Learning from Neighboring Communities: A Spatial Analysis of Improved Seed Adoption

1 Introduction

The adoption of new technologies have become crucial in tackling climate change, especially for those regions facing socioeconomic adversities. Sub-Saharan Africa (SSA) continues to face profound challenges as the most food-insecure region in the world, exacerbated by the limited adoption of advanced technological interventions (Arslan et al., 2016). Smallholder agriculture remains the primary driver of economic growth in most SSA countries, yet it is also the most vulnerable to the adverse effects of climate change (Barrios et al., 2008). Adopting improved seed varieties have shown to increase agricultural productivity accounting for about 50–90% of global crop yield increase (Kafle, 2010). The possibility of SSA farmers to adopt to new improved seed varieties and be aware of new agricultural technology is still very limited. Indeed, information, or the lack thereof, is often found to be the most important factor limiting rural development among farmers and communities (Fang and Richards, 2018).

Our study primary centers in Tanzania, a representative nation within the Sub-Saharan Africa (SSA) region, well-known for its heavy dependence on smallholder farmers and agriculture as the main source of employment and livelihood (Rowhani et al., 2011). The objective of this study is to evaluate how spatial interactions with neighboring communities influence proportion of land dedicated to improved seed varieties, using community georeferenced panel data for Tanzania from 2008 to 2013. The primary contribution of this study is to provide deeper insights into the learning dynamics among rural communities, where the proportion of land dedicated to improved seed varieties may be influenced by information shared with neighboring communities. Furthermore, despite the growing literature on this topic, there remains a significant gap in understanding the spatial dependence of seed adoption in Tanzania. This study fills that gap by being the first to investigate how spatial interactions between farmers and communities affect the adoption of improved seed varieties in the region. Additionally, it is one of the first studies to explore how spatial dependence impacts the proportion of land dedicated to improved seed varieties, using available plot-level data.

2 Literature review

The decision to adopt improved seed varieties is heavily influenced by both farmers and community characteristics. According to research, factors such as access to seeds, extension services, labor availability, and the location of farm households play a crucial role in determining adoption rates(Martey et al., 2020). Additionally, characteristics like farm size and membership in farmer organizations have been found to positively impact the adoption of improved seed varieties. Membership in these organizations enhances access to information and facilitates collective action, which strengthens the dissemination of agricultural technologies and encourages broader adoption of new practices (Kalinda et al., 2014). Furthermore, seed scaling programs implemented through community institutions have been shown to significantly increase adoption rates compared to non-collective approaches (Hossain et al., 2024), highlighting the importance of community structures in the adoption process.

The influence of neighboring farmers and communities plays a pivotal role in the adoption of improved seed varieties. Social learning, or learning by observing others, has been identified as a key mechanism that increases the likelihood of adoption (Fang and Richards, 2018; Li et al., 2021). Farmers tend to cluster around similar practices, where the decisions of neighboring farmers significantly impact their own choices regarding seed adoption. In addition, Fang and Richards (2018) founded that local networks act as important agents for information exchange, where farmers rely on their immediate neighbors for recommendations, and weigh the neighbors' opinions heavily. They also go beyond immediate neighbors to exploit extended social networks for gathering information. This spatial dependency is critical, as interactions within these networks facilitate the diffusion of information, leading to higher adoption rates.

The interaction effect of social learning offers valuable insights into how neighboring communities influence decisions regarding the proportion of land allocated to improved seed varieties. Based on the literature, I anticipate a positive spatial lag effect when a community follows the practices of its neighbors, and a negative spatial lag effect when it diverges from them. It is expected that social learning operates through interactions within each network, indicating that these interactions are inherently spatial.

3 Methodology: The Spatial Model

3.1 Spatial Weight Matrix & Correlation Test

In spatial econometrics, assuming independence between observations is highly restrictive, as the observed independent values y for a given observation i are often statistically similar to those in nearby locations. This phenomenon is known as spatial dependence (Ahumada et al., 2018). To estimate it, a starting point is to define an appropriate weights matrix, $\mathbf{W}$. Based on Tobler (1970) "everything is related to everything else, but near things are more related than distant things", this weights matrix represents the independence and interactions strength between the spatial observations.

