The rejection of no spatial dependence means that we have to estimate a SSFA model to account for the dependence. Panel (b) of Fig. 1 presents the Moran's plot for the SSFA model accounting for spatial dependence. As shown in Panel (b), the presence of spatial dependence (i.e., $\rho = 0$) is rejected ($\rho = 0.188$, LR = 57.74, $p\text{-value} = 0.000$) at 1% level of significance. The rho is positive and statistically significant, suggesting that at global level, 19% of the farmer's inefficiency depends on

the efficiency of the farmers from whom they seek farming advice. In terms of model fit, the LR (likelihood ratio) shows that the SSFA model outperformed the standard SFA, indicating that accounting for spatial dependence significantly improves the fit of the farmer's production function. The corresponding Moran's statistic (Moran's $I = 0.018$, $p\text{-value} = 0.220$) is not statistically significant at any conventional level, indicating that the SSFA model has sufficiently accounted for the spatial dependence. The distribution of the residuals in Panel (b) also shows that, the gap between the mean plot (i.e., the solid line) is now very close to the line of origin (i.e., the dash line), compared to Panel (a), suggesting that spatial dependence has been sufficiently addressed.

4.2 Spatial Heterogeneity in Efficiency

In Table 3, we present estimates based on Eq. 3 of the empirical specifications in three panels B, C, and D, representing transitivity, degree-centrality and eccentricity interventions, respectively, assigned to the three virtual network communities (i.e., T0, T1 and T2). Each panel contains three models, with each model representing a specification with a different network-specific structure used to construct the adjusted weighting matrix (i.e., \tilde{W}) employed in the estimation of the model. The criteria for identification is that, after accounting for spatial heterogeneity, the spatial dependence structure as captured by ρ should vary according to the changing network structure (Anselin 2010). In addition, after accounting for spatial

Fig. 2 Comparing the effect of spatial dependence on correlations of residuals distribution. SFA and SSFA represent stochastic frontier analysis and spatial stochastic frontier analysis models, respectively; LR represents the statistic of the likelihood ratio test



heterogeneity the local level spatial dependence as captured by the local¹¹ Moran's I statistic becomes statistically zero, such that the SSFA and the SFA models' parameters converge and the model is consistently estimated (Fusco and Vidoli 2013, Vidoli et al. 2016).

Panel A in Table 3 presents the global model for comparison. The global model only accounts for spatial dependence (as discussed in the previous section above) and not spatial heterogeneity. Hence, serves as a benchmark for detection of observed changes in the spatial dependence structure due to changes in the network-specific structure. We also present the local Moran's and kernel density plots of the model residuals, as well as the efficiency scores predicted at the production frontier in Figs. 2 and 3, respectively. The local Moran's plot illustrates the effect of spatial heterogeneity on the distribution of the model residuals, while the kernel density plot illustrates the effect of accounting for spatial heterogeneity on the estimated farmers' technical efficiency. In the interest of brevity, we report only the model parameters that are common to this

discussion, and place the full estimates of all the models in Table A2 in the Appendix.

Panel B of Table 3 contains three models 1, 2, and 3, each representing estimates obtained from Eq. 3, based on the adjusted matrices (\bar{W}_1 , \bar{W}_2 , and \bar{W}_3 , respectively) for virtual communities $T0$, $T1$ and $T2$, respectively, characterized by high transitivity. The results in Panel B show that the coefficient of the spatial dependence parameter ρ , across all the three models are negative, compared to the positive coefficient in the global model. The negative coefficient suggests that changes in the network-specific structure results in changes in the spatial dependence structure. In particular, the Moran's I statistic is also negative and not statistically significant, implying that all forms of observed and unobserved spatial heterogeneity have been addressed. The models (i.e., Models 1–3) that account for spatial heterogeneity in terms of the LR also perform better than the global model. The negative signs in all the models of both the SFA and SSFA indicate that the parameters are also consistently estimated. The implication of the negative spatial dependence structure suggests that less technically efficient (or highly inefficient) farmers are more likely to depend on more technically efficient (or less inefficient) farmers in their information network for farming

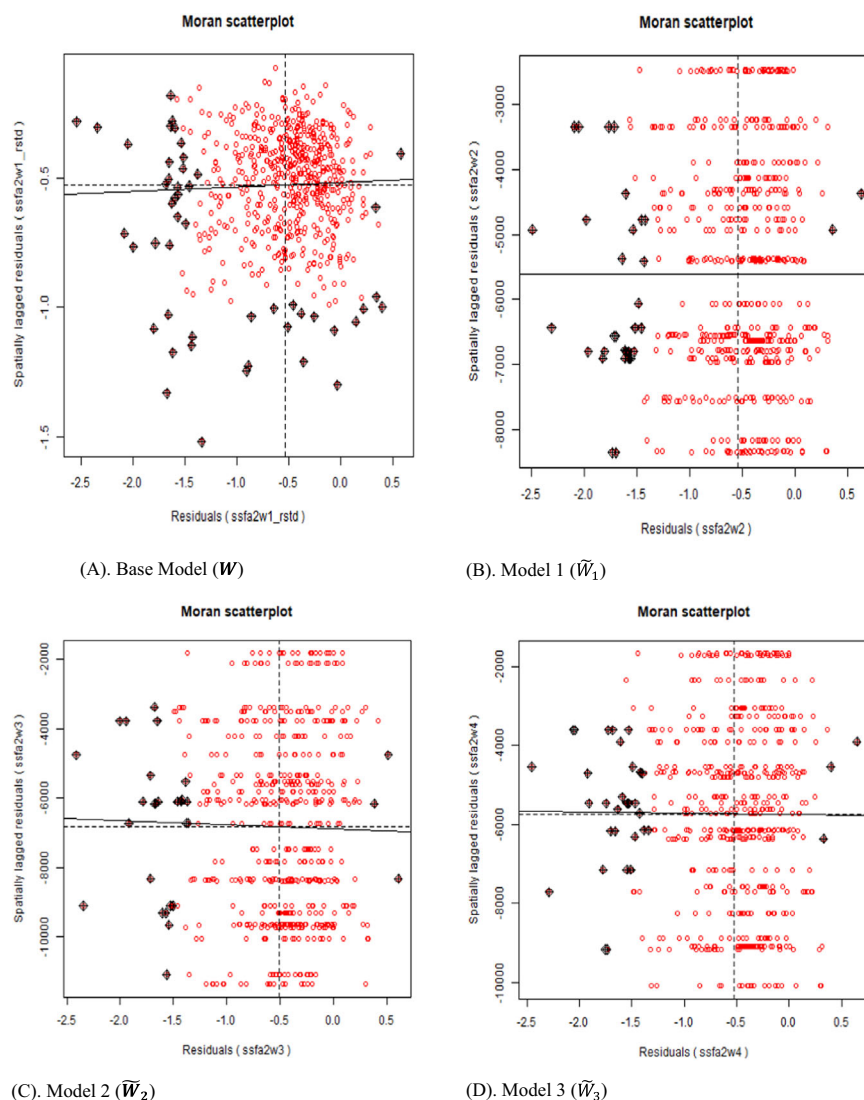
¹¹ The local Moran's I statistic converges to the Local Indicators for Spatial Association (LISA) statistic after accounting for local or network level spatial dependence structure (Anselin 2010).

