Entrega Final - Seminario de Tesis

Maestría en Economía

Profesores
Franco Nuñez
Matías Gomez Seeber

Alumno
Juan Segundo Zapiola

23 de Diciembre 2024

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).

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 quantity of improved seed varieties used. 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.

Smallholder farming systems are accustomed to coping with many challenges, including climate variability (Nordhagen and Pascual, 2013). However, climate change may pose formidable novel challenges for which traditional livelihood strategies may not be fully suited (FAO, 2008). Some of the key adaptation strategies that have been studied in the literature include switching toward crop varieties resistant to heat and drought, development and adoption of new cultivars, changing the farm portfolio of crops and livestock, integrating the use of forecasts into cropping decisions, uses on fertilizers and irrigation, increased storage of food/feed or reliance on imports, and many more (Nordhagen and Pascual, 2013). Under drought shocks, farmer's seed security becomes profoundly influenced. Seed security is often defined by the three parameters of availability (seed being available in space and time), access (physical and economic access), and utilization (seed quality meets user's needs and preferences) (Makate et al., 2023; Sperling et al., 2008; McGuire and Sperling, 2016). Seed purchasing and the quantity of what seed variety to use enables the farmer and farmer communities to respond to negative factors that result in chronic and temporary seed insecurity, like drought shocks. Some studies have found that drought shock exposure in prior seasons increases seed purchasing for both improved and local seeds in Malawi and Tanzania while encouraging (discouraging) local (improved) seed purchases in Ethiopia (Makate et al., 2023). Others, found that having experienced an adverse weather-related shock had a significantly positive impact on the likelihood of farmers purchasing local seeds (not improved) (Nordhagen and Pascual, 2013). Following this line, a study in Kenya found that frequent past climatic shocks, as manifested by drought incidence, reduce the maize area share per farm allocated to hybrid seeds (Bozzola et al., 2018). Building on this literature, I anticipate that a drought shock during the previous rainy season, when crops are grown, will increase the likelihood of farmers purchasing local seeds. However, there is a gap in understanding how such drought shocks might alter the influence of neighbors on your decision regarding the quantity of seed varieties to use.

Our study primary centers in Malawi, a representative nation within the Sub-Saharan Africa (SSA) region, where the agriculture sector plays a pivotal role in Malawi's socioeconomic wellbeing and poverty alleviation (Stevens and Madani, 2016). The study aims to: (1) assess how spatial interactions with neighboring communities influence the quantity of proportion of improved seed varieties used, and (2) examine how drought shocks shape the dynamics of peer effects on the adoption of improved seed varieties, using georeferenced community panel data from Malawi (2010–2019). 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 Malawi. This study addresses this gap by being the first to examine how spatial interactions between communities influence the adoption of improved seed varieties in the region. Furthermore, it is among the first to explore how spatial dependence affects the quantity of improved seeds used and how drought shocks influence peer dynamics in decisions about the quantity of improved seeds to adopt, using available plot-level data.

2 Data

This study utilizes two georeferenced datasets: the Malawi Integrated Household Panel Survey, which provides information on household, community, and agricultural characteristics, as well as data on improved seed usage; and the Climatic Research Unit (CRU) SPEI database, which supplies the drought shock index. These datasets are further described, along with the presentation of descriptive statistics.

2.1 Malawi Integrated Household Panel Survey

This study will use the Malawi Integrated Household Panel Survey (IHPS). This is a nationally representative panel data of household, community, fishery and agriculture surveys implemented by the Government of Malawi through the National Statistical Office. The data was collected by World Bank's Living Standards Measurements Survey (LSMS) program. It is a georeferenced database, meaning that each enumeration area (EA)/community has it's longitude and latitude coordinates. Given the increasing numbers of households to be tracked, as well as budget/resource constraints, starting in 2016, the IHPS target household sample was adjusted as the households that have been associated with 102 out of 204 baseline EAs. The IHPS collects data about health, education, labor, crop production, fertilizers, weather conditions, among others. The survey conducts interviews over a year, encompassing four waves occurring biennially.

From this study, we are going to obtain the quantity of improved seed varieties used at the community level for the 2010, 2013, 2016 and 2019 year waves from the agriculture questionary focused on the main rainy season. We will focus exclusively on households that own at least one agricultural plot. Accordingly, our dataset includes 95 out of the 102 georeferenced enumeration areas (EAs) or communities, comprising 885 households per round. 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.