Since the community's proportion of land dedicated to improved seed varieties are more likely to be influenced by their nearby neighbour communities rather than the remotes ones, the inverse distance between communities' locations is usually used as weights to measure the proximity between communities and neighboring communities (Elhorst, 2014; LeSage and Pace, 2009). To construct the distance-based weight matrix $\mathbf{W_d}$, following the studies of Zheng et al. (2021); Fang and Richards (2018); Li et al. (2021), it is measured as

$$w_{ij} = \begin{cases} 0 & \text{if } i = j \\ 1/d_{ij} & \text{if } d_{ij} \leq D \\ 0 & \text{if } d_{ij} > D \end{cases} \tag{1}$$

where d_{ij} is the arc-distance between the centroids of community i and the neighbouring community j. D is the critical distance cut-off, it would be assumed that there is no spatial relationship between communities that are beyond the distance D from each other, as indicated in equation (1). The critical distance cut-off D to be used would be 100 km. An extra weight matrix would be defined for comparison to emphasize on immediate neighbours. This weight matrix is going to be based on rook continuity, which only observations that are adjacent to the focal observation in rook fashion are considered neighbors (Fang and Richards, 2018). Both matrices would be row standardised.

To test whether there is spatial dependence on the proportion of land dedicated to improved seed varieties, the Moran's I test would be employed for every year in the panel:

$$\text{Moran's I} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{S^2 \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \tag{2}$$

Where $S^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2, \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$, y_i indicates the observed value in the ith communities, w_{ij} is the spatial weight of the link between two communities i and j (as defined above), and n refers to the total number of communities.

3.2 Modeling spatial panel data

According to the theoretical framework, the proportion of land dedicated to improved seed varieties y in a community i can be explained by it's characteristics and it's information acquisition. This can be specified as a non-spatial pooled regression model:

$$ln(y_{it}) = \beta_1 X_{it} + \beta_2 G_{it} + u_{it} \tag{3}$$

where y is the proportion of land dedicated to improved seed varieties in a community i in the year t. X_{it} represents a matrix of the community's characteristics, and G_{it} a matrix of community's information variables. β_1 and β_2 are vectors of parameters to be estimated, and u_{it} refers to the error term.

Based on the literary review, the proportion of land dedicated to improved seed varieties y in a community i, is also affected by their neighbouring communities proportion of land dedicated to improved seed varieties y_j, its characteristics X_j and information acquisition G_j. A general nested pooled model [1] that incorporates all types of spatial interaction effects takes the form,

$$ln(y_{it}) = \rho W ln(y_{it}) + \beta_1 X_{it} + \beta_2 G_{it} + W X_{it}\theta_1 + W G_{it}\theta_2 + u_{it}; \quad u_{it} = \lambda W u_{it} + \epsilon_{it} \tag{4}$$

the endogenous spatial lag is specified as Wy (substantive spatial dependence), the exogenous spatial lag is specified as WX, and the residual spatial lag is specified as Wu (residual spatial dependence).

To capture spatial correlation, one of the standard approaches in most spatial analysis is to start with a general model containing, nested within it as special cases, a series of simpler models that ideally should represent all the alternative economic hypothesis requiring consideration (Elhorst, 2014). LeSage and Pace recommend using the Spatial Durbin Model (SDM) as a starting point, as it nests the largest number of simpler models. Additionally, Elhorst suggests that if the SDM cannot be reduced to a simpler model, one should also consider starting from the Spatial Durbin Error Model (SDEM) to compare the predictive performance of both models. Following the strategy from a general to a specific model (GETS), the Likelihood Ratio Tests (LR) would be used. This is because, when estimating more complex models, the subsequent models will be nested within them. LR tests allow for the comparison of the differences in log-likelihoods between the more complex model and the simpler one, helping to determine which is more appropriate. In addition, as (Elhorst, 2014) recommends, a fixed effect model would be applied as spatial units are located in unbroken study areas, in our case regions of Tanzania. To estimate the spatial models Maximum Likelihood (ML) would be applied as estimates using Ordinary Least Squares (OLS) are biased and inconsistent.

Following the analysis of an static panel spatial model, a dynamic time-spatial panel model is

[1] This Cliff-Ord model has an identification problem, making it impossible to estimate all the spatial coefficients. The most complete model can only incorporate two over the three possible spatial effects. It is shown as a reference model too illustrate all the spatial interaction effects.

constructed by adding the spatial-time lag value of the proportion of land dedicated to improved seed varieties:

$$\ln(y_{it}) = \gamma \ln(y_{i,t-1}) + \rho W \ln(y_{it}) + \eta W \ln(y_{i,t-1}) + \beta_1 X_{it} + \beta_2 G_{it} + W X_{it}\theta_1 + W G_{it}\theta_2 + u_{it} \quad (5)$$

$$u_{it} = \lambda W u_{it} + \epsilon_{it}$$

This model will assess how relevant time and spatial-time lag value of the proportion of land dedicated to improved seed varieties is. In presence of a dynamic model, the short and long run marginal effects can be obtained.