Table 3 Summary estimates from the stochastic frontier models

Model	Matrix			SFA			SSFA								
	\tilde{W}	$\sigma^2_{u_{sfa}}$	$\sigma^2_{v_{sfa}}$	λ_{sfa}	I_{sfa}	\bar{E}_{sfa}	LR_{sfa}	$\sigma^2_{u_{sfa}}$	σ^2_u	$\sigma^2_{v_{sfa}}$	λ_{ssf}	I_{ssf}	ρ	\hat{E}_{ssf}	LR_{ssf}
Panel A: Adjacency Matrix															
Global Model	W_0	0.470***	0.039***	3.488	0.092***	0.630	51.96***	0.463***	0.1%	0.039***	11.97	0.018	0.188	0.63	57.74***
Network-Specific Structure Models															
Panel B: Transitivity															
Model 1	\tilde{W}_1	0.470***	0.039***	3.488	−0.002	0.630	51.96***	0.470***	0.3%	0.037***	12.54	−0.001	−0.000	0.63	53.94***
Model 2	\tilde{W}_2	0.470***	0.039***	3.488	−0.004	0.630	51.96***	0.418***	0.1%	0.047***	8.88	−0.002	−0.000	0.64	65.08***
Model 3	\tilde{W}_3	0.470***	0.039***	3.488	−0.004	0.630	51.96***	0.453***	0.1%	0.041***	11.00	−0.006	−0.000	0.63	56.56***
Panel C: Degree-Centrality															
Model 4	\tilde{W}_4	0.470***	0.039***	3.488	−0.007	0.630	51.96***	0.376***	0.2%	0.058***	6.50	−0.002	−0.000	0.66	63.43***
Model 5	\tilde{W}_5	0.470***	0.039***	3.488	0.002	0.630	51.96***	0.461***	0.1%	0.037***	12.41	0.001	−0.000	0.63	62.62***
Model 6	\tilde{W}_6	0.470***	0.039***	3.488	−0.009	0.630	51.96***	0.419***	0.2%	0.046***	9.09	−0.002	−0.000	0.64	67.07***
Panel D: Eccentricity															
Model 7	\tilde{W}_7	0.470***	0.039***	3.488	−0.014	0.630	51.96***	0.471***	0.1%	0.038***	12.41	−0.014	−0.000	0.63	52.71***
Model 8	\tilde{W}_8	0.470***	0.039***	3.488	0.009	0.630	51.96***	0.473***	0.2%	0.036***	13.01	0.001	−0.000	0.63	57.17***
Model 9	\tilde{W}_9	0.470***	0.039***	3.488	0.009	0.630	51.96***	0.399***	0.3%	0.054***	7.38	0.004	−0.000	0.65	52.10***

*, ** and *** are 10, 5 and 1% level of significance. The Table presents estimates from Eqs. 2 and 3 in Panel (A) and Panels (B, C and D), respectively. Panel A, presents estimates of the global model employed to account for spatial dependence, while the estimates in Panels B, C and D, present estimates on the network properties (transitivity, degree-centrality and eccentricity, respectively) that characterized the social ties among farmers within the egocentric information network community. Each of the three models represents a specific network community (T0, T1 and T2, respectively) employed to account for spatial heterogeneity. SFA and SSFA are the stochastic frontier analysis and the spatial stochastic frontier analysis, respectively

Fig. 3 Comparing the effect of spatial heterogeneity on the distribution of residuals. The Base Model accounts for spatial dependence, while Models 1–3 account for spatial heterogeneity using the adjusted social contiguity matrices \tilde{W}_1 , \tilde{W}_2 and \tilde{W}_3 , representing matrices constructed using the pairing of the random walk modularity (i.e., Modwalk) and transitivity, the eigenvector modularity (i.e., Modleiden) and transitivity, as well as the betweenness modularity (i.e., Modbetwn) and transitivity in the adjusted social contiguity matrix



advice. Intuitively, inefficient farmers tend to seek farming advice from highly efficient peers, since quality advice contributes to better performance. This finding is consistent with the lead farmer concept employed in farmer-to-farmer extension delivery systems (see Kondylis et al. 2017; Shikuku et al. 2019). We also observe that all the network communities have similar effects on the spatial dependence structure, meaning that no matter the nature of influence (i.e., social importance or power) of the most central farmer in the network community, the effect will be the same, as long as there is social cohesion (i.e., high transitivity) among farmers in the network community.

However, in terms of the proportion of the individual farmer's inefficiency variance (i.e., σ_u^2) that is attributable to the inefficiency of the farmers in their information network, the randomly chosen network communities (i.e., Model 1), account for higher variance (0.3%), compared to Models 2 and 3, respectively. The finding is an indication that

randomly chosen lead farmers increase the level of inefficiency within the network, compared to those chosen on the basis of social importance or power within the network community. This observation may be due to the fact that randomly chosen lead farmers, unlike others who have social importance or recognition to maintain, may require some material motivation in order to spend time to share quality information with peers, hence, the higher inefficiency observed among farmers in that network community. This observation is also in line with Shikuku's et al. (2019) finding that information sharing of randomly chosen lead farmers with their peers is weak, because of the absence of a private motivating factor, compared to lead farmers chosen based on the farmers' social importance in the village.

In terms of average efficiency score, all the models have similar efficiency scores with Model 2 being marginally higher, compared to Models 1 and 3. This indicates that the lead farmers chosen on the basis of social importance of the

farmer may slightly increase efficiency among farmers in the village. Intuitively, choosing best performing or award-winning farmers within a community as lead farmers for extension delivery could marginally enhance learning and performance by other farmers, a finding that is in line with Shikuku et al. (2019) and Fafchamps et al. (2021).