2.2 Weather Data & Drought Index

To address for drought shocks we utilized the Climatic Research Unit (CRU) SPEI database[1]. This base is a global scale high-resolution drought index developed from a combination of precipitation and potential evapotranspiration datasets for the Hydro-JULES project (Solomon Gebrechorkos, 2023). The drought index is developed using the Standardized Precipitation Evapotranspiration Index (SPEI). SPEI classifies drought based on precipitation data for 3-, 6-,12-, 24-, and 48-month timescales. However, what sets SPEI apart from other drought indices is its consideration of reference evapotranspiration, which represents the amount of water that would evaporate under reference conditions. SPEI considers climatic factors such as temperature, humidity, solar radiation, and wind, providing a more comprehensive measure of the available water (climatic water balance). By incorporating climatic factors, SPEI offers insights into the combined effects of precipitation and evapotranspiration on drought severity in different locations and time periods (Beguería et al., 2014).

This high-resolution global scale drought index is available from 1981-2022 at a monthly and 5km spatial resolution (0.05 degree). The SPEI index is available from 1-48 months timescales. Selecting a specific timescale depends on assessing drought impacts on different water resources. For example, meteorological drought requires a 1- or 2-month SPEI, while agricultural drought analysis utilizes SPEI ranging from 1 to 6 months. For hydrological drought investigations and applications, longer timescales, such as six months up to 24 months or more, are typically employed (McKee et al., 1993; Svoboda et al., 2012; Santiago Beguería, 2023). For this study, a 3-month SPI is calculated for each grid cell to exploit the drought variability during the first three months of planting during the main rainy season (December-February).

3 Methodology: The Spatial Model

3.1 Spatial Weight Matrix & Spatial Heterogeneity

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

[1] SPEI Dataset: https://catalogue.ceda.ac.uk/uuid/ac43da11867243a1bb414e1637802dec/

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 $S^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2$, $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$, 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. In addition to this, a GWR estimation would be implemented to determine if there is spatial heterogeneity on improved seed adoption in Malawi. It is important that each community persists across each wave so that $\mathbf{W}$ is constant over time and the panel is balanced (Elhorst, 2014).

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 neighboring 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 [2] 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 Malawi. 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 constructed by adding the spatial-time lag value of the proportion of land dedicated to improved seed varieties:

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

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

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.

[2] 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.

4 Pasos a Seguir

- Mentor: Marcos Herrera.
- Diciembre/Enero: la idea es terminar de armar la base final para poder empezar a correr las regresiones y ver que modelo espacial es el que debería utilizar. Para la base estoy terminando de agregar las caracteristicas de la comunidad y tengo que agregar el shock de sequía para cada grid cell.
- Febrero: Mes de regresiones, ver si agrego dinamísmo al modelo final, empezar a pensar testeos de robustez.
- Marzo: Terminar testeos de robustez, empezar redacción de resultados y posible primer entrega.
- Abril: posible entrega final.

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.

Beguería, S., Vicente-Serrano, S. M., Reig, F., and Latorre, B. (2014). Standardized precipitation evapotranspiration index (spei) revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. *International Journal of Climatology*, 34(10):3001–3023.

Bozzola, M., Smale, M., and Di Falco, S. (2018). Climate, shocks, weather and maize intensification decisions in rural kenya. In *Agricultural Adaptation to Climate Change in Africa*, pages 107–128. Routledge.

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.

FAO (2008). Climate change adaptation and mitigation: Challenges and opportunities for food security. *Report from high-level conference on world food security: The challenges of climate change and bioenergy.*

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.

Makate, C., Angelsen, A., Holden, S. T., and Westengen, O. T. (2023). Rainfall shocks and inequality have heterogeneous effects on farmers' seed purchase decisions in east africa. *Climate Risk Management*, 42:100556.

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.

McGuire, S. and Sperling, L. (2016). Seed systems smallholder farmers use. *Food security*, 8:179–195.

McKee, T. B., Doesken, N. J., Kleist, J., et al. (1993). The relationship of drought frequency and duration to time scales. In *Proceedings of the 8th Conference on Applied Climatology*, volume 17, pages 179–183. California.

Nordhagen, S. and Pascual, U. (2013). The impact of climate shocks on seed purchase decisions in malawi: implications for climate change adaptation. *World Development*, 43:238–251.

Santiago Beguería, Borja Latorre, F. R. S. M. V.-S. (2023). About the SPEI. Accessed: 23 December 2024.

Solomon Gebrechorkos, Jing Peng, E. D. D. G. M. (2023). Hydro-jules: Global high-resolution drought datasets from 1981-2022. *NERC EDS Centre for Environmental Data Analysis*. 07 July 2023.

Sperling, L., Cooper, H. D., and Remington, T. (2008). Moving towards more effective seed aid. *The Journal of Development Studies*, 44(4):586–612.

Stevens, T. and Madani, K. (2016). Future climate impacts on maize farming and food security in malawi. *Scientific reports*, 6(1):36241.

Svoboda, M., Hayes, M., and Wood, D. (2012). Standardized precipitation index: user guide.

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.