The Institute for Agricultural Economics
of the Christian-Albrechts-Universität Kiel

Farmer Organizations, Spatial Effects, and Farm Household Performances: Econometric Evidence from Senegal

Dissertation
submitted for the Doctoral Degree
awarded by the Faculty of Agricultural and Nutritional Sciences
of the
Christian-Albrechts-Universität Kiel

submitted
M.Sc. Kougblénou Christophe Adjin
Born in Benin

Kiel, 2020

Dean:	Prof. Dr. Karl H. Mühling
1. Examiner:	Prof. Dr. Dr. Christian H.C.A. Henning
2. Examiner:	Assoc. Prof. Dr. Arne Henningsen
Day of Oral Examination:	17. November 2020

Gedruckt mit der Genehmigung der Agrar-und Ernährungswissenschaftlichen

Fakultät der Christian-Albrechts-Universität zu Kiel

Dedication

A ma femme Sèjolo, et à mes deux filles Houéfa and Fènou! Pour tout le soutien et les sacrifices consentis.

Acknowledgement

D'abord, Merci au Seigneur pour avoir veillé sur moi et m'avoir donné le courage et une santé robuste durant ces quatre années!

Waoooh, four years, that was fast! It's like it was yesterday. Even with a pandemic! But this four years journey would have not been possible without the support and assistance of several people and institutions. I would like to take this opportunity to sincerely thank them.

I want to express my utmost and sincere gratitude to my supervisor, Prof. Dr. Dr. Christian H.C.A. Henning, for his valuable support and guidance. Professor, you gave me the opportunity to complete a Ph.D. under your supervision and I am extremely grateful. I would also like to express my profound gratitude to my second examiner Dr. Arne Henningsen for his support and assistance whenever I needed it. I am grateful to Prof. Dr. Awudu Abdulai and Prof. Dr. Ulrich Orth for accepting to be members of the examination committee, and for their valuable feedbacks.

I am also grateful to other faculty members of the program "Third Ways of Feeding The World", Prof. Dr. Friedhelm Taube, Prof. Dr. Konrad Ott, and Dr. Lieske Voget-Kleschin, and to all other Ph.D. candidates, for their treasured contributions during our research colloquiums and workshops.

My special gratitude goes to my late parents, Leonard and Emilienne, for their last encouragement and prayers. Vous avez tous les deux décidé de quitter cette terre avant la fin de l'aventure. Merci beaucoup. Où que vous soyez, priez pour vos enfants. Que vos âmes reposent en paix.

I would like to thank my wife and daughters for the unwavering support, and for enduring my long periods of absence. Je vous aime plus que tout au monde!

My great appreciation goes to Evangelisches Studienwerk (from the Haus Villisgt Foundation) for the financial support throughout the study period.

I extend my gratitude to all my colleagues at the Chair of Agricultural Policy, especially Eric Sessou, Boipelo Freude, Edmond Kanu, Peron A. Collins-Sowah, Sherin Khalifa, and Christine Kornher, for all the valuable comments made during our workshops and of course for the wonderful times we shared together. Special thanks to Peron A. Collins-Sowah for all his contributions (also as a co-author) and for proofreading the early version of this thesis. I would also like to thank all my co-authors Dr. Anatole Goundan and Saer Sarr, for their great contributions.

Contents

List of Figures

List of Tables

Abstract

In sub-Saharan Africa, agriculture is a significant source of economic growth and the sector has the largest contribution to poverty reduction. But its development is challenged by the need for institutional innovations to solve problems such as market failures or access to improved technologies. Meanwhile, for decades, collective action groups were considered as policy institutional tools to address these challenges and improve agricultural performance. However, despite the growing interest in these organizations in recent years, impact evaluations of the contribution of farmer organizations are still limited. This study, therefore, attempts to fill in the gap by providing more comprehensive insights on the role of farmer organizations, neighbourhood, and spatial heterogeneity in farm performances. Several methodological approaches were applied, and the main data used for empirical analyses come from a survey conducted in 2017 in Senegal which randomly sampled 4480 rain-fed cereal producing households.

The dissertation is a collection of five essays. The first essay examines the empirical causal relationship between membership in farmer organizations and food availability. It applied a generalized spatial two-stage least squares method to control for selection biases and spatial heterogeneity. The results showed a positive and significant association between membership in farmer organizations and households' levels of food availability. The second essay analysed the impact of membership in farmer organizations on household land productivity and income. It applied the Endogenous Switching Regression model to derive treatment effects of membership in farmer organizations. The results showed positive, significant and heterogeneous effects of membership in farmer organizations. The third essay analyses the impact of membership in farmer organizations on rice farms technical efficiency. The essay combined the propensity score matching method with the sample selection stochastic frontier model and the stochastic meta-frontier approach, to mitigate selection biases in the sample and to account for technology heterogeneity. Findings mainly showed that members of farmer organizations do not perform better than non-members. The fourth essay explored the roles and complementarity of neighbourhood and membership in farmer organizations on the adoption of two productivity-enhancing technologies. After applying a Bayesian Spatial Durbin Probit model, the results reveal that close neighbouring farmers show similar choice behaviour regarding productivity-enhancing technologies, and membership in farmer organizations affects significantly and positively the choice of farmers and of their neighbours. The last essay aimed to provide empirical evidence on the Senegalese farmers' technical efficiency in a context of climate variability and spatial heterogeneity. Using simulated data, the paper first evaluated the newly developed spatial stochastic frontier estimation technique based on skew-normal distributions. Moreover, empirical findings reveal that farm technical efficiency appears to be significantly affected by unobserved spatial features.

The findings of this dissertation induced some implications for policy and future research. First, support for farmer organizations in Senegal should take into account

the spatial distribution of farmers. Second, policymakers when designing programs for rural areas, should consider the social links created by both farmer organizations and farmers neighbourhood. Third, policymakers should encourage more the design and dissemination of agricultural technologies that are very adaptable to specific spatial conditions of farmers. Finally, in the field of spatial stochastic frontier modelling, future studies should continue investigating the performances of the skew-normal approach.

Zusammenfassung

In Afrika südlich der Sahara ist die Landwirtschaft eine bedeutende Quelle des Wirtschaftswachstums, und der Sektor leistet den größten Beitrag zur Armutsbekämpfung. Probleme wie Marktversagen und fehlender Zugang zu verbesserten Technologien stellen seine Entwicklung jedoch vor eine Herausforderung, und um sie zu lösen bedarf es institutioneller Innovationen. Jahrzehntelang wurden kollektive Aktionsgruppen als politische institutionelle Instrumente zur Bewältigung dieser Herausforderungen und zur Verbesserung der landwirtschaftlichen Leistung angesehen. Trotz des wachsenden Interesses an diesen Organisationen in den letzten Jahren gibt es immer noch wenige Auswertungen zum Beitrag der Bauernorganisationen. Diese Studie versucht daher, diese Lücke zu schließen, indem sie umfassendere Einblicke in die Rolle von Bauernorganisationen, die Nachbarschaft und die räumliche Heterogenität bei der Leistung von landwirtschaftlichen Betrieben liefert. Es wurden mehrere methodische Ansätze angewendet, und die wichtigsten Daten, die für empirische Analysen verwendet wurden, stammen aus einer 2017 im Senegal durchgeführten Umfrage, bei der 4480 regengespeiste Getreide produzierende Haushalte nach dem Zufallsprinzip befragt wurden.

Die Dissertation ist eine Sammlung von fünf Aufsätzen. Der erste Aufsatz untersucht den kausalen Zusammenhang zwischen der Mitgliedschaft in Bauernorganisationen und der Verfügbarkeit von Nahrungsmitteln. Es wurde eine verallgemeinerte räumliche zweistufige Methode der kleinsten Quadrate angewendet, um Selektionsverzerrungen und räumliche Heterogenität zu kontrollieren. Die Ergebnisse zeigten einen positiven und signifikanten Zusammenhang zwischen der Mitgliedschaft in Bauernorganisationen und der Verfügbarkeit von Nahrungsmitteln in den Haushalten. Der zweite Aufsatz analysierte die Auswirkungen der Mitgliedschaft in Bauernorganisationen auf die Produktivität und das Einkommen der Haushalte. Es wendete das Endogenous Switching-Regression Model an, um Treatment-Effekte der Mitgliedschaft in Bauernorganisationen abzuleiten. Die Ergebnisse zeigten positive, signifikante und heterogene Auswirkungen bei einer Mitgliedschaft in Bauernorganisationen. Der dritte Aufsatz analysiert die Auswirkungen der Mitgliedschaft in Bauernorganisationen auf die technische Effizienz von Reisfarmen. Der Aufsatz kombinierte die Propensity-Score-Matching-Methode mit dem stochastischen Grenzmodell der Stichprobenauswahl und dem stochastischen Meta-Frontier-Ansatz, um Auswahlverzerrungen in der Stichprobe abzuschwächen und die technologische Heterogenität zu berücksichtigen. Die Ergebnisse zeigten hauptsächlich, dass Mitglieder von Bauernorganisationen nicht besser abschneiden als Nichtmitglieder. Der vierte Aufsatz untersuchte die Rolle und Komplementarität von Nachbarschaft und Mitgliedschaft in Bauernorganisationen bei der Einführung von zwei produktivitätssteigernden Technologien. Nach Anwendung eines Bayesian Spatial Durbin Probit-Modells zeigen die Ergebnisse, dass nahe benachbarte Landwirte ein ähnliches Auswahlverhalten in Bezug auf produktivitätssteigernde Technologien zeigen und die Mitgliedschaft in Bauernorganisationen die Auswahl der Landwirte und ihrer Nachbarn erheblich und positiv beeinflusst. Der letzte Aufsatz zielte darauf ab, empirische Belege für die technische Effizienz der senegalesischen Landwirte im Kontext von Klimava-

riabilität und räumlicher Heterogenität zu liefern. Unter Verwendung simulierter Daten bewertete die Arbeit zunächst die neu entwickelte räumliche stochastische Grenzschätzungstechnik basierend auf Schrägnormalverteilungen. Darüber hinaus zeigen empirische Ergebnisse, dass die technische Effizienz der landwirtschaftlichen Betriebe offenbar erheblich von unbeobachteten räumlichen Merkmalen beeinflusst wird.

Die Ergebnisse dieser Studie führten zu einigen Implikationen für die Politik und die zukünftige Forschung. Bei der Unterstützung von Bauernorganisationen im Senegal sollte vor allem die räumliche Verteilung der Landwirte berücksichtigt werden.

Chapter 1
General introduction

1.1 Problem statement

Due to its contribution to economic growth and poverty reduction, the agricultural sector is important for most developing countries including Senegal. However, its development has been challenged by the need for institutional innovations to solve issues such as market failures or low technology adoption and productivity (World Bank, 2007). Therefore, for decades, policymakers have promoted agricultural cooperatives as rural institutions with the main purpose of enhancing smallholder farmers participation in markets, and improve their access to production inputs and technologies (World Bank, 2007; Nganwa *et al.*, 2010; Feyaerts *et al.*, 2020). Such objectives were assigned to agricultural cooperatives, for the reason that, the collective actions of farmers (such as pooling of resources or selling of products) are expected to allow more economies of scale (Valentinov, 2007; Cazzuffi and Moradi, 2012), to contribute in reducing individual transaction costs (Staatz, 1987; Cazzuffi and Moradi, 2012; Fanasch and Frick, 2018), and to increase members bargaining power (Cakir and Balagtas, 2012; Grashuis and Su, 2019).

In some developing countries, agricultural cooperatives constituted the basis for rural development policies. For example, the Agricultural Services and Producer Organizations Projects implemented by the World Bank in Chad, Mali, and Senegal during the period 2000-2011, were mainly based on the development of farmers organizations, with the expectation that these farmers organizations could influence and improve agricultural development and performances. Although there is a relative lack of empirical evidence concerning farmers cooperatives (Grashuis and Su, 2019), in countries like Ethiopia, Rwanda, Kenya, Costa Rica or China, where agricultural cooperatives have been widely established by governments and/or international donors, the scientific community seemed to have acknowledged their importance as mechanisms for driving agricultural and rural development. Empirical findings have shown that membership in a cooperative or farmer organizations affects positively product' prices received by farmers (Wollni and Zeller, 2007; Bernard *et al.*, 2008b; Bernard and Spielman, 2009), commercialization rates (Bernard and Spielman, 2009; Francesconi and Heerink, 2010; Chagwiza *et al.*, 2016; Fischer and Qaim, 2012), technologies adoption (Abebaw and Haile, 2013; Ma *et al.*, 2018a), and households welfare (Fischer and Qaim, 2012; Mojo *et al.*, 2017; Ahmed and Mesfin, 2017; Verhofstadt and Maertens, 2015; Ito *et al.*, 2012; Ma and Abdulai, 2016). Evidence also shows that membership in cooperatives may also induce technical and managerial spillovers through experience and knowledge sharing among farmers (Feyaerts *et al.*, 2020).

However, some issues related to cooperatives, such as the input-output efficiency of members, have received little attention in the literature. The question of whether membership in cooperatives contributes to improving members' technical efficiency has not been amply examined. Very few studies, including those of Abate *et al.* (2014); Abdul-Rahaman and Abdulai (2018) and Ma *et al.* (2018b) have attempted answering this question, yet their methodologies have only partially taken into account the estimation biases arising from the potential heterogeneity of the production technology in use. Moreover, though the empirical literature on agricultural cooperatives or farmers organizations is growing fast, no consensus has been reached yet regarding the direction of the impacts of these organizations on their members' performances. Some evidence reveals that membership in agricultural cooperatives does not always have a positive effect on farmers welfare or state of food security. In Ethiopia, Getnet and Anullo (2012) did not find any significant positive relationship between membership in cooperatives and total household income. In the same vein, the study of Shumeta and D'Haese (2018) could not confirm any positive effect of membership in Ethiopian coffee cooperatives on household food expenditure and income. Similarly, Hoken and Su (2015) did not observe any significant difference in the received income between members and non-members of rice cooperatives in suburban China. Given the knowledge gap and the conflicting results regarding the effectiveness of farmer organizations, some important questions need to be answered: (i) are farmers organizations effective in enhancing their members' technical efficiency, and (ii) do they improve household income and food security?

Beyond the estimation of the impact that membership in a farmer organisation can have on farmers performance, spatial effects are also an important component when conducting such an analysis. Spatial effects can affect both the distribution of cooperatives and the various outcomes of interest. As observed by Ma and Abdulai (2016) clustering factors such as soil conditions and regions features may have influenced farmers' decisions to belong to cooperatives in China. Soil conditions and locations represent spatial characteristics that are generally integrated into economic modelling as discrete factors. However, such an approach is not always efficient because it ignores the form of the appropriate relationship between the dependent variable and the spatial covariates which may vary among observations in a continuous rather than a discrete way (Vidoli *et al.*, 2016). Most studies in agricultural economics generally ignore the presence of spatial effects in their analysis with the risk of obtaining biased estimates and misleading inferences (Anselin, 1988; LeSage and Pace, 2009). Given such considerations, additional research questions are raised: (i) do spatial features affect the impact that farmers organizations might have on farmers performances? (ii) literature has amply proved the importance of farmers organizations in the adoption of agricultural technologies, however, the effect of the farmers' neighbourhood has been barely taken into account, therefore in addition to farmers organization effects what is the contribution of spatial dependence or farmers neighbourhood?

Moreover, the Senegal context offers a unique case study that necessitates being examined. Senegal is one the world top food importers[1]. Its agriculture faces a myriad of issues including low productivity and climate change. Hence, since 2012 the Government[2] has given priority to agriculture and a sort of revival of the sector could be observed, as noted by the substantial increase of the sector's contribution to the national GDP from 12% in 2011 to 16% in 2017 (World Bank, 2019). Additionally, because of several institutional changes in the sector, there is a kind of renaissance of the cooperative movement during the last decade (Reed and Hickey, 2016). Furthermore, there is no comprehensive quantitative study in the country regarding the actual contribution of farmers organizations to agriculture. Therefore, whether or not and how the revival of farmers organizations movement has contributed to the Senegalese agriculture performances still need to be investigated.

The present dissertation tried to fill in those research gaps and contribute to the empirical literature by investigating the causal relationship between farmers organizations, farm households spatial interactions, and farms performances in Senegal, with a focus on cereals producers. The next section presents an overview of Senegalese agriculture and the evolution of its cooperative movement.

[1]With an import volume of 1.25 million metric tons in 2018, Senegal is the 10^{th} rice importing country in the world. `https://www.statista.com/statistics/255948/top-rice-exporting-countries-worldwide-2011/`. Accessed 20/06/2020.

[2]Plan Senegal Emergent, the policy framework of the current government, has made agriculture an important component of Senegal economic growth. `http://www.presidence.sn/en/pse`. Accessed 24/06/2020

1.2 Background

1.2.1 Senegalese agriculture

Senegal is the westernmost country on the African continent, with a population of about 16 million inhabitants. In recent years, the country has experienced steady economic growth. From 2014 to 2018, the average GDP growth rate is about 6.6%, with a peak of 7.08% in 2017, the highest since 1982[3]. However, poverty and food insecurity are still prevalent. Figures show that the average poverty rate for the whole country is about 46.7% (Republique du Senegal, 2013) and results from the National Food Security Survey of 2016 (Republique du Senegal, 2016) report that about 12% of the population had limited and unsatisfactory food consumption.

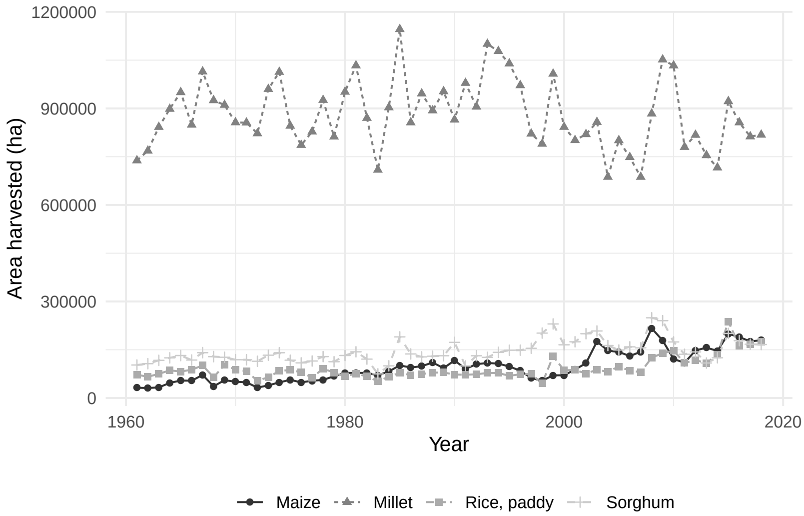
The Senegalese economy relies heavily on agriculture. In 2017, the sector accounted for approximately 32% of the country's total employment, and more than 16% of the national GDP (World Bank, 2019), while at the same time accounting for 21%[4] of the total country's exports (Republique du Senegal, 2018a). The contribution of the agricultural sector to GDP formation has gradually declined over the last forty years. However, since 2011, the trend has changed, from 12% in 2011 it reached 17% in 2018 (World Bank, 2019). Indeed, since 2012, Senegalese agricultural policies and strategies (Programme d'Accélération de la Cadence de l'Agriculture Sénégalaise, PRACAS) has emphasized more on a competitive, diversified, and sustainable agriculture (Republique du Senegal, 2018b).

Agriculture in Senegal is mainly seasonal with nearly 9 out of 10 households practising rain-fed agriculture (Republique du Senegal, 2014). The dominant agricultural products include groundnut, cereals (such as millet, rice, maize and sorghum), cotton, and horticultural crops (Republique du Senegal, 2018b). Cereals represent more than 50% of the total cultivated land in 2018, followed by oil-crops (37%, mainly groundnut), pulses (6%), and horticultural crops (1%). In 2018, the total cultivated land for cereals was 1.34 million hectares. Cereals constitute the main staple foods for rural households in Senegal. This dissertation put the focus on cereals producers, with an emphasis on rice production in chapter 4. Figures 1.1 and 1.2 show the evolution of the harvested area and yield for the main cereals crops since 1961. The cultivated land for cereals has greatly fluctuated over the last decades. These fluc-

[3]Calculation based on World Bank data (World Bank, 2019)

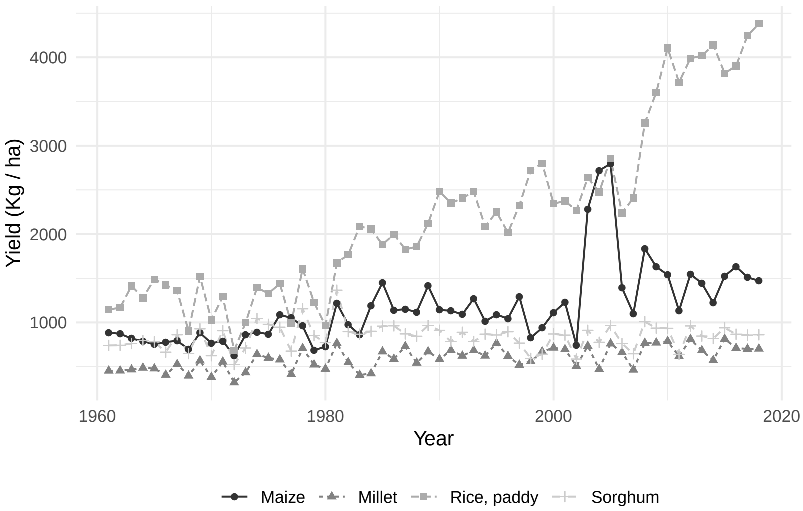
[4]Calculation is based on values of exportation figures in Republique du Senegal (2018a) and includes exports of fishes products, groundnut products, and cotton and cotton fabrics.

tuations are obvious in the case of millet, where the land dedicated to the crop has only increased from 738,100 hectares in 1961 to 817,901 hectares in 2018. Regarding the yields, the same patterns can be observed for most of the cereal crops. However, paddy rice has experienced important growth since 1960. The rice yield which was only 963 kg/ha in 1980 has greatly increased to 4,381 kg/ha in 2018.



Source: Compiled using FAOSTAT data

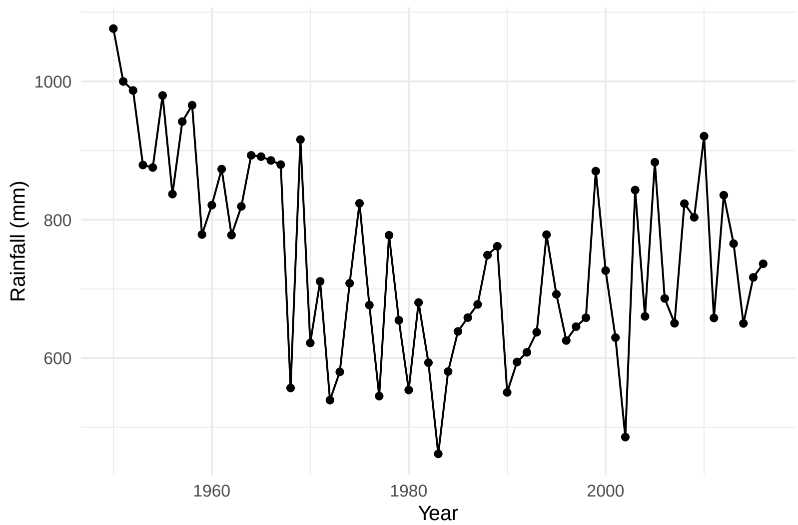
Figure 1.1: Harvested area of main cereals in Senegal (1961-2018)



Source: Compiled using FAOSTAT data

Figure 1.2: Yields of main cereals in Senegal (1961-2018)

Due to its heavy reliance on rain-fed cultivation, the agricultural sector is subject to persistent effects of climate variability which significantly affects yearly crops productions and rural household's food security. The climate in Senegal is typically Sahelian, with one rainy season. Figure 1.3 depicts the inter-annual variation of rainfall observed in the country from 1950 to 2016. Rainfall in Senegal appears to vary constantly with large fluctuations from year to year. According to experts projections, during the coming decades, the whole Senegalese agriculture will experience a loss in productivity due to consistent decline in rainfall and increase in temperature. Therefore, in chapter 6, we include in our modelling climatic data to account for climate variability.



Source: Compiled with data from `https://climateknowledgeportal.worldbank.org/`

Figure 1.3: Annual rainfall in Senegal from 1950 to 2016

Besides the irregular rainfall effects, the sector also faces a myriad of constraints, such as limited access to inputs and credit (Fall, 2016), lack of agricultural infrastructures (Fall, 2016), low technology adoption rates, and low productivity (Hathie *et al.*, 2017). For instance, in Sub-Saharan Africa, Senegal has one of the lowest productivity level in terms of cereals yields and agricultural value-added per worker (Hathie *et al.*, 2017). The combination of climate change effects and low productivity has made Senegal one of the world's top food importers[5]. For example, the average annual quantity of milled rice imported during the period 2000-2016 is valued at 315 million US dollars (FAO, 2019) and it represents a serious burden on the country's trade and foreign exchange balance.

Results of the general census of the population of 2013 (Republique du Senegal, 2014) show that the agricultural sector is dominated by smallholder farmers. These farmers are mainly organized into rural producer organizations which are mostly regarded, as a means to solve the problems of employment security and social insurance (Fall,

[5]With an import volume of 1.25 million metric tons in 2018, Senegal is the 10^{th} rice importing country in the world. `https://www.statista.com/statistics/255948/top-rice-exporting-countries-worldwide-2011/`. Accessed 20/06/2020.

2008).

1.2.2 Cooperative movement in Senegal

Cooperatives are defined as autonomous associations of persons united voluntarily to meet their common economic, social, and cultural needs and aspirations through a jointly-owned and democratically-controlled enterprise. They are organised and operated following cooperative principles[6]. In the rural development context, agricultural cooperatives play an important role in supporting and empowering economically and socially smallholder farmers. Over time, several factors and influences have shaped the development of cooperatives, including economic conditions, farmer organizations, and public policy (Ortmann and King, 2007). In the Francophone area, such as Senegal, the established cooperative model was the one that brought together people with common social and economic objectives (Wanyama *et al.*, 2009). In Senegal, 70% of rural households are members of farmers organizations, and these rural institutions have expanded rapidly during the last decades (Bernard *et al.*, 2015). Although it is difficult to have an accurate picture of the evolution of the farmer organizations, their types and numbers through years, it is evident that they have been shaped through time resulting in various legal and institutional forms.

The journey of the Senegalese cooperative movement started with the colonial ruler, where the 'societies indigenes de prévoyance' were introduced to increase cash crop production for export markets and to control economic activity in rural areas. From 1960, with the independence of the country, agricultural cooperatives were created and controlled by the government which mainly encouraged their development. These cooperatives served as the main vehicles and mechanisms (credit granting, inputs distribution, prices fixing) through which agricultural products were collected and purchased, putting, therefore, farmers under the dependence of the State (Gagné *et al.*, 2008). By the beginning of the 1980s, a sharp decline in the support offered by the State was observed, following the imposition of Structural Adjustment Policies by the World Bank. This situation resulted in a sort of inertia in the cooperative movement (Gagné *et al.*, 2008), and following the reform of the cooperative system in 1983 ("Nouvelle Politique Agricole"), non-governmental organizations other than cooperatives began to emerge especially the Economic Interest Groups (Groupement d'Interet Economiques, GIE)(Gaye, 1994). GIEs were viewed as an alternative solu-

[6]International Cooperative Alliance. `https://www.ica.coop/en/cooperatives/cooperative-identity`. Accessed 24/06/2020.

tion to failing cooperatives, emphasizing their economic aspects to the purely social considerations of cooperatives (Gaye, 1990). From the 1990s, Government disengagement contributed to the rapid development of cooperatives (Fall, 2008; Gagné *et al.*, 2008), which are seen as a means for vulnerable populations to solve their problems of job security and social insurance(Fall, 2008), and for governments and donors as a major channel to reach the rural poor (Bernard *et al.*, 2008a).

Since 2009, the cooperative movement is experiencing a revival, with the introduction of the Agro-Sylvo-Pastoral Law which provided both legal and financial support for farmer-based organizations development (Reed and Hickey, 2016). This historical development led to various types of farmer' organizations, which differ in their legal forms, their functions, and the way they are organized. Recent data obtained from the Ministry of agriculture of Senegal indicate that in 2010, 4903 of farmer organizations were registered representing about 1.67 million farmers. Table 1.1 shows the type and numbers of organization, and the size of their members for agriculture, livestock and forestry sub-sectors.

Table 1.1: Registered Rural Organizations in 2010, by sub-sector

Sub-sector	Number of Organizations	Number of Members	Share (%)
Agriculture	4,903	1,666,050	95%
Livestock	166	29,250	3%
Forestry	89	27,225	2%
Total	5,158	1,722,525	

Source: Compiled using data collected from Senegalese Ministry of Agriculture

1.3 Research objectives

The overall objective of this thesis is to understand the causal relationship between farmer organizations, farm households' spatial interactions, and farm productivity and efficiency. More specifically, this dissertation uses the case study of Senegal to:

1. Illustrate the empirical causal relationship between membership in farmer organizations and food availability.
2. Determine the effect of membership in farmer organizations on household land productivity and net income.

3. Examine the effect of membership in farmer organizations on rice farming households' technical efficiency.

4. Analyse the complementary roles of neighbourhood and membership in farmer organizations on the adoption of two productivity-enhancing technologies.

5. Assess Senegalese farmers' technical efficiency in a context of climate variability and spatial heterogeneity.

6. Identify based on the established evidence some policy recommendations in order to improve Senegalese farm productivity and efficiency, to reduce rural food insecurity and poverty.

1.4 Relevance of the study

This dissertation contributes to the literature in several ways. Most studies in agricultural economics ignore the spatial interdependence of sampled units, leading to biased or inefficient results. Chapters 2, 5, and 6 used recent spatial econometrics methods to derive unbiased and efficient estimates. Such an approach is important and it helps to derive strong evidence which is necessary when one wants to design appropriate policies. Moreover, in the field of the spatial stochastic frontier, researchers are asking for more simulations studies (Glass *et al.*, 2016). Therefore, in chapter 6, we contribute to literature by extending the simulations works of de Graaff (2020). Furthermore, the relationship between neighbourhood and technology adoption is hardly studied in the literature, we therefore provide empirical evidence in chapter 5. Most of all, the present study, would be the first quantitative work that focuses on farmers organization in Senegal. Therefore, it would help to explore the various policies implications that could help policy-makers to fight poverty and food insecurity in rural areas.

1.5 Data

The data used in this study comes from a cross-sectional survey conducted in Senegal, which randomly sampled 4480 households that mainly produce rain-fed cereals. The survey was carried out within the framework of the Agricultural Policy Support Project (PAPA), which is funded by USAID. The project, implemented by the

Senegalese Government, focused on several commodity value chains (cereals, horticulture), and inputs value chains such as seeds and fertilizers. The Senegalese National Agricultural Research Institute conducted the survey in 2017 with the support of the International Food Policy Research Institute (IFPRI). A multi-stage sampling procedure was applied to select survey units. A structured household questionnaire was used to collect information. This questionnaire included several modules and gathered information on a range of topics such as crop production, membership in farmer-based organizations, household assets, access to rural infrastructure and institutions, and household demographic and socio-economic characteristics. Data collection also included production inputs used, markets prices and household use of agricultural technologies such as fertilizers and improved seeds. The survey covers the production season of 2016/2017 and households located in all six agro-ecological zones. Additional information such as soil types, post-harvest losses, climate data, and population density was retrieved from online publicly available databases.

1.6 Thesis structure

This dissertation is organized as a collection of research papers. Chapter 2 examines the empirical causal relationship between membership in farmer organizations and the ability of farmers to produce food. It applied various econometrics estimations techniques that control for selection biases and spatial heterogeneity. Findings show a positive and significant association between organizations membership and household levels of food availability. Chapter 3 investigates the effectiveness of membership in farmer organizations on household land productivity and net income. The paper combined the Propensity Score Matching method with the Endogenous Switching Regression model to derive treatment effects of membership in farmer organizations. Complementary to previous chapters, chapter 4 analysed the extent to which membership in farmer organizations affects farm technical efficiency. This study combined the propensity score matching method with the sample selection stochastic frontier model (Greene, 2010) and the stochastic meta-frontier approach (Huang *et al.*, 2014), to mitigate biases from observable and unobservable variables in the sample and also to account for technology heterogeneity. Chapter 5 explored the complementary roles of neighbourhood and membership in farmer organizations on the adoption of two productivity-enhancing technologies (improved seeds and inorganic fertilizers). After applying a Bayesian Spatial Durbin Probit model, this paper reveals that close neighbouring farmers show similar choice behaviour regarding productivity-enhancing technologies. Chapter 6 provides empirical evidence on Senegalese farmers' technical efficiency in a context of climate variability and spatial heterogeneity. Primarily using simulated data, this last paper evaluated the newly developed spatial stochastic frontier estimation technique based on skew-normal distributions. Secondly, an empirical analysis is conducted for Senegalese farm households. Finally, chapter 7 summarizes the main results of this dissertation and presents some policy implications. It also discusses some of the dissertation's limitations and offers suggestions for future research.

References

Abate, G. T., Francesconi, G. N. and Getnet, K. (2014) Impact of agricultural cooperatives on smallholders' technical efficiency: empirical evidence from Ethiopia, *Annals of Public and Cooperative Economics*, **85**, 257–286.

Abdul-Rahaman, A. and Abdulai, A. (2018) Do farmer groups impact on farm yield and efficiency of smallholder farmers? Evidence from rice farmers in northern Ghana, *Food Policy*, **81**, 95 – 105.

Abebaw, D. and Haile, M. G. (2013) The impact of cooperatives on agricultural technology adoption: Empirical evidence from Ethiopia, *Food Policy*, **38**, 82 – 91.

Ahmed, M. H. and Mesfin, H. M. (2017) The impact of agricultural cooperatives membership on the wellbeing of smallholder farmers: empirical evidence from eastern Ethiopia, *Agricultural and Food Economics*, **5**, 1–20.

Anselin, L. (1988) *Spatial Econometrics: Methods and Models*, NATO Asi Series. Series E, Applied Sciences, Springer Netherlands.

Bernard, T., Collion, M.-H., de Janvry, A., Rondot, P. and Sadoulet, E. (2008a) Do village organizations make a difference in African rural development? A study for Senegal and Burkina Faso, *World Development*, **36**, 2188 – 2204.

Bernard, T., Frölich, M., Landmann, A., Unte, P. N., Viceisza, A. and Wouterse, F. (2015) Building trust in rural producer organizations in Senegal: Results from a randomized controlled trial, Tech. rep., IZA Discussion Papers.