4 Data

This study will use the Tanzania National Panel Survey (TNPS). This is a nationally representative panel data of household, community, and agriculture surveys collected by Tanzania's National Bureau of Statistics (NBS). It is a georeferenced database, meaning that each enumeration area (EA)/community has it's longitude and latitude coordinates. We account for 409 georeferenced EA/-communities in Tanzania. The TNPS collects data about health, education, labor, crop production, fertilizers, weather conditions, among others. The survey conducts interviews over a year, from October of the initial year to October of the subsequent year, encompassing three waves occurring biennially.

From this study, we are going to obtain the proportion of land cultivated with the improved seed varieties adopted at the community level for the 2008/2009, 2010/2011, and 2012/2013 year waves. Additionally, it will collect data on community characteristics and the information they have acquired regarding agricultural practices. From the panel dataset, those household members who migrated or lost track from the original household during the waves would be erased. It is important that each community persists across each wave so that **W** is constant over time and the panel is balanced (Elhorst, 2014).

5 Final Remarks

This study highlights the importance of understanding the spatial interactions that influence the adoption of improved seed varieties in Tanzania. By examining how neighboring communities and farmers share information and interact, this research provides valuable insights into the mechanisms of social learning in agricultural practices. The analysis of spatial dependencies shows that the proportion of land dedicated to improved seed varieties is not only shaped by a community's characteristics but also can be significantly influenced by the practices and decisions of nearby communities.

Regarding the results, I mainly anticipate that neighbouring communities that adopt a bigger proportion of land dedicated to improved seed varieties have a positive effect over the proportion of land of the community. In addition, the characteristics and information adopted by neighbouring communities will influence positively over a community's proportion of land dedicated to improved seed varieties.

In addition to addressing a critical gap in the literature on spatial dependence causality in seed adoption, this study offers practical insights for policymakers and development programs. While a farmer-level analysis would provide a more granular comparison with neighboring communities, georeferenced data on households and farming plots could not be accessed due to confidentiality restrictions. Nonetheless, by identifying the influence of neighboring communities on adoption decisions, future interventions can be better tailored to promote collective action and information sharing across regions. The findings highlight the importance of implementing spatially informed strategies, such as seed scaling programs through community institutions, to enhance agricultural productivity and resilience in Tanzania.

References

Ahumada, H. A., Gabrielli, M. F., Herrera Gomez, M. H., and Sosa Escudero, W. (2018). Una nueva econometría: Automatización, big data, econometría espacial y estructural.

Arslan, A., Belotti, F., Asfaw, S., Karfakis, P., and Lipper, L. (2016). Welfare impacts of climate shocks: evidence from Tanzania.

Barrios, S., Ouattara, B., and Strobl, E. (2008). The impact of climatic change on agricultural production: Is it different for Africa? *Food policy*, 33(4):287–298.

Elhorst, J. P. (2014). Linear spatial dependence models for cross-section data. *Spatial econometrics: From cross-sectional data to spatial panels*, pages 5–36.

Fang, D. and Richards, T. J. (2018). New maize variety adoption in Mozambique: a spatial approach. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, 66(3):469–488.

Hossain, M., Nayak, S., Mohapatra, S., Trivedi, P., and Waza, S. A. (2024). Women-led community institutions as a potential vehicle for the adoption of varieties and improved seed practices: an impact case from India. *CABI Agriculture and Bioscience*, 5(1):53.

Kafle, B. (2010). Determinants of adoption of improved maize varieties in developing countries: A review. *International Research Journal of Applied and Basic Sciences*, 1(1):1–7.

Kalinda, T., Tembo, G., Kuntashula, E., and Lusaka, Z. (2014). Adoption of improved maize seed varieties in southern Zambia. *Asian Journal of Agricultural Sciences*, 6(1):33–39.

LeSage, J. and Pace, R. K. (2009). *Introduction to spatial econometrics*. Chapman and Hall/CRC.

Li, F., Feng, S., Lu, H., Qu, F., and D'Haese, M. (2021). How do non-farm employment and agricultural mechanization impact on large-scale farming? a spatial panel data analysis from Jiangsu Province, China. *Land Use Policy*, 107:105517.

Martey, E., Etwire, P. M., and Kuwornu, J. K. (2020). Economic impacts of smallholder farmers' adoption of drought-tolerant maize varieties. *Land use policy*, 94:104524.

Rowhani, P., Lobell, D. B., Linderman, M., and Ramankutty, N. (2011). Climate variability and crop production in Tanzania. *Agricultural and forest meteorology*, 151(4):449–460.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic geography*, 46(sup1):234–240.

Zheng, H., Ma, W., and Li, G. (2021). Learning from neighboring farmers: Does spatial dependence affect adoption of drought-tolerant wheat varieties in China? *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, 69(4):519–537.