Panel C in Table 3 also presents the results of Models 4, 5, and 6, each representing estimates based on the matrices (\bar{W}_4 , \bar{W}_5 , and \bar{W}_6 , respectively) for virtual communities T0, T1, and T2, respectively, characterized by farmers with high degree-centrality. The results show similar negative ρ coefficients and statistically insignificant Moran's I statistic across all the three models, compared to the global model, indicating that changes in the network-specific structures lead to changes in the spatial dependence structure. The negative spatial dependence emphasizes the earlier findings that less efficient farmers learn from more efficient farmers in order to improve their performance.

In terms of its impact on extension delivery organization, the choice of the lead farmer is important. For example, we observe that Model 4, which assumes choosing the lead farmer randomly, gives higher average efficiency (i.e., \hat{E}_{ssfa}) of 66%, compared to any other model, suggesting that choosing a lead farmer randomly in a network community characterized by high proportion of popular farmers (i.e., *degree-centrality*) leads to higher performance. This observation highlights synergies in information sharing among farmers, needed for tacit social learning at the local level. The finding that the proportion of popular farmers within the network community structure has positive effect on efficiency is consistent with the social theory in diffusion studies, whereby the proportion of adopters of a new technology or behavior in a network influences the adoption decisions of other network members (see, Granovetter 1973).

Also informative is the positive sign of the local Moran's I statistic in Model 5, which chooses a lead farmer based on social importance within the network community. The positive sign suggests that highly efficient farmers also share agricultural information or farming advice with peers who are equally technically efficient, to maintain their performance level (Kondylis et al. 2017). This observation is an indication of mutual relationship in information sharing among farmers in an information network.

Panel D of Table 3 presents the results of Models 7, 8, and 9, representing estimates based on the matrices (\bar{W}_7 , \bar{W}_8 , and \bar{W}_9 respectively) for virtual communities T0, T1, and T2, respectively, characterized by high eccentricity (i.e., shorter social distances or close proximity). The results in Panel D are consistent with that of panels B and C, in terms of the negative coefficients in the spatial dependence structure, compared to the global model. The results of Model 9 show that, average efficiency score of farmers in

high *betweenness-centrality* communities is 67%, suggesting that in network communities characterized by powerful farmers, shorter social distances among all network members increase efficiency. This indicates that close social proximity in farmers' information networks may have high influence on members' efficiency, due to effective communication. However, in terms of inefficiency variance, it also accounts for higher (0.3%) variation of the inefficiency among farmers in the network community, suggesting that increasing proximity could equally have greater consequences on inefficiency, in case the network is dominated by highly inefficient farmers.

We also observed a positive coefficient of the local Moran's I statistic in Models 8 and 9 respectively, suggesting that network communities formed around socially important and powerful farmers that maintain shorter social distance to all farmers within the network community generate more mutual information sharing among farmers, compared to randomly structured network communities.

Furthermore, Fig. 3 shows that the residual distributions in the SSFA models that account for spatial heterogeneity are now more even (i.e., Panels B, C and D), compared to the residual distribution in the global model (i.e., Panel A). We explore the effects of accounting for spatial heterogeneity on average efficiency score of the farmer. Figure 4 presents a kernel density plot of average efficiency scores predicted from all the nine models (i.e., Models 1–9), compared to the efficiency scores from the global model. The results in Fig. 4 reveal that, in estimating spatial dependence of technical efficiency, failure to account for spatial heterogeneity can lead to underestimation of the efficiency of high (i.e., efficiency score >0.6) performing farmers, while overestimating that of medium (i.e., efficiency scores ranging 0.36–0.5) and low (i.e., efficiency scores ranging 0.1–0.35) performing farmers.

4.3 Impact on efficiency gains and distributive mechanisms

We now examine the impact of information networks on farmers' productivity, in terms of technical efficiency improvement (or otherwise) in the production process. Figure 5 presents the productivity gains in classes of (both inter-class and intra-class) percentiles across all the models (i.e., *Models 1–9*), in comparison to the global model that accounts for spatial dependence and not spatial heterogeneity. Generally, Fig. 5 reveals strong heterogeneity in both inter-class and intra-class distribution of productivity gains among farmers within and across each information network, suggesting that farmers' benefits differ according to individual influence in the network, as well as the structural characteristics of the network community.

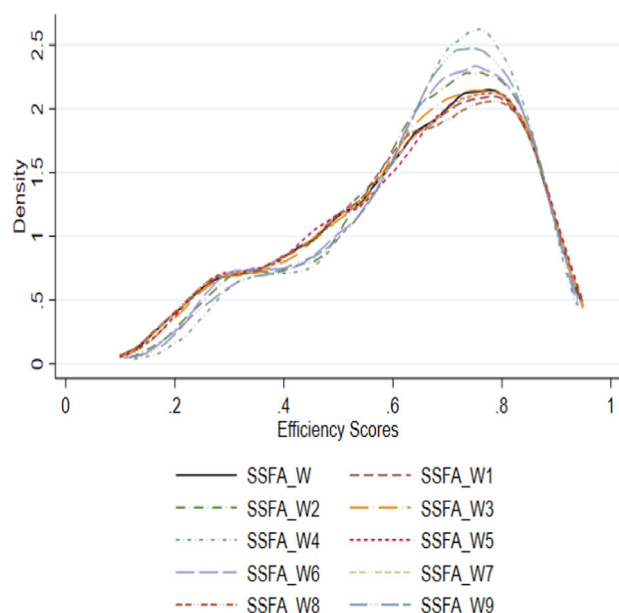


Fig. 4 Effect of spatial heterogeneity on farmers technical efficiency scores. SSFA represents spatial stochastic frontier analysis model, while the Ws represent the social contiguity matrices used in the estimation of the models to account for spatial effects. So, SSFA_W presents the base model estimated using \tilde{W} ; SSFA_W1 represents model 1 estimated using the adjusted matrix \tilde{W}_1 , SSFA_W2 for model 2 using \tilde{W}_2 , SSFA_W3 for model 3 using \tilde{W}_3 , SSFA_W4 for model 4 using \tilde{W}_4 , SSFA_W5 for model 5 using \tilde{W}_5 , SSFA_W6 for model 6 using \tilde{W}_6 , SSFA_W7 for model 7 using \tilde{W}_7 , SSFA_W8 for model 8 using \tilde{W}_8 , SSFA_W9 for model 9 using \tilde{W}_9

In particular, the figure shows that when the information network is characterized by high *transitivity* (i.e., social cohesion) productivity gains are higher (10th to 60th percentiles) in *eigenvector* centrality communities (i.e., *dW2*), compared to the *betweenness* centrality communities (10th to 50th) (i.e., *dW3*). However, the productivity gains are much lower (10th to 30th) in random communities (i.e., *dW1*), compared to the *edge-eigenvector* and *betweenness* centrality communities, respectively, suggesting that benefits differ according to the network structure (Beaman and Dillon 2018).