Bernard, T. and Spielman, D. J. (2009) Reaching the rural poor through rural producer organizations? A study of agricultural marketing cooperatives in Ethiopia, *Food Policy*, **34**, 60 – 69.

Bernard, T., Taffesse, A. S. and Gabre-Madhin, E. (2008b) Impact of cooperatives on smallholders' commercialization behavior: evidence from Ethiopia, *Agricultural Economics*, **39**, 147–161.

Cakir, M. and Balagtas, J. V. (2012) Estimating market power of U.S. dairy cooperatives in the fluid milk market, *American Journal of Agricultural Economics*, **94**, 647–658.

Cazzuffi, C. and Moradi, A. (2012) Membership size and cooperative performance: Evidence from Ghanaian cocoa producers' societies, 1930–36, *Economic History of Developing Regions*, **27**, 67–92.

Chagwiza, C., Muradian, R. and Ruben, R. (2016) Cooperative membership and dairy performance among smallholders in Ethiopia, *Food Policy*, **59**, 165–173.

de Graaff, T. (2020) On the estimation of spatial stochastic frontier models: an alternative skew-normal approach, *The Annals of Regional Science*, **64**, 267–285.

Fall, A. S. (2008) *The Senegalese co-operative movement: Embedded in the social economy*, ILO and The World Bank Institute, Geneva, pp. 330–365.

Fall, C. S. (2016) *Impact de la liberalisation commerciale au Senegal: Evaluation de l'Accord de Partenariat Economique sur l'agriculture et les menages Senegalais*, phdthesis, Université de Pau et des Pays de l'Adour.

Fanasch, P. and Frick, B. (2018) What makes cooperatives successful? identifying the determinants of their organizational performance, *Journal of Wine Economics*, **13**, 282–308.

FAO (2019) FAOSTAT, Food and agriculture data.

Feyaerts, H., Van den Broeck, G. and Maertens, M. (2020) Global and local food value chains in africa: A review, *Agricultural Economics*, **51**, 143–157.

Fischer, E. and Qaim, M. (2012) Linking smallholders to markets: Determinants and impacts of farmer collective action in Kenya, *World Development*, **40**, 1255–1268.

Francesconi, G. N. and Heerink, N. (2010) Ethiopian agricultural cooperatives in an era of global commodity exchange: Does organisational form matter?, *Journal of African Economies*, **20**, 153–177.

Gagné, M., Carré, G. and Fall, M. (2008) Le mouvement coopératif au Sénégal: Comprendre les enjeux de son développement, Tech. rep., Société de Coopération pour le Développement International.

Gaye, M. (1990) Les structures coopératives sénégalaises face aux mutations institutionnelles, *Annals of Public and Cooperative Economics*, **61**, 125–134.

Gaye, M. (1994) L'emergence des petits groupements de producteurs ruraux au Senegal, *Annals of Public and Cooperative Economics*, **65**, 507–528.

Getnet, K. and Anullo, T. (2012) Agricultural Cooperatives and Rural Livelihoods: Evidence from Ethiopia, *Annals of Public and Cooperative Economics*, **83**, 181–198.

Glass, A. J., Kenjegalieva, K. and Sickles, R. C. (2016) A spatial autoregressive stochastic frontier model for panel data with asymmetric efficiency spillovers, *Journal of Econometrics*, **190**, 289 – 300, endogeneity Problems in Econometrics.

Grashuis, J. and Su, Y. (2019) A review of the empirical literature on farmer cooperatives: Performance, ownership and governance, finance, and member attitude, *Annals of Public and Cooperative Economics*, **90**, 77–102.

Greene, W. (2010) A stochastic frontier model with correction for sample selection, *Journal of Productivity Analysis*, **34**, 15–24.

Hathie, I., Seydie, B., Samaké, L. and Sakho-Jimbira, S. (2017) Ending Rural Hunger: The Case of Senegal, Tech. rep., The Brookings Institution.

Hoken, H. and Su, Q. (2015) Measuring the effect of agricultural cooperatives on household income using PSM-DID : a case study of a rice-producing cooperative in China, IDE Discussion Papers 539, Institute of Developing Economies, Japan External Trade Organization(JETRO).

Huang, C. J., Huang, T.-H. and Liu, N.-H. (2014) A new approach to estimating the metafrontier production function based on a stochastic frontier framework, *Journal of Productivity Analysis*, **42**, 241–254.

Ito, J., Bao, Z. and Su, Q. (2012) Distributional effects of agricultural cooperatives in China: Exclusion of smallholders and potential gains on participation, *Food Policy*, **37**, 700–709.

LeSage, J. and Pace, R. (2009) *Introduction to Spatial Econometrics*, Statistics, textbooks and monographs, CRC Press.

Ma, W. and Abdulai, A. (2016) Does cooperativemembership improve household welfare? Evidence from apple farmers in China, *Food Policy*, **58**, 94–102.

Ma, W., Abdulai, A. and Goetz, R. (2018a) Agricultural Cooperatives and Investment in Organic Soil Amendments and Chemical Fertilizer in China, *American Journal of Agricultural Economics*, **100**, 502–520.

Ma, W., Renwick, A., Yuan, P. and Ratna, N. (2018b) Agricultural cooperative membership and technical efficiency of apple farmers in China: An analysis accounting for selectivity bias, *Food Policy*, **81**, 122 – 132.

Mojo, D., Fischer, C. and Degefa, T. (2017) The determinants and economic impacts of coffee farmer cooperatives: Recent evidence from rural Ethiopia, *Journal of Rural Studies*, **50**, 84–94.

Nganwa, P., Lyne, M. and Ferrer, S. (2010) What will south africa's new cooperatives act do for small producers? an analysis of three case studies in kwazulu-natal, *Agrekon*, **49**, 39–55.

Ortmann, G. F. and King, R. P. (2007) Agricultural cooperatives i: History, theory and problems, *Agrekon*, **46**, 18–46.

Reed, G. and Hickey, G. M. (2016) Contrasting innovation networks in smallholder agricultural producer cooperatives: Insights from the Niayes Region of Senegal, *Journal of Co-operative Organization and Management*, **4**, 97 – 107.

Republique du Senegal (2013) Rapport définitif de la deuxième Enquête de Suivi de la Pauvreté au Sénégal (ESPS II-2011), Tech. rep., Agence Nationale de la Statistique-ANSD.

Republique du Senegal (2014) Rapport Defintif Recensement Général de la Population et de l'Habitat, de l'Agriculture et de l'Elevage - RGPHAE 2013, Tech. rep., Agence Nationale de la Statistique et de la Démographie, ANSD.

Republique du Senegal (2016) Enquête nationale de Sécurité alimentaire au Sénégal 2016 (ENSAS, 2016), Tech. rep., Primature. Secrétariat Exécutif du Conseil national de Sécurité alimentaire (SECNSA).

Republique du Senegal (2018a) Note d'Analyse du Commerce Exterieur, Edition 2017, Tech. rep., Agence Nationale de la Statistique et de la Démographie, ANSD.

Republique du Senegal (2018b) Programme National d'Investissement Agricole pour la Sécurité Alimentaire et la Nutrition. PNIASAN Sénégal 2018-2022, Tech. rep., Ministère de l'Agriculture et de l'Équipement Rural.

Shumeta, Z. and D'Haese, M. (2018) Do coffee farmers benefit in food security from participating in coffee cooperatives? evidence from southwest ethiopia coffee cooperatives, *Food and Nutrition Bulletin*, **39**, 266–280, pMID: 29669426.

Staatz, J. M. (1987) Farmers' incentives to take collective action via cooperatives: a transaction cost approach, *Cooperative theory: New approaches*, **18**, 87–107.

Valentinov, V. (2007) Why are cooperatives important in agriculture? an organizational economics perspective, *Journal of Institutional Economics*, **3**, 55–69.

Verhofstadt, E. and Maertens, M. (2015) Can Agricultural Cooperatives Reduce Poverty? Heterogeneous Impact of Cooperative Membership on Farmers' Welfare in Rwanda, *Applied Economic Perspectives and Policy*, **37**, 86–106.

Vidoli, F., Cardillo, C., Fusco, E. and Canello, J. (2016) Spatial nonstationarity in the stochastic frontier model: An application to the italian wine industry, *Regional Science and Urban Economics*, **61**, 153 – 164.

Wanyama, F. O., Develtere, P. and Pollet, I. (2009) Reinventing the wheel? African cooperatives in a liberalized economic environment, *Annals of Public and Cooperative Economics*, **80**, 361–392.

Wollni, M. and Zeller, M. (2007) Do farmers benefit from participating in specialty markets and cooperatives? the case of coffee marketing in Costa Rica, *Agricultural Economics*,

37, 243–248.

World Bank (2007) *World Development Report 2008 : Agriculture for Development*, World Bank Washington, DC.

World Bank (2019) World bank data, world Bank Data Base.

Chapter 2

A cooperative way to more food. An analysis of the contribution of farmer organizations to food security in Senegal[1]

K. Christophe Adjin, Christian H.C.A. Henning.

Abstract

Food insecurity remains a major concern in most Sub-Saharan African countries. This paper, therefore, investigates the empirical causal relationship between membership of farmer organizations and food availability in Senegal. Using a unique national data of cereal farming households, and applying various econometrics estimations techniques that control for selection biases and spatial heterogeneity, the study found a positive and significant association between organization membership and households' levels of food availability. Findings are consistent across estimations methods. Being a member of a farmer organization increases at least the cereal production by 19% and the daily per adult equivalent food calories by 13%. These results suggest once again the importance of farmer organizations in the fight against rural food insecurity. In addition, other factors such as access to extension services, fertilizer subsidies and the rainfall appear to significantly determine households' food availability. Furthermore, results also reveal the relevance of incorporating spatiality in the analysis of the agricultural sector in developing countries.

Keywords: Farmer organizations, impact evaluation, spatial heterogeneity.

JEL Codes: Q13, D04, C21.

[1]Part of this chapter has been published as a working paper: *Farmers organizations and food availability in Senegal: An impact analysis using a spatial econometrics approach.* Agricultural Policy Working Papers WP2019- 07. Institute of Agricultural Economics, University of Kiel. 24p. `https://www.econstor.eu/handle/10419/213605`

2.1 Introduction

Food insecurity remains a serious concern in Sub-Saharan Africa (SSA). According to FAO *et al.* (2019), the hunger prevalence in the region in 2018 is estimated at 22.8% and about 240 million people are affected. As one of the food unsecured countries in the region, Senegal with approximately a population of about 16 million, has experienced significant and steady economic growth since 2014, with average GDP growth of 6.64% from 2014 to 2018 (World Bank, 2019). However, according to the 2019 Human Development Index, the country is ranked 166 out of 189, indicating a low human development level in 2018. In addition, poverty rates are still high, 53.2% of the population are considered multidimensionally poor (UNDP, 2019). Moreover, during the period 2016-18, about 11.3 % (1.8 million people) of the Senegalese population have suffered from hunger (FAO *et al.*, 2019), and figures from the 2013 census indicated that poverty is mostly prevalent in rural areas where the primary source of income and food is agriculture (Republique du Senegal, 2014).

Food security is commonly defined as the situation "when all people, at all times, have physical, social and economic access to sufficient, safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life" (FAO *et al.*, 2019, p.186). This generally accepted definition implies four dimensions: food availability, economic and physical access to food, food utilization, and stability over time. The first dimension involves primarily substantial food production at the domestic level. However, in the context of Senegal, food production which is regularly affected by climatic shocks remains at low levels.

According to McArthur and McCord (2017), agricultural productivity can play a strong role in driving structural change. Increasing farm productivity and agricultural production could, therefore, constitute a primary way of ensuring and improving food security and living conditions in rural areas. However, due to inadequate access to improved technologies, Senegalese agriculture is still at the subsistence state. Figures show that in 2015, Senegal was far below the Sub-Saharan average of cereal yields and agricultural value-added per worker (Hathie *et al.*, 2017).

Accessing production inputs and technologies are general challenges for agricultural sectors in most developing countries (World Bank, 2007). Nevertheless, producers organizations could help to alleviate such burdens, by playing their expected role, as mechanisms of reduction of transaction cost (Latynskiy and Berger, 2016). According to Bernard *et al.* (2015), in developing countries, these farmers-based organizations can provide smallholders with better access to production inputs. In addition, results from previous empirical studies show that farmers' collective action groups improve significantly commercialization rates (Barham and Chitemi, 2009; Chagwiza *et al.*, 2016), technologies adoption levels (Abebaw and Haile, 2013; Ma *et al.*, 2018), households welfare (Fischer and Qaim, 2012), and food security (Zeweld *et al.*, 2015). Nonetheless, other empirical studies such as those by Bernard *et al.* (2008), Francesconi and Heerink (2010), and Hoken and Su (2015) do not find any positive association between farmer organizations and commercialization rates or farms productions.

Furthermore, previous studies failed to control for potential biases stemming from the spatial features of farmer's specific location. In general, the potential effect of proximity among farmers is usually ignored in impact studies. However, the magnitude of these effects might be significant in farming settings in developing countries. To highlight the importance of spatiality, let consider the situation of extension services in the African context. With the aim of reaching most farmers, extension services usually target progressive farmers (Diagne, 2006), who are therefore better aware of new technologies and have better access to them. Such a strategy tends to favour some villages or communities than others. Therefore, in the same country, for one reason or another, some regions might be fully provided with active extension agents, while others would barely be covered. Such a situation, which would reinforce the gap between regions' levels of technology adoption, would probably lead to significant differences in farmers productions.

The same argument holds as well for the difference in the other agricultural infrastructures or facilities among regions (road, markets, credits institutions, research institutes). Some communities might live closer to roads or markets that help them to have access to inputs and technologies with a certain ease. Meanwhile, other communities struggle to reach these purchasing points[2]. Moreover, agro-environmental features (e.g. temperature, rainfall, soil fertility) of the location of each farmer constitute structural conditions that might affect their technology choices and therefore their levels of productivity and food production. For instance, some Senegalese regions experience recurrent environmental shocks that constantly threaten or hamper agricultural production and therefore exacerbate household food security. According to Hathie *et al.* (2017), in Senegal, geography plays an important role in food security. Some regions in the country despite their natural resource endowments and economic potential are more prone to food insecurity, due to either the lack of,

[2]See Wanmali and Islam (1997) and Jouanjean (2013) for discussions on the impact of differential access to rural infrastructures on agriculture in developing countries.

or poor quality of transport infrastructure.

Such spatial heterogeneity that influences farmers yields in most cases is not observable to the researcher. As pointed out by LeSage and Pace (2009), amenities and characteristics of the location of a farmer usually constitute unobservable factors that might affect the performance of farmers, and it is difficult to find explanatory variables that capture easily and completely all types of these latent effects. Past studies on the impact of cooperatives membership assume independence between outcome variables in neighboring farmers, without controlling for spatial heterogeneity. This approach presents some limitations that lead to biased or inefficient results and inadequate policy recommendations that followed.

In this paper, we apply a spatial econometric approach to determine the impact of farmer organizations membership on household food production in Senegal. The results contribute to a better understanding of the contribution of collective action groups and have several implications for policy recommendations that would take into account spatial heterogeneity in farming in Senegal. The remainder of this paper is organized as follows. The next section describes the empirical framework. The last sections sequentially present, discuss and summarize the estimation' results.

2.2 Empirical framework

We assume that there is an association between the farm household food availability and their membership in a rural producer organization or farmer organization (or farmer group membership)[3]. We specified the following model of food availability as:

$$Y = f(C, E, X, W), \tag{2.1}$$

where Y represents the food availability of a household and depends on producer organization membership (C), access to extension services E, other household characteristics X including environmental factors, and geographical proximity W. To estimate this model, we consider in our empirical strategy four specifications. In the first two, we assume that $W = 0$, therefore we applied an ordinary least squares technique and a two-stage least squares instrumental variables method. For the last two specifications, we included, spatial heterogeneity between observations via W.

[3]We use the three expressions alternatively.

2.2.1 Estimation strategy

We first consider a linear regression model, specified as:

$$Y = \alpha + \gamma C + \theta E + \beta X + \epsilon, \tag{2.2}$$

where Y denotes a measure of the household food availability indicator; C is a binary variable for farmer organization membership; E is a binary variable for access to extension services; X is a k-dimensional vector of other explanatory variables; α, γ , θ, and β are the parameters to be estimated and respectively associated with organization (or group) membership, extension and the control variables; and ϵ is the error term.

Assuming that $E[\epsilon|C, E, X] = 0$ (i.e. the errors are uncorrelated with any of the right-hand side variables), we can apply the ordinary least squares technique (referred as OLS) to estimate all parameters mentioned above. Therefore, for any randomly selected household, the parameter of interest γ, would be interpreted as the average effect of farmer-based organizations membership on household food availability.

However, prior literature on farmer organizations and access to extension services have demonstrated the possible endogeneity of these two variables when estimating their effects on farm production or incomes (Francesconi and Ruben, 2012; Wossen *et al.*, 2017; Ma and Abdulai, 2016). Therefore, an OLS technique will provide inconsistent estimates for these parameters and especially for the one of interest γ. The endogeneity of these variables is sourced in farmer's self-selectivity in producer organizations or in accessing extension services. Farmers who are members of groups mostly self-select themselves to be members, rendering membership non-random. Farmers might be members of organizations or participate in extension services, due to some unobservables characteristics e.g. motivation; that is not controlled for in OLS regression. Furthermore, membership in groups or the access to extension might be driven by a farmer level of food production. Concurrently, a farm household with a low-average food productivity might join a farmer organization with the motivation of improving his/her level of food production. At the same time, a farmer who have high-average food production might join cooperatives because of her/his high level of food production. To control for biases that stem from observable factors, we could include in the OLS specification as many as justifiable exogenous control variables. However, for selection biases arising from unobservables, we could address it by the means of the instrumental variable method. The usual instrumental variables (IV) regression is a two-stage estimation approach. However, empirically and similarly to Adams *et al.* (2009), the IV method adopted in this paper follows Wooldridge (2010, p: 937-942) three-step approach of IV estimation with an endogenous dummy variable (referred as 2SLS-IV). In a first step, we estimate a probit model for each endogenous dummy variable (group membership and extension services) as functions of the respective instruments (solar grids and extension needs) and other control variables. In the second step, we regress each endogenous dummy variable (group membership and extension services) on the predicted probabilities from the first step of the endogenous variables ($\hat{C}$ and $\hat{E}$) and X. In the third step, we regress Y on the predicted values of the second step and the covariates (X). In other words, after the first step, the fitted probabilities are used as instruments for the endogenous dummies in a usual two Stages Least Squares IV estimation of equation (2.2). This estimation procedure exploits better the binary nature of our endogenous variables and produces more precise estimates. In addition, the usual 2SLS standard errors and test statistics are asymptotically valid (Wooldridge, 2010).

The 2SLS-IV technique requires at least valid instruments at the first stage of estimation. A valid instrument (Z) has to fulfil two important conditions: (i) the relevance condition i.e. it has to be significantly correlated with the endogenous variable (group membership or extension services) and, (ii) the exclusion restriction i.e. this instrument has to affect the food availability of farmers only through the endogenous variable. We use household ownership of solar grids as instrument for membership in producer organizations. Thus, from the question: "what is your main source of fuel for lightning?", we created a dummy variable "solar grids" which takes the value 1, if the household uses solar grids as lightning fuel and the value 0, otherwise. The use of solar grids expresses the inner motivation of farmers towards new technologies, predisposition to learn, to invest in innovations or to take a risk. Farmers who use solar grids are expected to participate actively in farmers' groups. However, using solar grids as lightening fuel is not supposed to directly affect the household food indicators, but only through group membership. Access to extension services is instrumented by the self-expression of farmers for the need for support. Similarly to the first instrument, from the two questions: "do you need extension services?" and "what do you need extension services for?", we created a dummy variable "extension needs" which takes the value 1, if the household responds that he needs extension first and he needs supports and the value 0, otherwise. Farmers who need to be supported are expected to have access to extension services, or at least exploring ways to have access to it.

To check for the validity of these instruments, we run separately, probit models of the endogenous binary variables C and E on Z and X (previously described as the first stage). This was followed by OLS regressions of the outcomes on group membership, extension services, covariates X, and the instruments, and we checked the significance of the instrument coefficients. As argued by Adams *et al.* (2009), the IV approach does not require the probit specifications to be correct, its only requires the designed instrument to be correlated with the endogenous variable. Results of the probit estimations, in table 2.5, show that the suggested instruments are positively correlated respectively with group membership ($z = 2.848$, p-value < 0.01) and extension services ($z = 4.258$, p-value < 0.01). Therefore, we can conclude that our instruments are relevant. The instruments affect the respective endogenous variables in the right and predicted direction, and they are also strongly correlated to these endogenous variables. In addition, OLS regressions reveal that the instruments are not directly correlated with the outcome variables ($F = 0.623(2)$, p-value $= 0.536$, $F = 0.558(2)$, p-value $= 0.572$).

The implication of using the instrumental variables technique is that in this model, γ measures the local effect of group membership. This means that IV estimates of γ measure the impact of group membership for households that are affected in their choice to be members by the instrument (i.e. the use of solar grids).

As motivated previously, in the presence of spatial heterogeneity in cross-sectional data, non-spatial regression models violate the classical assumptions of the independence between observations. The error terms ϵ in equation 2.2 are no longer identically and independently distributed, therefore the obtained estimates are biased and inconsistent (LeSage, 2008). Once, the conventional assumption is relaxed, one has to find ways to model the structure of the dependence between observations. When unobserved and unobservable spatial features affect observations, the spatial heterogeneity of these features leads to spatially correlated errors. We assume therefore that $W \neq 0$, and estimate the spatial error model (referred to as SEM). The SEM model accounts for spatial heterogeneity between farmers food availability outcome, it is specified as:

$$Y = \alpha + \gamma C + \theta E + \beta X + \epsilon \quad \text{with} \quad \epsilon = \lambda W \epsilon + \xi, \tag{2.3}$$

where Y, C, E, and X are defined as previously; α, γ, θ, β, and λ are parameters to be estimated respectively for group membership, extension services, the other control variables, and the spatial error lag; λ measures the spatial autocorrelation and falls between a value of -1 and 1; ϵ and ξ are the error terms; and W is a pre-specified

$(N \times N)$ exogenous spatial weights matrix. Since ordinary least squares are assumed to produce non consistent estimates for spatial models (LeSage, 2008). The SEM model is estimated using the maximum likelihood estimation method (Ord, 1975; Anselin, 1988).

The spatial weight matrix W is a symmetric matrix, where its elements w_{ij} express closeness or proximity of a household i with a household j. As common practice, to enable interpretation of model coefficients, W is row standardized so that the sum of the row elements equals one. In addition, the diagonal elements w_{ii} are set to zero, in order to prevent the effect of the i household from directly predicting itself. Many specifications of weight matrices have been used in the literature, and specifying the weight matrix is arbitrary. However, prior knowledge of the study population and economic theory can help guide in the specification of these matrices. We consider in our study only one specification, the exact inverse distance matrix, which expresses the geographical proximity of farmers. Neighbors in this specification have different weights, and those with higher weights are closer in distance. In an inverse distance matrix W, elements w_{ij} are defined as $1/d_{ij}$, where d_{ij} is the Euclidean distance between households i and j.

Finally, we also estimate the same SEM model (equation 2.3) by accounting this time for the endogeneity of group membership and access to extension services (referred to as SEM-IV). As demonstrated by Betz *et al.* (2019), the widely used IV models generally ignore the spatial patterns of the outcome variables, leading in asymptotically biased estimates even when instruments are randomly assigned. Furthermore, if the instrument exhibits spatial patterns similar to that of the outcome (as in many popular instruments that are not randomly distributed across space), the bias in IV estimates increases, and sometimes, they are greater than that of ordinary least squares (Betz *et al.*, 2019). The Generalized Spatial Two-Stage Least Squares method (GSTSLS) of Kelejian and Prucha (1998, 1999) was used to estimate the SEM-IV model. GSTSLS estimators are justified in our estimation strategy, due to the presence of two endogenous explanatory variables, group membership and access to extension, that need to be instrumented (Elhorst, 2010). Appendix A.2.1 presents in-depth details on the Generalized Spatial Two-Stage Least Squares approach.

Before implementing the spatial error models, the spatial autocorrelation index or Moran's I was employed to test whether there is a spatial correlation between farmers' food availability outcome. Based on the spatial weights, Moran's I statistic is computed as:

$$I = (\sum_i \sum_j w_{ij}(Y_i - \bar{Y})(Y_j - \bar{Y}))/(\sum_i (Y_i - \bar{Y})^2), \tag{2.4}$$

where w_{ij} is a spatial weight between households i and j; Y_i represents the food availability outcome of household i; and $\bar{Y}$ is the mean of the food availability outcome. The range of Moran's I is $(-1, 1)$, with 1 indicating perfect spatial similarity (or positive spatial correlation), 0 indicating no spatial correlation, and -1 indicating perfect dispersion (or negative correlation). If we observe a significant spatial autocorrelation based on Moran's I statistic, spatial regressions models should be used to correct for the spatial autocorrelation errors. In addition, using residuals from the non-spatial OLS model, we also computed the standard Lagrange Multiplier (LM_{error}) test for spatial error correlation (Anselin *et al.*, 1996). The standard LM_{error} test is specified as:

$$LM_{error} = [e'We/(e'e/N)]^2/[tr(W^2 + W'W)] \tag{2.5}$$

where e denotes the estimated residual from the non-spatial model; N is the number of farmers; and W are defined as previously. The Maximum Likelihood method in the R package *spdep* (Bivand and Wong, 2018) was used to estimate the SEM model. Following Betz *et al.* (2019), the Generalized Spatial 2SLS built in the R package *sphet* Piras (2010) was used to estimate the SEM-IV.

2.2.2 Data sources and variables description

The data used in this paper is primarily derived from a cross-sectional survey conducted in Senegal, which randomly sampled 4480 households that mainly produce rainfed or dry cereals (millet, sorghum, maize, fonio, rainfed rice). The data was collected in 2017 in the framework of the Agricultural Policy Support Project (Projet d'Appui aux Politiques Agricoles, PAPA in French) funded by USAID. The Senegalese National Agricultural Research Institute (ISRA) conducted the survey, with the support of the International Food Policy Research Institute (IFPRI). A multi-stage sampling procedure was applied for the selection of households. Data covered the main agricultural season of 2016/2017. A structured household questionnaire was used and the collected information included crop production, rural producer organizations membership, household assets, access to infrastructure, access to rural institutions, use of agricultural technologies, and household demographic and socioeconomic characteristics, inputs use information, markets prices of both inputs and outputs, and climatic shocks. Surveyed households were located in all six agro-ecological zones of Senegal.

Although the survey was directed towards cereals farming, many of the surveyed households did not produce (or did not report) cereals harvests for the season, we, therefore, restrained our sample to households that produce the main five types of cereal including millet, maize, rainfed rice, sorghum, and fonio. These crops are also the principal elements of rural households' diets. The final sample comprises then 3939 farm households.

Table 2.1 presents the definition and summary statistics of the variables used in the analysis. Two outcome variables were used as proxy for household food production: the total production of cereal crops during the whole season (referred hereafter as cereal production), and the total quantity of calories produced per day and per equivalent adult (referred hereafter as daily food calories). The dependent variable cereal production, expressed in West African Franc (FCFA)[4] represents the gross value of all cereal productions valued at the market prices. The considered cereals are millet, maize, rice, sorghum, and fonio. This approach is more suitable to compare farmers, since most cereals productions are not marketed by farmers, and their weights are not valued at the same market price. In addition, cereals grains represent a large proportion of the dietary energy supply, especially in rural areas. Farmers in our sample produce an average of 1315 kg of cereals (of which 622 kg is millet, 298 kg is rice, 279 kg is maize and 115 kg is sorghum) representing a total average market value of 222,720 FCFA. Households also produce beans and roots crops, for instance, they harvested about 45 kg of cowpea and 12 kg of cassava on average.

Concerning the dependent variable daily food calories, it is expressed in terms of per adult equivalent food calorie available per day. This variable is computed using the gross production of all five cereals previously mentioned, plus the productions of legumes (cowpea, Bambara groundnut), and of roots and tuber crops (cassava, sweet and yellow potatoes, and taro). First, the farm-gate productions were converted into available food crop productions by assigning to them post-harvest losses ratios [5].

[4] 1 FCFA=0.0017 USD as at December 2019.

[5] We used APHILIS data post-harvest loss ratios retrieved at https://www.aphlis.net/. The post-harvest losses include losses at the harvesting or field drying, platform drying, threshing and shelling, winnowing, transport and storage. The values were aggregated at the regional level and applied to farmers' production. Table 2.6 in the appendix shows the aggregated post-harvest loss for each cereal crop per region. For rice production, the milled rice ratio of 0.67 was first applied.

Second, the derived available food crop productions are converted into calories (kcal) using the West African Food Composition (Stadlmayr *et al.*, 2012). The conversion ratios from kilograms to kilo-calories are presented in the appendix in table 2.7. Third, the total amount of food calories was divided by the total household size which was previously converted into adult equivalent following Claro *et al.* (2010)[6]. Finally, the obtained value is divided by 365 to have the daily food available per adult equivalent. The daily per adult-equivalent food availability approach is used because it can determine the capacity of each household to provide proper food energy to its members during a whole calendar year. Households in the sample produce on average, 1358.71 kcal per day and per adult equivalent. These two dependent variables were logged in the econometric estimations.

Following the definition of Bernard *et al.* (2015), our variable of interest "group membership" is referred to membership in a rural producer organization that provides farmers with farming and farm-related services including access to inputs, markets and credit, collective sales, and capacity development. Eight types of farmer organizations were mentioned by the surveyed units: Producers Groups, Economic Interest Groups, Rural Associations, Cooperatives, Women Producers Groups, Federations, Unions, and Networks. The variable "group membership" is binary, coded as 1 if a member of the household belongs to any of this group, and 0 otherwise. In some households, several family members expressed their belonging to these groups, with a maximum of 7 members. However, on average only one family member belongs to a group. About 9% of the households in the sample have at least one person belonging to a group. The main organizations, which gather most household members, are the Economic Interest Groups (43.6%), Rural Associations (17.3%), Producers Groups (16.7%), and Cooperatives (15.3%).

Several control variables have been included in the models, notably the household and its heads socioeconomic characteristics, household's assets and access to rural institutions, ecological conditions, and some environmental risks factors. Household socio-economic characteristics variables include gender, age, active household size, dependents, education, and migration. Gender is a dummy variable for the gender of the household head, with value 1 for males and 0 otherwise. The households in our sample are predominantly male-headed, with more than 94% of males as household heads. Meanwhile, the age of household head ranges from 16 to 96 years, with an average of 53 years. The household size is a continuous variable that was

[6]Reference and conversion ratios are presented in table 2.8 in appendix

categorized into 2 groups: active and non-actives members[7]. The average active household size in the sample is around 6 indicating the existence of enough family labor for agricultural tasks. We also include a dummy variable for the migration status of the household head. This variable serves as a proxy for involvement in off-farming activities. Education is a binary variable coded as 1 if the farmer has attended at least primary school and 0 if he has no formal education. Most farmers in the sample are not formally educated (more than 60%), they can not read nor write. Household assets included equipment and total land area owned. Equipment represents the total value in FCFA of all agricultural implements owned by the household. On average, households in the sample owns about 130,000 FCFA of agricultural implements and about 5.93 hectares of farm land. Variables related to access to infrastructure and institutions include distance to the nearest road and access to extension services. Only around 11% of the farmers in our sample have access to extension services. Ecological condition variables include rainfall, the percentages of clay in soils[8]. Dummy variables for agro-ecological zones are also included due to the expected spatial heterogeneity in farmers' conditions to produce food. Most of the households are located in the Groundnut, Casamance and South East agro-ecological zones and this, constitute, more than 88% of farmers in the sample. The included environmental risks variables, faced by the farmer and that could have affected food production during the season, are drought, crop diseases, and the early stop of rain.

[7]The first category comprises active members, aged between 15 and 65 years, and second regroups dependents i.e. members aged below 15 years and more than 65 years.

[8]Rainfall data was retrieved from publicly available databases of the Climate Hazards Center of the University of California (https://www.chc.ucsb.edu/data), using the surveyed households location coordinates. Soil percentages of clay, silt and sand were also retrieved from publicly available databases from International Soil Reference and Information Centre (ISRIC – World Soil Information) at `https://data.isric.org/` using the geographical coordinates of each household. The Database uses machine learning and data collected in 2017 and 2018.

Table 2.1: Description of variables

Variables	**Description and measurement**	**Pooled** (1)	**Members** (2)	**Non-Members** (3)	**P-values** (4)
Groups Membership	Membership in farmers groups (1=yes, 0=no)				
Cereals production	Cereals production (1.000 FCFA)	222.72 (416.80)	415.82 (1112.27)	203.71 (255.59)	<0.01
Daily food calories	Daily food availability (kcal/adult equivalent)	1358.71 (2848.97)	2235.73 (6655.18)	1272.38 (2117.50)	0.01
	Cereal, cowpea and cassava Production				
Cereals	Total Cereal Production (Kg)	1315.60 (2991.97)	2852.84 (8770.41)	1164.28 (1423.09)	<0.01
Millet	Millet Production (Kg)	622.37 (1005.91)	436.44 (874.42)	640.67 (1016.21)	<0.01
Maize	Maize Production (Kg)	279.96 (740.76)	425.21 (1144.43)	265.66 (686.93)	0.01
Rice	Rice Production (Kg)	298.23 (2705.96)	1800.20 (8756.23)	150.38 (520.68)	<0.01
Sorghum	Sorghum Production (Kg)	115.04 (431.55)	190.99 (588.21)	107.57 (412.29)	0.01
Fonio	Fonio Production (Kg)	4.82 (79.11)	3.80 (52.16)	4.92 (81.29)	0.72
Cowpea	Cowpea Production (Kg)	45.43 (238.08)	16.74 (124.63)	48.26 (246.27)	<0.01
Cassava	Cassava Production (Kg)	12.16 (281.98)	0.00 (0.00)	13.36 (295.51)	0.01
	Household and Head characteristics				
Gender	Household head is a male (1=yes, 0=no)	0.94 (0.24)	0.95 (0.23)	0.94 (0.25)	0.40
Age	Age of household head (years)	53.00 (13.47)	51.03 (12.24)	53.19 (13.57)	<0.01
Education	Formal education (1=yes, 0=no)	0.37 (0.48)	0.51 (0.50)	0.35 (0.48)	<0.01
Active members	Active family members	6.10 (3.31)	6.72 (3.57)	6.04 (3.27)	<0.01
Dependents	Non-active family members	4.02 (3.25)	4.82 (3.86)	3.94 (3.18)	<0.01
Migrant	Household head is a migrant (1=yes, 0=no)	0.23 (0.70)	0.26 (0.72)	0.23 (0.70)	0.38
	Household Assets				
Equipment	Agricultural Equipment (1.000.000 FCFA)	0.13 (0.58)	0.18 (0.48)	0.13 (0.58)	0.05
Area Owned	Land size owned by household (ha)	5.93 (8.46)	5.78 (6.33)	5.94 (8.64)	0.65
	Access to infrastructure				
Distance to road	Distance to nearest all-weather road (km)	10.41 (14.57)	10.76 (14.16)	10.38 (14.61)	0.63
Extension	Access to extension services (1=yes, 0=no)	0.11 (0.31)	0.45 (0.50)	0.07 (0.26)	<0.01
Seed subsidy	Access to subsided seeds	0.39 (0.49)	0.47 (0.50)	0.39 (0.49)	<0.01
Fertilizer subsidy	Access to subsided Fertilizers	0.33 (0.47)	0.63 (0.48)	0.30 (0.46)	<0.01
	Agro-ecological zones				
Groundnut AEZ	Groundnut agro-ecological zone (1=yes, 0=no)	0.49 (0.50)	0.29 (0.45)	0.51 (0.50)	<0.01
Casamance AEZ	Casamance agro-ecological zone (1=yes, 0=no)	0.27 (0.44)	0.33 (0.47)	0.26 (0.44)	0.02
South-East AEZ	South East agro-ecological zone (1=yes, 0=no)	0.12 (0.32)	0.16 (0.36)	0.11 (0.32)	0.03
Other AEZ	Other agro-ecological zones (1=yes, 0=no)	0.12 (0.32)	0.16 (0.36)	0.11 (0.32)	0.03
	Ecological conditions				
Rainfall	Rainfall 2016 (m)	0.70 (0.29)	0.70 (0.35)	0.70 (0.28)	0.65
Clay	Percentage of clay (%)	20.23 (7.21)	23.65 (5.81)	19.90 (7.24)	<0.01
Drought	Drought (1=yes, 0=no)	0.08 (0.27)	0.07 (0.26)	0.08 (0.27)	0.56
Early Rain Stop	Early rain stop (1=yes, 0=no)	0.37 (0.48)	0.37 (0.48)	0.37 (0.48)	0.74
Crop disease	Crop disease (1=yes, 0=no)	0.07 (0.26)	0.07 (0.25)	0.07 (0.26)	0.72
	Instrumental Variables				
Solar grids	Use of solar grids as lightening (1=yes, 0=no)	0.10 (0.30)	0.15 (0.35)	0.10 (0.30)	0.01
Extension needs	Express need for support (1=yes, 0=no)	0.17 (0.38)	0.14 (0.35)	0.17 (0.38)	0.15
N	Number of Observations	3939	353	3586	3939

2.3 Results and discussion

2.3.1 Comparative descriptive analysis

Columns 2, 3 and 4 of table 2.1 shows the descriptive statistics of producer organizations members and non-members, with the associated p-values of computed differences between means. When comparing farmer group members to non-members, statistically significant differences can be observed for some of their characteristics. Group members tend to have larger household size than non-members and they appear to be on average more educated. Group members have better access to rural institutions such as extension and subsidies than non-members. For the two indicators of food availability, groups members seem to produce more food than non-members and the differences are significant (p-value < 0.01). These differences suggest that farmer organizations might play an important role in enhancing farmers' ability to produce more food and improve locally food security. However, this result does not allow one to make inferences about the impact that farmer group membership might have on farmers' food availability. These comparisons of mean differences do not account for confounding factors such as observed household and farm-level characteristics and unobserved factors (e.g., farmers' innate skills, perception and motivations of membership decision).

2.3.2 Econometric estimations

Tables 2.3 and 2.4 present econometric estimations respectively for cereal productions and daily food calories. In each table, columns (1), (2), (3) and (4) refer to the OLS model, the final step of the 2SLS-IV model, SEM specification, and SEM-IV model respectively. After the results are presented, they are compared and discussed in a separate section.

The OLS regressions for the two outcomes show an adjusted R^2 value of of 0.174 and 0.205. The computed root mean square errors (RMSE) are low compared to the food availability indicators values (1.026 and 1.039 versus the means values of 12.313 and 7.214). The estimated coefficients of our variable of interest, i.e., group membership, are positive and significantly different from 0 for the two indicators. Estimates of the effect show that group membership improves farm households' food security indicators in terms of cereal production and daily food calories by 18.7% and 15.1% respectively. This would mean that when one controls for the observed characteristics of farmers, being a member of farmer organizations affects positively and significantly the quantity of food available for the farm household.

Results of our IV model exhibit significant F tests for weak instruments, suggesting that the predicted probabilities obtained from the first step are sufficiently strong instruments, corroborating previous justifications of the used instruments (i.e solar grids use and the need for extension services). In addition, the Wu-Hausman test reports an F-statistic = 12.03 with an associated a p-value <0.01 for cereals productions and an F-statistic = 8.67 and p-value <0.01 for daily food calories, indicating that IV model results are more consistent than OLS, and supporting the use of instrumental variables technique. If we assume that IV estimation techniques are unbiased, the coefficient of our variable of interest can be interpreted as the local average treatment effect. The IV estimates show that the coefficient of group membership exhibits positive and statistically significant value at 1%, indicating that membership in a producer organization has a strong and positive effects of 28.5% and 20.5% on respectively households' ability to produce cereals and food calories when one controls for bias stemming from households observable and unobservable characteristics. These results are congruent to those obtained from OLS.

As stated previously, before implementing the spatial models, we computed two spatial autocorrelation tests. Table 2.2 shows that the Moran's I statistic, for the two outcome variables, are positive and highly significant ($p < 0.01$), indicating that there is a strong positive spatial correlation between farmers food production indicators. Farmers with relatively high food availability seem to live close to other farmers with a high level of food production, and households with relatively low food production also tend to live near households with low food availability. These results suggest that controlling for spatial auto-correlation should be considered in our analysis. Furthermore, using residuals from OLS estimates, we computed the standard Lagrange Multiplier error test (Anselin, 1988). The null hypotheses were rejected mostly at $p < 0.01$, indicating that at least spatial auto-correlation should be incorporated in the models. These spatial correlation tests' results are corroborated by the Moran plots presented in figure 2.1, where clouds of points could be observed in some of the plot quadrants, indicating that household food production outcomes are strongly and positively correlated to their neighbours' ones.

Table 2.2: Moran I and Lagrange Multiplier tests

	Cereals production	**Daily Food calories**
Moran I	0.240***	0.278***
Standard LM Error	602.9***	628.92***

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

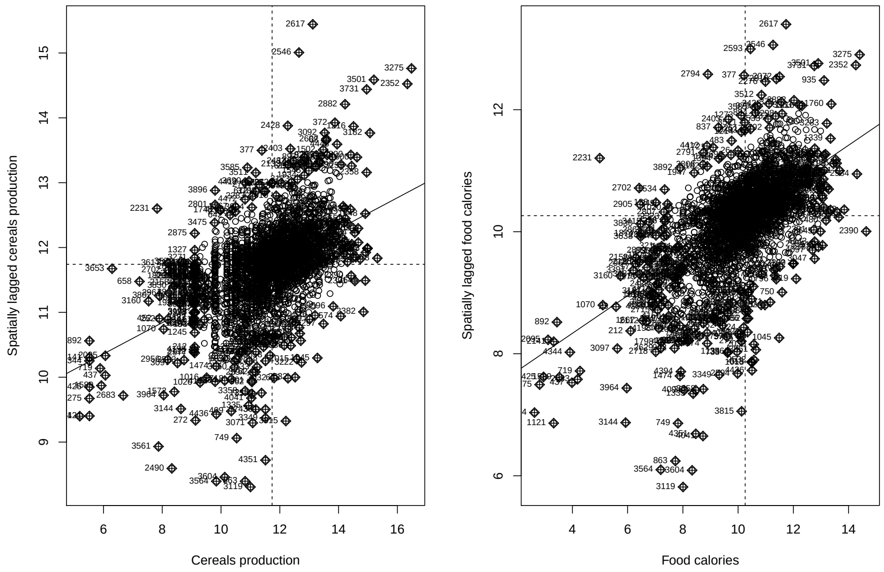


Figure 2.1: Moran Plots

Columns 3 of tables2.3 and 2.4 present the results from the SEM model. For both outcome variables, the likelihood ratio test statistic of 527.214 is significant at 1%, suggesting that the spatial error model fits better than a simple linear model. In addition, the Akaike Infomation Criterion (AIC) values of 10877.855 and 10964.202 obtained for the spatial error models are lower than that of the linear models (11403.063 and 11507.246), indicating a better model fit for the SEM models. For both food security indicators, the spatial lag error coefficients λ are positive, significantly different from 0 and with high values of 0.599 and 0.604, suggesting a high spatial correlation of farmers' food availability indicators. This would mean that there is high spatial heterogeneity in food production indicators, due to spatial observable and unobservable characteristics. Furthermore, the variable of group membership shows a positive sign, and this is statistically different from 0 at 1% and 5% levels of significance. These results suggest that, when one controls for spatial heterogeneity, that is, the role that geography plays in Senegalese food availability, belonging to rural producer organization influences significantly and positively the farmers ability to produce food, with an increase of 18% and 15% percentage points respectively for cereals and daily food calories production.

Table 2.3: Models Estimations: Cereal Productions

	OLS (1)	2SLS-IV (2)	SEM (3)	SEM-IV (4)
Intercept	10.643 (0.277)***	10.576 (0.277)***	10.216 (0.305)***	10.195 (0.331)***
Group Membership	0.187 (0.064)***	0.285 (0.072)***	0.180 (0.061)***	0.19 (0.074)**
Gender	0.357 (0.068)***	0.356 (0.068)***	0.252 (0.063)***	0.254 (0.066)***
Age	−0.005 (0.008)	−0.005 (0.008)	−0.002 (0.007)	−0.002 (0.007)
Age squared	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Education	−0.006 (0.036)	−0.013 (0.036)	0.006 (0.035)	0.002 (0.035)
Active members	0.038 (0.006)***	0.037 (0.006)***	0.046 (0.005)***	0.046 (0.007)***
Dependents	0.007 (0.006)	0.007 (0.006)	0.011 (0.005)**	0.011 (0.005)**
Migrant	−0.121 (0.048)**	−0.127 (0.048)***	−0.020 (0.045)	−0.028 (0.045)
Equipment	0.102 (0.029)***	0.100 (0.029)***	0.066 (0.026)**	0.066 (0.026)**
Area owned	0.031 (0.002)***	0.031 (0.002)***	0.029 (0.002)***	0.029 (0.007)***
Distance to road	−0.002 (0.001)	−0.002 (0.001)	−0.002 (0.002)	−0.001 (0.002)
Extension	0.370 (0.058)***	0.461 (0.070)***	0.233 (0.058)***	0.354 (0.07)***
Seeds subsidy	−0.027 (0.041)	−0.027 (0.041)	0.050 (0.040)	0.046 (0.038)
Fertilizer subsidy	0.321 (0.042)***	0.301 (0.043)***	0.209 (0.042)***	0.203 (0.04)***
Clay	0.009 (0.004)**	0.007 (0.004)*	0.006 (0.005)	0.005 (0.006)
Rainfall	0.510 (0.452)	0.785 (0.456)*	1.516 (0.651)**	1.63 (0.743)**
Rainfall squared	−0.736 (0.244)***	−0.867 (0.246)***	−1.192 (0.352)***	−1.252 (0.403)***
Groundnut AEZ	0.185 (0.082)**	0.189 (0.082)**	0.162 (0.105)	0.164 (0.125)
Casamance AEZ	0.281 (0.112)**	0.263 (0.112)**	0.228 (0.148)	0.222 (0.162)
South-East AEZ	0.537 (0.105)***	0.522 (0.106)***	0.448 (0.144)***	0.443 (0.152)***
Drought	0.019 (0.062)	0.020 (0.062)	0.057 (0.061)	0.056 (0.06)
Early rain stop	−0.120 (0.036)***	−0.123 (0.036)***	−0.064 (0.037)*	−0.07 (0.037)*
Crop diseases	−0.073 (0.065)	−0.071 (0.065)	−0.106 (0.064)	−0.106 (0.064)*
λ (Spatial error lag)			0.599 (0.023)***	0.630 (0.027)***
Adj. R^2	0.174	0.173		
RMSE	1.026	1.026		
LR test			527.214***	
AIC	11403.069		10877.855	
Weak Instruments (Group membership)		7641.24***		
Weak Instruments (Extension)		4797.36***		
Wu-Hausman test		12.030***		
N	3939	3939	3939	3939

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$. Standard errors are in parentheses

Table 2.4: Models Estimations: Daily Per Adult Equivalent Food Calories

	OLS	2SLS-IV	SEM	SEM-IV
	(1)	(2)	(3)	(4)
Intercept	7.058 (0.281)***	7.000 (0.281)***	6.670 (0.309)***	6.638 (0.334)***
Group Membership	0.151 (0.064)**	0.205 (0.073)***	0.153 (0.062)**	0.131 (0.075)*
Gender	0.320 (0.069)***	0.319 (0.069)***	0.198 (0.064)***	0.199 (0.068)***
Age	−0.011 (0.008)	−0.012 (0.008)	−0.009 (0.007)	−0.009 (0.007)
Age squared	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Education	0.023 (0.036)	0.017 (0.036)	0.035 (0.035)	0.033 (0.036)
Active members	−0.060 (0.006)***	−0.060 (0.006)***	−0.050 (0.006)***	−0.051 (0.007)***
Dependents	−0.061 (0.006)***	−0.061 (0.006)***	−0.057 (0.005)***	−0.056 (0.005)***
Migrant	−0.115 (0.048)**	−0.122 (0.048)**	−0.020 (0.045)	−0.027 (0.045)
Equipment	0.135 (0.030)***	0.133 (0.030)***	0.100 (0.026)***	0.099 (0.032)***
Area owned	0.034 (0.002)***	0.034 (0.002)***	0.031 (0.002)***	0.031 (0.007)***
Distance to road	−0.003 (0.001)*	−0.002 (0.001)*	−0.002 (0.002)	−0.002 (0.002)
Extension	0.327 (0.059)***	0.438 (0.071)***	0.211 (0.058)***	0.355 (0.072)***
Seeds subsidy	0.008 (0.042)	0.009 (0.042)	0.060 (0.041)	0.058 (0.038)
Fertilizer subsidy	0.275 (0.043)***	0.258 (0.043)***	0.200 (0.042)***	0.193 (0.041)***
Clay	0.001 (0.004)	−0.001 (0.004)	−0.002 (0.005)	−0.003 (0.006)
Rainfall	0.038 (0.458)	0.281 (0.462)	1.029 (0.662)	1.175 (0.737)
Rainfall squared	−0.563 (0.247)**	−0.681 (0.249)***	−1.012 (0.358)***	−1.088 (0.403)***
Groundnut AEZ	0.002 (0.083)	0.005 (0.083)	−0.028 (0.106)	−0.026 (0.122)
Casamance AEZ	0.210 (0.113)*	0.197 (0.113)*	0.127 (0.150)	0.121 (0.16)
South-East AEZ	0.461 (0.107)***	0.448 (0.107)***	0.341 (0.146)**	0.335 (0.151)**
Drought	0.047 (0.063)	0.046 (0.063)	0.065 (0.062)	0.064 (0.061)
Early rain stop	−0.135 (0.036)***	−0.140 (0.036)***	−0.083 (0.037)**	−0.09 (0.038)**
Crop diseases	−0.046 (0.066)	−0.046 (0.066)	−0.086 (0.065)	−0.088 (0.066)
λ (Spatial error lag)			0.604 (0.022)***	0.653 (0.028)***
Adj. R^2	0.205	0.204		
RMSE	1.039	1.040		
LR test			545.045***	
AIC	11507.246		10964.202	
Weak Instruments (Group membership)		7641.242***		
Weak Instruments (Extension)		4797.357***		
Wu-Hausman test		8.671***		
N	3939	3939	3939	3939

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors are in parentheses

Concerning the SEM-IV models, we find that, the spatial lag error coefficients λ are positive and statistically at the 1% level under both outcomes. With the high relative values of 0.630 and 0.653, this shows that high spatial heterogeneity exists in households food availability outcomes. Farmer group membership also exhibits positive and statistically significant at 5% and 10% levels. These results would indicate that, when controlling for spatial heterogeneity and selection bias, farmer organization membership influences significantly and positively food availability indicators by 19% and 13.1%, respectively for the production of cereals and the production of daily food calories. We called these estimated values the "spatial local average treatment", similarly to the standard local average treatment.

Besides membership in a farmer organizations, other explanatory variables were also significantly associated with the food availability outcomes. The other factors that affect significantly food production in the households are gender of household head, number of active household members, assets owned, access to extension services and to fertilizer subsidies, level of annual rainfall, South-East agro-ecological region, and early stop of rain.

2.3.3 Discussion

The different results obtained with the various estimation techniques indicate that belonging to a producer or farmer organization increases, in general, the level of household food availability. If we consider that food availability and especially domestic food production is an important factor in fighting food insecurity, then our estimates suggest that farmer organizations are effective at improving farm households' food security. Similar results were observed in Ethiopia by Zeweld *et al.* (2015) who used the total expenditure per adult equivalent as a proxy for household food security level and Heckman selection model. Our findings also corroborate recent results in the growing literature on farmer-based organizations in developing countries, where most scholars observed a positive correlation between farmer group membership and farm performances (Verhofstadt and Maertens, 2014) and farm households economic welfare (Ma and Abdulai, 2016).

The good performances of farmer group members in Senegal could be explained first by the differential impact of the adoption of agricultural technologies such as fertilizers. Our comparative t-test analysis shows significant differences between members and non-members, in access to seed and fertilizers subsidies. Subsidies, in general, improve and encourage the use and intensity of use of technologies, and therefore enhance farm productions and households' food security. In addition, the fertilizer subsidies variable is positively and significantly correlated with food productions in all regressions. As shown by Abebaw and Haile (2013), membership in farmer organizations such as cooperatives affects positively the use of technologies (e.g. fertilizer). The adoption of fertilizers is also induced by the differential ease that members of organizations might have to access them, because of being members. For example, the study by Ajah (2015) in Nigeria found that in general farmer-based organizations members had significantly better access to farm inputs than non-members. Furthermore, previous studies have also shown that membership in farmer organizations is motivated by the reduction of transaction costs and therefore improving the access of members to farm inputs and technologies compared to non-members.

Earlier, we argued about the importance of the differential effects of rural infrastructure and institutions. The performance of farmer organizations could also be explained by the differential access to extension services (significant at $p < 0.01$). Farmer group members tend to have better access to extension and therefore they are more prone to have access to the necessary knowledge and new technologies to increase farm productivity and food productions. In addition, the social networks within organization members can be an important channel to diffuse and receive knowledge and enhance therefore their level of food productions.

Regarding the other drivers of household food production, male-headed households seem to produce more food. A plausible interpretation would be that in rural areas male-headed households compared to female-headed, are more likely to have better access to production inputs (such as labor and secured land) and agricultural modern technologies, therefore they are able to produce more food crops. Owned assets also play an important role in ensuring food production, farmers who have more assets are generally more capable of producing more food crops. We also observed that the relationship between rainfall and food availability indicators exhibits an inverted U-shape behaviour, suggesting that farm household food security in Senegal is sensitive to rainfall. Being in the South-East zone also improves food production. South-East is the predilection zone of cereals production in Senegal. Farmers living in that area are therefore more prone to produce efficiently and sufficiently cereals and other food crops. Meanwhile, the early stop of rain impedes significantly food production in Senegal, backing the previous finding regarding the effect of annual rainfall.

Furthermore, our results showed that farmers' food productions are affected by location spatial features. Although the regressions tried to incorporate most of the geographic and physical variables that are observable, the estimated spatial correlation coefficients are still high, denoting the presence of unobserved and unobservable spatial characteristics that seriously affect Senegalese household's food availability. The non-inclusion of such strong spatiality into regressions would have led to overestimated coefficients.

2.4 Concluding remarks

Producer organizations can constitute the main vehicle for access to farm inputs and therefore enhance farm productivity and farm household's food production. However, despite the growing literature on collective action groups' importance in developing countries, no quantitative study on the impact of Senegalese farmers organizations has been done. This paper aimed to fill in the gap and contribute to the literature by applying various estimations techniques including a spatial econometric approach, on a country scale survey data, to derive quantitative effects of membership in farmer organizations on household food availability.

Estimations results revealed that farmer organization membership affects significantly and strongly farm households food production. In particular, belonging to a farmer organization improves significantly cereals productions and daily food calories available for the household. In addition, results show that households' food availability indicators are also positively and significantly correlated with the characteristics of the household (gender of the head, active and dependents members, the possession of agricultural assets), access to extension services and to fertilizers subsidies, and early stop of rainfall. These results were robust to changes in estimations techniques. Furthermore, farmers' food production is also driven by spatial features. These findings support the idea that rural producer organizations have the potential to benefit rural households' food security levels by providing conditions and the necessary social networks for access to technologies, knowledge and production inputs.

In terms of policy implications, future support for farmer organizations in Senegal should take into account the spatial distribution of farmers. Future research should investigate the geographic distribution of farmer organizations and their impacts on farm households' performances. In addition, as demonstrated, group members perform better compared to non-members, however, their efficiency is questionable in regard to the high level of production inputs used. Further analyses are then necessary to derive the effects of membership on members' technical efficiency.

References

Abebaw, D. and Haile, M. G. (2013) The impact of cooperatives on agricultural technology adoption: Empirical evidence from Ethiopia, *Food Policy*, **38**, 82 – 91.

Adams, R., Almeida, H. and Ferreira, D. (2009) Understanding the relationship between founder–CEOs and firm performance, *Journal of Empirical Finance*, **16**, 136 – 150.

Ajah, J. (2015) Comparative analysis of cooperative and non-cooperative farmers' access to farm inputs in Abuja, Nigeria, *European Journal of Sustainable Development*, **4**, 39–50.

Anselin, L. (1988) *Spatial Econometrics: Methods and Models*, NATO Asi Series. Series E, Applied Sciences, Springer Netherlands.

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

Barham, J. and Chitemi, C. (2009) Collective action initiatives to improve marketing performance: Lessons from farmer groups in Tanzania, *Food Policy*, **34**, 53 – 59.

Bernard, T., Frölich, M., Landmann, A., Unte, P. N., Viceisza, A. and Wouterse, F. (2015) Building trust in rural producer organizations in Senegal: Results from a randomized controlled trial, Tech. rep., IZA Discussion Papers.

Bernard, T., Taffesse, A. S. and Gabre-Madhin, E. (2008) Impact of cooperatives on smallholders' commercialization behavior: evidence from Ethiopia, *Agricultural Economics*, **39**, 147–161.

Betz, T., Cook, S. J. and Hollenbach, F. M. (2019) Spatial interdependence and instrumental variable models, *Political Science Research and Methods*, p. 1–16.

Bivand, R. S. and Wong, D. W. S. (2018) Comparing implementations of global and local indicators of spatial association, *TEST*, **27**, 716–748.

Chagwiza, C., Muradian, R. and Ruben, R. (2016) Cooperative membership and dairy performance among smallholders in Ethiopia, *Food Policy*, **59**, 165–173.

Claro, R. M., Levy, R. B., Bandoni, D. H. and Mondini, L. (2010) Per capita versus adult-equivalent estimates of calorie availability in household budget surveys, *Cadernos de saude publica*, **26**, 2188–2195.

Diagne, A. (2006) Diffusion and adoption of Nerica rice varieties in Cote d'Ivoire, *The Developing Economies*, **44**, 208–231.

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

FAO, IFAD, UNICEF, WFP and WHO (2019) *The State of Food Security and Nutrition in the World 2019: Safeguarding against economic slowdowns and downturns*, The State of Food Security and Nutrition in the World (SOFI), FAO.

Fischer, E. and Qaim, M. (2012) Linking smallholders to markets: Determinants and impacts of farmer collective action in Kenya, *World Development*, **40**, 1255–1268.

Francesconi, G. N. and Heerink, N. (2010) Ethiopian agricultural cooperatives in an era of global commodity exchange: Does organisational form matter?, *Journal of African Economies*, **20**, 153–177.

Francesconi, G. N. and Ruben, R. (2012) The hidden impact of cooperative membership on quality management: A case study from the Dairy Belt of Addis Ababa, *Journal of Entrepreneurial and Organizational Diversity*, **1**, 85–103.

Hathie, I., Seydie, B., Samaké, L. and Sakho-Jimbira, S. (2017) Ending Rural Hunger: The Case of Senegal, Tech. rep., The Brookings Institution.

Hoken, H. and Su, Q. (2015) Measuring the effect of agricultural cooperatives on household income using PSM-DID : a case study of a rice-producing cooperative in China, IDE Discussion Papers 539, Institute of Developing Economies, Japan External Trade Organization(JETRO).

Jouanjean, M.-A. (2013) Targeting infrastructure development to foster agricultural trade and market integration in developing countries: An analytical review, Tech. rep., Overseas Development Institute.

Kelejian, H. H. and Prucha, I. R. (1998) A generalized spatial two stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances, *Journal of Real Estate Finance and Economics*, **17**, 99–121.

Kelejian, H. H. and Prucha, I. R. (1999) A generalized moments estimator for the autoregressive parameter in a spatial model, *International Economic Review*, **40**, 509–533.

Latynskiy, E. and Berger, T. (2016) Networks of Rural Producer Organizations in Uganda: What can be cone to make them work better?, *World Development*, **78**, 572 – 586.

LeSage, J. and Pace, R. (2009) *Introduction to Spatial Econometrics*, Statistics, textbooks and monographs, CRC Press.

LeSage, J. P. (2008) An Introduction to Spatial Econometrics, *Revue d'économie industrielle*, **123**, 19–44.

Ma, W. and Abdulai, A. (2016) Does cooperativemembership improve household welfare? Evidence from apple farmers in China, *Food Policy*, **58**, 94–102.

Ma, W., Abdulai, A. and Goetz, R. (2018) Agricultural Cooperatives and Investment in Organic Soil Amendments and Chemical Fertilizer in China, *American Journal of Agricultural Economics*, **100**, 502–520.

McArthur, J. W. and McCord, G. C. (2017) Fertilizing growth: Agricultural inputs and their effects in economic development, *Journal of Development Economics*, **127**, 133 – 152.

Ord, K. (1975) Estimation methods for models of spatial interaction, *Journal of the American Statistical Association*, **70**, 120–126.

Piras, G. (2010) sphet: Spatial models with heteroskedastic innovations in r, *Journal of Statistical Software, Articles*, **35**, 1–21.

Republique du Senegal (2014) Rapport Defintif Recensement Général de la Population et de l'Habitat, de l'Agriculture et de l'Elevage - RGPHAE 2013, Tech. rep., Agence Nationale de la Statistique et de la Démographie, ANSD.

Stadlmayr, B., Charrondiere, U., Enujiugha, V., Bayili, G., Fagbohoun, E., Samb, B., Addy, P., Barikmo, I., Ouattara, F., Oshaug, A., Akinyele, I., Annor, G., Bomfeh, K., Ene-Obong, H., Smith, I., Thiam, I. and Burlingame, B. (2012) *West African Food Composition Table/ Table De Composition Des Aliments D'afrique De L'ouest.*

UNDP (2019) *Human Development Report 2019. Inequalities in Human Development in the 21st Century. Briefing note for countries on the 2019 Human Development Report. Senegal*, Human Development Report, United Nations Development Programme.

Verhofstadt, E. and Maertens, M. (2014) Smallholder cooperatives and agricultural performance in Rwanda: do organizational differences matter?, *Agricultural Economics*, **45**, 39–52.

Wanmali, S. and Islam, Y. (1997) Rural Infrastructure and Agricultural Development in Southern Africa: A Centre-Periphery Perspective, *The Geographical Journal*, **163**, 259–269.

Wooldridge, J. (2010) *Econometric Analysis of Cross Section and Panel Data*, The MIT

Press, MIT Press.
World Bank (2007) *World Development Report 2008 : Agriculture for Development*, World Bank Washington, DC.
World Bank (2019) World bank data, world Bank Data Base.
Wossen, T., Abdoulaye, T., Alene, A., Haile, M. G., Feleke, S., Olanrewaju, A. and Manyong, V. (2017) Impacts of extension access and cooperative membership on technology adoption and household welfare, *Journal of Rural Studies*, **54**, 223–233.
Zeweld, W., Van Huylenbroeck, G. and Buysse, J. (2015) Household food security through cooperative societies in northern Ethiopia, *International Journal of Development Issues*, **14**, 60–72.

Appendix

Table 2.5: Instruments Checking

Dependent Variable	Group Membership (1)	Extension (2)	Cereals Productions (3)	Daily Food Calories (4)
Group Membership		1.050 (0.084)***	0.187 (0.064)***	0.151 (0.065)**
Intercept	−1.182 (0.559)**	−1.663 (0.519)***	10.625 (0.278)***	7.042 (0.281)***
Sex	−0.113 (0.140)	0.093 (0.128)	0.358 (0.068)***	0.320 (0.069)***
Age	0.010 (0.017)	0.029 (0.016)*	−0.005 (0.008)	−0.011 (0.008)
Age squared	−0.000 (0.000)	−0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Education	0.317 (0.068)***	0.120 (0.064)*	−0.007 (0.036)	0.022 (0.036)
Active members	0.031 (0.011)***	−0.013 (0.011)	0.038 (0.006)***	−0.060 (0.006)***
Dependents	0.032 (0.010)***	−0.009 (0.010)	0.007 (0.006)	−0.061 (0.006)***
Migrant	0.005 (0.092)	0.298 (0.080)***	−0.125 (0.048)***	−0.120 (0.049)**
Equipment	−0.048 (0.047)	0.075 (0.038)**	0.101 (0.029)***	0.134 (0.030)***
Area owned	0.003 (0.004)	0.002 (0.004)	0.031 (0.002)***	0.034 (0.002)***
Distance to road	−0.009 (0.003)***	−0.009 (0.003)***	−0.002 (0.001)	−0.003 (0.001)*
Extension	0.995 (0.081)***		0.367 (0.059)***	0.324 (0.059)***
Seeds subsidies	0.069 (0.083)	0.036 (0.076)	−0.028 (0.041)	0.007 (0.042)
Fertilizers subsidies	0.487 (0.079)***	0.333 (0.074)***	0.321 (0.042)***	0.276 (0.043)***
Clay	0.044 (0.008)***	0.039 (0.007)***	0.009 (0.004)**	0.001 (0.004)
Rainfall	−4.312 (0.836)***	−4.693 (0.748)***	0.521 (0.453)	0.046 (0.459)
Rainfall squared	1.802 (0.443)***	2.432 (0.397)***	−0.740 (0.244)***	−0.567 (0.248)**
Groundnut AEZ	−0.083 (0.173)	−0.106 (0.146)	0.187 (0.082)**	0.003 (0.083)
Casamance AEZ	0.651 (0.231)***	0.054 (0.201)	0.280 (0.112)**	0.209 (0.113)*
South-East AEZ	0.358 (0.216)*	0.339 (0.185)*	0.540 (0.106)***	0.464 (0.107)***
Drought	−0.011 (0.122)	0.220 (0.105)**	0.020 (0.062)	0.048 (0.063)
Early rain stop	−0.048 (0.073)	0.343 (0.066)***	−0.119 (0.036)***	−0.134 (0.036)***
Crop diseases	−0.240 (0.133)*	0.100 (0.115)	−0.072 (0.065)	−0.045 (0.066)
Solar grids	0.280 (0.098)***		0.030 (0.056)	0.035 (0.056)
Extension Need		0.323 (0.076)***	0.043 (0.045)	0.039 (0.045)
AIC	1849.489	2177.603		
Log Likelihood	-900.745	-1064.801		
Adj. R^2			0.174	0.205
RMSE			1.026	1.039
N	3939	3939	3939	3939

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 2.6: Post-harvest loss ratios per crop and per region (%)

Regions	Maize	Rice	Sorghum	Millet	Fonio
Dakar	18.44	0.00	11.56	8.25	0.00
Diourbel	22.72	0.00	12.79	21.30	0.00
Fatick	22.72	11.39	12.79	8.50	0.00
Kaffrine	32.60	11.29	23.31	21.30	0.00
Kaolack	22.72	11.29	11.56	8.50	12.00
Kédougou	29.24	11.89	11.67	11.09	25.10
Kolda	29.24	23.79	13.01	23.89	24.95
Louga	18.44	11.29	11.56	8.25	0.00
Matam	18.44	11.39	11.56	8.25	0.00
Saint-Louis	18.44	11.34	11.56	8.57	0.00
Sédhiou	29.24	23.79	23.42	11.09	25.10
Tambacounda	18.44	11.34	23.31	8.25	12.00
Thies	28.32	11.29	23.31	21.30	0.00
Ziguinchor	19.36	24.39	11.67	11.09	0.00

Source: APHLIS, https://www.aphlis.net/

Table 2.7: Converting rations of kilograms to kilocalories

N°	Description	Edible conversion factor	Kcal/100g
	Cereals		
	Maize, yellow, whole kernel, dried, raw	1.00	353
	Rice, white, polished, raw	1.00	353
	Sorghum, whole grain, raw	1.00	344
	Millet, whole grain, raw	1.00	348
	Fonio, husked grains, raw (bran removed)	1.00	347
	Beans		
	Cowpea, dried, raw	1.00	316
	Bambara groundnut, dried, raw (Vigna subterranea)	1.00	376
	Roots and Tubers		
	Cassava, tuber, raw	0.84	153
	Sweet potato, pale yellow, raw	0.84	115
	Sweet potato, pale yellow, raw	0.84	115
	Potato, raw	0.84	80
	Taro, tuber, raw	0.86	92

Source: Stadlmayr *et al.* (2012)

Table 2.8: Adult-equivalent conversion factors for estimated calorie requirements according to age and gender

Age (years)	Calories (kcal)	Adult-equivalent conversion factor
Newborns		
0-1	750	0.29
Children		
01-03	1,300	0.51
04-06	1,800	0.71
7-10	2,000	0.78
Men		
11-14	2,500	0.98
15-18	3,000	1.18
19-24	2,900	1.14
25-50	2,900	1.14
51+	2,300	0.90
Women		
11-14	2,200	0.86
15-18	2,200	0.86
19-24	2,200	0.86
25-50	2,200	0.86
51+	1,900	0.75

Source: Claro *et al.* (2010)

Chapter 3

Quantifying the impact of membership in farmer organizations on land productivity and household income in Senegal

K. Christophe Adjin, Anatole Goundan, Christian H. C. A. Henning, Saer Sarr

Abstract

The recent renaissance of the Senegalese cooperative movement coupled with the revival of the agricultural sector motivated this study, which mainly aims to analyse the impact of farmer-based organization membership on household land productivity and net income. We combined the Propensity Score Matching (PSM) method with an Endogenous Switching Regression (ESR) model to derive treatment effects of membership in these farmer organizations using a national household-level survey data. Results exhibit consistency across estimations techniques. Estimates of both ESR and PSM models showed that membership in farmer organizations affects positively and significantly the household land productivity and net income. Moreover, findings show that membership has a heterogeneous impact. Households with the lowest probability to be member of farmer organizations have the highest impact. The effect of membership depends also on the specific type of organization.