Also, the intra-class distribution of productivity gains follows similar patterns within the network community, suggesting that farmers' benefits differ according to individual influence or position within the network. This observation is also consistent with the literature on the distribution of economic benefits in embedded social relationships (Tan and Reddy 2021; Beaman and Dillon 2018). The highest population of farmers are within the (10th percentile) productivity class across all network communities, with the population of farmers decreasing as the productivity class size increases, suggesting that productivity gains may be higher among smaller groups of farmers, compared to larger groups. This finding is in line with Vidoli et al. (2016) and Di Falco et al. (2018), who

found an inverse relationship between productivity gains and size of the farmer's network.

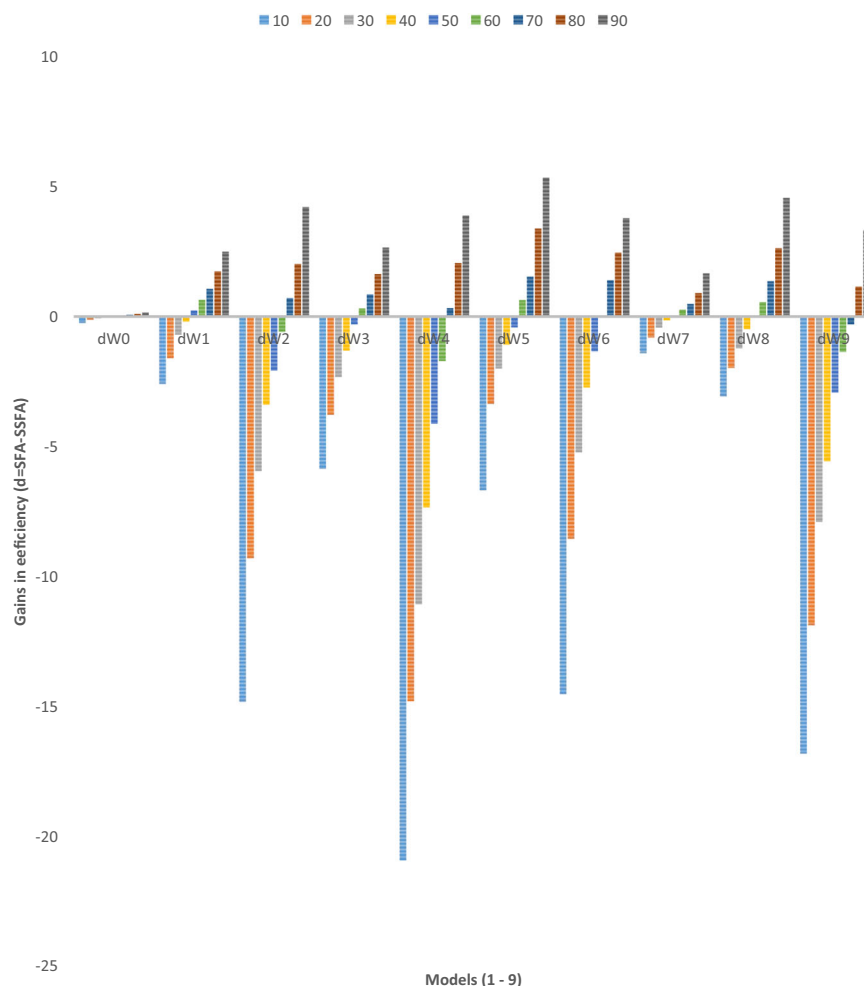
On the other hand, in high *degree-centrality* (i.e., high popularity) networks, productivity gains are higher (10th to 60th percentile) in random communities (i.e., *dW4*), compared to *betweenness-centrality* (i.e., *dW6*) and *eigenvector-centrality* (i.e., *dW5*) communities, respectively. This finding suggests that in random communities, the distribution of benefits may depend more on the individual farmer's characteristics than the structure of the information network. By intuition, in communities with high number of influential farmers, productivity gains from the information network could be evenly distributed among farmers with at least a weakest link to the influential farmer, compared to other network communities. This finding is consistent with Beaman and Dillon (2018), who found high composing knowledge in randomly structured network communities of male farmers, with high centrality influence in the network, compared to other farmer network structures in Mali.

Furthermore, Fig. 5 reveals that in high *eccentricity* (i.e., shortest distance) networks, productivity gains are higher (10th–70th percentiles) in *betweenness* centrality communities (i.e., *dW9*), compared to *eigenvector* centrality (i.e., *dW8*) and random (i.e., *dW7*) communities, respectively. This finding indicates that in high *eccentricity* network communities, an individual farmer's benefits from the network depends on their close proximity to farmers who serve as the information bridges to other farmers in the network community. This observation is intuitive as farmers with more knowledge on a technology are more likely to devote more attention to very close relations during information sharing, compared to any other farmers. Hence, effective communication occurs leading to high efficiency gains (Beaman and Dillon 2018; Akerlof 1997).

In addition, Fig. 5 reveals that failure to account for spatial heterogeneity confounds farmers' productivity gains, as the global model (i.e., *dW0*) suggests equal productivity gains (10%) for all farmers in the information network, contrary to the heterogeneous classes of gains observed across all models. This finding supports recent literature (e.g., Shikuku and Melesse 2020; Shikuku et al. 2019; Kondylis et al. 2017) criticizing the lead farmer concept of extension delivery, where all farmers in the community are assumed to benefit equally from the lead farmer, which thus, implicitly overlook the fact that benefits may differ according to the lead farmer's social influence in the community and the mode by which the lead farmer was chosen for the community¹².

¹² See Appendix 7 for more discussion on the distributive mechanisms.