Keywords: Farmer organizations, impact evaluation, land productivity, household income, Senegal

JEL Codes: Q13, D04, Q15, Q12

3.1 Introduction

Agriculture is the main economic sector in Sub-Saharan Africa. However, its performances are challenged by many factors mainly the access to production inputs and technologies (World Bank, 2007). For decades, policymakers regarded collective action groups, such as farmers-based organizations or agricultural cooperatives, as important tools to address these challenges and improve agricultural performance (Salifu *et al.*, 2010). According to Schwettmann (2014), Sub-Saharan Africa cooperatives experienced several stages of development from the colonial era to post-structural adjustment programs or contemporaneous era. The contemporary cooperatives are less structured and economically less powerful compared to their predecessors, however, they are more diverse, more efficient and better adapted to local circumstances (Schwettmann, 2014).

The development approach based on farmers' collective action groups still prevails in many developing countries. For example, the Agricultural Services and Producer Organizations Projects implemented by the World Bank in Chad, Mali, and Senegal during the period 2000-2011, were mainly based on the development of farmer organizations, with the expectation that these farmer groups could influence and improve agricultural development and performances in these countries. Fortunately, nowadays, such an approach is increasingly supported by quantitative studies, in which scholars try to estimate the effective contribution of agricultural cooperatives membership to various agricultural indicators (technology adoption, commercialization, and marketing processes, farm performances, farmer welfare, etc.).

The literature on these studies in developing countries reveals that several factors are associated with the membership in farmer-based organizations, such as gender and age of farmers, assets possessed or wealth level (Bernard and Spielman, 2009; Fischer and Qaim, 2012; Abebaw and Haile, 2013; Mojo *et al.*, 2017), access to various rural institutions such as extension, credits and even cooperatives (Abebaw and Haile, 2013; Mojo *et al.*, 2017), off-farm activities, leadership, farming experience, geographic location (Abebaw and Haile, 2013), family size and social networks that farmers belong to, and education level (Mojo *et al.*, 2017). Meanwhile, membership in a cooperative or farmer organizations mostly affects positively and significantly the prices received by farmers (Wollni and Zeller, 2007; Bernard *et al.*, 2008; Bernard and Spielman, 2009), commercialization rates (Barham and Chitemi, 2009; Bernard and Spielman, 2009; Francesconi and Heerink, 2010; Fischer and Qaim, 2012; Chagwiza *et al.*, 2016), technologies adoption (Abebaw and Haile, 2013; Ma *et al.*, 2018), households welfare (Fischer and Qaim, 2012; Ito *et al.*, 2012; Verhofstadt and Maertens, 2015; Ma and Abdulai, 2016; Mojo *et al.*, 2017; Ahmed and Mesfin, 2017; Mishra *et al.*, 2018).

Despite all these interesting findings, according to the recent study of Mojo *et al.* (2017), impact evaluation of the contribution of farmer-based organizations are still limited. Furthermore, some scholars have found no effect of membership in farmer-based organizations in their empirical work. Hoken and Su (2015) for instance did not observe any significant difference in received income between members and non-members of rice-producing cooperatives in suburban China. In addition, farmer-based organizations performance and impacts may vary across countries and regions even within the same agricultural sub-sector or across commodities (Bernard and Taffesse, 2012; Mojo *et al.*, 2017). Moreover, as pointed out by Verhofstadt and Maertens (2015) studies on cooperative organizations usually focus on a single co-operative or on multiple cooperatives in a single sub-sector. This study aims to contribute to this growing literature by taking advantage of an original country-wide survey data set collected in Senegal to quantify the effect of membership in farmer-based organizations[1] on farmers' land productivity and household incomes. The sample data used for the analysis comprises 4245 farmers located in all six Senegalese agro-ecological zones. Looking at the effect of the farmers-based organizations at a country level gives a broader perspective of analysis, which is necessary for policy design.

Moreover, the Senegalese case study is of particular interest for several reasons. First, as argued by Reed and Hickey (2016), during the last decade, there has been a renaissance of cooperative movement due to several institutional changes. Second, since 2012, a sort of revival of the entire Senegalese agriculture is also observed, noted by the substantial increase of the sector's contribution to the national GDP from 12% in 2011 to 16% in 2017 (World Bank, 2017). Finally, in regards to quantitative analysis of the contribution of farmers organizations to agriculture, very little studies have been carried out in the case of Senegal. The remainder of this paper is organized as follows. The following sections describe the econometric framework and the data used. The last sections present, discuss and summarize the results of the estimations.

[1]We will use alternatively the expressions farmer organizations, farmer-based organizations (FBO), producer organizations, or agricultural collective action groups alternatively to define farmer-based organizations. Farmer organizations in our study include therefore all forms of organizations that provide farmers with farm or farm-related services as we will conceptualize later in the paper.

3.2 Econometric framework

3.2.1 Estimation strategy

Generally, a farmer decides to become a member of a farmer organization for the services provided by such a collective action group regarding access to credit, farm inputs, technologies, information, or marketing facilities. Therefore, an assumed rational farmer would choose to be a member of a farmer organization if the expected utility from this organization membership (M_1) is greater than that of from non-membership (M_0) . This utility gain from membership in a farmer-based organization ($M^* = M_1 - M_0$) can be expressed as a function of an observable vector of covariates (Z) in a latent model as follows:

$$M_i^* = \alpha Z_i + \eta_i, \quad M_i = 1 \text{ if } \; M_i^* > 0, \tag{3.1}$$

where M_i is a binary variable that equals 1 if household i is a member of a farmer organization and zero otherwise; α is a vector of parameters to be estimated and Z_i is a vector of household demographics, socio-economic, and farm-level characteristics; and η_i is a random error term assumed to be normally distributed. Membership in a farmer organization is expected to affect various outcome variables at the farm or household level including land productivity and household income. Assuming that the outcome variable (land productivity or household income) is a linear function of a vector of exogenous variables X_i and endogenous membership in farmers organization M_i such that:

$$Y_i = \beta X_i + \delta M_i + \epsilon_i, \tag{3.2}$$

where Y_i represents the outcome variables (land productivity and agricultural income); M_i is defined as previously; β and δ are parameters to be estimated, and ϵ_i is the error term. However, farmers may self-select into FBOs, rather than being randomly selected. Therefore, estimating equation 3.2 using ordinary least square (OLS) might produce biased estimates. We explored then the propensity score matching and endogenous switching regression models to produce unbiased and consistent estimates. The PSM controls for selection bias through controlling for observable confounding factors. However, an important shortcoming of the PSM method is its inability to deal with biases resulting from unobservable characteristics of sampled units. The endogenous switching regression addresses the endogeneity of membership in farmers' organizations by accounting for both observed and unobserved sources of bias (Lokshin and Sajaia, 2004). Both are used to analyse the consistency of the obtained results across the estimation techniques

3.2.2 Propensity Score Matching (PSM)

The propensity score matching method (PSM) is a quasi-experimental technique often used in observational causal studies. PSM uses observable characteristics of observation units in the sample to generate a control group that is comparable to the treated group conditional on identified exogenous factors, but different regarding the intervention status, here membership in farmers organization (Rosenbaum and Rubin, 1983). PSM works under two main assumptions. The first is the conditional independence or unconfoundedness, stating that observable characteristics must be independent of potential outcomes, which implies that the membership decision is only based on observable characteristics of households. The second is the common support condition that needs to be satisfied, i.e. the distributions of observable characteristics between members of farmer organizations and non-members have to overlap (Jelliffe *et al.*, 2018). Empirically, in a first step, we regressed the membership of farmers organizations on a vector of observable variables Z (as in equation 3.1) to generate the propensity scores using a probit estimation (Hirano *et al.*, 2003). The estimated propensity scores ($PS_i = \text{Prob}\,(M_i = 1 \mid Z_i)$) represent the probability of a farmer to belong to a farmer-based organization, and the marginal effects express the impact of variables in Z on this probability. We included in Z a large set of conditioning factors in order to minimize omitted variables bias. Secondly, the generated propensity scores (PS) are used to match farmers who are members of FBOs to non-members. Numerous algorithms can be applied to match members and non-members of similar propensity scores. Furthermore, PSM methods are sensitive to a particular specification and matching method (Imbens, 2004; Caliendo and Kopeinig, 2008). Therefore, we use three different common matching techniques: the nearest neighbor matching, the kernel matching, and the radius matching. The nearest neighbor matching (NNM) algorithm was implemented with a caliper of 0.01. In a third step, we examined the extent of overall covariates balancing property and the overlap over the common support. The fourth step consisted of calculating the Average treatment on treated ATT, which is the mean difference between the two matched groups (Dehejia and Wahba, 2002; Imbens, 2004). Specifically, the estimated ATT is:

$$ATT(Z) = E\left[Y_1 \mid M = 1, \text{Prob}\,(Z)\right] - E\left[Y_0 \mid M = 1, \text{Prob}\,(Z)\right], \qquad (3.3)$$

where, Y_1 represents the outcome indicator of the members of farmers organizations, Y_0 is the outcome indicator of non-members; M is defined as previously. Finally, we checked the robustness of our estimates by using the Rosenbaum (2002) bounding approach. The main assumption behind matching is selection on observables. However, if there are unobserved variables that affect both membership and the outcome variable, a hidden bias might arise and affect the estimates of matching estimators (Rosenbaum, 2002). In particular, the hidden bias could lead to both positive and negative unobserved selection. Rosenbaum's method is based on the sensitivity parameter Γ that measures the degree of departure from random assignment of treatment. Two households with the same observed characteristics may differ in the odds of belonging to farmers organizations by at most a factor of Γ. Considering the upper bounds, the factors Γ are incrementally computed until the threshold of 10% of p-values is reached. The relatively higher is the Γ factor; the more robust is our model regarding hidden bias due to unobserved confounders. This sensitivity analysis is based on the Wilcoxon sign rank test. PSM analyses were conducted using STATA 14. Although we conducted these robustness checks, PSM only controls for selection biases from observed characteristics. We then applied an Endogenous switching regression analysis that has the potential to mitigate biases from both observable and unobservable factors.

3.2.3 Endogenous Switching Regression (ESR)

Under the Endogenous Switching Regression (ESR) framework, the impact of membership in farmer organizations on land productivity (and household income) is estimated in two stages: the first stage concerns the decision to join agricultural collective action groups (equation 3.1), and the second stage consists in the estimation of two regimes outcomes equations: one for members and another one for non-members (equations 3.4 and 3.5) represented as follows:

$$\text{Regime 1}: \quad Y_{1i} = \beta_1 X_i + \epsilon_{1i} \quad \text{if} \quad M_i = 1 \quad \text{(Members)} \tag{3.4}$$

$$\text{Regime 2}: \quad Y_{2i} = \beta_2 X_i + \epsilon_{2i} \quad \text{if} \quad M_i = 0 \quad \text{(Non}-\text{Members)}, \tag{3.5}$$

where Y_1 and Y_2 represent the outcome respectively for farmer organization members (regime 1) and non-members (regime 2); X_i represents the vector of covariates of farmer i; β_1 and β_2 are parameters to be estimated; and ϵ_{1i} and ϵ_{2i} are errors terms associated with the outcomes variables. In the ESR framework, the error terms in the three equations (3.1, 3.5 and 3.4) are assumed to have a trivariate normal distribution, with zero mean and covariance matrix of the following form:

$$\operatorname{cov}(\eta, \epsilon_1, \epsilon_2) = \begin{bmatrix} \sigma_\eta^2 & \sigma_{\eta 1} & \sigma_{\eta 2} \\ \sigma_{1\eta} & \sigma_1^2 & . \\ \sigma_{2\eta} & . & \sigma_2^2 \end{bmatrix}, \tag{3.6}$$

where σ_η^2 is the variance of the error term in the selection equation (3.1); σ_1^2 and σ_2^2 are the variances of the error terms in the outcome equations (3.5 and 3.4); $\sigma_{1\eta}$ and $\sigma_{2\eta}$ are the covariances of η, ϵ_{1i} and ϵ_{2i}. Covariance between ϵ_{1i} and ϵ_{2i} is not defined since Y_1 and Y_2 are not observed simultaneously (Maddala *et al.*, 1986). The expected values of ϵ_{1i} and ϵ_{2i} conditional on the sample selection are non-zero, because the error term of equation 3.1 is correlated with the error terms of the outcome equations 3.5 and 3.4:

$$E[\epsilon_{1i} \mid M = 1] = \sigma_{1\eta} \frac{(\phi(Z_i\alpha))}{(\Phi(Z_i\alpha))} = \sigma_{1\eta}\lambda_{1i} \tag{3.7}$$

$$E[\epsilon_{2i} \mid M = 0] = \sigma_{2\eta} \frac{(\phi(Z_i\alpha))}{(1 - \Phi(Z_i\alpha))} = \sigma_{2\eta}\lambda_{2i} \tag{3.8}$$

where $\phi(.)$ is the standard normal probability density function; $\Phi(.)$ is the standard normal cumulative density function; and λ_{1i} and λ_{2i} are the inverse Mills Ratios (IMR) computed from equation 3.1 with $\lambda_{1i} = \frac{(\phi(Z_i\alpha))}{(\Phi(Z_i\alpha))}$ and $\lambda_{2i} = \frac{(\phi(Z_i\alpha))}{(1-\Phi(Z_i\alpha))}$, and included in equations 3.4 and 3.5 to correct for selection biases resulting from unobservable factors. Therefore, we have:

$$Y_{1i} = \beta_1 X_i + \sigma_{1\eta}\lambda_{1i} + \delta_{1i} \quad \text{if} \quad M_i = 1 \quad (\text{Members}) \tag{3.9}$$

$$Y_{2i} = \beta_2 X_i + \sigma_{2\eta}\lambda_{2i} + \delta_{2i} \quad \text{if} \quad M_i = 0 \quad (\text{Non} - \text{Members}), \tag{3.10}$$

where δ_{1i} and δ_{2i} are error terms with conditional zero means. The full information maximum likelihood (FIML) method was applied to have consistent estimates (Greene, 2000; Lokshin and Sajaia, 2004). Furthermore, appropriate identification of ESR requires at least one variable in Z that does not appear in X. This variable represents the exclusion restriction necessary to fully estimate the model. The estimation of the selection equation (3.1) thus includes two potential instruments. A valid instrument is required to influence the farmer's choice of membership but does not have any direct effect on the outcomes of interest. The first potential instrument that we use is whether farmers receive information on sales. Thus, from the question "do you receive information on sales", we created a dummy variable "Information on sales" which takes a value of 1, if the farmer receives information on sales and the value 0, otherwise. This instrument is supposed to correlate significantly with the membership in FBOs. Those farmers who receive information on sales have a higher probability to belong to farmer organizations. Farmers could join these organizations with the motivation to be more informed on sales and the associated better prices. However, receiving this information is not supposed to directly affect the outcome variables of interest since only receiving information does not directly improve or decreases the land productivity nor the total household incomes (but indirectly affects both outcomes through membership in the organization). The second potential instrument is the main type of water source used by the household. Similarly to the first instrument, from the question: "what is your main source of drinking water ?", we created a dummy variable "water source" that takes the value of 1, if the household uses tap water and the value of 0, otherwise. The use of tap water is an asset variable that expresses the capacity of the household to be a member of farmer organizations, the capacity to afford membership fees.

To check for the validity of these instruments, we ran a probit model for the equation 3.1 and OLS regressions for outcome equations (3.4 and 3.5) separately and checked in which equation these variables are effectively significant. The results are presented in appendix table 3.11. The positive coefficients of variable "Information on sales" and "Source of water" confirms the expectation that households who have access to information on sales and use tap water are more likely to be members of farmer organizations. The designed instruments significantly influence the membership in FBOs but not the non-members farmers' land productivity (F = 0.084 (2), p-value = 0.920) and household net income (F = 0.838 (2), p-value = 0.433).

From the assumptions on the distribution of the error terms (3.6), the derived log-likelihood function is specified as:

$$\ln L = \sum_{i=1}^{N} \left\{ A_i \left[\ln\phi \left(\frac{\epsilon_{1i}}{\sigma_1} \right) - \ln\sigma_1 + \ln\Phi\left(\theta_{1i}\right) \right] + \right. \tag{3.11}$$

$$\left. (1 - A_i) \left[\ln\phi \left(\frac{\epsilon_{2i}}{\sigma_2} \right) - \ln\sigma_2 + \ln\left(1 - \Phi\left(\theta_{2i}\right)\right) \right] \right\}, \tag{3.12}$$

where $\theta_{ji} = \frac{(Z_i\alpha + (\rho_j \epsilon_{ji})\sigma_j)}{\sqrt{1-\rho_j^2}}$, with $j = 1, 2$ and ρ_j ($\rho_1 = \frac{\sigma_{1\nu}^2}{\sigma_\nu}\sigma_1$ and $\rho_2 = \frac{\sigma_{2\eta}^2}{\sigma_\eta}\sigma_2$) being the correlation coefficients between the error term η_i of the selection equation (3.1) and respectively the error terms ϵ_{1i} and ϵ_{2i} of the outcome equations (3.4 and 3.5). If one of the estimates of correlation coefficients ρ_1 or ρ_2 is statistically significant, this would indicate the existence of a selectivity bias due unobserved factors (Abdulai and Huffman, 2014). Then, the endogenous switching regression model would be appropriate. When $\rho_1 > 0$, this implies a negative selection bias, indicating that farmers who have below than average outcomes are more likely to choose to be members of farmer organizations, whereas with $\sigma_{1\nu} < 0$, this would suggest a positive selection bias. Moreover, if ρ_1 or ρ_2 have alternate signs, then farmers choose to be members of producer organizations based on their comparative advantage: members have above-average outcomes from membership status and the non-members have above-average outcomes from being non-members. If these correlation coefficients have the same sign, it would mean a hierarchical sorting: members have above-average outcomes whether they are members or not, but they are better off being members, while non-members have below-average outcomes in either case, but they are better off not being members. The coefficients from the ESR model allow one to derive the average treatment effect on the treated (ATT). Specifically, the observed and unobserved counterfactual outcomes for farmer organization members can be computed as:

$$E\left[Y_{1i} \mid M=1\right] = \beta_1 X_i + \sigma_{1\eta}\lambda_{1i} \tag{3.13}$$

$$E\left[Y_{2i} \mid M=0\right] = \beta_2 X_i + \sigma_{2\eta}\lambda_{2i} \tag{3.14}$$

$$E\left[Y_{2i} \mid M=1\right] = \beta_2 X_i + \sigma_{2\eta}\lambda_{1i} \tag{3.15}$$

$$E\left[Y_{1i} \mid M=0\right] = \beta_1 X_i + \sigma_{1\eta}\lambda_{2i} \quad . \tag{3.16}$$

Equation 3.13 computes the observed outcome (a) for organization members and equation 3.14 calculates the observed outcome (b) for non-members. The expected outcome (c) in equation 3.15 represents the counterfactual for the observed outcome (a) in equation 3.13. This counterfactual expresses what would have happened had the farmers decided to be member of the organizations. Similarly the equation 3.16 is a counterfactual outcome (d) for the observed outcome (b) in equation 3.14. It represents the scenario in which farmers decided to be members of producers organizations. Using these expected outcomes (equations 3.13 to 3.16) we derive unbiased treatment effects: the average treatment effect on treat (ATT, which is the difference between equation 3.13 and 3.15 that is a − c), and the average treatment effect on untreated (ATU, which is the difference between equation 3.16 and 3.14 that is d − b).

$$ATT = E\left[Y_{1i} \mid M=1\right] - E\left[Y_{2i} \mid M=1\right] = \left(\beta_1 - \beta_2\right) X_i + \lambda_{1i}\left(\sigma_{1\nu} - \sigma_{2\nu}\right) \tag{3.17}$$

$$ATU = E\left[Y_{1i} \mid M=0\right] - E\left[Y_{2i} \mid M=0\right] = \left(\beta_1 - \beta_2\right) X_i + \lambda_{2i}\left(\sigma_{1\nu} - \sigma_{2\nu}\right). \tag{3.18}$$

3.2.4 Addressing other empirical issues

For the empirical specification of the first stage of the ESR model (estimation of the selection equation), several factors are associated with membership in producer organizations. These factors which include personal details of household head (gender, age, education), household characteristics (e.g. household size, agricultural assets, land size), access to agricultural extension services, and the geographic location of the household. It is worth noting however, that households could have better access to extension due to their membership in collective action groups, rendering the access to extension services variable potentially endogenous in the modeling of the choice to belong to farmer organizations and leading then to biased estimates. We, therefore, corrected this endogeneity issue with the two-stage control function approach suggested by Wooldridge (2015). In a first stage, we estimated separately, the access to extension services and the membership in organizations on the same independent variables plus an instrument, here the farmer's expressed needs for extension services, using a probit model. The instrument, "need for extension"[2], significantly influences the access to extension services (χ^2 (1) = 3.613, p-value = 0.057) but not directly the household decision to belong to organizations (χ^2 (1) = 0.647, p-value = 0.421, see table 3.9 in the appendix). In the second-stage probit estimation, the access to extension services variable and their generalized residuals predicted from the first-stage are included in the selection equation. Moreover, this variable "extension needs" is not correlated to the other instruments used in the rest of the analysis, such as information on sales (Pearson's correlation = 0.011, t = 0.741 (4243), p-value = 0.459) or the use of tap water for drinking (Pearson's correlation = −0.005, t = −0.353 (4243), p-value = 0.724)

3.2.5 Heterogeneous treatment effects analysis

Following (Abebaw and Haile, 2013) and (Verhofstadt and Maertens, 2015), we analyse how the estimated outcome effects of organizations membership vary within members. Therefore, we used the estimates of ATT as a dependent variable and run ordinary least squares (OLS) to regress it on farm household characteristics. In

[2] From the two questions: "do you need extension services?" and "what do you need extension services for?", we created a dummy variable "extension needs" which takes the value 1, if the household responds "yes" to the first question and states technology diffusion services in the second question and the value 0, otherwise. Farmers who expressed a need for technologies in their activities are expected to have access to extension services, or at least exploring ways to have access to it.

addition, we plotted OLS regressions of estimated ATT on the propensity score, and on some farm characteristics (i.e. age, education, gender, size of the household, and distance to nearest road) to derive smoothed curves. Such graphical and statistical analyses help to find out which type of households the impact of membership in farmer organizations is the most important.

3.3 Data sources and descriptive statistics

3.3.1 Data sources

The data used for the analysis derived from a survey conducted in Senegal, which randomly sampled 4480 households that mainly produce dry cereals (or rainfed cereals). The survey was done under the Agricultural Policy Support Project (Projet d'Appui aux Politiques Agricoles, PAPA)[3], which is an initiative of the Government of Senegal funded by USAID-Senegal as part of the "Feed The Future" initiative, and implemented for a period of 3 years (2015 - 2018) by the Senegalese Ministry of Agriculture and Rural Facilities with technical support from the International Food Policy Research Institute (IFPRI). A multistage sampling procedure was applied for the selection of households and a structured household questionnaire was used to collect information. This questionnaire included several modules and gathered information on a range of topics such as household demographic and socioeconomic characteristics, farmer organization membership, household assets, crop production, livestock revenues, income and expenditures, access to infrastructure, access to institutions, commercialization, and production shocks and risk management strategies. Besides information on crop production and inputs used, data collection also included market prices and households' adoption of agricultural technologies during the main agricultural season of 2016/2017. After the data cleaning and the removal of observations with no information on the different outcome variables, a final sample of 4245 households was used for the analysis. This sample includes farmers located in all six Senegalese agro-ecological zones.

[3]Official website of the project is http://www.papa.gouv.sn/.

3.3.2 Descriptive statistics

Table 3.1 presents the definition and summary statistics of the variables used in the analysis. It also reports the comparative descriptive statistics of these variables based on farmer organization membership status. Following the definition of Bernard *et al.* (2015), our variable of interest "organization membership" is referred to as membership in a rural producer organization that provides farmers with farming and farm-related services including access to inputs, markets and credits, collective sales, and capacities reinforcement. Eight types of farmers organizations were mentioned by the surveyed households: Producer Groups, Economic Interest Groups, Rural Associations, Cooperatives, Women Producers Groups, Federations, Unions, and Networks. Therefore, the variable "organization membership" is binary, coded as 1 if a member of the household belongs to any of this farmer-based organization, and 0 otherwise. In some households, several family members expressed their belonging to these organizations, with a maximum of 7 members. However, on average only one family member belongs to a group. About 9% of the households in the sample have at least one person belonging to a group. The main organizations, which gather most of the household family members, are Economic Interest Groups (44.1%), Rural Associations (16.7%), Producer Groups (16.1%), and Cooperatives (15.3%).

Regarding the outcome variables, land productivity is measured as the net value in FCFA[4] of all crop outputs valued at the market prices per unit of land area. This approach is more suitable since most cereals productions are not marketed by farmers. The net value of all crop production represents the value of all crop production after the deduction of all crop production costs, such as seeds cost, fertilizer costs, all other costs, and hired labour. Farmers in the sample have on average a land productivity of 130,050 FCFA per hectare. The household income was generated by adding to the net value of all crop production, the livestock income received by the farmer during the last 12 months, and all off-farm incomes [5]. On average, the sampled households receive 592,100 FCFA as net total income. The two outcomes variables are log specified.

[4] 1 FCFA=0.0017 USD as at December 2019.

[5] Crafts, hunting, forestry, fishing, small business, farm products processing, transport

Table 3.1: Description of variables

Variables	**Description and measurement**	**Pooled** (1)	**Members** (3)	**Non-Members** (2)	**P-values** (4)
Organization Membership	Membership in farmer organizations (1=yes, 0=no)	0.088 (0.28)			
	Outcome variables				
Land Productivity	All crop production per hectare (1000 FCFA/ha)	130.05 (301.62)	255.75 (631.57)	117.97 (244.60)	<0.01
Household income	Total net household income (1000 FCFA)	592.10 (878.51)	844.09 (1299.84)	567.90 (823.01)	<0.01
	Household and Head characteristics				
Gender	Household head is a male (1=yes, 0=no)	0.93 (0.25)	0.95 (0.22)	0.93 (0.25)	0.14
Age	Age of household head (years)	53.07 (13.44)	51.09 (12.13)	53.27 (13.55)	<0.01
Education	Formal education (1=yes, 0=no)	0.37 (0.48)	0.51 (0.50)	0.36 (0.48)	<0.01
Active members	Active family members	5.72 (3.15)	6.37 (3.43)	5.66 (3.12)	<0.01
Dependents	Non-active family members	4.28 (3.31)	5.01 (3.82)	4.21 (3.25)	<0.01
Migration	Household head is a migrant (1=yes, 0=no)	0.14 (0.35)	0.15 (0.36)	0.14 (0.35)	0.51
	Household Assets				
Equipment	Agricultural Equipment (1.000.000 FCFA)	0.13 (0.56)	0.17 (0.46)	0.13 (0.57)	0.08
Area Owned	Land size owned by household (ha)	5.82 (8.37)	5.62 (6.24)	5.84 (8.54)	0.52
	Access to infrastructures				
Distance to road	Distance to nearest all-weather road (km)	10.15 (14.15)	10.32 (13.90)	10.14 (14.18)	0.81
Extension	Access to extension services (1=yes, 0=no)	0.11 (0.31)	0.42 (0.49)	0.08 (0.27)	<0.01
	Agro-ecological zones				
Groundnut AEZ	Groundnut agro-ecological zone (1=yes, 0=no)	0.50 (0.50)	0.28 (0.45)	0.52 (0.50)	<0.01
Casamance AEZ	Casamance agro-ecological zone (1=yes, 0=no)	0.25 (0.43)	0.31 (0.46)	0.25 (0.43)	0.01
South-East AEZ	South East agro-ecological zone (1=yes, 0=no)	0.11 (0.31)	0.15 (0.36)	0.11 (0.31)	0.02
Other AEZ	Other agro-ecological zones (1=yes, 0=no)	0.14 (0.35)	0.25 (0.44)	0.13 (0.34)	<0.01
	Instrumental Variables				
Information on Sales	Information on Sales (1=yes, 0=no)	0.01 (0.11)	0.05 (0.22)	0.01 (0.10)	<0.01
Tap water	Use of tap water for drinking (1=yes, 0=no)	0.35 (0.48)	0.36 (0.48)	0.35 (0.48)	0.65
Extension needs	Express need for technologies (1=yes, 0=no)	0.01 (0.09)	0.01 (0.07)	0.01 (0.09)	0.37
N	Number of Observations	4245	372	3873	4245

Note: Standard deviations are in parenthesis

Following the literature on land productivity and agricultural household incomes, we have included in the models, several control variables, such as household and its heads characteristics (gender, age, education, active household size, dependents[6], and migration status[7]), household assets (the total value of possessed agricultural equipment and the land area owned), household access to rural institutions (extension services, distance to the nearest road), and agro-ecological zones dummies. About 93% of the households in the sample are predominantly male-headed. The sampled households heads are generally old with an average age of 53 years and without any formal education. Besides farming activities, households also get revenues from off-farm activities (33.8%). On average, the household includes ten family members and owns about 130,000 FCFA of agricultural implements and about 5.82 ha of farming land with 4.47 ha dedicated to crop cultivation. More than 85% of farm households in the sample are located in the Groundnut basin, Casamance and South East agro-ecological zones.

When comparing members of farmer organizations to non-members, significant differences in means can be observed for outcome indicators as for most of the control variables. Members of farmer organizations tend to have larger households (11 persons) than non-members (9 to 10 persons), and they appear averagely to be more educated. They also have better productivity per hectare and receive higher incomes than non-members. These significant differences in means between members and non-members suggest that farmer-based organizations might play an important role in enhancing farmers' adoption of technologies and permitting them to have a higher level of productivity and incomes. However, these results do not permit making inferences about the effect that membership in farmers organizations might have on farmers' incomes. These comparisons of means do not account for confounding factors such as observed household and farm-level characteristics and unobserved factors (e.g. perception and motivations of membership choice).

3.4 Results and discussion

This section reports first the identified factors that drive membership in farmers' organizations using the probit regression model. Then, it is followed by the results

[6]Active members are aged between 15 and 65 years and dependents regroup members aged below 15 years and more than 65 years.

[7]It is a dummy variable for the migration status of the household head. This variable also serves as a proxy for involvement in off-farming activities.

of the impact of organization membership on land productivity and income using the PSM and ESR models. Finally, the heterogeneous effects are analyzed and discussed.

3.4.1 Membership in farmers organizations

Factors that influence households' decision to belong to farmer organizations are presented in table 3.2 with their marginal effects. The likelihood ratio test shows that the model estimates are significant at 1% level ($\chi^2 = 445.49$ (17), $p < 0.01$). Results of estimation of equation 3.1 indicate that membership in farmer organizations is significantly influenced by the education level of the household head, household's size (number of active persons living in the household and the dependents), distance of the household to the nearest road, access to information on sales, the existence of tap water in the household and the location of the household in different agro-ecological zones (Groundnut basin, Casamance and South-East).

Formal education significantly and positively affects the probability for a household to be a member of an agricultural collective action group. Households with an educated head are about 4% more likely to join agricultural collective action groups. The household family size has also a positive and significant effect on membership in farmer organizations. These results support those of Bernard and Spielman (2009) and Ma and Abdulai (2016). For instance, households that have more active persons in the household have a higher probability (0.4%) to be members of producer organizations. With more active people, households have a better chance that one of their family members could belong to a farmer-based organization.

Geographic location and agro-climatic conditions of the households also have significant effects on the decision of farmers to be members or not. Results reveal that farmers who live closer to all-weather roads are respectively better prone to participate in groups actions with a 0.1% probability for each additional kilometre. These results suggest a clustering of farmers' organization members, due to spatial non-observable factors such as climate, institutions, and infrastructure. These findings corroborate those of Abebaw and Haile (2013) and Ma and Abdulai (2016). According to Ma and Abdulai (2016), in China variables representing soil types and regions have significant cluster effects.

Gender of the household head and the different assets owned by the household such as the value of agricultural implements and the land area do not appear to have any significant effect on membership, contradicting with some of the previous studies by Abebaw and Haile (2013) and Mojo *et al.* (2017). In addition, access to extension services affect positively but not significantly the farmers' probability to be members of collective action organizations. However, the effect appears significant in the ESR regressions. Access to various institutions e.g. agricultural extension services (Abebaw and Haile, 2013) and credit (Abdul-Rahaman and Abdulai, 2018) and even the access to farmer organizations Mojo *et al.* (2017) are, in previous literature associated with membership.

Table 3.2: Probit Estimation of Membership in Farmers Organizations

	Coefficients	Marginal Effets
Intercept	-1.868 (0.450)***	
Gender	0.060 (0.133)	0.007 (0.014)
Age	0.024 (0.017)	0.003 (0.002)
Age Squared	-0.000 (0.000)**	-0.000 (0.000)**
Education	0.314 (0.078)***	0.040 (0.010)***
Active persons	0.033 (0.010)***	0.004 (0.001)***
Dependents	0.027 (0.010)***	0.003 (0.001)***
Migration	0.042 (0.101)	0.005 (0.013)
Equipment	-0.009 (0.050)	-0.001 (0.006)
Area owned	0.002 (0.004)	0.000 (0.000)
Distance to road	-0.008 (0.003)***	-0.001 (0.000)***
Extension	0.286 (0.832)	0.040 (0.136)
Groundnut AEZ	-0.892 (0.154)***	-0.109 (0.020)***
Casamance AEZ	-0.318 (0.123)***	-0.033 (0.011)***
South-East AEZ	-0.280 (0.118)**	-0.028 (0.010)***
Information on Sales	0.801 (0.292)***	0.162 (0.085)*
Tap water	0.196 (0.076)***	0.024 (0.010)**
Extension residuals	0.437 (0.437)	
Log Likelihood	-1038.128	-1038.128
LR Test	445.49***	
Num. obs.	4245	4245

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

3.4.2 Impact of FBO membership: PSM results

This section presents the treatment effects estimated from the PSM models. Based on the probit estimation of equation 3.1, propensity scores were obtained for the matching. The validation of PSM models depends on the quality of the matching. Table 3.10 in appendix provides the overall covariate balancing test. Results show that the standardized mean difference for all covariates used for the matching reduce from 23.9% before matching to 2.7% after matching. Moreover, the likelihood ratio test indicates that the null hypothesis of the joint significance of all covariates could be rejected before matching p $> \chi^2 = 0.000$. Conversely, after the matching, with the same test the joint significance of all covariates could not be rejected p $> \chi^2 = 0.997$. These results indicate that the required balancing property of the distribution of propensity scores is satisfied. Furthermore, Figure 3.2 in the appendix shows the common support between the two groups. Most of farmers organizations members and non-members had a common support region, only seven members were outside the common support region and therefore dropped from the matched sample.

Table 3.3 reports the average treatment effect on the treated from the PSM models. The robust standard errors of these estimates were calculated by bootstrapping using 50 replications. As stated previously three matching methods were used: the nearest neighbor matching, the kernel matching, and the radius matching. The average treatment effects on the treated for land productivity and household income are all positive and statically significant. For instance, with the Nearest Neighbour matching method, the effects of membership are evaluated at 28% for land productivity and 14.4% for household income. The estimated values of the effect of membership of producer organizations are quite close across the alternative matching specifications. From these results, one can conclude that that in the absence of observable selection bias, membership in a collective action group affects positively and significantly farmers' land productivity and household income. Our findings are similar to those of other studies that empirically reported a significant and positive relationship between membership in farmer-based organizations and farm productivity and household welfare, in China (Ma and Abdulai, 2016) and in Rwanda (Verhofstadt and Maertens, 2014).

Table 3.3: ATT of FBO membership: PSM Estimates

Outcomes	Matching Methods		
	Nearest Neighbor	Kernel	Radius
Land Productivity	0.280 (0.072)***	0.323 (0.067)***	0.331 (0.050)***
Household Income	0.144 (0.081)*	0.182 (0.073)**	0.183 (0.070)***

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Robust standard errors are in parentheses.

To check for the robustness of our PSM model results, as mentioned previously, we calculated the Rosenbaum bounds (Becker and Caliendo, 2007) and reported in table 3.4 the upper bounds results with their p-values. The Rosenbaum bounds were computed for treatment effects that are significantly different from zero. Considering the significance level of 10%, the lowest value of Γ in all PSM specifications was $1.10 - 1.15$ obtained with the nearest neighbour technique and the largest value was $1.70 - 1.75$ observed for a kernel matching. For instance, when considering the impact of membership on land productivity (for PSM Nearest Neighbour), the sensitivity analysis implies that at a level of $\Gamma = 1.50$, the causal inference may be viewed critically. This would mean that if farmers with similar covariates differ in their odds of being members of farmer-based organizations by a factor of 50%, the significance of membership effect on land productivity might be questionable. This value is relatively low. Considering the threshold of 80% for Γ, which is generally used in social sciences. These results suggest that the positive and significant impact estimates of organization membership on land productivity and household incomes are at some levels sensitive to unobservables or hidden-bias. Therefore, we considered the endogenous switching regression approach that accounts for both observed and unobserved factors.

Table 3.4: Rosenbaum Γ bounds sensitivity analysis for hidden bias

Outcomes	Matching Methods		
	Nearest Neighbor	Kernel	Radius
Land Productivity	$1.45 - 1.50$ $(0.066 - 0.109)$	$1.70 - 1.75$ $(0.089 - 0.131)$	$1.65 - 1.70$ $(0.084 - 0.127)$
Household Income	$1.10 - 1.15$ $(0.055 - 0.109)$	$1.20 - 1.25$ $(0.059 - 0.110)$	$1.20 - 1.25$ $(0.090 - 0.157)$

Notes: P-values are in parenthesis

3.4.3 Impact of FBO membership: ESR results

Results from the endogenous switching regression models are presented in tables 3.5 and 3.6. The ESR models were estimated using the FIML approach which derives both the selection and outcome equations jointly. The first stages of the estimation of ESR regressions are presented in columns (1) while the second stages of the estimation, i.e. estimation of separate outcome equations for organizations members and non-members, are reported in columns (2) and (3).

Except for the variables access to extension services and information on sales, the estimation results of the selection equation are similar, in terms of signs and significance, to the estimation of the probit estimation of equation 3.1 discussed previously. The exclusion restriction variable, access to information on sales, is statistically significant only for the household income model. Meanwhile, the second stage of the FIML shows that the estimated coefficients of the correlation ρ between farmer organizations membership and both land productivity and household income are all negative, but statistically significant only for members, implying that the hypothesis of absence of sample selectivity bias, in both models may be rejected. These findings suggest that both observed and unobserved factors influence the decision to belong to farmer organizations and both land productivity and household income given the membership. Moreover, ρ_1 (members correlation coefficients) in both outcome models have a negative sign, indicating a positive selection bias and implying that households with above average land productivity and household income are more likely to belong to farmer-based organizations. Furthermore, ρ_1 and ρ_2 have the same sign, suggesting that members have above-average land productivity and household income whether they are members or not, but they are better off being members, while non-members have below-average outcomes in either case, but they are better off not being members.

Table 3.5: ESR Regression of Land Productivity

	Selection Equation	Members	Non-Members
	(1)	(2)	(3)
Intercept	-1.808 (0.441)***	15.829 (0.894)***	11.035 (0.213)***
Gender	0.037 (0.131)	-0.027 (0.242)	0.090 (0.060)
Age	0.017 (0.016)	-0.049 (0.028)*	-0.002 (0.007)
Age Squared	-0.000 (0.000)*	0.001 (0.000)**	0.000 (0.000)
Education	0.245 (0.073)***	-0.364 (0.120)***	0.120 (0.033)***
Active persons	0.034 (0.010)***	-0.004 (0.020)	0.016 (0.006)***
Dependents	0.029 (0.010)***	-0.036 (0.018)**	-0.003 (0.005)
Migration	-0.023 (0.096)	-0.130 (0.153)	-0.063 (0.044)
Equipment	-0.042 (0.052)	0.173 (0.131)	0.079 (0.027)***
Area owned	0.001 (0.004)	-0.019 (0.010)*	-0.007 (0.002)***
Distance to road	-0.007 (0.003)***	-0.004 (0.005)	-0.004 (0.001)***
Extension	1.399 (0.702)**	-0.712 (0.207)***	0.250 (0.066)***
Groundnut AEZ	-0.762 (0.140)***	-0.597 (0.203)***	-0.232 (0.056)***
Casamance AEZ	-0.255 (0.116)**	-0.764 (0.177)***	0.361 (0.056)***
South-East AEZ	-0.283 (0.118)**	-1.008 (0.208)***	0.358 (0.068)***
Information on Sales	0.287 (0.254)		
Tap water	0.180 (0.066)***		
Extension residuals	-0.158 (0.369)		
σ_1		1.375 (0.127)***	
σ_2			0.936 (0.011)***
ρ_1		-0.868 (0.046)***	
ρ_2			-0.057 (0.093)
Log Likelihood	-6754.896		
Num. obs.	4245	372	3873

Note:$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors are in parentheses

Outcomes equations from the ESR regressions show that members land productivity

Table 3.6: ESR Regression of Household Income

	Selection Equation (1)	Members (2)	Non-Members (3)
Intercept	−1.864 (0.444)***	14.675 (0.985)***	11.454 (0.255)***
Gender	0.053 (0.132)	0.700 (0.254)***	0.528 (0.072)***
Age	0.020 (0.017)	−0.039 (0.030)	0.003 (0.009)
Age Squared	−0.000 (0.000)*	0.000 (0.000)	−0.000 (0.000)
Education	0.267 (0.075)***	−0.196 (0.132)	−0.025 (0.039)
Active persons	0.035 (0.010)***	0.021 (0.021)	0.055 (0.007)***
Dependents	0.027 (0.010)***	−0.032 (0.019)	0.020 (0.006)***
Migration	0.011 (0.098)	−0.106 (0.160)	−0.100 (0.052)*
Equipment	−0.027 (0.051)	0.315 (0.153)**	0.135 (0.033)***
Area owned	0.003 (0.004)	0.061 (0.011)***	0.037 (0.002)***
Distance to road	−0.007 (0.003)***	−0.007 (0.005)	−0.007 (0.001)***
Extension	1.001 (0.756)	−0.791 (0.243)***	0.248 (0.084)***
Groundnut AEZ	−0.807 (0.146)***	0.605 (0.228)***	0.013 (0.068)
Casamance AEZ	−0.267 (0.119)**	0.028 (0.188)	0.135 (0.066)**
South-East AEZ	−0.278 (0.119)**	0.140 (0.218)	0.256 (0.081)***
Information on Sales	0.581 (0.270)**		
Tap water	0.180 (0.070)**		
Extension residuals	0.057 (0.397)		
σ_1		1.349 (0.150)***	
σ_2			1.117 (0.013)***
ρ_1		−0.796 (0.081)***	
ρ_2			−0.054 (0.113)
Log Likelihood	−7473.406		
Num. obs.	4245	372	3873

Note:***$p < 0.01$, **$p < 0.05$, *$p < 0.1$. Standard errors are in parentheses

is significantly determined by the age, education, number of dependents, area of land owned, access to extension, and the household agro-ecological zone (Groundnut basin, Casamance and South East). For non-members, the main variables that affect significantly their land productivity are education, number of active family people, the value of agricultural equipment, area of land owned, distance to the nearest road, access to extension services, the residence in agro-ecological zones (Groundnut basin, Casamance and South East).

Results of the ESR also exhibit some differences in the determinants of household income for members and non-members. Variables such as gender, value of agricultural equipment, land area owned, access to extension services, and the residence in Casamance agro-ecological zone, affect significantly members household income. Meanwhile, the household income of non-members is influenced by gender, household size, migration, agricultural equipment, area of land owned, distance to road, access to extension services, residence in the Casamance and South-East agro-ecological regions.

The ESR model produces mean outcomes on treated household and corresponding counterfactual outcomes i.e. what would have been the outcome had the treated group not received the treatment. The average treatment effect on treated (ATT) is the net difference between these two outcomes. Similarly, the model also produces the mean outcome of the control group (non-members) and its counterfactual i.e. what would have been the mean outcome had the control group received the treatment. The difference between these last two outcomes produces the average treatment effect on untreated (ATU). These average outcomes and the estimated ATT and ATU are presented in table 3.7. The estimates reveal that the treatment effect for membership in farmer-based organizations on land productivity and household income are positive and significantly different from zero. The ATT are 2.405 and 1.959 for land productivity and household income, respectively. Membership in producer organizations significantly improves the log of land productivity and household income by 19.3% and 14.1%, respectively. Had non-members decided to be members of farmer-based organizations, the log value of their land productivity would have been increased by 24.5% and their income by 20%.

Table 3.7: ATT and ATU of FBO membership: ESR Estimates

Outcomes	Mean outcomes		Treatment Effect	Effect(%)
	Members	Non-Members		
Land Productivity	14.842 (0.724)	12.438(0.265)	$ATT = 2.405$ (0.769)***	19.3
	13.416 (0.588)	10.774(0.260)	$ATU = 2.643$ (0.692)***	24.5
Household Income	15.899 (0.977)	13.940 (0.486)	$ATT = 1.959$ (0.686)***	14.1
	14.629 (0.856)	12.194 (0.455)	$ATU = 2.435$ (0.601)***	20.0

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors are in parentheses.

Our results suggest that farmer organizations in Senegal are effective at enhancing farmers' land productivity and welfare. These results are in line with those of Ma and Abdulai (2016) in China, Mishra *et al.* (2018) in Nepal, and Francesconi and Ruben (2012) in Ethiopia, who found that members of farmer-based organizations generally experience better crop yields than non-members. Our findings are also consistent with the results of the growing literature on farmer-based organizations in developing countries, where most scholars observed a positive correlation between membership and economic welfare (Fischer and Qaim, 2012; Verhofstadt and Maertens, 2015; Wossen *et al.*, 2017).

3.4.4 Heterogeneous treatment effects

For the rest of the analysis, we focus on the evaluation of the heterogeneity of the effect of membership in farmer organizations, using graphical and regressions techniques. Figure 3.1 shows how the treatment effect on land productivity and household income (estimated from ESR models) vary over the propensity scores. The results show that the ATT on both outcomes indicators varies significantly with the propensity score and that the slope is negative, suggesting that the effects of farmers based organization membership on land productivity and household income are stronger for households with the lowest probability to belong to a farmers organization and these effects decrease with the propensity of membership. The slopes coefficients of the graphs are estimated at 2.3 and 3.4 respectively for land productivity and household income. This would mean that with every 1 percentage point increase in the likelihood of membership in farmer organizations, the effect of membership on land productivity and household income would reduce respectively by 2.3% and 3.4%. The household income effect of membership in farmer organizations even becomes zero in the upper end of the propensity score distribution. These results to some extent are similar to those observed in Rwanda by Verhofstadt and Maertens (2015). As stated by these authors farmers who would take most from membership in producer organizations are the ones who face entry constraints (human or physical) and therefore are less keen to become members.

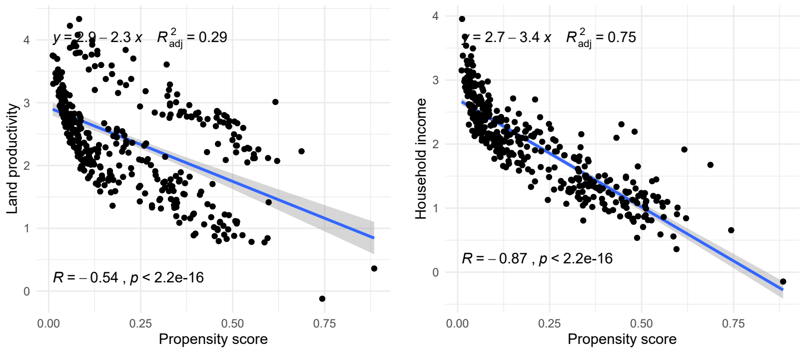


Figure 3.1: Heterogeneity over propensity scores

The OLS regression of the estimates of ATT, from the ESR models, on some of the characteristics of organization members are presented in table 3.8. Results show that the effect of membership in FBO on farm land productivity and household income appears to be different for each member. The impact of membership for both outcomes decreases significantly with the number of active persons living in the household and this effect is less important for those households who have access to extension services. Moreover, the effect has a U-shape relation with the age of the household head, implying that the effect of membership decreases with age for younger household heads and increases after a certain age. The impact of membership on land productivity also increases significantly with distance to the nearest road. Furthermore, this effect is less important for formally educated members. With regard to the household income, statistically significant differential effects are observed for other characteristics such as gender and the area of land possessed. The effect is larger for male members than for female members and it increases with the area of land possessed. OLS regressions results are at some extent corroborated by figures (3.3, 3.4, 3.5, 3.6 and 3.7) in the appendix. Moreover, membership effect appears to be determined by the specific type of organizations that households belong to. Results show that the impact of membership on land productivity is stronger for households who belong to the Economic Interest Groups. Meanwhile, the effect of membership on household income is more important for households who are members of Cooperatives. Furthermore, for both outcomes, this effect is less significant for households who are members of Rural Associations.

Table 3.8: Heterogeneous treatment effects:OLS regressions

	OLS without types of Farmer Organizations		OLS with types of Farmer Organizations	
	Land Productivity	Household Income	Land Productivity	Household Income
Intercept	3.714 (0.382)***	3.122 (0.256)***	3.847 (0.374)***	3.172 (0.251)***
Gender	0.021 (0.121)	0.202 (0.081)**	−0.039 (0.119)	0.181 (0.080)**
Age	−0.042 (0.014)***	−0.041 (0.009)***	−0.046 (0.014)***	−0.041 (0.009)***
Age Squared	0.001 (0.000)***	0.001 (0.000)***	0.001 (0.000)***	0.001 (0.000)***
Education	−0.476 (0.055)***	−0.060 (0.037)	−0.457 (0.054)***	−0.056 (0.036)
Active members	−0.040 (0.009)***	−0.060 (0.006)***	−0.037 (0.009)***	−0.059 (0.006)***
Migration	−0.101 (0.075)	0.060 (0.050)	−0.109 (0.073)	0.055 (0.049)
Equipment	−0.008 (0.065)	0.006 (0.044)	−0.026 (0.064)	0.008 (0.043)
Area owned	−0.008 (0.005)	0.038 (0.003)***	−0.007 (0.005)	0.036 (0.003)***
Distance to road	0.015 (0.002)***	0.000 (0.001)	0.014 (0.002)***	0.000 (0.001)
Extension	−0.901 (0.056)***	−1.074 (0.038)***	−0.944 (0.055)***	−1.098 (0.037)***
Types of Farmer Organizations				
Economic Interest Groups			0.181 (0.070)***	−0.032 (0.047)
Rural Associations			−0.208 (0.082)**	−0.162 (0.055)***
Producer Groups			−0.019 (0.084)	0.034 (0.057)
Cooperatives			0.125 (0.084)	0.157 (0.057)***
Adj. R^2	0.562	0.753	0.589	0.767
Num. obs.	372	372	372	372

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors are in parentheses.

3.5 Conclusion

In Senegal, recent renaissance of the cooperative movement coupled with the revival of the agricultural sector led us to explore the impact of farmer-based organization membership on cereals producing farm household land productivity and net incomes. Results were derived using a nationally represented household cross-sectional data collected in all agro-ecological regions and two econometric estimation techniques that control for selection bias arising from both observed and unobserved factors.

We find that the education of the household head, household size, distance to the nearest road, access to extension and to information on sales, living conditions of the household proxied by water source and the location of the household in various agro-ecological zones are the most important factors influencing households decision to belong to a producer organization. Additionally, findings suggest that membership in farmers' collective action groups is a key component of farm households' land productivity and income, and obtained results appear to be consistent throughout the two estimation methods. In particular, results from our preferred model, the Endogenous Switching Regressions, show that being a member of an organization helps to increase land productivity by almost twenty percent and household income by at least fourteen percent. Furthermore, membership in farmer organizations exhibits heterogeneous effects over the propensity score and over household characteristics. The estimated treatment effects are negatively correlated with households' likelihood to belong to a farmer-based organization, implying that the effect of membership is stronger for households with the lowest propensity to become members, meanwhile suggesting the possible existence of entry barriers that might face some farmers.

These results support once again the idea that farmer organizations have the potential to benefit farmers by increasing their incomes through the provision of conditions and the necessary social networks for access to technologies, knowledge, and production inputs. These collective action groups would, therefore, induce better farm productivity for improved incomes and then contribute to reducing rural poverty.

References

Abdul-Rahaman, A. and Abdulai, A. (2018) Do farmer groups impact on farm yield and efficiency of smallholder farmers? Evidence from rice farmers in northern Ghana, *Food Policy*, **81**, 95 – 105.

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

Abebaw, D. and Haile, M. G. (2013) The impact of cooperatives on agricultural technology adoption: Empirical evidence from Ethiopia, *Food Policy*, **38**, 82 – 91.

Ahmed, M. H. and Mesfin, H. M. (2017) The impact of agricultural cooperatives membership on the wellbeing of smallholder farmers: empirical evidence from eastern Ethiopia, *Agricultural and Food Economics*, **5**, 1–20.

Barham, J. and Chitemi, C. (2009) Collective action initiatives to improve marketing performance: Lessons from farmer groups in Tanzania, *Food Policy*, **34**, 53 – 59.

Becker, S. and Caliendo, M. (2007) Sensitivity analysis for average treatment effects, *Stata Journal*, **7**, 71–83.

Bernard, T., Frölich, M., Landmann, A., Unte, P. N., Viceisza, A. and Wouterse, F. (2015) Building trust in rural producer organizations in Senegal: Results from a randomized controlled trial, Tech. rep., IZA Discussion Papers.

Bernard, T. and Spielman, D. J. (2009) Reaching the rural poor through rural producer organizations? A study of agricultural marketing cooperatives in Ethiopia, *Food Policy*, **34**, 60 – 69.

Bernard, T. and Taffesse, A. S. (2012) Returns to scope? Smallholders' commercialisation through multipurpose cooperatives in Ethiopia, *Journal of African Economies*, **21**, 440–464.

Bernard, T., Taffesse, A. S. and Gabre-Madhin, E. (2008) Impact of cooperatives on smallholders' commercialization behavior: evidence from Ethiopia, *Agricultural Economics*, **39**, 147–161.

Caliendo, M. and Kopeinig, S. (2008) Some practical guidance for the implementation of propensity score matching, *Journal of Economic Surveys*, **22**, 31–72.

Chagwiza, C., Muradian, R. and Ruben, R. (2016) Cooperative membership and dairy performance among smallholders in Ethiopia, *Food Policy*, **59**, 165–173.

Dehejia, R. and Wahba, S. (2002) Propensity score-matching methods for nonexperimental causal studies, *The Review of Economics and Statistics*, **84**, 151–161.

Fischer, E. and Qaim, M. (2012) Linking smallholders to markets: Determinants and impacts of farmer collective action in Kenya, *World Development*, **40**, 1255–1268.

Francesconi, G. N. and Heerink, N. (2010) Ethiopian agricultural cooperatives in an era of global commodity exchange: Does organisational form matter?, *Journal of African Economies*, **20**, 153–177.

Francesconi, G. N. and Ruben, R. (2012) The hidden impact of cooperative membership on quality management: A case study from the Dairy Belt of Addis Ababa, *Journal of Entrepreneurial and Organizational Diversity*, **1**, 85–103.

Greene, W. (2000) *Econometric Analysis*, Prentice Hall.

Hirano, K., Imbens, G. W. and Ridder, G. (2003) Efficient estimation of average treatment effects using the estimated propensity score, *Econometrica*, **71**, 1161–1189.

Hoken, H. and Su, Q. (2015) Measuring the effect of agricultural cooperatives on household income using PSM-DID : a case study of a rice-producing cooperative in China, IDE Discussion Papers 539, Institute of Developing Economies, Japan External Trade Organization(JETRO).

Imbens, G. (2004) Nonparametric estimation of average treatment effects under exogeneity: A review, *The Review of Economics and Statistics*, **86**, 4–29.

Ito, J., Bao, Z. and Su, Q. (2012) Distributional effects of agricultural cooperatives in China: Exclusion of smallholders and potential gains on participation, *Food Policy*, **37**, 700–709.

Jelliffe, J. L., Bravo-Ureta, B. E., Deom, C. M. and Okello, D. K. (2018) Adoption of High-Yielding Groundnut Varieties: The Sustainability of a Farmer-Led Multiplication-Dissemination Program in Eastern Uganda, *Sustainability*, **10**, 1–21.

Lokshin, M. and Sajaia, Z. (2004) Maximum likelihood estimation of endogenous switching regression models, *Stata Journal*, **4**, 282–289.

Ma, W. and Abdulai, A. (2016) Does cooperativemembership improve household welfare? Evidence from apple farmers in China, *Food Policy*, **58**, 94–102.

Ma, W., Abdulai, A. and Goetz, R. (2018) Agricultural Cooperatives and Investment in Organic Soil Amendments and Chemical Fertilizer in China, *American Journal of Agricultural Economics*, **100**, 502–520.

Maddala, G., Chesher, A. and Jackson, M. (1986) *Limited-Dependent and Qualitative Variables in Econometrics*, Econometric Society Monographs, Cambridge University Press.

Mishra, A. K., Kumar, A., Joshi, P. K. and D'Souza, A. (2018) Cooperatives, contract farming, and farm size: The case of tomato producers in Nepal, *Agribusiness*, **34**, 865–886.

Mojo, D., Fischer, C. and Degefa, T. (2017) The determinants and economic impacts of coffee farmer cooperatives: Recent evidence from rural Ethiopia, *Journal of Rural Studies*, **50**, 84–94.

Reed, G. and Hickey, G. M. (2016) Contrasting innovation networks in smallholder agricultural producer cooperatives: Insights from the Niayes Region of Senegal, *Journal of Co-operative Organization and Management*, **4**, 97 – 107.

Rosenbaum, P. (2002) *Observational Studies*, Springer Series in Statistics, Springer.

Rosenbaum, P. R. and Rubin, D. B. (1983) The central role of the propensity score in observational studies for causal effects, *Biometrika*, **70**, 41–55.

Salifu, A., Francesconi, G. N. and Kolavalli, S. (2010) A review of collective action in rural Ghana, Tech. Rep. 998, International Food Policy Research Institute (IFPRI).

Schwettmann, J. (2014) Cooperatives in Africa: Success and Challenges, in *A Contribution to the International Symposium on Cooperatives and Sustainable Development Goals: The Case of Africa, Berlin, 2 September 2014.* (Ed.) I. L. O. I. Geneva.

Verhofstadt, E. and Maertens, M. (2014) Smallholder cooperatives and agricultural performance in Rwanda: do organizational differences matter?, *Agricultural Economics*, **45**, 39–52.

Verhofstadt, E. and Maertens, M. (2015) Can Agricultural Cooperatives Reduce Poverty? Heterogeneous Impact of Cooperative Membership on Farmers' Welfare in Rwanda, *Applied Economic Perspectives and Policy*, **37**, 86–106.

Wollni, M. and Zeller, M. (2007) Do farmers benefit from participating in specialty markets and cooperatives? the case of coffee marketing in Costa Rica, *Agricultural Economics*, **37**, 243–248.

Wooldridge, J. (2015) Control function methods in applied econometrics, *Journal of Human Resources*, **50**, 420–445.

World Bank (2007) *World Development Report 2008 : Agriculture for Development*, World Bank Washington, DC.

World Bank (2017) World bank data, world Bank Data Base.

Wossen, T., Abdoulaye, T., Alene, A., Haile, M. G., Feleke, S., Olanrewaju, A. and Manyong, V. (2017) Impacts of extension access and cooperative membership on technology adoption and household welfare, *Journal of Rural Studies*, **54**, 223–233.

Appendix

Table 3.9: Addressing potential endogeneity of extension variable

	Membership	Extension
Intercept	-1.809 (0.430)***	-1.850 (0.396)***
Gender	0.063 (0.125)	0.043 (0.111)
Age	0.027 (0.016)*	0.028 (0.014)**
Age Squared	-0.000 (0.000)**	-0.000 (0.000)*
Education	0.307 (0.060)***	0.255 (0.056)***
Active members	0.029 (0.010)***	-0.004 (0.010)
Dependents	0.024 (0.009)***	-0.007 (0.009)
Migration	0.067 (0.081)	0.272 (0.071)***
Equipment	0.002 (0.004)	0.003 (0.003)
Area owned	-0.002 (0.043)	0.080 (0.037)**
Distance to road	-0.007 (0.002)***	-0.006 (0.002)***
Groundnut AEZ	-0.917 (0.091)***	-0.656 (0.084)***
Casamance AEZ	-0.350 (0.094)***	-0.360 (0.090)***
South-East AEZ	-0.288 (0.111)***	-0.078 (0.103)
Information on Sales	0.842 (0.182)***	0.892 (0.176)***
Tap water	0.212 (0.071)***	0.083 (0.064)
Extension needs	-0.305 (0.359)	0.507 (0.238)**
Log Likelihood	-1152.478	-1371.070
Num. obs.	4245	4245

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 3.10: Propensity score matching quality test

	Before Matching	After Matching
Pseudo R2	0.177	0.005
LR χ^2	445.49	5.33
P-value ($p > \chi^2$)	0.000	0.997
Mean standardized bias	23.9	2.7

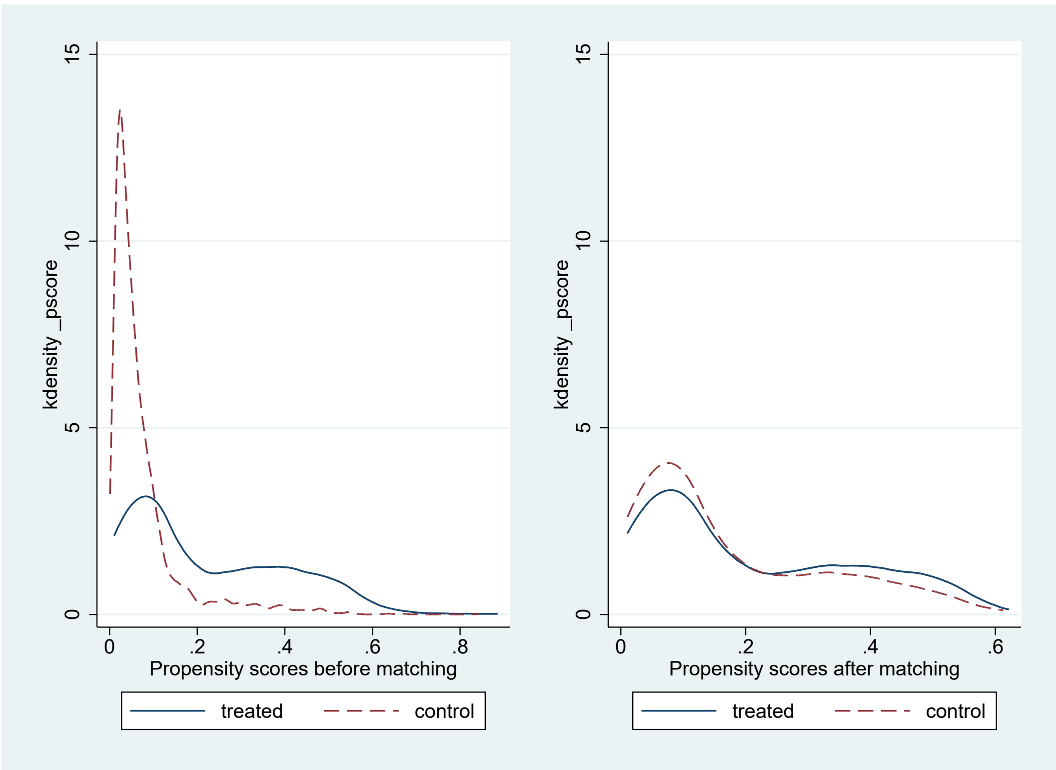


Figure 3.2: Kernel density of propensity scores

Table 3.11: Instrumental variable checking for ESR regressions

	Membership	Land Productivity		Household Income	
		Non-Members	Members	Non-Members	Members
Intercept	-1.87 (0.45)***	11.04 (0.21)***	12.88 (0.70)***	11.47 (0.26)***	12.27 (0.76)***
Gender	0.06 (0.13)	0.09 (0.06)	0.10 (0.22)	0.53 (0.07)***	0.77 (0.24)***
Age	0.02 (0.02)	-0.00 (0.01)	-0.03 (0.03)	0.00 (0.01)	-0.02 (0.03)
Age Squared	-0.00 (0.00)**	0.00 (0.00)	0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)
Education	0.31 (0.08)***	0.12 (0.03)***	-0.12 (0.10)	-0.02 (0.04)	0.04 (0.11)
Active members	0.03 (0.01)***	0.02 (0.01)***	0.03 (0.02)	0.06 (0.01)***	0.05 (0.02)**
Dependents	0.03 (0.01)***	-0.00 (0.01)	-0.01 (0.02)	0.02 (0.01)***	-0.00 (0.02)
Migration	0.04 (0.10)	-0.06 (0.04)	-0.16 (0.14)	-0.10 (0.05)*	-0.14 (0.15)
Equipment	-0.01 (0.05)	0.08 (0.03)***	0.16 (0.12)	0.13 (0.03)***	0.29 (0.13)**
Area owned	0.00 (0.00)	-0.01 (0.00)***	-0.02 (0.01)*	0.04 (0.00)***	0.05 (0.01)***
Distance to road	-0.01 (0.00)***	-0.00 (0.00)***	-0.01 (0.00)	-0.01 (0.00)***	-0.01 (0.00)**
Extension	0.29 (0.83)	0.27 (0.06)***	0.36 (0.10)***	0.27 (0.07)***	0.14 (0.11)
Groundnut AEZ	-0.89 (0.15)***	-0.24 (0.05)***	-1.23 (0.17)***	0.00 (0.06)	0.03 (0.19)
Casamance AEZ	-0.32 (0.12)***	0.36 (0.06)***	-0.82 (0.17)***	0.12 (0.07)*	-0.14 (0.18)
South-East AEZ	-0.28 (0.12)**	0.35 (0.07)***	-1.10 (0.19)***	0.24 (0.08)***	-0.02 (0.20)
Information on Sales	0.80 (0.29)***	0.09 (0.16)	0.15 (0.23)	0.26 (0.19)	0.53 (0.25)**
Tap water	0.20 (0.08)***	-0.00 (0.04)	0.23 (0.12)*	-0.02 (0.04)	0.16 (0.13)
Extension residuals	0.44 (0.44)				
Adj. R^2		0.10	0.29	0.15	0.25
Log Likelihood	-1038.13				
Num. obs.	4245	3873	372	3873	372

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

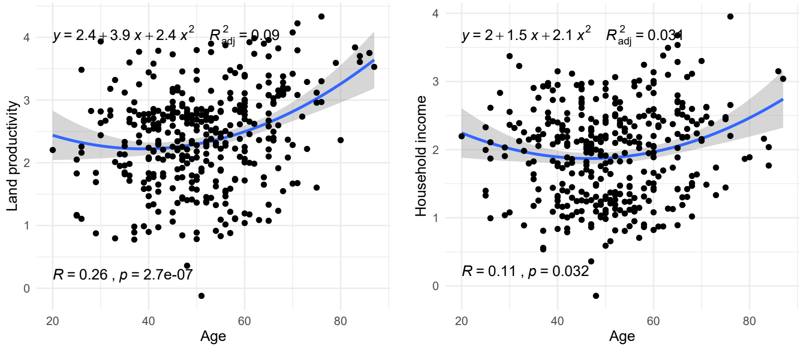


Figure 3.3: Heterogeneity over household head age

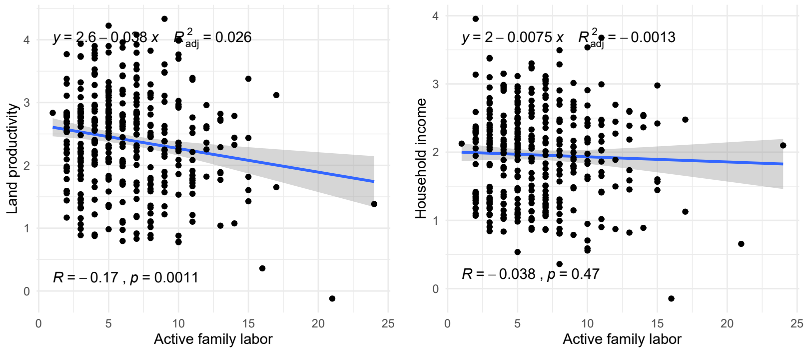


Figure 3.4: Heterogeneity over active family labour

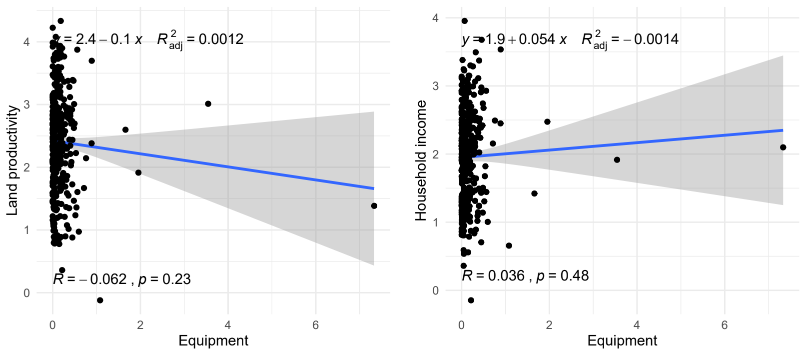


Figure 3.5: Heterogeneity over agricultural equipment ownership

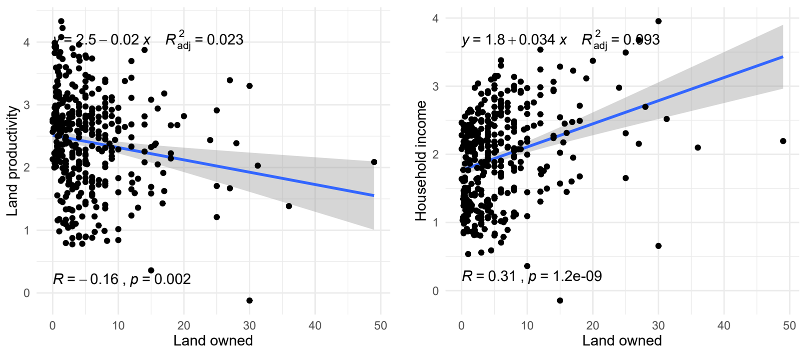


Figure 3.6: Heterogeneity over land ownership

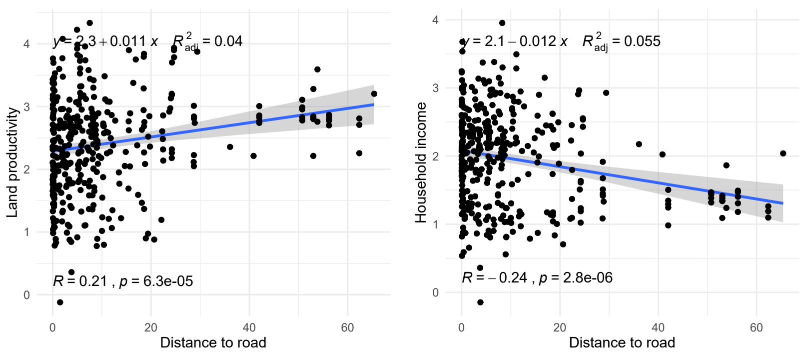


Figure 3.7: Heterogeneity over access to road

Chapter 4

A stochastic meta-frontier approach to estimating the impact of membership in farmer organizations on rice producers' efficiency: contrasting results from Senegal[1]

K. Christophe Adjin, Christian H. C. A. Henning

Abstract

Previous empirical literature mostly suggest that membership in farmer organizations helps farmers to improve their access to production inputs and technology. However, the question of whether members of organizations are efficient in using these inputs still need to be answered. Using a cross-sectional data from 835 rice-farming households in Senegal, we therefore, investigated the extent to which membership in farmer organizations affects farm technical efficiency. To do so, we combine the propensity score matching method with the sample selection stochastic frontier model and the stochastic meta-frontier approach. The propensity score matching and the sample selection stochastic frontier framework help in mitigating selection biases in the production frontier. Applying the meta-frontier approach, farmers' technical efficiency were estimated and compared. Results show that membership in farmer organizations contributes significantly to improving rice production. However, when considering group-specific frontiers (farmers operating in their own benchmark), organization members do not technically perform better than non-members. Furthermore, when considering the meta-frontier estimates, significant differences in technical efficiency between members and non-members can still be observed in favour of non-members.