Fig. 5 Inter-class and Intra-class distributions of average efficiency gains in egocentric network communities. The estimates are based on Eq. 4. The dW0 denotes gains calculated based on the global model, while dW1–dW9 denote gains calculated based on Models 1–9, respectively. The bars at each model indicates percentiles, while comparison within a model is intra-class and between models is inter-class. In Figure 5, the bars above zero indicate efficiency loss by farmers, whereas the bars below zero indicate efficiency gains by farmers



4.4 Determinants of productivity gains in farmer information networks

In this section, we discuss the control variables from the Spatial Cox survival model specified in Eq. 5, showing the factors that influence the likelihood of farmers' information network to contribute to productivity of members of the network. The coefficients discussed here are log hazard ratios from the spatial survival model. We therefore interpret coefficients that are close to 1 as non-contribution to productivity gains, far less than 1 as positive contribution to productivity gains and greater than 1, means negative contribution to productivity gains (i.e. productivity losses) (Sullivan 2021). It is significant to note that the estimates discussed here are not determinants of (in)efficiency as in the spatial stochastic frontier analysis (SSFA) model. Due to space constraints, we attach the estimates of the SSFA models in the appendix (see Table A2 in the Appendix), since their contribution to explaining the aggregate network behavior in this context is less important. The purpose of this discussion is to identify

factors influencing productivity gains from farmers' *ego-centric* networks, in order to inform extension delivery policies that leverage on such networks for technology transfer and diffusion to farmers.

Table 4 reports estimates from the spatial Cox proportional hazard model. The table presents estimates from all the nine models (Models 1–9), compared to the global model. For brevity, we focus the discussion on the network level (i.e., the village level) factors that determine productivity gains from the network, since that is the target unit for policy action. We report the individual farmer level factors and district fixed effects in Table A3 in the Appendix.

Table 4 shows that the coefficient of average age (Vage) of farmers in all models is far less than 1 and statistically significant for Models 1–5 and global model, suggesting that age density at the network level contributes positively to explaining productivity gains from the network. This implies that benefits distribution in egocentric networks is mutual for all ages of farmers, who constitute the network. This observation is intuitive, as

Table 4 Spatial Cox proportional hazard estimates

Variable	Global Model (\tilde{W}_0)	Model 1 (\tilde{W}_1)	Model 2 (\tilde{W}_2)	Model 3 (\tilde{W}_3)	Model 4 (\tilde{W}_4)	Model 5 (\tilde{W}_5)	Model 6 (\tilde{W}_6)	Model 7 (\tilde{W}_7)	Model 8 (\tilde{W}_8)	Model 9 (\tilde{W}_9)
	Coeff. (S.E)	Coeff. (S.E)	Coeff. (S.E)	Coeff. (S.E)	Coeff. (S.E)	Coeff. (S.E)	Coeff. (S.E)	Coeff. (S.E)	Coeff. (S.E)	Coeff. (S.E)
Vage	0.074* (0.039)	0.108** (0.047)	0.056* (0.032)	0.074* (0.036)	0.079*** (0.033)	0.133*** (0.036)	0.059 (0.036)	0.040 (0.041)	-0.033 (0.038)	-0.043 (0.032)
Vgender	-0.036 (0.502)	1.762*** (0.580)	0.461 (0.406)	0.629 (0.439)	0.315 (0.393)	-1.265*** (0.484)	0.711* (0.416)	-0.184 (0.507)	-1.093** (0.520)	1.393*** (0.420)
Vedu	-0.036 (0.064)	-0.471*** (0.086)	-0.190*** (0.059)	-0.229*** (0.066)	-0.137*** (0.056)	-0.105* (0.062)	-0.165*** (0.059)	-0.024 (0.058)	-0.228*** (0.067)	-0.111** (0.050)
Vhsize	0.556*** (0.178)	-0.385* (0.222)	0.439*** (0.144)	0.665*** (0.171)	0.241* (0.134)	-0.182 (0.152)	0.367*** (0.154)	-0.519*** (0.172)	0.540*** (0.174)	0.615*** (0.153)
Vextcont	0.259 (0.435)	2.735*** (0.700)	0.080 (0.364)	1.497*** (0.371)	0.115 (0.330)	0.441 (0.336)	0.415 (0.328)	0.782* (0.474)	0.380 (0.422)	0.475 (0.326)
Vimpvar	-0.901* (0.533)	-2.623*** (0.579)	-0.192 (0.436)	-0.445 (0.495)	-0.334 (0.416)	-0.967** (0.486)	0.225 (0.443)	-1.880*** (0.523)	0.345 (0.506)	0.162 (0.424)
Vinouse	-2.555** (1.202)	2.061 (1.771)	-2.500*** (1.014)	-6.297*** (1.145)	-3.320*** (1.017)	-6.611*** (1.143)	-4.786*** (1.092)	-0.824 (1.476)	-2.500** (1.273)	-1.985** (0.937)
Velgrid	0.921 (1.843)	-2.795 (2.565)	2.074 (1.539)	7.487*** (1.784)	3.343* (1.599)	5.705*** (1.895)	5.147*** (1.728)	2.304 (2.335)	4.338* (1.879)	1.514 (1.406)
Vdistmkt	0.170*** (0.072)	0.207*** (0.074)	0.086 (0.059)	0.058 (0.069)	0.035 (0.060)	0.250*** (0.066)	0.098 (0.063)	0.071 (0.072)	0.166*** (0.070)	0.017 (0.057)
Vfarmsize	-0.062 (0.053)	-0.004 (0.047)	-0.040 (0.039)	-0.057 (0.046)	-0.069* (0.040)	-0.056 (0.053)	-0.096** (0.048)	-0.102** (0.051)	-0.146*** (0.056)	-0.041 (0.041)
Vsoil	5.623*** (1.928)	11.290*** (2.234)	4.921*** (1.697)	7.142*** (2.006)	6.903*** (1.785)	4.121** (1.770)	6.782*** (1.893)	6.704*** (1.904)	1.744 (1.899)	3.058* (1.550)
Vrain	-0.047* (0.027)	0.019 (0.029)	-0.034 (0.022)	-0.013 (0.023)	-0.013 (0.022)	-0.005 (0.026)	-0.015 (0.024)	0.003 (0.028)	-0.001 (0.027)	0.014 (0.021)
LRT (df=570)	56.45**	164.7**	64.25**	131.9**	70.61**	144.7**	95.47**	76.62**	66.2**	50.11**
Mean	45.50%	44.67%	64.83%	53.67%	68.33%	54.17%	59.83%	50.00%	49.50%	71.50%
Beneficiaries										
Observ.(N)	600	600	600	600	600	600	600	600	600	600

*, ** and *** are 10, 5 and 1% level of significance. The Table presents estimates from Eq. 5. Column 1, presents estimates of the global model employ to account for spatial dependence, while each of Columns **1–9** respectively, presents estimates on the network properties (transitivity, degree-centrality and eccentricity, respectively) that characterizes the social ties among farmers within the egocentric information network community. Each of the three models represents a specific network intervention community (**T0**, **T1** and **T2**, respectively) employed to account for spatial heterogeneity

people who share information tend to benefit from each other's pool of diverse experiences.