Keywords: Farmer organizations, technical efficiency, Senegal.

JEL Codes: Q13, D24.

[1]A earlier version of this chapter has been published as a contributed paper: *A stochastic meta-frontier approach to estimating the impact of cooperatives membership on rice farmers' efficiency: Contrasting results from Senegal.* Contributed Paper prepared for presentation at the 94th Annual Conference of the Agricultural Economics Society, KULeuven, Belgium. 15-17/04/20 (Cancelled). `https://ageconsearch.umn.edu/record/303700/`

4.1 Introduction

According to FAO statistics (FAO, 2019), rice is the most important cereal in terms of production in sub-Saharan Africa. Rice has become an economically important crop and the main staple food of millions of people. Indeed, due to population growth and a shift in consumption habits in favour of rice, the relative growth of demand for rice is faster in sub-Saharan Africa than anywhere else in the world (Balasubramanian *et al.*, 2007; Seck *et al.*, 2010).

In Senegal, rice occupies an important position nationally both in terms of consumption and production. Average annual consumption over the last decade (2007–2016) was more than 1.2 million[2] of tonnes of milled rice while the average yearly production over the same period was only 358,357 tonnes[3]. There is, therefore, a significant gap between production and consumption, which is filled with large-scale rice importation every year. Milled rice imports have increased from 536,870 tonnes in 2000 to more than 973,000 tonnes in 2016, at an average annual cost of about 315 million US dollars (FAO, 2019). This gap between domestic production and consumption denotes a real food security problem in Senegal. The country was hardly hit by the 2007–2008 food crisis, with violent riots observed during the crisis (Seck *et al.*, 2010; Diagne *et al.*, 2013). In addition, the heavy dependence on imports represents a serious burden on the country's trade and foreign exchange balance. Reducing Senegal's dependence on imported rice and meeting the population's demand for rice are real challenges for the Senegalese government. Hence, since 2009, priority has been given by the government to the domestic rice sector because of its potential to provide national food security, support economic growth, and alleviate poverty (République du Sénégal, 2009). The 2014-2017 revised National Program for Self-Sufficiency in Rice ("Programme National d'Autosuffisance en riz - PNAR") intended to increase rice production in the country to reach self-sufficiency in 2017[4].

To achieve this goal and for the rice sector to express its full potential, rice farmers need to have access to production inputs and technologies, which constitute however, general challenging factors for the agricultural sector in most developing countries (World Bank, 2007). According to Salifu *et al.* (2010), to overcome these challenges and improve agricultural performance, policymakers during the past decades, have

[2]Consumption here refers to apparent consumption and it is computed using data from FAOSTAT. Consumption equals to paddy rice produced converted into milled rice using a ratio of 0.67 plus imports and net of export

[3]Statistics compiled using paddy rice production data of FAOSTAT (FAO, 2019), using a paddy to a milled rice conversion factor of 0.67 (Soullier and Moustier, 2018)

[4]`http://sakss.sn/programme-national-dautosuffisance-en-riz-pnar`

relied on collective action groups, such as cooperatives and farmer organizations. Nowadays, this agricultural development approach based on farmer organizations prevails. However, such an approach is increasingly supported by quantitative studies in which scholars try to evaluate the effective contribution of membership in farmer-based organizations to various agricultural indicators. Therefore, during the last decade, an important body of literature was dedicated to the analyses of the impact of farmer organizations on farm household welfare (Fischer and Qaim, 2012b; Ito *et al.*, 2012; Verhofstadt and Maertens, 2015; Ma and Abdulai, 2016; Mojo *et al.*, 2017; Ahmed and Mesfin, 2017; Mishra *et al.*, 2018), farm products commercialization and marketing (Wollni and Zeller, 2007; Bernard *et al.*, 2008; Bernard and Spielman, 2009; Barham and Chitemi, 2009; Bernard and Spielman, 2009; Francesconi and Heerink, 2010; Fischer and Qaim, 2012b; Chagwiza *et al.*, 2016), or agricultural technology adoption (Abebaw and Haile, 2013; Ma *et al.*, 2018a). However, regarding the association between membership in these organizations and farms' productivity and efficiency, very few research have been done, especially concerning technical efficiency analysis.

A review of studies shows that membership in farmer organizations has a positive impact on farm yields and productivity (Ma *et al.*, 2018a; Mishra *et al.*, 2018; Francesconi and Ruben, 2012). Regarding the impact of membership on technical efficiency, the causal relationship between membership and technical efficiency is not straightforward and not conclusive. Abate *et al.* (2014) showed that farmer organizations through the mechanism of easing access to productive inputs contribute significantly to members' technical efficiency. Contrarily, the study by Addai *et al.* (2014) in Ghana found no significant impact of membership in maize farmer groups on technical efficiency. These two authors used a combination of a matching technique and a frontier approach with the assumption of similar technology for members and non-members. However, this assumption generally cannot hold. Mostly farmers join collective action groups to have access to improved technologies and to increase their productivity and technical efficiency. The membership in a farm-related organization becomes then endogenous. Therefore, it is crucial to take into account biases that arise from endogenous self-selection and from technological heterogeneity. These potential biases have been considered in two recent papers by Abdul-Rahaman and Abdulai (2018) and Ma *et al.* (2018b), where they used a propensity score matching technique and the sample-selection stochastic frontier approach (Greene, 2010) to control for selection biases in the production frontiers. They then designed two groups frontiers (members and non-members) and compared farmers' technical efficiencies from their respective groups' frontiers. By doing so, these authors found that cooperative members are more technically efficient than non-members. However, the two groups of farmers are operating against two different benchmarks, comparing their technical efficiency estimates does not permit one to estimate the real difference in the productivity of the two groups of farmers (Villano *et al.*, 2015; Henningsen *et al.*, 2015). In addition to the group frontiers, a meta-frontier approach to evaluate the technical efficiency of farmers would have been a more robust approach to compare farmers' technical efficiencies (Villano *et al.*, 2015; Henningsen *et al.*, 2015).

This paper therefore seeks to investigate the causal relationship between membership in farmer-based organizations and farms technical efficiency in Senegal, by using a methodology that combines three different approaches. First, in order to correct for possible selection biases, we used a propensity score matching approach. Secondly, to take into account both technology heterogeneity and membership in farmer organizations when comparing the efficiency between members and non-members, we use Greene (2010) sample selection frontier to estimate two separated stochastic production frontiers. Finally, following Huang *et al.* (2014) we built a stochastic meta-frontier that works as a benchmark against which the performances of different farms could be compared to. The remainder of this paper is organized as follows. The next section describes the suggested econometric framework. The third section presents the used data. The following section presents the estimation results and their discussion. In the final section, results are summarized with some policy recommendations.

4.2 Econometric framework

The main objective of this paper is to investigate the impact of membership in farmer organizations on farm households' technical efficiency. To do so, we adopted an econometric framework that combines three main approaches. To address the issue of selection bias arising from observables, we used the propensity matching technique. Greene (2010) sample selection stochastic frontier model helped us to address the selection bias resulting from unobservables in the designed frontiers. To take into account the technology heterogeneity that could result from membership in farmer organizations, we used the meta-frontier approach of Huang *et al.* (2014) to derive the technology gaps and to compare efficiencies between members and non-members.

4.2.1 Modelling membership in farmer organizations

Membership in farmer organizations can be modelled within the random utility framework. Here, a household chooses to be a member of an organization if the expected utility gained from membership (M_{i1}) is larger than the one from non-membership (M_{i0}). This would mean that a household becomes a member of an organization if the expected net utility ($M_{i1} - M_{i0}$) is greater than zero, which can be specified as a function of observed covariates (Z) in a latent variable model as follows:

$$M_i^* = \alpha' z_i + w_i, \ M_i = 1 \text{ if } M_i^* > 0, \tag{4.1}$$

where M_i is a binary variable that takes the value 1 for a household i that is a member of a farmer organization and 0 otherwise; α is a vector of parameters to be estimated; z_i is a vector of exogenous farm and household characteristics, and w_i is an error term. Several empirical works have shown that participating in a farm-related collective action group increases the adoption level of new agricultural technologies, through various mechanisms (see e.g. Fischer and Qaim (2012a,b); Abebaw and Haile (2013). Thus, the frontier production function might differ between members of such groups and non-members due to technology accessibility and adoption. It becomes therefore intuitive to design two different production functions for farmer-based organizations members and non-members, and statistically compare and test their parameters. However, proceeding so is complicated because of the self-selection of membership and the following resulting choice of technology (Mayen *et al.*, 2010).

4.2.2 Stochastic frontier approach

We adopt the stochastic frontier analysis (SFA) framework to estimate the production frontiers and measure the technical efficiency of farmers. The standard stochastic production frontier model is specified as:

$$y_i = f(x_i, \beta) \exp(v_i - u_i), \tag{4.2}$$

where y_i denotes the output for the i^{th} farm ($i = 1, ..., N$), x_i is a vector of inputs, β are parameters to be estimated, v_i is a two-sided stochastic term that accounts for statistical noise, u_i is a non-negative stochastic term representing inefficiency, and ε_i ($\varepsilon_i = v_i - u_i$) is the composite error term. Generally, it is assumed that v_i and u_i are identically and independently distributed, u_i follows a half-normal distribution with variance σ_u^2 and v_i follows a normal distribution with variance σ_v^2. This model is usually estimated using the maximum likelihood estimator as suggested by Aigner *et al.* (1977). After the estimation of the frontier model, following Jondrow *et al.* (1982)(JLMS thereafter) one can calculate the farm-specific technical efficiency.

4.2.3 Correcting for selection bias

Following Ma *et al.* (2018b) and Abdul-Rahaman and Abdulai (2018), to correct for selection bias in our estimates, we first use the Propensity Score Matching method to match members with non-members in the sample. Then we use the sample selection frontier approach to correct for selectivity bias in the production frontier. Propensity Score Matching (PSM) uses observable characteristics of units in the sample to generate a control group that is as similar to the treated group as possible except for the treatment status, herein membership in farmer organizations (Rosenbaum and Rubin, 1983). PSM works under two main assumptions. The first is the conditional independence or unconfoundedness, stating that observable characteristics must be independent of potential outcomes, which implies that the decision of membership is only based on observable characteristics of households. The second is the common support or overlap condition that needs to be satisfied, i.e. the distributions of observable characteristics between members of organizations (the treated) and non-members (the untreated) have to overlap (Jelliffe *et al.*, 2018).

Sample selection bias arises when there is a correlation of the unobservables in the production function equation with those in the selection equation. In recent years, the literature reveals two main alternative applications of the sample selection modelling in the stochastic frontier model. Kumbhakar *et al.* (2009) suggested a model framework in which the selection mechanism operates through the one-sided error (u_i). Greene (2010) proposed a framework where the selection mechanism is operated through the error term v_i. The model by Kumbhakar *et al.* (2009) requires computationally demanding log likelihood functions (Villano *et al.*, 2015). Therefore, in this paper, we follow Greene (2010) approach, and design for farmers two simultaneous equations: a selection equation and a production function equation. The specification of this model is derived as follows:

$$\text{Selection equation}: M_i = 1\left[\alpha' z_i + w_i > 0\right],\ w_i \sim N\left[0, 1\right]$$

$$\begin{aligned} \text{SFP function}: y_i = f(x_i, \beta) + \varepsilon_i,\ \varepsilon_i \sim\ N\left[0, \sigma_\varepsilon^2\right] \\ (y_i, x_i) \text{ observed only when } M_i = 1 \end{aligned} \tag{4.3}$$

$$\begin{aligned}
\text{Error structure} : \varepsilon_i &= v_i - u_i \\
u_i &= |\sigma_u U_i| = \sigma_u |U_i| \text{ where } U_i \sim N[0,1] \\
v_i &= \sigma_v V_i \text{ where } V_i \sim N[0,1] \\
(w_i, v_i) &\sim N_2\left[(0,1),(1,\rho\sigma_v,\sigma_v^2)\right],
\end{aligned}$$

where M_i is a dummy variable that takes the value of 1 for organization members and 0 for non-members, z_i is a vector of covariates in the selection equation, w_i is the error term of the selection equation, and y_i, x_i, v_i, u_i and ε_i are as defined previously. The inefficiency term u_i is assumed to follow a half-normal distribution with variance σ_u^2 and w_i and v_i are assumed to follow a bivariate normal distribution with variances 1 and σ_v^2, respectively, and a correlation coefficient of ρ. ρ, α and β are parameters to be estimated. Non-zero values of ρ indicate the presence of selection bias and when $\rho = 0$, the model reduces to that of the standard stochastic frontier model.

Following Greene (2010), a two-step estimation procedure is used. In the first step we modelled membership in farmer organizations with the selection equation (4.1), using a probit model. Consistent maximum likelihood estimates of α are obtained and used to derive the conditional simulated log-likelihood function of the combination of equations 4.1 and 4.3 (for more details, see Greene (2010)).

Empirically, for the selection model, the variable M_i is a dummy representing the likelihood that the farmer belongs to a farm-related organization, taking the value of 1 if the farmer is a member and 0 otherwise, z is defined as previously. Similarly to Bravo-Ureta *et al.* (2012) and Abdul-Rahaman and Abdulai (2018), we estimated two stochastic frontier models, one for organization members and one for non-members. Once the two stochastic frontier models are estimated, one can derive the group-specific technical efficiency estimates for both members and non-members. To do so, we used the JLMS approach and then compared these efficiency scores against each benchmark.

However, our methodological framework still has one limitation. It is not possible to compare directly, the estimated technical efficiency between members and non-members since those scores pertain to each group's own frontier (González-Flores *et al.*, 2014; Villano *et al.*, 2015; Henningsen *et al.*, 2015). Therefore, in order to address this issue, we used a meta-frontier approach that enables us to estimate and compare the technical efficiency of production units regrouped in different types of technology.

4.2.4 Meta-frontier approach

Considering that all farmers are in J groups ($j = 1, 2$) and farmers in each group operate under a group-specific technology, with group-specific frontiers defined as $f^j(x_{ji})$ and $f(.)$ a specified functional form. Commonly, the meta-frontier production function $f^M(x_{ji})$ that envelops all different groups' frontiers $f^j(x_{ji})$ is expressed as:

$$f^j(x_{ji}) = f^M(x_{ji})\exp(-u_{ji}^M),\ \forall\ j, i, \tag{4.4}$$

where $u_{ji}^M \geq 0$, therefore $f^M(.) \geq f^j(.)$ and the relationship of the j^{th} production frontier to the meta-frontier is defined as the meta technology ratio (MTR), which expresses the difference in efficiency due to the choice of a particular technology, and it is between zero and one. To estimate the metafrontier, we follow Huang *et al.* (2014) approach that has the main advantage to allow statistical interpretations. In a first step, the standard maximum likelihood (ML) estimation is used to estimate group-specific frontiers. In a second step, a stochastic frontier model (as in equation 4.5) is formulated and estimated by the maximum likelihood to obtain the estimates of the meta-frontier:

$$\hat{f}^j(x_{ji}) = f^M(x_{ji})\exp(v_{ji}^M - u_{ji}^M),\ \forall i, j = 1, 2. \tag{4.5}$$

This equation is similar to the traditional stochastic frontier, where $\hat{f}^j(x_{ji})$ represents the estimates of the group-specific frontier, u_{ji}^M ($u_{ji}^M \geq 0$) is the technological gap and is assumed to follow a truncated-normal distribution with the mode μ^M and independent from v_{ji}^M, and v_{ji}^M is assumed to follow a normal distribution with zero mean, but non independently and identically distributed. Additionally, the mode $\mu^M(q_{ji})$ is a function of "industry" environmental variables q_{ji}. As described by Huang *et al.* (2014), at any given level of input, an associated farm household output level y_{ji} with respect to the meta-frontier $f^M(x_{ji})$ has three components: the meta technology ratio $MTR_i^j = \frac{f^j(x_{ji})}{f^M(x_{ji})}$, the group specific technical efficiency of each production unit $TE_i^j = \frac{y_{ji}}{f^j(x_{ji})\exp(v_{ji})} = \exp(-u_{ji})$, and the technical efficiency of each farmer regarding the meta-frontier $MTE_i^j = \frac{y_{ji}}{f^M(x_{ji})\times\exp(v_{ji})} = MTR_i^j \times TE_i^j$.

4.2.5 Empirical strategy

As stated previously, to correct for selectivity bias, we use first the propensity score matching method and then a sample selection stochastic frontier approach. Following Ma *et al.* (2018b) and Abdul-Rahaman and Abdulai (2018), in a first step, we generated the propensity score of belonging to a farmer organization, using a probit model, by regressing the membership variable on farm households observable characteristics (see table 4.1). In the PSM approach, numerous algorithms can be applied to match members and non-members of similar propensity scores. We use the most common matching technique: the nearest neighbour matching with five neighbours and caliper of 0.01. By doing so, a total of 788 matched farmers were obtained including 106 members and 682 non-members with a similar range of observable characteristics. The balancing test results are also presented in table 4.12. Results show that the standardized mean difference for all covariates used for the matching reduces from 42.3% before matching to 8.9% after matching. Moreover, the likelihood ratio test indicates that the null hypothesis of the joint significance of all covariates could be rejected before matching (p $> \chi^2 = 0.000$). Conversely, after the matching, with the same test the joint significance of all covariates could not be rejected (p $> \chi^2 = 0.795$). The results indicate that the required balancing property of the distribution of propensity scores is satisfied. In addition, the common support condition is also satisfied, as shown in figure 4.1.

Once the matched sample were obtained, we estimated the sample-selection stochastic frontier model. Here, the first stage is the estimation of the selection equation (4.1) as a standard probit model. Several factors are associated with membership in farmer organizations (Fischer and Qaim, 2012b; Abebaw and Haile, 2013; Tolno *et al.*, 2015; Mojo *et al.*, 2017), including personal details of household head (gender, age, education, migration status) and household characteristics (e.g. household size, agricultural equipment, land size), access to rural institutions (e.g. agricultural extension services, market), and the specific agro-ecological location of the household. Based on previous studies, in our empirical specification, we assume that the probability that a household belongs to a farmer-based organization is a function of these main selected variables. However, it is worth noting that households could have better access to extension due to membership in farmer organizations, rendering the access to extension services variable potentially endogenous in the modelling of membership, and leading then to biased estimates. We, therefore, corrected this endogeneity issue with the two-stage control function approach suggested by Wooldridge (2015)[5]. The variables used to model the membership in farmers'

[5]In a first stage, we estimated separately, the access to extension services and the organization membership on the same independent variables plus an instrument (here the farmer's expressed needs for support) using a probit model. The instrument "extension needs" significantly influences the access to extension services but does not directly influence the household membership status (see table 4.11 in the appendix). From the two questions: "do you need extension services?" and

organizations are presented in table 4.1.

The second stage of the sample selection stochastic frontier model is the estimation of the production function. To do so, from preliminary comparisons using the pooled unmatched data, a maximum likelihood ratio test led to the rejection of the Cobb-Douglas (CD) in favour of the translog (TL) functional form ($\chi^2 = 192.41, p < 0.01$), which has the main advantage to add the effects of interactions between inputs. In addition, the Akaike Information Criterion of the translog ($AIC = 2145.698$) was less than that of the Cobb Douglas ($AIC = 2318.107$). Therefore, we used the translog specification which is expressed as:

$$y_i = f(x_i, \beta) + \delta D_i + \gamma G_i + \varepsilon_i,\ \varepsilon_i \sim\ N\left[0, \sigma_\varepsilon^2\right], \tag{4.6}$$

where y_i represents the natural logarithm of the output of the i^{th} farmer, x_i denote vectors of the natural logarithm of production inputs; D represent dummy variables; G represents other contextual variables; β, δ and γ are parameters to be estimated; ε_i is the composite error term as defined previously and comprising v_i and u_i The output here is the total rice production (in kilograms). The four inputs included in the models are the land cultivated (in hectare), the quantity of total seeds (in kilograms), the quantity of total labour (in working-days equivalent), and the quantity of total fertilizers (in kilograms). The dummy variables are organization membership, the use of improved seeds, and the non-use of fertilizers. The other environmental variables are the percentages of clay elements in soils, and the rainfall of the survey year 2016 (in millimetres). We follow Battese (1997) approach, to account for zero values of fertilizer use by including a dummy for the non-use of fertilizer, such that the logarithm of the fertilizer with zero values is taken only if it is positive, and zero otherwise.

To identify whether it is necessary to estimate separate frontiers for members and non-members, we first estimated a pooled stochastic production frontier including a dummy variable for farmer organization membership. Then, two separate stochastic production frontier models for members and non-members are estimated. Finally, using a likelihood ratio test, we checked if there is a difference in technologies used by the two groups of farmers (Bravo-Ureta *et al.*, 2012). Specifically, the estimated

"what do you need extension services for?", we created a dummy variable "extension needs" which takes the value 1, if the household reports that its needs extension first and its needs support and the value 0, otherwise. Farmers who need to be supported are expected to have access to extension services, or at least exploring ways to have access to it. In the second-stage probit estimation, the access to extension services variable, and their generalized residuals predicted from the first-stage are included in the organization membership equation and estimated likelihood ratio (LR) can be estimated as follows:

$$LR = -2 \times (\ln Lp - (\ln L1 + \ln L0)), \tag{4.7}$$

where $\ln Lp$, $\ln L1$, and $\ln L0$ respectively denote the log-likelihood values for the pooled stochastic production frontier model, the members' production frontier and the non-members' production frontier. Where the null hypothesis is that members and non-members use the same rice production technology. For the estimation of the meta-frontier, the second stage environmental variables or industry-specific variables (that are supposed to impact the group-specific technology gap ratio) included in the four meta-frontier models are farm households specific characteristics (age, household size, agricultural equipment, and the distance to the main road) and variables for agro-ecological zones of Casamance and Delta. The choice of these industry-specific variables was based on the effects that they have on the membership decision (after modelling the membership in farmer organizations). The estimation of the conventional stochastic production frontier for both matched and unmatched samples was performed using the R software, while the PSM was conducted in STATA software and NLOGIT 6 was used to estimate the sample selection stochastic production frontier models.

4.3 Data and variables

4.3.1 Data sources

The data used for the analysis derived from a survey conducted in Senegal, which randomly sampled 4480 households that mainly produce dry cereals (or rainfed cereals). The survey was done under the Agricultural Policy Support Project (Projet d'Appui aux Politiques Agricoles, PAPA)[6], which is an initiative of the Government of Senegal funded by USAID-Senegal as part of the "Feed The Future" initiative, and implemented for a period of 3 years (2015 - 2018) by the Senegalese Ministry of Agriculture and Rural Facilities with technical support of the International Food Policy Research Institute (IFPRI). A multistage sampling procedure was applied for the selection of households and a structured household questionnaire was used to collect information. This questionnaire included several modules and gathered information on a range of topics such as household demographic and socioeconomic characteris-

[6] Official website of the project is `http://www.papa.gouv.sn/`.

tics, farmer organization membership, household assets, crop productions, livestock revenues, income and expenditures, access to infrastructures, access to institutions, commercialization, and production shocks and risk management strategies. Besides crop production and the inputs information, data collection also included market prices and households' adoption of agricultural technologies during the main agricultural season of 2016/2017. After the data cleaning and after removing observations with no information on the different outcomes variables, we retrieved a set the farmers who produced rice during the 2016/2017 season. The final sample comprised therefore 835 farmers located in four agro-ecological zones. Using geographical coordinates, the variables rainfall and percentages of clay in soil were retrieved from publicly available databases of the Climate Hazards Center of the University of California (`https://www.chc.ucsb.edu/data`) and of the International Soil Reference and Information Centre (`https://data.isric.org/`), respectively.

4.3.2 Variables descriptive statistics

Table 4.1 presents the definition and summary statistics of the variables used in the analysis. It also reports the comparative descriptive statistics of these variables based on farmers' organization membership status. Following the definition of Bernard *et al.* (2015), our variable of interest "organization" is referred to as membership in a rural producer organization that provides farmers with farming and farm-related services including access to inputs, markets and credit, collective sales, and capacities reinforcement. About 18% of the households in the sample have at least one person belonging to a farmers' organization. The households in the sample are predominantly male-headed, i.e. about 90%. The household heads have an average of 53 years and with about 47% having formal education. On average, the household includes more than nine family members and owns about 3.55 hectares of agricultural land. Regarding the production variables, the farmers produce on average 1407 kg of paddy rice. However, the standard deviation shows that there is a huge variation in the production output. To produce rice, farmers dedicate an average of 0.9 hectares, 52 kg of seeds, and 86 kg of fertilizers. Most of the farmers however do not use fertilizers (61.8%). In addition, around 232 equivalent working days are devoted to rice plots during the season.

When comparing members of farmer organizations to non-members, significant differences are observed between members and non-members mostly with the unmatched sample. Organization members tend to have larger households (10 persons) than non-members (9 persons). They possess more valued agricultural equipment compared to non-members. Moreover, they have better access to rural institutions (extension and improved seeds). Furthermore, they use more agricultural production inputs (land, labor, seeds, and fertilizers) and produce much more quantities of rice compared to non-members.

4.4 Results and discussion

4.4.1 Determinants of membership in farmers organizations

Factors that determine households' decision to belong to a farmer organization are presented in table 4.2 with their marginal effects. The likelihood ratio test shows that the model estimates are significant at 1% level ($\chi^2(14) = 264.79$; $p < 0.01$). The coefficient of the residual from the first-stage of the access to extension services variable is not statistically significant, suggesting that the access to extension services is not endogenously correlated to the household's decision to belong to a farmer organization.

The results of the estimation of equation (4.1) suggest that the main factors that have a significant influence on whether the rice producing farmer decides to be a member of a farmer-related organization are age of the household head, household size, value of agricultural equipment, distance to main road, and the agro-ecological locations of Casamance and Delta. The decision to belong to an organization is negatively and significantly correlated with the age of the household head. Households with younger heads are therefore better prone to join farmer organizations than older ones, with a 0.1% probability. These results are in contrast with the findings of Mojo *et al.* (2015) and Abdul-Rahaman and Abdulai (2018). The household size has a positive and significant effect on organization membership. These results support those of Bernard and Spielman (2009) and Ma and Abdulai (2016). Those households that have more members have higher probability (0.8%) to be members of farmer organizations. With more members, these households have a better chance that one of their members could belong to an organization.

Table 4.1: Description of variables

Variables	Description and measurement	Unmatched Sample: Pooled	Members	Non-Members	P-val.	Matched Sample: Pooled	Members	Non-Members	P-val.
Organization	Membership in Farmers Organization (1=yes, 0=no)	0.18 (0.39)				0.13 (0.34)			
	Household and Head characteristics								
Male	Head is male (1=yes, 0=no)	0.90 (0.30)	0.93 (0.26)	0.89 (0.31)	0.12	0.89 (0.31)	0.91 (0.29)	0.89 (0.31)	0.61
Age	Age of household head (years)	53.59 (12.73)	52.58 (12.32)	53.82 (12.81)	0.27	53.60 (12.88)	52.23 (13.29)	53.82 (12.81)	0.25
Household size	Number of family members	9.65 (5.20)	10.33 (5.05)	9.49 (5.22)	0.07	9.66 (5.27)	10.75 (5.46)	9.49 (5.22)	0.03
Migrant	Head is a migrant (1=yes, 0=no)	0.13 (0.34)	0.12 (0.33)	0.13 (0.34)	0.72	0.13 (0.34)	0.13 (0.34)	0.13 (0.34)	0.94
Education	Formal education (1=yes, 0=no)	0.47 (0.50)	0.42 (0.49)	0.48 (0.50)	0.19	0.48 (0.50)	0.48 (0.50)	0.48 (0.50)	0.93
	Assets & living conditions								
Land owned	Total land size owned (ha)	3.55 (5.75)	3.00 (4.89)	3.68 (5.92)	0.14	3.62 (5.83)	3.23 (5.24)	3.68 (5.92)	0.43
Equipment	Agricultural Equipment (m. FCFA)	0.07 (0.19)	0.16 (0.38)	0.06 (0.09)	<0.01	0.06 (0.09)	0.08 (0.11)	0.06 (0.09)	0.02
	Location								
Distance to road	Distance to nearest road (km)	11.14 (12.90)	17.16 (17.66)	9.79 (11.15)	<0.01	10.00 (11.50)	11.38 (13.48)	9.79 (11.15)	0.25
Distance to market	Distance to nearest market (km)	15.96 (12.80)	17.09 (15.15)	15.71 (12.21)	0.30	15.57 (12.43)	14.64 (13.80)	15.71 (12.21)	0.46
	Access to institutions								
Extension	Extension services (1=yes, 0=no)	0.20 (0.40)	0.61 (0.49)	0.10 (0.31)	<0.01	0.15 (0.36)	0.45 (0.50)	0.10 (0.31)	<0.01
Improved seeds	Use of improved seeds (1=yes, 0=no)	0.21 (0.41)	0.56 (0.50)	0.13 (0.34)	<0.01	0.17 (0.38)	0.43 (0.50)	0.13 (0.34)	<0.01
Extension needs	Need for support (1=yes, 0=no)	0.16 (0.36)	0.19 (0.39)	0.15 (0.36)	0.23	0.15 (0.36)	0.19 (0.39)	0.15 (0.36)	0.32
	Ecological conditions								
Casamance AEZ	Casamance zone (1=yes, 0=no)	0.76 (0.43)	0.41 (0.49)	0.84 (0.37)	<0.01	0.80 (0.40)	0.56 (0.50)	0.84 (0.37)	<0.01
Delta AEZ	Delta zone (1=yes, 0=no)	0.10 (0.30)	0.41 (0.49)	0.03 (0.16)	<0.01	0.05 (0.22)	0.19 (0.39)	0.03 (0.16)	<0.01
Rainfall	Annual rainfall 2016 (mm)	1046.28 (386.95)	673.73 (535.05)	1129.86 (285.12)	<0.01	1094.04 (338.53)	863.54 (519.38)	1129.86 (285.12)	<0.01
Clay	Percentage of clay (%)	27.59 (4.09)	26.49 (4.75)	27.84 (3.89)	<0.01	27.65 (4.11)	26.40 (5.20)	27.84 (3.89)	0.01
	Production inputs								
Land	Total area cultivated (ha)	0.90 (1.10)	1.09 (1.93)	0.86 (0.80)	0.15	0.86 (0.80)	0.89 (0.81)	0.86 (0.80)	0.73
Labor	Total labor size (work-days.)	232.12 (258.24)	244.40 (203.18)	229.37 (269.10)	0.44	229.70 (262.88)	231.81 (219.74)	229.37 (269.10)	0.92
Seeds	Total seeds (KG)	52.07 (62.96)	70.50 (96.08)	47.93 (51.98)	0.01	50.01 (59.00)	63.34 (91.44)	47.93 (51.98)	0.09
Fertilizers	Total fertilizers (KG)	84.79 (387.56)	278.16 (832.28)	41.41 (138.05)	<0.01	56.14 (155.36)	150.93 (216.41)	41.41 (138.05)	<0.01
No Fertilizers	Non use of fertilizers (1=yes, 0=no)	0.62 (0.49)	0.27 (0.45)	0.70 (0.46)	<0.01	0.65 (0.48)	0.39 (0.49)	0.70 (0.46)	<0.01
	Outcome variable								
Rice Production	Total crops productions (KG)	1406.87 (5745.67)	4153.39 (12950.56)	790.71 (959.23)	<0.01	988.39 (1824.11)	2260.27 (4133.39)	790.71 (959.23)	<0.01
N	Number of Observations	835	153	682		788	106	682	

The value of agricultural equipment has a positive and significant effect on the decision of farmer organization membership. One additional million FCFA (1600 USD) of equipment value increases the probability of rice-producing households to be members of farmer organizations by 36.7%. This result suggests that those rice-producing households with higher levels of welfare are more likely to belong to farmer organizations. The location of household and climatic conditions also have significant effects on their decision to be members of farmer organizations. Farmers who live closer to a main all-weather road are respectively more prone to participate in farm-related organization activities with a 0.3% of marginal effect for a less additional kilometer distance. Living in an agro-ecological zone such as the Delta region (with a marginal effect of 49%) also appears to be a very important factor for rice-producing households in their decision to belong to a farmer organization. These results suggest a clustering of farmer organization members, due to potential spatial non-observables factors such as climate, institutions, and infrastructure. These findings corroborate those of Abebaw and Haile (2013) and Ma and Abdulai (2016). The later author found in China that variables representing soil types and regions have significant cluster effects.

Table 4.2: Probit Estimates of Organizations Membership: Unmatched Sample

	Coefficients	Marginal Effects
Intercept	-0.942 (0.443)**	
Male	-0.023 (0.220)	-0.005 (0.047)
Age	-0.009 (0.005)*	-0.002 (0.001)*
Household size	0.038 (0.011)***	0.008 (0.002)***
Migrant	-0.026 (0.201)	-0.005 (0.041)
Education	0.058 (0.132)	0.012 (0.028)
Area owned	-0.006 (0.012)	-0.001 (0.003)
Equipment	1.740 (0.690)**	0.367 (0.146)**
Distance to road	-0.013 (0.007)*	-0.003 (0.001)*
Distance to market	-0.009 (0.006)	-0.002 (0.001)
Casamance AEZ	-0.325 (0.184)*	-0.075 (0.046)
Delta AEZ	1.508 (0.603)**	0.491 (0.221)**
Extension	1.183 (0.819)	0.342 (0.284)
Extension residuals	-0.085 (0.455)	
Log Likelihood	-265.284	-265.284
Num. obs.	835	835

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

4.4.2 Production frontiers estimates

The conventional and sample selection frontier models estimates are respectively presented in tables 4.3 and 4.4 for the original unmatched data, and in tables 4.5 and 4.6 for the matched data. From the estimation of the stochastic frontier models with the pooled data, the likelihood ratio (LR) test suggests that members and non-members of farmer organizations are employing heterogeneous technologies. The LR test rejected the null hypothesis of homogeneous technology between members of farmer organizations and non-members for both the unmatched data ($\chi^2(22) = 45.276$, $p < 0.01$) and for the matched data ($\chi^2(22) = 33.098$, $p < 0.1$) justifying the identification strategy of two separate production frontiers for members and non-members. Furthermore, in the pooled data estimation with both unmatched and matched data, the positive and significant effect of farmer organization membership dummy on the frontier estimates suggests that agricultural organizations membership contribute significantly to rice production in Senegal ($\chi^2(1) = 9.792$, $p < 0.01$). Abdul-Rahaman and Abdulai (2018) and Ma *et al.* (2018b) observed similar results in Ghana and China, respectively. These results can be explained by the fact that farmer-based organization members, in general, have better access to farm inputs and technologies through their social networks, and therefore could increase their productions.

For most of the estimated production frontier models, the inefficiency dispersion parameters σ_u are significant, suggesting that most of the farm households are producing below the production frontier. In addition, the terms σ_u in most of the models are much larger for the members of farmer organizations than non-members, suggesting that the members are more affected by inefficiency than non-members. Results from the sample selection production frontier model show that the estimated sample selectivity term ρ for members is negative for both the unmatched and matched data, but not statistically significant. For non-members, the estimated ρ is positive for both matched and unmatched data, and only statistically significant in the case of the unmatched data, indicating the presence of selectivity bias from unobserved factors. These results support the use of the sample selectivity production frontier framework (Greene, 2010).

Table 4.3: Conventional Estimates of Translog Production Frontier: Unmatched sample

	Pooled	Members	Non-Members	Metafrontier
Intercept	5.190 (0.673)***	3.854 (3.107)	5.024 (0.723)***	5.085 (0.217)***
Land	0.082 (0.208)	−0.741 (1.022)	0.195 (0.221)	0.082 (0.067)
Seeds	0.055 (0.214)	1.019 (0.747)	−0.079 (0.239)	0.107 (0.070)
Fertilizers	0.088 (0.188)	0.011 (0.888)	0.246 (0.207)	0.021 (0.060)
Labor	0.102 (0.116)	0.477 (0.338)	−0.000 (0.134)	0.112 (0.037)***
Land^2	−0.200 (0.042)***	−0.334 (0.285)	−0.199 (0.043)***	−0.196 (0.014)***
Seeds^2	0.064 (0.053)	0.064 (0.169)	0.050 (0.057)	0.056 (0.017)***
Fertilizers^2	0.150 (0.041)***	0.113 (0.164)	0.104 (0.052)**	0.171 (0.013)***
Labor^2	0.038 (0.019)**	0.165 (0.064)**	0.018 (0.021)	0.041 (0.006)***
Land×Seeds	0.083 (0.048)*	0.088 (0.159)	0.096 (0.051)*	0.078 (0.016)***
Land×Fertilizers	−0.011 (0.021)	0.048 (0.068)	−0.062 (0.027)**	−0.034 (0.007)***
Land×Labor	0.012 (0.022)	0.159 (0.127)	−0.014 (0.024)	0.015 (0.007)**
Seeds×Fertilizers	−0.023 (0.015)	−0.009 (0.047)	−0.015 (0.018)	−0.016 (0.005)***
Seeds×Labor	−0.020 (0.025)	−0.225 (0.106)**	0.019 (0.027)	−0.026 (0.008)***
Fertilizers×Labor	−0.044 (0.010)***	−0.053 (0.024)**	−0.050 (0.012)***	−0.038 (0.003)***
No Fertilizers	0.448 (0.396)	−0.309 (2.257)	0.682 (0.416)	0.533 (0.124)***
Improved Seeds	0.013 (0.090)	0.150 (0.200)	−0.123 (0.100)	0.109 (0.030)***
Rainfall	−0.001 (0.000)***	−0.001 (0.000)***	−0.000 (0.000)*	−0.001 (0.000)***
Clay	0.032 (0.008)***	0.016 (0.017)	0.032 (0.009)***	0.039 (0.003)***
Organization	0.283 (0.091)***			
Intercept				−23.875 (26.501)
Age				0.064 (0.076)
Household size				−0.329 (0.412)
Equipment				−16.378 (19.458)
Distance to road				0.509 (0.550)
Casamance AEZ				−43.814 (45.660)
Delta AEZ				−16.209 (17.871)
σ_u	0.898 (0.078)***	1.180 (0.134)***	0.654 (0.143)***	2.273 (1.278)*
σ_v	0.638 (0.038)***	0.468 (0.080)***	0.703 (0.050)***	0.230 (0.007)***
$\rho(w, v)$	–	–	–	–
Log-Likelihood	-1030.055	-187.379	-820.038	-70.097
Num. obs.	835	153	682	835

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 4.4: Conventional Estimates of Translog Production Frontier: Matched sample

	Pooled	Members	Non-Members	Metafrontier
Intercept	4.839 (0.682)***	−2.821 (4.773)	5.024 (0.723)***	5.002 (0.140)***
Land	0.122 (0.212)	−0.898 (1.155)	0.195 (0.221)	0.135 (0.042)***
Seeds	0.062 (0.221)	0.971 (0.692)	−0.079 (0.239)	0.033 (0.047)
Fertilizers	0.199 (0.199)	1.580 (1.484)	0.246 (0.207)	0.139 (0.041)***
Labor	0.084 (0.120)	0.561 (0.361)	−0.000 (0.134)	0.076 (0.024)***
Land^2	−0.195 (0.042)***	−0.399 (0.358)	−0.199 (0.043)***	−0.191 (0.008)***
Seeds^2	0.050 (0.054)	0.212 (0.191)	0.050 (0.057)	0.059 (0.011)***
Fertilizers^2	0.118 (0.048)**	−0.149 (0.282)	0.104 (0.052)**	0.137 (0.010)***
Labor^2	0.032 (0.020)*	0.193 (0.062)***	0.018 (0.021)	0.037 (0.004)***
Land×Seeds	0.094 (0.049)*	0.084 (0.189)	0.096 (0.051)*	0.089 (0.010)***
Land×Fertilizers	−0.047 (0.023)**	−0.046 (0.084)	−0.062 (0.027)**	−0.063 (0.005)***
Land×Labor	0.002 (0.023)	0.197 (0.125)	−0.014 (0.024)	0.005 (0.005)
Seeds×Fertilizers	−0.017 (0.016)	−0.013 (0.053)	−0.015 (0.018)	−0.013 (0.003)***
Seeds×Labor	−0.011 (0.025)	−0.303 (0.103)***	0.019 (0.027)	−0.015 (0.005)***
Fertilizers×Labor	−0.046 (0.010)***	−0.036 (0.026)	−0.050 (0.012)***	−0.044 (0.003)***
No Fertilizers	0.677 (0.413)	4.513 (3.812)	0.682 (0.416)	0.674 (0.080)***
Improved Seeds	0.004 (0.091)	0.468 (0.258)*	−0.123 (0.100)	0.079 (0.021)***
Rainfall	−0.001 (0.000)***	−0.001 (0.000)**	−0.000 (0.000)*	−0.001 (0.000)***
Clay	0.032 (0.009)***	0.043 (0.021)**	0.032 (0.009)***	0.039 (0.002)***
Organization	0.298 (0.094)***			
Intercept				−0.824 (0.852)
Age				−0.012 (0.010)
Household size				−0.006 (0.023)
Equipment				2.436 (1.288)*
Distance to road				0.035 (0.017)**
Casamance AEZ				−2.647 (1.459)*
Delta AEZ				−0.309 (0.398)
σ_u	0.759 (0.108)***	0.518 (0.705)	0.654 (0.143)***	0.534 (0.157)***
σ_v	0.681 (0.043)***	0.711 (0.190)***	0.703 (0.050)***	0.127 (0.005)***
$\rho(w, v)$	–	–	–	–
Log-Likelihood	-960.217	-123.630	-820.038	316.387
Num. obs.	788	106	682	788

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 4.5: Sample Selection Estimates of the Translog Production Frontier: Unmatched sample

	Members	Non-Members	Metafrontier
Intercept	3.355 (5.946)	5.003 (0.946)***	5.005 (0.256)***
Land	−1.004 (1.340)	0.199 (0.264)	0.079 (0.081)
Seeds	1.094 (1.151)	−0.070 (0.262)	0.130 (0.083)
Fertilizers	0.103 (1.546)	0.259 (0.290)	0.036 (0.069)
Labor	0.516 (0.619)	0.007 (0.168)	0.121 (0.044)***
Land^2	−0.397 (0.388)	−0.198 (0.056)***	−0.195 (0.016)***
Seeds^2	0.066 (0.236)	0.049 (0.059)	0.054 (0.020)***
Fertilizers^2	0.111 (0.287)	0.102 (0.066)	0.169 (0.015)***
Labor^2	0.164 (0.104)	0.017 (0.023)	0.041 (0.007)***
Land×Seeds	0.120 (0.227)	0.096 (0.057)*	0.079 (0.019)***
Land×Fertilizers	0.056 (0.086)	−0.053 (0.032)	−0.032 (0.008)***
Land×Labor	0.179 (0.175)	−0.013 (0.031)	0.016 (0.009)*
Seeds×Fertilizers	−0.024 (0.056)	−0.019 (0.017)	−0.018 (0.006)***
Seeds×Labor	−0.233 (0.184)	0.017 (0.032)	−0.029 (0.010)***
Fertilizers×Labor	−0.055 (0.048)	−0.050 (0.015)***	−0.037 (0.004)***
No Fertilizers	−0.173 (3.964)	0.647 (0.694)	0.560 (0.142)***
Improved Seeds	0.153 (0.294)	−0.152 (0.107)	0.118 (0.035)***
Rainfall	−0.001 (0.001)*	−0.000 (0.000)	−0.001 (0.000)***
Clay	0.021 (0.028)	0.030 (0.010)***	0.041 (0.003)***
Intercept			−7.140 (5.586)
Age			0.030 (0.041)
Household size			−0.138 (0.140)
Equipment			−7.632 (7.730)
Distance to road			0.185 (0.141)
Casamance AEZ			−15.882 (9.095)*
Delta AEZ			−6.121 (4.978)
σ_u	1.242 (0.151)***	0.729 (0.126)***	1.335 (0.491)***
σ_v	0.493 (0.123)***	0.694 (0.044)***	0.270 (0.008)***
$\rho(w, v)$	−0.188 (0.593)	0.449 (0.254)*	
Log-Likelihood			-189.444
Num. obs.	153	682	835

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 4.6: Sample Selection Estimates of the Translog Production Frontier: Matched sample

	Members	Non-Members	Metafrontier
Intercept	−0.969 (6.928)	5.015 (0.945)***	4.820 (0.248)***
Land	−0.976 (1.318)	0.196 (0.262)	0.132 (0.078)*
Seeds	0.992 (1.238)	−0.071 (0.263)	0.091 (0.082)
Fertilizers	1.256 (1.674)	0.258 (0.288)	0.134 (0.070)*
Labor	0.545 (0.720)	0.008 (0.167)	0.096 (0.043)**
Land^2	−0.451 (0.371)	−0.198 (0.055)***	−0.192 (0.015)***
Seeds^2	0.187 (0.289)	0.050 (0.059)	0.055 (0.019)***
Fertilizers^2	−0.082 (0.320)	0.102 (0.066)	0.151 (0.017)***
Labor^2	0.176 (0.116)	0.018 (0.023)	0.038 (0.007)***
Land×Seeds	0.101 (0.230)	0.097 (0.057)*	0.083 (0.018)***
Land×Fertilizers	−0.026 (0.095)	−0.054 (0.032)*	−0.067 (0.009)***
Land×Labor	0.192 (0.185)	−0.013 (0.030)	0.008 (0.009)
Seeds×Fertilizers	−0.026 (0.068)	−0.019 (0.017)	−0.013 (0.006)**
Seeds×Labor	−0.283 (0.176)	0.017 (0.031)	−0.023 (0.009)**
Fertilizers×Labor	−0.041 (0.051)	−0.050 (0.015)***	−0.044 (0.004)***
No Fertilizers	3.474 (4.406)	0.658 (0.689)	0.771 (0.148)***
Improved Seeds	0.428 (0.340)	−0.150 (0.107)	0.149 (0.035)***
Rainfall	−0.001 (0.001)	−0.000 (0.000)	−0.001 (0.000)***
Clay	0.035 (0.033)	0.030 (0.010)***	0.040 (0.003)***
Intercept			−0.835 (1.116)
Age			−0.007 (0.014)
Household size			−0.037 (0.042)
Equipment			2.064 (1.889)
Distance to road			0.044 (0.023)*
Casamance AEZ			−6.642 (5.322)
Delta AEZ			−0.634 (0.624)
σ_u	0.721 (0.468)	0.734 (0.129)***	0.606 (0.202)***
σ_v	0.707 (0.165)***	0.693 (0.044)***	0.269 (0.008)***
$\rho(w, v)$	−0.497 (0.445)	0.456 (0.293)	
Log-Likelihood			-147.160
Num. obs.	106	682	788

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

4.4.3 Predicted frontiers

Table 4.7 presents the means of the predicted frontiers for all models (pooled, members, and non-members) and the differences between the predicted frontiers of farmer organization members and those of non-members. Results of this table reveal that members have higher production frontiers than non-members and the differences are statistically significant. In the estimates with the matched data, being a member of a farmer organization increases the production of rice by around 10.0% in the conventional stochastic production frontier estimates and when selectivity bias is taken into account, the increase is about 19.5%. These figures confirm the previous results that membership in farmer organizations increases rice production. These results corroborate those observed by Abdul-Rahaman and Abdulai (2018) in the rice sector in Ghana, where the participation in farmers groups significantly enhances rice farming yield. Figure 4.2 in the appendix plots the kernel distribution of predicted frontiers.

Table 4.7: Predicted frontiers

SF Models	Pooled	Members	Non-Members	Difference
Unmatched				
Conventional	6.963 (1.079)	7.985 (1.170)	6.589 (0.968)	1.396***
Sample Selection	6.858 (1.155)	8.088 (1.140)	6.582 (0.963)	1.507***
Matched				
Conventional	6.775 (1.012)	7.251 (1.151)	6.589 (0.968)	0.662***
Sample Selection	6.760 (1.069)	7.874 (1.041)	6.587 (0.965)	1.286***

$^{***}p < 0.01$

4.4.4 Technical efficiency scores and meta-technology ratios

Tables 4.8, 4.9 and 4.10 present the means of technical efficiency scores (TE) , the meta-technology ratios (MTR), and the meta-frontier technical efficiency derived from the estimated different production frontiers (i.e pooled, groups-level frontiers and meta frontier models). Figures 4.3, 4.4 and 4.5 in appendix display respectively their kernel distributions. Considering the pooled data estimates with the unmatched sample, on average farmer organization members and non-members have similar mean technical efficiency scores. Members have a mean technical efficiency score of 55.62% (sd = 15.67%) [7] while those of non-members is 55.61% (sd = 14.22%),

[7] sd is the standard deviation

with no statistically differences observed between the two technical efficiency score ($t = 0.012, p = 0.990$). The non-statistical difference is also observed for the pooled matched sample ($t = -0.217, p = 0.829$).

When considering that members and non-members are operating with different technologies, the mean technical efficiency (TE) estimates for non-members, which varies from 57.4% to 63.6% are in the case of the unmatched sample, significantly higher than that of members (45.5% to 69.1%). In the matched sample, the difference in technical efficiency reduces, however, it not significant when we controlled for selection bias on unobservable. These results suggest that after controlling for biases arising from observable and unobservable differences between members and non-members of farmer organizations in the production frontiers, there is no difference in the performances of members and non-members within their own frontier. Therefore, one can conclude that considering the group-specific frontiers, membership in a farmer organization does not really affect farmers technical efficiency. These results are in contrast with those recently obtained by Abdul-Rahaman and Abdulai (2018) and Ma *et al.* (2018b), who stopped their analysis at this stage of our methodological framework and found that members in cooperatives are more technically efficient in their own frontiers than non-members. As stated previously, comparing farmers technical efficiencies from their own benchmark could bias the results. Technical efficiency estimates of organization members and non-members are measured against different production frontiers.

The results from the meta-frontier estimates show that the meta-technology ratios of members in most of the models (ranging from 89.7% to 93.6%) are significantly higher than those of non-members (ranging from 88.5% to 91.3%), suggesting that members of farmer organizations operate more closely to the meta-frontier than non-members. Therefore, one can conclude that membership in a farm-related organization affects strongly and positively the output of rice farming, confirming the previous result.

After combining the meta-technology ratios and the group-specific technical efficiencies, the obtained mean meta-frontier technical efficiencies estimates of the member groups varies between 41.4% (matched selectivity corrected) and 62.1% (matched conventional). These MTE estimates in most of the models for members are significantly lower than those of non-members. These results confirm some of the previous findings and mainly suggest that after correcting for selectivity bias and technology heterogeneity, belonging to a cooperative does not enhance farm efficiency.

Table 4.8: Levels of technical efficiency

SF Models	Pooled	Members	Non-Members	Difference
Unmatched				
Conventional	0.556 (0.145)	0.491 (0.193)	0.636 (0.101)	-0.145^{***}
Sample Selection	0.554 (0.144)	0.455 (0.187)	0.576 (0.121)	-0.122^{***}
Matched				
Conventional	0.599 (0.121)	0.691 (0.076)	0.636 (0.101)	0.055^{***}
Sample Selection	0.575 (0.123)	0.580 (0.126)	0.574 (0.122)	0.006

$^{***}p < 0.01$

Table 4.9: Levels of meta-technology ratios

SF Models	Pooled	Members	Non-Members	Difference
Unmatched				
Conventional	0.896 (0.092)	0.917 (0.047)	0.891 (0.099)	0.026^{***}
Sample Selection	0.890 (0.099)	0.913 (0.045)	0.885 (0.106)	0.028^{***}
Matched				
Conventional	0.904 (0.096)	0.897 (0.100)	0.905 (0.095)	-0.008
Sample Selection	0.917 (0.105)	0.936 (0.043)	0.913 (0.111)	0.022^{***}

$^{***}p < 0.01$

Table 4.10: Levels of meta-frontier technical efficiency

SF Models	Pooled	Members	Non-Members	Difference
Unmatched				
Conventional	0.545 (0.132)	0.450 (0.179)	0.567 (0.109)	-0.117^{***}
Sample Selection	0.492 (0.138)	0.414 (0.173)	0.509 (0.122)	-0.095^{***}
Matched				
Conventional	0.582 (0.109)	0.621 (0.099)	0.576 (0.109)	0.045^{***}
Sample Selection	0.526 (0.127)	0.543 (0.122)	0.524 (0.127)	0.019

$^{***}p < 0.01$

4.5 Conclusion

Farmers in developing countries are characterized by remarkably low levels of productivity and efficiency, mainly due to the lack of access to inputs and improved technologies. Therefore, farmers collective action groups can constitute the vehicle for access to farm inputs and therefore enhance farm productivity. However, despite the growing literature on the importance of cooperative-like organizations in developing countries, very few studies have investigated the impact that farmer organizations can have on farm households' technical efficiency. This paper aimed to fill in the gap by evaluating the quantitative effects of membership in farmer organizations on technical efficiency of rice-producing households in Senegal, where access to modern technologies, productivity and efficiency in the rice sector are crucial issues.

Applying an econometric framework that combines a propensity score matching (PSM) method with the selection corrected stochastic production frontier model and a meta-frontier approach, on a cross-sectional data of 835 individuals, we derived for two groups of farmers (members and non-members of farmer organizations) their group-specific technical efficiency scores, the meta-technology ratios and the meta-frontier technical efficiency. The PSM method enables us to match organization members with non-members, addressing the biases from observed variables. With the selectivity-corrected stochastic production frontier model, the biases arising from unobserved factors were controlled. The meta-frontier approach helps to compare the technical efficiency score of both groups.

Estimation results confirmed that selection bias was present, and the two groups are using two different technologies for rice production, therefore justifying the combined framework that we used. The analysis shows that belonging to a farmer organization affects positively and significantly the production of rice in Senegal, confirming the importance of cooperative-like organizations in developing countries and their roles in enhancing farms productions. However, non-members of farmer organizations seem to be technically more efficient than members when each group operate in its own frontier, contradicting recent studies. The rest of the analysis shows that members have higher meta-technology ratios, meaning that they are operating much closer to the meta-frontier than non-members. At the meta-frontier, significant differences in technical efficiency are observed between members and non-members, with non-members appearing to be more efficient.

These results have some policy implications. Farmer organizations are still good policy instrument to enhance farm productions in developing countries, by easing farm inputs and modern technologies access. However, not all farmers benefit from being members, as shown in the group-specific inefficiency scores. Therefore, policymakers could exploit social networks structures of farmer organizations to enable farmers to have better access to technical knowledge in order to increase farm productivity and efficiency, not only for members but also for non-members through their "natural" social networks (family, religion, geographic, etc). Further research could also investigate the spillovers effects of membership on non-members productivity and efficiency.

References

Abate, G. T., Francesconi, G. N. and Getnet, K. (2014) Impact of agricultural cooperatives on smallholders' technical efficiency: empirical evidence from Ethiopia, *Annals of Public and Cooperative Economics*, **85**, 257–286.

Abdul-Rahaman, A. and Abdulai, A. (2018) Do farmer groups impact on farm yield and efficiency of smallholder farmers? Evidence from rice farmers in northern Ghana, *Food Policy*, **81**, 95 – 105.

Abebaw, D. and Haile, M. G. (2013) The impact of cooperatives on agricultural technology adoption: Empirical evidence from Ethiopia, *Food Policy*, **38**, 82 – 91.

Addai, K. N., Owusu, V. and Danso-Abbeam, G. (2014) Effects of farmer-based-organization on the technical efficiency of maize farmers across various agro-ecological zones of Ghana, *Journal of Economics*, **2**, 141–161.

Ahmed, M. H. and Mesfin, H. M. (2017) The impact of agricultural cooperatives membership on the wellbeing of smallholder farmers: empirical evidence from eastern Ethiopia, *Agricultural and Food Economics*, **5**, 1–20.

Aigner, D., Lovell, C. and Schmidt, P. (1977) Formulation and estimation of stochastic frontier production function models, *Journal of Econometrics*, **6**, 21 – 37.

Balasubramanian, V., Sie, M., Hijmans, R. and Otsuka, K. (2007) Increasing rice production in Sub-Saharan Africa: Challenges and opportunities, in *Advances in Agronomy* (Ed.) D. L. Sparks, Academic Press, vol. 94 of *Advances in Agronomy*, pp. 55 – 133.

Barham, J. and Chitemi, C. (2009) Collective action initiatives to improve marketing performance: Lessons from farmer groups in Tanzania, *Food Policy*, **34**, 53 – 59.

Battese, G. E. (1997) A note on the estimation of Cobb-Douglas production functions when some explanatory variables have zero values, *Journal of Agricultural Economics*, **48**, 250–252.

Bernard, T., Frölich, M., Landmann, A., Unte, P. N., Viceisza, A. and Wouterse, F. (2015) Building trust in rural producer organizations in Senegal: Results from a randomized controlled trial, Tech. rep., IZA Discussion Papers.

Bernard, T. and Spielman, D. J. (2009) Reaching the rural poor through rural producer organizations? A study of agricultural marketing cooperatives in Ethiopia, *Food Policy*, **34**, 60 – 69.

Bernard, T., Taffesse, A. S. and Gabre-Madhin, E. (2008) Impact of cooperatives on smallholders' commercialization behavior: evidence from Ethiopia, *Agricultural Economics*, **39**, 147–161.

Bravo-Ureta, B. E., Greene, W. and Solís, D. (2012) Technical efficiency analysis correcting for biases from observed and unobserved variables: an application to a natural resource management project, *Empirical Economics*, **43**, 55–72.

Chagwiza, C., Muradian, R. and Ruben, R. (2016) Cooperative membership and dairy performance among smallholders in Ethiopia, *Food Policy*, **59**, 165–173.

Diagne, M., Demont, M., Seck, P. A. and Diaw, A. (2013) Self-sufficiency policy and irrigated rice productivity in the Senegal River Valley, *Food Security*, **5**, 55–68.

FAO (2019) FAOSTAT, Food and agriculture data.

Fischer, E. and Qaim, M. (2012a) Gender, agricultural commercialization, and collective action in Kenya, *Food Security*, **4**, 441–453.

Fischer, E. and Qaim, M. (2012b) Linking smallholders to markets: Determinants and impacts of farmer collective action in Kenya, *World Development*, **40**, 1255–1268.

Francesconi, G. N. and Heerink, N. (2010) Ethiopian agricultural cooperatives in an era of global commodity exchange: Does organisational form matter?, *Journal of African Economies*, **20**, 153–177.

Francesconi, G. N. and Ruben, R. (2012) The hidden impact of cooperative membership on quality management: A case study from the Dairy Belt of Addis Ababa, *Journal of Entrepreneurial and Organizational Diversity*, **1**, 85–103.

González-Flores, M., Bravo-Ureta, B. E., Solís, D. and Winters, P. (2014) The impact of high value markets on smallholder productivity in the Ecuadorean Sierra: A stochastic production frontier approach correcting for selectivity bias, *Food Policy*, **44**, 237 – 247.

Greene, W. (2010) A stochastic frontier model with correction for sample selection, *Journal of Productivity Analysis*, **34**, 15–24.

Henningsen, A., Mpeta, D. F., Adem, A. S., Kuzilwa, J. A. and Czekaj, T. G. (2015) A meta-frontier approach for causal inference in productivity analysis: The effect of contract farming on sunflower productivity in Tanzania, 2015 AAEA & WAEA Joint Annual Meeting, July 26-28, San Francisco, California 206200, Agricultural and Applied Economics Association.

Huang, C. J., Huang, T.-H. and Liu, N.-H. (2014) A new approach to estimating the metafrontier production function based on a stochastic frontier framework, *Journal of Productivity Analysis*, **42**, 241–254.

Ito, J., Bao, Z. and Su, Q. (2012) Distributional effects of agricultural cooperatives in China: Exclusion of smallholders and potential gains on participation, *Food Policy*, **37**, 700–709.

Jelliffe, J. L., Bravo-Ureta, B. E., Deom, C. M. and Okello, D. K. (2018) Adoption of High-Yielding Groundnut Varieties: The Sustainability of a Farmer-Led Multiplication-Dissemination Program in Eastern Uganda, *Sustainability*, **10**, 1–21.

Jondrow, J., Lovell, C. K., Materov, I. S. and Schmidt, P. (1982) On the estimation of technical inefficiency in the stochastic frontier production function model, *Journal of Econometrics*, **19**, 233 – 238.

Kumbhakar, S. C., Tsionas, E. G. and Sipiläinen, T. (2009) Joint estimation of technology choice and technical efficiency: an application to organic and conventional dairy farming, *Journal of Productivity Analysis*, **31**, 151–161.

Ma, W. and Abdulai, A. (2016) Does cooperativemembership improve household welfare? Evidence from apple farmers in China, *Food Policy*, **58**, 94–102.

Ma, W., Abdulai, A. and Goetz, R. (2018a) Agricultural Cooperatives and Investment in Organic Soil Amendments and Chemical Fertilizer in China, *American Journal of Agricultural Economics*, **100**, 502–520.

Ma, W., Renwick, A., Yuan, P. and Ratna, N. (2018b) Agricultural cooperative membership and technical efficiency of apple farmers in China: An analysis accounting for selectivity bias, *Food Policy*, **81**, 122 – 132.

Mayen, C. D., Balagtas, J. V. and Alexander, C. E. (2010) Technology adoption and technical efficiency: Organic and conventional dairy farms in the united states, *American Journal of Agricultural Economics*, **92**, 181–195.

Mishra, A. K., Kumar, A., Joshi, P. K. and D'Souza, A. (2018) Cooperatives, contract farming, and farm size: The case of tomato producers in Nepal, *Agribusiness*, **34**, 865–886.

Mojo, D., Fischer, C. and Degefa, T. (2015) Social and environmental impacts of agricultural cooperatives: evidence from Ethiopia, *International Journal of Sustainable Development & World Ecology*, **22**, 388–400.

Mojo, D., Fischer, C. and Degefa, T. (2017) The determinants and economic impacts of coffee farmer cooperatives: Recent evidence from rural Ethiopia, *Journal of Rural Studies*, **50**, 84–94.

Rosenbaum, P. R. and Rubin, D. B. (1983) The central role of the propensity score in observational studies for causal effects, *Biometrika*, **70**, 41–55.

République du Sénégal (2009) Programme National d'Autosuffisance en Riz, Tech. rep., Ministère de l'Agriculture.

Salifu, A., Francesconi, G. N. and Kolavalli, S. (2010) A review of collective action in rural Ghana, Tech. Rep. 998, International Food Policy Research Institute (IFPRI).

Seck, P. A., Tollens, E., Wopereis, M. C., Diagne, A. and Bamba, I. (2010) Rising trends and variability of rice prices: Threats and opportunities for sub-Saharan Africa, *Food Policy*, **35**, 403 – 411.

Soullier, G. and Moustier, P. (2018) Impacts of contract farming in domestic grain chains on farmer income and food insecurity. Contrasted evidence from Senegal, *Food Policy*, **79**, 179 – 198.

Tolno, E., Kobayashi, H., Ichizen, M., Esham, M. and Balde, B. S. (2015) Economic analysis of the role of farmer organizations in enhancing smallholder potato farmers' income in Middle Guinea, *Journal of Agricultural Science*, **7**, 123–137.

Verhofstadt, E. and Maertens, M. (2015) Can Agricultural Cooperatives Reduce Poverty? Heterogeneous Impact of Cooperative Membership on Farmers' Welfare in Rwanda, *Applied Economic Perspectives and Policy*, **37**, 86–106.

Villano, R., Bravo-Ureta, B., Solís, D. and Fleming, E. (2015) Modern rice technologies and productivity in the philippines: Disentangling technology from managerial gaps, *Journal of Agricultural Economics*, **66**, 129–154.

Wollni, M. and Zeller, M. (2007) Do farmers benefit from participating in specialty markets and cooperatives? the case of coffee marketing in Costa Rica, *Agricultural Economics*, **37**, 243–248.

Wooldridge, J. (2015) Control function methods in applied econometrics, *Journal of Human Resources*, **50**, 420–445.

World Bank (2007) *World Development Report 2008 : Agriculture for Development*, World Bank Washington, DC.