The results in Table 4 also show that the coefficient of average education (*Vedu*) of farmers in the network is negative (i.e., far less than 1) across all models and statistically significant (at 1% level), suggesting that density of educated persons in *egocentric* networks contributes positively to productivity gains of members of the network. The implication is that, the more educated persons in the farmer's *egocentric* network, lead to more productive networks and vice versa. This is intuitive, as a network with high density of educated farmers means high cognitive proximity among members, a major requirement needed for accurate communication and effective information sharing in embedded social networks (Bochma 2005).

In Table 4, gender (i.e., male = 1) distribution in the network shows a mixed effect. For instance, in network communities with high transitivity (i.e., Models 1–3), the density of male farmers in the *egocentric* networks does not contribute to explaining productivity gains of members in the network, compared to the individual farmer level. A similar observation is made in high *degree-centrality* networks, particularly, in random and *edge-betweenness* network communities. However, in high *eccentric* network communities, the coefficient of gender is negative and statistically significant, suggesting that male farmers are more likely to obtain productivity improvement in the network, compared to female farmers in the network. This implies that the distribution of gender in information networks may have distributional inequalities, due to difference in social distances between male and female farmers, which is likely to affect the close proximity required for effective communication and information sharing.

Furthermore, Table 4 shows that the density of average number of extension contacts (*Vextcont*) of farmers has less positive contribution to explaining productivity gains at the network level, compared to the number of contacts at the individual farmer level (see Panel A in Table A3 in the Appendix). This is intuitive, as farmers may not have to depend on the information network for farming advice, once all farmers have equal access to extension services.

The results in Table 4 also show that, the coefficient of farmers using improved technologies such as improved crop variety (*Vimpvar*) and yield enhancement inputs (e.g., rhizobia inoculant–*Vinouse*) are far less than 1 (being negative in almost all models) and statistically significant, suggesting the density of farmers using improved technologies in the farmers' *egocentric* information networks have positive contribution to productivity gains for farmers in the network. This observation suggests that targeting *egocentric* networks in technology adoption programs will not only enhance diffusion of the technology, but will also improve the performance of the technology, due to the potential of

farmers to learn from the experiences of other farmers in their network communities.

In Table 4, the coefficient of farm size is negative across all models (also statistically significant in most models), suggesting that at the network level, farm size tends to have positive impact on productivity gains from the network. This observation is in line with Simpson (2020), who also observed positive relationship between farm size and productivity gains among rice farmers in *egocentric* networks in China.

On the other hand, Table 4 also shows that, the coefficient of soil quality (i.e., *Vsoil*) and lack of amenities such as availability of electricity (i.e., *Velgrid*) at the village level far exceed 1 (also statistically significant in most models), suggesting that poor soil conditions as well as lack of social amenities such as lack of electricity connectivity negatively impact on productivity gains (i.e., productivity losses) across the network communities.

Furthermore, the bottom row of Table 4 also reports the mean population of farmers that will be affected, in terms of productivity gains, due to extension policy based on each of the network community and the centrality measure, compared to the global model. The results in the table reveal that in network communities characterized by high *transitivity* (i.e., Models 1–3), *eigenvector-centrality* communities have larger (65%) impacts, compared to *betweenness-centrality* 54% and random (45%) communities, respectively. These findings suggest that in *egocentric* networks with high social cohesion among farmers, organizing extension delivery program around farmers with high *eigenvector-centrality* (i.e., most successful farmers) in the community will be more beneficial to majority of farmers in the community, compared to high *betweenness-centrality* farmers (i.e., powerful farmers). However, randomly chosen farmers in the community for extension delivery program will be less beneficial to majority of farmers in the community, compared to the two *centrality* measures.

In network communities with high *degree-centrality* (i.e., Models 4–6), the impact is larger in random communities (68%), compared to *betweenness-centrality* communities (60%) and *eigenvector-centrality* communities (54%), respectively. This finding suggests that in *egocentric* networks with highly popular farmers, organizing extension delivery program around randomly chosen farmers in the community will be more beneficial to majority of farmers in the community, compared to the two *centrality* measures. However, choosing either farmers with high *betweenness-centrality* or *eigenvector-centrality* will still benefit more than half of the population of farmers in the community.

However, the largest (72%) impact occurs in *betweenness-centrality* communities characterized by high *eccentricity*, compared to all network communities. This finding suggests that in *egocentric* networks of farmers with shorter social

distances, organizing extension delivery program around farmers with high *betweenness-centrality* may provide higher outreach to almost all farmers in the community, compared to both *eigenvector-centrality* and random farmers. However, choosing farmers either randomly or based on *eigenvector-centrality* will still benefit about half of the population of farmers in the community.

4.5 Robustness checks

As robustness checks for the spatial heterogeneity observed in this study, we ignore the assumption that the information network is structured around either farmers with high *eigenvector-centrality* or *betweenness-centrality* in the community and assume a random community structure. Panels E and F of Table A4 in the Appendix present the results of two models (i.e., Models 10 and 11) as robustness checks. Panel E reports estimates of Model 10, which accounts for spatial heterogeneity based on adjusted weighting matrix (\widehat{W}_{10}), focusing on *transitivity* and *degree-centrality*, while Panel F reports that of Model 11 estimated based on (\widehat{W}_{11}), focusing on *degree-centrality* and *eccentricity* of a random community structure. The results show that our findings are robust, as both the coefficient of the spatial dependence structure and the LISA statistic both have the negative signs, which is consistent with the random communities in Models 1, 4, and 7 of panels B, C and D, respectively. The *LR* statistics are also statistically significant at the 1% level, suggesting that the SSFA models accounting for spatial heterogeneity provide better fit of the farmers' production function.

In Figure A2, we present kernel density plots of average efficiency scores predicted from models 10 and 11 (i.e., Models 10–11), compared to the efficiency scores from the global model, as robustness check of the effects of failure to account for spatial heterogeneity on farmers' efficiency scores as observed in this study. The results in Figure A2 reveal the same patterns of underestimating high performing farmers, while overestimating that of low and medium performing farmers. Indicating that the findings as observed in this study is consistent and robust. In the interest of brevity, estimates on the robustness checks are reported in the appendix (see the online supplementary materials).

5 Conclusions

In this study, we examined the impact of farmers' *egocentric* information networks on technical efficiency in the production functions of farmers and its distributive mechanisms in the networks. Using community detection algorithms in a data-driven approach, based on observed real-world *egocentric* networks data of 600 soybean farmers

from Ghana, we estimate spatial dependence of farmers' technical efficiency on their *egocentric* information networks, while accounting for unobserved spatial heterogeneity in the network structures of the information networks, within the social network space.

The empirical results generally revealed that farmers' technical (in)efficiency strongly correlate with that of farmers in their *egocentric* networks, suggesting that farmers who share farming information with inefficient farmers are more likely to be inefficient, compared to those who share information with highly efficient farmers. This finding is consistent with the notion that farmers tend to learn from high performing peers in their *egocentric* networks in order to improve their own performance.

The results also showed that the *egocentric* network level of influence on technical (in)efficiency of farmers is network-specific and differ according to the nature of the social ties or influence between farmers in the network. We also found that network communities formed around farmers with social importance increases efficiency among farmers in the community through information sharing with highly efficient farmers in the network. Another quite interesting result is the finding that in networks of farmers with high *degree-centrality*, randomly structured relationship in the network have greater impact on efficiency, compared to any other network community. These network communities generate synergies in information sharing among farmers, needed for tacit social learning at the local level.

In terms of organizing extension delivery around farmers' *egocentric* networks, the findings from this study show that in highly social cohesive networks, organizing extension delivery around farmers with high *eigenvector-centrality* in the community will increase the efficiency of majority of farmers, compared to farmers with high *betweenness-centrality*. Furthermore, we find that organizing extension delivery around randomly chosen farmers in highly cohesive networks decreases the number of farmers that will benefit from the network.

The findings further reveal that in *egocentric* networks with highly popular farmers, organizing extension delivery around randomly chosen farmers' increases efficiency of majority of the farmers in the community, though choosing either *betweenness-centrality* or *eigenvector-centrality* farmers still benefits more than half of the population of farmers in the community.

Finally, we find that in *egocentric* networks of farmers with shorter social distances, organizing extension delivery programs around farmers with high *betweenness-centrality* increases the efficiency of almost the entire population of farmers in the community, though choosing farmers either randomly or based on *eigenvector-centrality* also benefits about half of the population of farmers in the community.

The study concludes that identifying central farmers' in *egocentric* networks and improving their technical knowledge in a farmer-to-farmer extension organization can leverage the limited extension agents to improve productivity of many farmers.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Code availability

The codes are available from the corresponding author upon reasonable request.

Acknowledgements The first author acknowledge funding from the Deutscher Akademischer Austauschdienst (DAAD) as part of the Ghanaian-German Postgraduate Training Programme, 2017. Funding ID No. (57344816).

Funding Open Access funding enabled and organized by Projekt DEAL.

Compliance with ethical standards

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abdulai A, Huffman W (2014) The adoption of soil and water conservation technology: an endogenous switching regression application. *Land Econ* 90:26–43
- Abdul Mumin Y, Abdulai A (2022) Social networks, adoption of improved variety and household welfare: evidence from Ghana. *Eur Rev Agric Econ* 49(1):1–32
- Akerlof GA (1997) Social distance and social decisions. *Econometrica* 65(5):1005–1027
- Anselin L (2010) Thirty years of spatial econometrics. *Pap Reg Sci* 89(1):3–25
- Areal FJ, Balcombe K, Tiffin R (2012) Integrating spatial dependence into stochastic frontier analysis. *Austr J Agric Resour Econ* 56(4):521–541
- Areal FJ, Pede VO (2021) Modeling spatial interaction in stochastic frontier analysis. *Front Sustain Food Syst* 5:673039. <https://doi.org/10.3389/fsufs.2021.673039>
- Badham J, Kee F, Hunter RF (2021) Network structure on simulated network interventions for behavior change. *Soc Netw* 64:55–62
- Bai L, Gillen DL, Bartell SM, Viera VM (2020) Mapping smoothed spatial effect estimates from individual-level data: MapGAM. *R J* 12(1):32–48
- Bandeira O, Rasul L (2006) Social networks and technology adoption in northern Mozambique. *Econ J* 116(514):896–902
- Banerjee A, Chandrasekhar AG, Duflo E, Jackson MO (2013) The diffusion of microfinance. *Science* 341(6144):1236–1239
- Banerjee S, Dey DK (2005) Semiparametric proportional odds models for spatially correlated survival data. *Life Time Data Anal* 11:175–191
- Beaman L, Dillon D (2018) Diffusion of agricultural information within social networks: evidence on gender inequalities from Mali. *J Dev Econ* 133:147–161
- Beaman L, Ben Yishay A, Magruder J, Mobarak AM (2021) Can network theory-based targeting increase technology adoption? *Am Econ Rev* 111(6):1918–1943
- Billé AG, Salvioni C, Benedetti R (2018) Modelling spatial regimes in farms technologies. *J Prod Anal* 49:173–185
- Blum ML, Szonyi J (2011) Investment requirements in extension to achieve zero hunger and adapt to climate change. Food and Agriculture Organization of the United Nations (FAO), Rome, Italy
- Bochma RA (2005) Proximity and innovation: A critical assessment. *Reg Stud* 39(1):61–74
- Cai J, De Janvry A, Sadoulet E (2015) Social networks and the decision to insure. *Am Econ J: Appl Econ* 7(2):81–108
- Chatterjee K, Dutta B (2016) Credibility and strategic learning in networks. *Int Econ Rev* 57(3):759–786
- Clauset A, Newman ME J, Moore C (2004) Finding community structure in very large networks. *Phys Rev E* 70(6):066111–116
- Cleveland WS, Devlin SJ (1988) Locally weighted regression: an approach to regression analysis by local fitting. *J Am Stat Assoc* 83(403):596–610
- Conley TG, Udry C (2010) Learning about a new technology: pineapple in Ghana. *Am Econ Rev* 100(1):35–69
- de Graaff T (2020) On the estimation of spatial stochastic frontier models: an alternative skew-normal approach. *Ann Reg Sci* 64:267–285
- De Janvry A, Macours K, Sadoulet E (eds) (2017) Learning for Adopting: Technology Adoption in Developing Country Agriculture: FERDI. Clermont-Ferrand, France
- Di Falco S, Feri F, Pin P, Vollenweider X (2018) Ties that bind: Network redistributive pressure and economic decisions in village economies. *J Dev Econ* 131:123–131
- Druska V, Horrace W (2004) Generalized moments estimation for spatial panel data: Indonesian rice farming. *Am J Agric Econ* 86(1):185–198
- Fafchamps M, Islam A, Malek A, Pakrashi D (2021) Mobilizing P2P diffusion for new agricultural practices: experimental evidence from Bangladesh. *World Bank Econ Rev* 00(0):1–26
- Foster AD, Rosenzweig MR (2010) Microeconomics of technology adoption. *Ann Rev Econ* 2:395–424
- Fusco E, Vidoli F (2013) Spatial stochastic frontier models: controlling spatial global and local heterogeneity. *Int Rev Appl Econ* 27(5):679–694
- Geng J, Bhattacharya A, Pati D (2019) Probabilistic community detection with unknown number of communities. *J Am Stat Assoc* 114(526):893–905
- Granovetter MS (1973) The strength of weak ties. *Am J Sociol* 78(6):1360–1380
- Gude A, Álvarez I, Orea L (2018) Heterogeneous spillovers among Spanish provinces: a generalized spatial stochastic frontier model. *J Prod Anal* 50:155–173
- Horrace WC, Jung H (2018) Stochastic frontier models with network selectivity. *J Prod Anal* 50:101–116

- Jondrow CAJ, Lovell K, Materov SI, Schmidt P (1982) On the estimation of technical inefficiency in the stochastic frontier production function model. *J Econom* 19(2–3):233–238
- Kondylis F, Mueller V, Zhu J (2017) Seeing is believing? Evidence from an extension network experiment. *J Dev Econ* 125:1–20
- Krivitsky PN, Morris M (2017) Inference for social network models from egocentrically sampled data, with application to understanding persistent racial disparities in HIV prevalence in the US. *Ann Appl Stat* 11(1):427–455
- Kutlu L, Tran KC, Tsionas MG (2020) A spatial stochastic frontier model with endogenous frontier and environmental variables. *Eur J Oper Res* 286(1):389–399
- Li X, Chen Y, Xu J (2021) Convex relaxation methods for community detection. *Stat Sci* 36(1):2–15
- Manski CF (1993) Identification of endogenous social effects: The reflection problem. *Rev Econ Stud* 60(3):531–542
- Marsden PV (2002) Egocentric and sociocentric measures of network centrality. *Soc Netw* 24:407–422
- Mohammed S, Abdulai A (2022) Do ICT based extension services improve technology adoption and welfare? Empirical evidence from Ghana. *Appl Econ* 54(23):2707–2726
- Newman ME, Girvan MJ (2004) Finding and evaluating community structure in networks. *Phys Rev E* 69:026113–115
- Orea L, Álvarez I (2019) A new stochastic frontier model with cross-sectional effects in both noise and inefficiency terms. *J Econom* 213:556–577
- Pavlyuk D (2019) Feature selection and extraction in spatiotemporal traffic forecasting: a systematic literature review. *Eur Transp Res Rev* 11(6):1–19
- Pede VO, Areal FJ, Alphonse Singbo A, McKinley J, Kajisa K (2018) Spatial dependency and technical efficiency: an application of a Bayesian stochastic frontier model to irrigated and rainfed rice farmers in Bohol, Philippines. *Agric Econ* 49:301–312
- Polzehl J, Spokoiny V (2000) Adaptive weights smoothing with applications to image segmentation. *J R Stat Soc* 62:335–354
- Pons P, Latapy M (2006) Computing communities in large networks. Using random walks. *J Graph Algorithms Appl* 10(2):191–218
- Qu X, Lee L (2015) Estimating a spatial autoregressive model with an endogenous spatial weight matrix. *J Econom* 184:209–232
- R Core Team (2017) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Schmidt AM, Moreira ARB, Helfand SM, Fonseca TCO (2009) Spatial stochastic frontier models: accounting for unobserved local determinants of inefficiency. *J Prod Anal* 31:101–112
- Schweinberger M, Krivitsky PN, Butts CT, Stewart JR (2020) Exponential-family models of random graphs: inference in finite, super and infinite population scenarios. *Stat Sci* 35(4):627–662
- Shikuku KM, Melesse MB (2020) Networks, incentives and technology adoption: evidence from a randomized experiment in Uganda. *Eur Rev Agric Econ* 47(5):1740–1775
- Shikuku KM, Pieters J, Bulte E, Läderach P (2019) Incentives and the diffusion of agricultural knowledge: Experimental evidence from Northern Uganda. *Am J Agric Econ* 101(4):1164–1180
- Simpson CR (2020) Farm size shapes friend choice amongst rice producers in China: Some evidence for the theory of network ecology. *Soc Netw* 61:107–127
- Sullivan L (2021) *Survival analysis*. SPH, Boston University, Access: <https://sphweb.bumc.bu.edu>
- Suri T (2011) Selection and comparative advantage in technology adoption. *Econometrica* 79(1):212–220
- Takahashi K, Muraoka R, Otsuka K (2020) Technology adoption, impact, and extension in developing countries: a review of recent literature. *Agricultural Economics* 51:31–45
- Tan YH, Reddy SK (2021) Crowdfunding digital platforms: Backer networks and their impact on project outcomes. *Soc Netw* 64:158–172
- Tsionas EG, Michaelides PG (2016) A spatial stochastic frontier model with spillovers: Evidence from Italian Regions. *Scott J Polit Econ* 63(3):243–257
- Valente TW (1996) Social network thresholds in the diffusion of innovations. *Soc Netw* 18(1):69–89
- Valente TW (2012) Network interventions. *Science* 337(6090):49–53
- Valente TW, Yon GGV (2020) Modeling social dynamics: diffusion/contagion processes on social networks. *Health & Educ Behav* 47(2):235–248
- Vidoli F, Cardillo C, Fusco L, Canello J (2016) Spatial non-stationarity in the stochastic frontier model: an application to the Italian wine industry. *Reg Sci Urban Econ* 61:153–164
- Wasserman S, Faust K (1994) *Structural analysis in social sciences. Social network analysis: Methods and application*. Cambridge University Press, USA
- Yen T-J, Fu Y-C, Hwang J-S (2016) Alters as species: Predicting personal network size from contact diaries. *Soc Netw* 45:78–88