Appendix

Table 4.11: Addressing potential endogeneity in extension variable

	Organization Membership	Extension
Intercept	-0.566 (0.353)	-0.748 (0.343)**
Sex	-0.077 (0.206)	-0.080 (0.192)
Age	-0.010 (0.005)**	-0.004 (0.005)
Household size	0.039 (0.011)***	0.018 (0.011)*
Migrant	0.126 (0.170)	0.326 (0.156)**
Education	0.074 (0.126)	0.020 (0.123)
Area owned	-0.009 (0.012)	-0.016 (0.014)
Equipment	2.154 (0.540)***	1.659 (0.545)***
Distance to road	-0.014 (0.006)**	-0.012 (0.006)**
Distance to market	-0.009 (0.005)*	-0.005 (0.005)
Extension Needs	0.111 (0.159)	0.561 (0.142)***
Casamance AEZ	-0.451 (0.151)***	-0.437 (0.150)***
Delta AEZ	2.041 (0.284)***	1.854 (0.277)***
Log Likelihood	-291.364	-313.185
Num. obs.	835	835

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 4.12: Propensity score matching quality test

	Before Matching	After Matching
Pseudo R2	0.333	0.030
LR χ^2	264.79	8.70
P-value ($p > \chi^2$)	0.000	0.795
Mean standardized bias	42.3	8.9

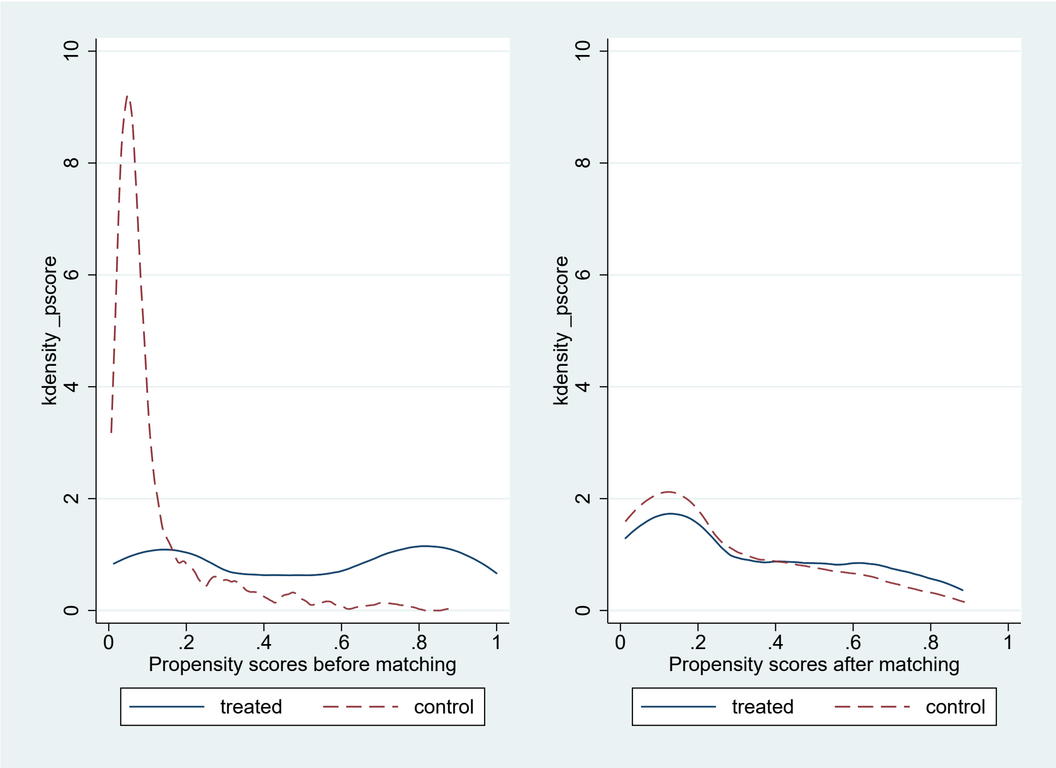


Figure 4.1: Kernel density of propensity scores

Table 4.13: Estimates of the Cobb-Douglas and translog production frontiers

	Cobb-Douglas	Translog
Intercept	6.583 (0.233)***	6.049 (0.585)***
Land	0.703 (0.043)***	0.179 (0.207)
Seeds	0.033 (0.041)	-0.088 (0.217)
Fertilizers	0.174 (0.015)***	-0.254 (0.092)***
Labor	0.058 (0.026)**	0.119 (0.119)
Land^2		-0.192 (0.041)***
Seeds^2		0.098 (0.055)*
Fertilizers^2		0.266 (0.023)***
Labor^2		0.024 (0.020)
Land×Seeds		0.062 (0.049)
Land×Fertilizers		-0.063 (0.020)***
Land×Labor		0.011 (0.023)
Seeds×Fertilizers		-0.014 (0.015)
Seeds×Labor		-0.013 (0.025)
Fertilizers×Labor		-0.047 (0.010)***
σ_u	0.846 (0.148)***	0.926 (0.079)***
σ_v	0.817 (0.056)***	0.658 (0.037)***
Log-Likelihood	-1152.053	-1055.849
Num. obs.	835	835

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

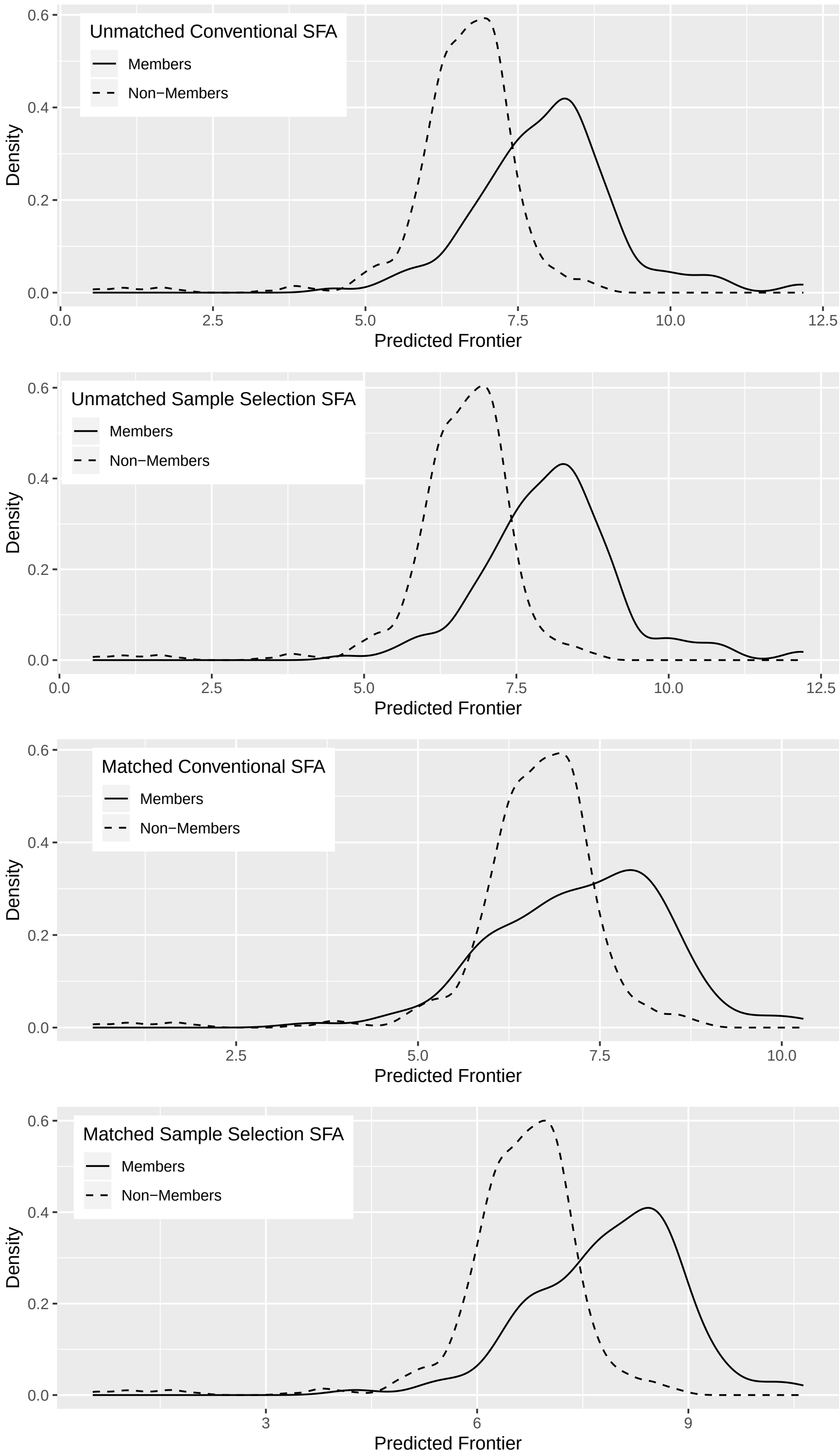


Figure 4.2: Kernel distributions of predicted frontiers

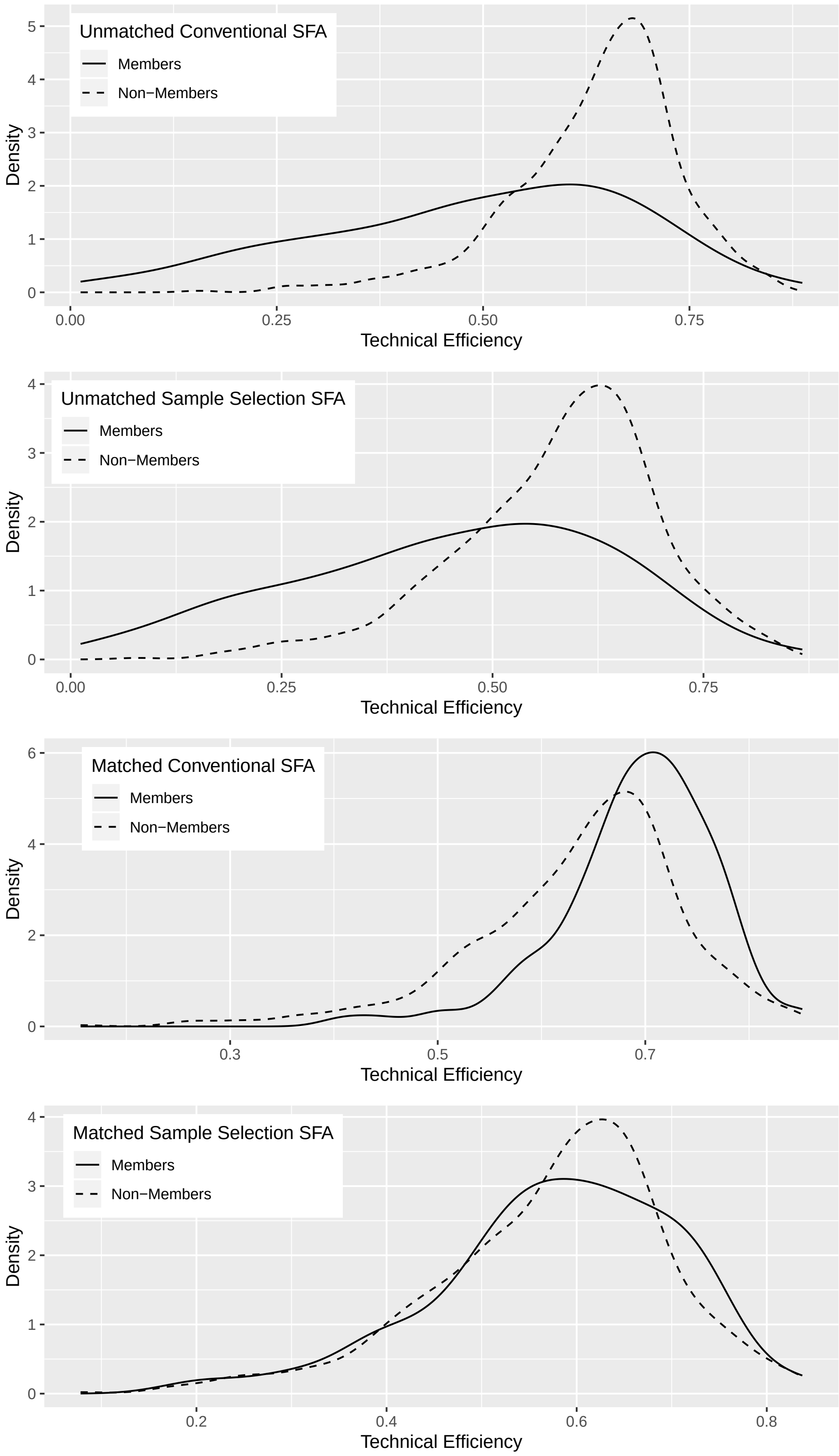


Figure 4.3: Kernel distributions of estimated efficiency scores

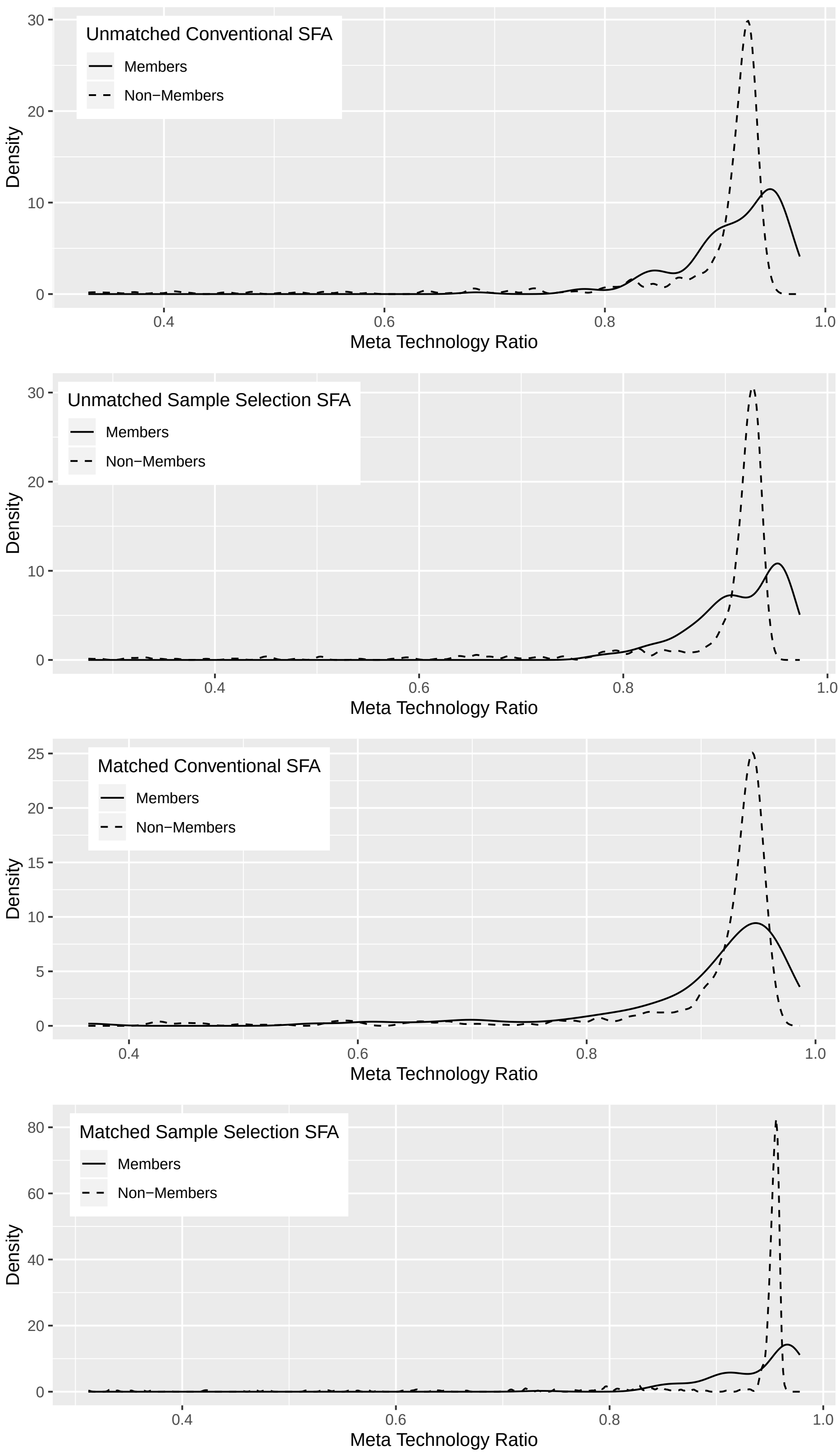


Figure 4.4: Kernel distributions of estimated meta-technology ratios

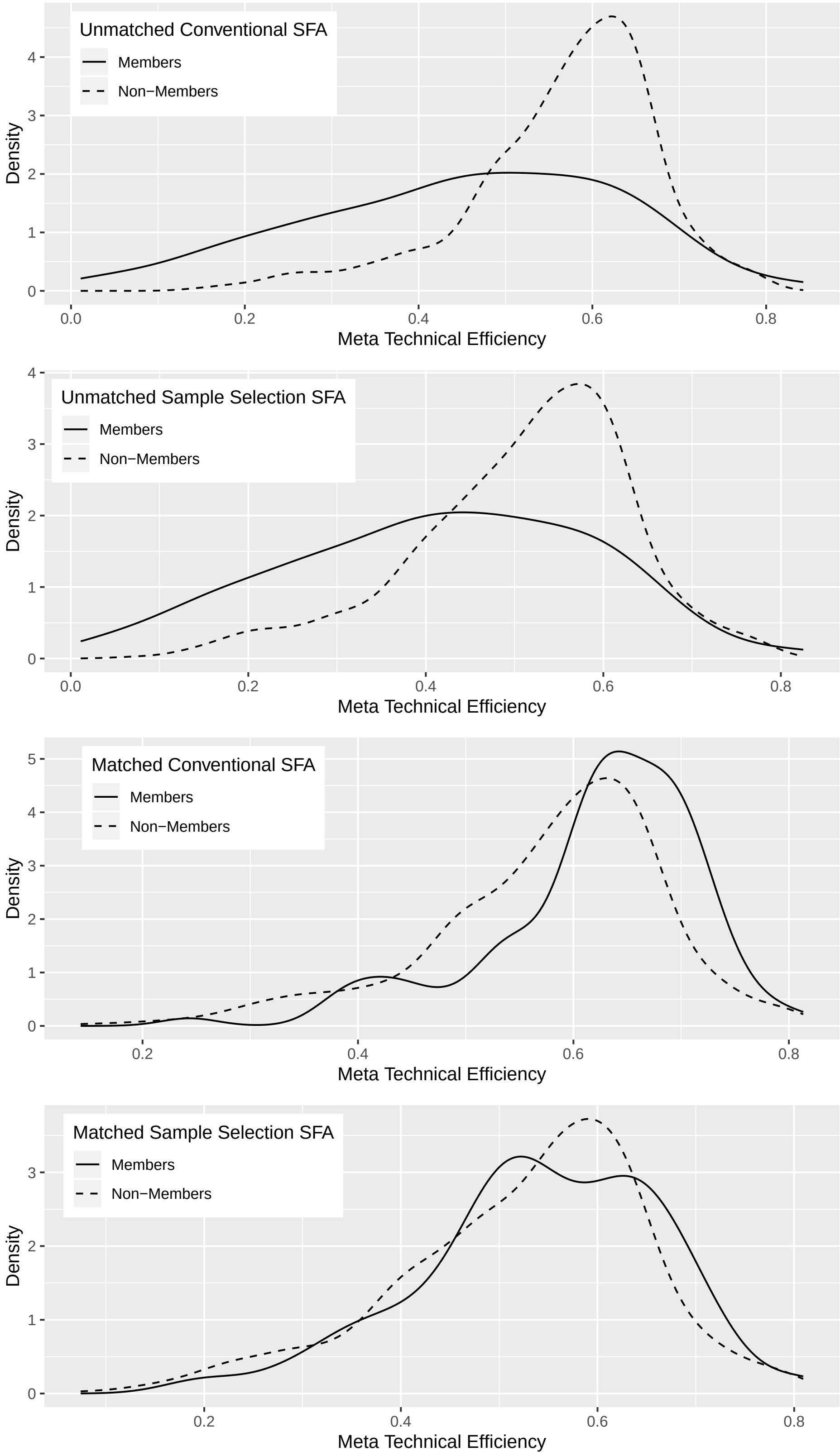


Figure 4.5: Kernel distributions of estimated meta-technology efficiency scores

Chapter 5

Farmer organizations or simple neighbourhood? In the pursuit of ways to push technology adoption in Senegal

K. Christophe Adjin, Peron A. Collins-Sowah, Christian H. C. A. Henning

Abstract

It is commonly accepted that social learning constitutes a relevant component of agricultural technology adoption, and techniques to incorporate social interactions into the analysis of farmers behaviours have greatly evolved during recent years. This paper uses a national and household-level data, to analyse the effects of rural producers organizations membership and neighbourhood, on the adoption of two productivity-enhancing technologies (improved seeds and inorganic fertilizers), among dry cereals farmers in Senegal. After applying a Bayesian Spatial Durbin Probit model, the results reveal that close neighbouring farmers show similar choice behaviour regarding productivity-enhancing technologies. Moreover, we find that membership in farmer-based organizations affects significantly and positively not only the choice of farmers who are members of such organization but also their neighbours choice to use productivity-enhancing technologies. The results imply important policy recommendations for Senegal, mainly when searching for ways to increase the diffusion of agricultural technologies.

Keywords: Farmer organizations, spatial dependence, technology adoption, Senegal

JEL Codes: Q13, C21, Q16, C11.

5.1 Introduction

Agricultural productivity can play a strong role in driving structural change (McArthur and McCord, 2017), and at the same time, increasing farm productivity represents a important channel of improving living conditions in rural areas. In the same vein, Barrett *et al.* (2017) argued that structural change in agriculture and in rural African economies generally is fundamental to end extreme poverty. Although significant progress has been observed in some Sub-Saharan African countries, on average, the agriculture sector productivity growth has remained low (Barrett *et al.*, 2017).

Such a situation particularly prevails in Senegal, where due to inadequate access to improved technologies by smallholder farmers, the agricultural sector has stayed for decades at the subsistence level. For instance, figures from 2015, show that Senegal was far below the Sub-Saharan average of cereals yield and agricultural value-added per worker (Hathie *et al.*, 2017). At the same time, several empirical studies have demonstrated that, in sub-Saharan Africa, adoption combined with adequate use of productivity-enhancing technologies (PET) such as high yielding seeds or inorganic fertilizers can significantly contribute to fighting poverty, by increasing and improving rural households welfare and income (Minten and Barrett, 2008; Cunguara and Darnhofer, 2011; Kabunga *et al.*, 2014; Khonje *et al.*, 2015; Verkaart *et al.*, 2017; Ahmed *et al.*, 2017; Abdoulaye *et al.*, 2018), farms productivity and profitability (Duflo *et al.*, 2008; Marenya and Barrett, 2009; Koussoubé and Nauges, 2017), farm production quality (Wopereis-Pura *et al.*, 2002), and household food security (Kassie *et al.*, 2014; Katungi *et al.*, 2018).

Several studies, using different approaches or paradigms, have analysed the factors that drive technology adoption in a number of developing contexts. Many authors have investigated agricultural technology adoption and linked it to farmers' risk preferences (Feder, 1980; Feder *et al.*, 1985; Knight *et al.*, 2003; Liu, 2013; Barham *et al.*, 2014) or behavioural attitudes (Duflo *et al.*, 2011). It is then argued that farmers do not adopt PETs because of their reluctance to take risks and invest in these technologies. Some authors have also pointed out that technology adoption is hampered by liquidity and credit constraints (Giné and Klonner, 2008; Abdulai *et al.*, 2008; Karlan *et al.*, 2014). At the same time, farmers might not adopt beneficial technologies due to their lack of information about the profitability of the technology and how to use it (Abdulai *et al.*, 2008; Matuschke and Qaim, 2008; Kabunga *et al.*, 2012). Simultaneously, various studies have shown the importance of learning and especially learning from others or peers in technology adoption (Foster and Rosenzweig, 1995; Bandiera and Rasul, 2006; Conley and Udry, 2010).

The process of learning from others, generally referred to in the literature, as social learning, highlight the relevance of social interactions in agricultural technologies adoption. Through social interactions with peers, that is, through their social networks and/or their neighbourhoods, farmers might acquire several relevant information about a technology and therefore, narrow the knowledge gaps that were preventing them from using particular technologies (Munshi, 2004; Matuschke and Qaim, 2009; Van den Broeck and Dercon, 2011; Magnan *et al.*, 2015). Although social network and peers or neighbours effects in agricultural technology adoption modelling is well known, few studies have really focused on the role of neighbourhood effects in technology adoption in developing countries, especially regarding the adoption of productivity-enhancing technologies. In particular, the neighbourhood offers opportunities for farmers to communicate and share information and experiences about technologies. In addition, farmers who live in close proximity might also exhibit similar technology adoption behaviour due to unobserved characteristics or structural climatic conditions (weather, soil quality).

Since the first study of Case (1992), there is a growing empirical literature, that account for neighbourhood effects on agricultural technology adoption (Holloway *et al.*, 2002; Wollni and Andersson, 2014; Läpple and Kelley, 2015; Läpple *et al.*, 2017). However, besides the studies of Langyintuo and Mekuria (2008); Krishnan and Patnam (2013); Tessema *et al.* (2016) and Fang and Richards (2018), little attention has been paid to the African context so far. Therefore, we use a nationality representative household survey in Senegal to explore the influence of neighbourhood in a setting of productivity-enhancing technologies adoption. Furthermore, in most developing countries, besides traditional channels of diffusion of technologies such as extension services, farmer organizations also represent a relevant source of information and knowledge transfer. In addition, farmer organizations ease the access to agricultural technologies. Previous empirical studies show that farmers collective action groups significantly improve technology adoption levels (Abebaw and Haile, 2013; Ma *et al.*, 2018). Therefore, this study put the emphasis on these two main potential sources (levers) of technology diffusion: the farmer's neighbourhood and his/her partial social network proxied by membership in farmer organizations.

This paper contributes to the literature in two ways: first, the combined effect of farmer organization membership and spatial effects has not been really scrutinized. Secondly, we use a national level data to take into account the entire Senegalese agriculture sector to examine this combined effect. The remainder of the paper is structured as follows. Section 5.2 presents the conceptual framework. In section 5.3, the theoretical framework is presented. Section 5.4 outlines the econometric framework and the estimation strategy. Section 5.5 describes the data and variables used. Estimation results are presented in Section 5.6 and the final section contains conclusions and some policy implications.

5.2 Concepts and context

Productivity-enhancing technologies (PETs) refer generally to technologies that increase crop yield. In the context of Senegal, most PETs include improved seeds or high yield crop varieties, inorganic fertilizers, pesticides and irrigation. A number of studies have demonstrated the importance of the adoption of PETs in the improvement of rural household welfare in the developing world. For instance, in rural Madagascar, Minten and Barrett (2008) results show that communities, which have adopted improved agricultural technologies, exhibit better welfare indicators. Cunguara and Darnhofer (2011) observed a similar positive significant impact of the adoption of improved technologies such as improved maize seeds, improved granaries, tractor mechanization, and animal traction on the income of Mozambican rural households. Meanwhile, in Burkina Faso, Koussoubé and Nauges (2017) found that fertilizer use increases farm profitability by forty percentage points. Despite the proven benefits of these technologies, many regions in Africa still exhibits low uptakes rates. A lot of empirical studies have tried to investigate reasons for such low uptakes rates.

Heterogeneity in agricultural technology adoption has been primarily attributed to differences in risk attitudes and uncertainties, credit constraints, and knowledge gaps (see Feder (1980); Knight *et al.* (2003); Liu (2013); Barham *et al.* (2014); Giné and Klonner (2008); Abdulai *et al.* (2008), and Karlan *et al.* (2014)). However, since the 1990s, a growing number of studies have extended previous models of technology adoption by incorporating the potential effects that farmers' peers could have on one's adoption behaviour. The work of Foster and Rosenzweig (1995) demonstrated the relevance of social learning in India. They observed that farmers, who have adopted high-yielding seed varieties, are more profitable when they have experienced neighbours. Bandiera and Rasul (2006) empirical results showed that the size of the social network of sunflower adopters positively affected the individual farmer's likelihood to adopt sunflower. Conley and Udry (2010), after defining for each farmer an information neighbourhood, demonstrated that in Ghana, farmers in the pineapple sector adjust their inputs use with those of their information neighbours, indicating a process of social learning. The study of Matuschke and Qaim (2009) also showed the important role of social networks on the adoption of modern seed technologies (hybrid seeds) among smallholder farmers in India, after considering a comprehensive data on farmer social interactions.

In the same line, several other empirical studies using similar approaches have pointed out the importance of including peers, neighbourhood effects or social influence in agricultural technology adoption (see for instance Isham (2002) and Van den Broeck and Dercon (2011)). Therefore, social peers effects in technology adoption modelling are well known. However, according to Manski (1993), a common identification challenge arises when predicting separately a farmer adoption choice from his/her peers' ones. Manski (1993) distinctively defined three types of peers influences that could affect an individual's outcome (e.g. in our case the adoption of PET) in a peers' group (e.g. farmers' social network): the endogenous effect, exogenous or contextual effects and the correlated effects. The first type of effect expresses the influence of peers' outcomes, the second type is related to the influence of exogenous peers' characteristics, and the third type refers to the tendency of a same group's individuals to behave similarly because they are alike or face the same environment. For Manski (1993) and from his "reflection problem", the separate identification of the three peers effects, is unfeasible if the researcher does not have prior information on the reference groups.

To address such difficulty, costly dynamic data or randomization approaches have been suggested by Foster and Rosenzweig (2010). Meanwhile, taking advantage of the development of spatial econometrics estimation techniques over the past decades, some authors such as Case (1992) and Holloway *et al.* (2002) have successfully identified the neighbourhood effects in technology adoption modelling. Bramoullé *et al.* (2009) later proposed a generalized approach.

In the developing world, Wollni and Andersson (2014) find that the adoption of organic agriculture in Honduras is strongly influenced by the availability of information in farmers' neighbourhood networks and social conformity. In Ethiopia, using farmers' spatial networks and panel data, Krishnan and Patnam (2013) investigated the impact of learning from extension services to the learning process from neighbours in the case of adoption of two modern technologies (improved seeds and fertilizer) and found that compared to extension, the effect of the learning from neighbours lasts over the years (over time). More recently, the study of Tessema *et al.* (2016)

revealed that neighbourhood effect is a significant driver of conservation tillage technology adoption in the northwest of Ethiopia. Langyintuo and Mekuria (2008) and Fang and Richards (2018) using spatial econometrics approaches, also found significant neighbourhood effects in the adoption of improved maize varieties in rural Mozambique.

This literature review highlight first, the importance of productivity-enhancing technologies in the fight against poverty and food insecurity, and their low levels of uptake in Africa. This study, therefore, put the emphasis on these two technologies, for which, policymakers are still searching for ways to increase their adoption rates. Second, it also points out the relevance of incorporating social learning in the adoption of agricultural technologies. However, besides spatial or social networks, it is also important to account for the specific social ties that could derive from the membership in various rural organizations. In this study, we consider social learning in two different and combined situations, the social learning that is derived from the farmers' peers due to neighbourhood and the one that comes from membership in farmer organizations. Third, the literature reveals a constant development of estimation approaches in order to meet the challenge of the reflection problem. Therefore, our study uses recent spatial econometrics, which is more suited for cross-sectional data.

5.3 Theoretical framework

Our theoretical framework is drawn from the works of Abdulai *et al.* (2008), Wollni and Andersson (2014) and Läpple and Kelley (2015). Productivity enhancing technology (PET) adoption is modelled, as an investment decision, within the random utility framework. It is assumed that a household chooses to adopt PET if the expected utility gained from using (U_A) the technology is larger than the utility from not using it (U_{NA}). A utility maximizing household therefore would adopt PET if the expected net utility is greater than zero ($U_A - U_{NA} > 0$). This expected utility, that accounts for several factors, can be defined as follows:

$$E\left[U^A\left(\pi^A, T, I^A, S\right)\right] - E\left[U^{NA}\left(\pi^{NA}\right)\right] > 0, \tag{5.1}$$

with

$$\pi^K = P^K.Q^K\left(F^K, E, I^K\right) - W^K.F^K, \tag{5.2}$$

where π^K $(K = A, NA)$ is the household profit either for adopters A and non-adopters NA; T represents the additional investments costs necessary for the use of the technology; I^K is the information availability related to the used technology, S is a deviation from the social norm; P^K are output prices, Q^K is the quantity produced, which is a function of input quantities used F^K, agro-ecological and other spatially correlated economic factors E; and W^K are inputs prices.

Assuming then that farmers adoption choice is affected by their neighbourhood, therefore a household i adoption utility U_i is affected by its neighbour's utility U_j (with $i \neq j$). Such influence could happen through several channels. First, farmers facing uncertainty may strategically be in search of information. According to Foster and Rosenzweig (2010), the limited information about the profitability of the technology could impede its uptake. Before investing additional substantial financing in PET, households could increase their information level about the technology profitability by updating their beliefs within their neighbourhood and/or their membership in farmer organizations. Such interaction would therefore result in information spillovers I^K that in turn can affect the transaction costs and productivity. Information spillovers are generally the results of contact with more experienced or knowledgeable farmers, early adopters or economically successful farmers. Secondly, neighbourhood effects could also arise from neighbours' characteristics. Farmers may adopt PET, for the simple reason that their neighbours are members of farmer organizations or not, male or female, and charismatic or not. Farmers would therefore conform or not through the deviation from social norm S. Finally, households may exhibit similar adoption behaviour due some unobservable location factors E, including geographical and economic conditions (e.g. extension services, markets, weather or quality of soils, etc.), that affect both their productivity and the gained utility.

5.4 Econometric framework

5.4.1 The model

To analyse households' engagement in productivity enhancing technologies, and since their decision is a binary choice variable, we applied a spatial probit model. We can therefore specify the utility gained from adopting PETs $(U_A - U_{NA})$ in a latent variable model as follows:

$$y = \begin{cases} 1 & \text{if} \quad y^* > 0 \\ 0 & \text{if} \quad y^* \leq 0, \end{cases} \tag{5.3}$$

where y^* is the $N \times 1$ vector of latent variable determining the outcome of the observed technology adoption status y, that is a binary variable that takes the value 1 for a household if the household decides to adopt the PET technology and 0 otherwise. Variable y^* is assumed to be dependent on exogenous farm household characteristics X, and also the neighbours' characteristics WX as explained earlier; with W being an $N \times N$ spatial weight matrix and X being a $N \times K$. To incorporate such spatial dependence in household adoption behaviour, we follow LeSage and Pace (2009) by employing a Spatial Durbin Probit model (SDM), therefore the latent variable y^* takes the following data generating process:

$$y^* = \rho W y^* + \beta X + \theta W X + \varepsilon, \tag{5.4}$$

where y^* and X are defined as previously; β and θ are parameters to be estimated, and ε is the $N \times 1$ vector of error term, which is assumed to be normally distributed $N(0, \sigma_\varepsilon^2 I_N)$ with I_N defined as an identity matrix. Parameter ρ is a scalar and indicates global spatial dependence. Wy^* represents the average outcome of the farmer's neighbours, excluding himself. When $\rho = 0$ and $\theta = 0$, then the model reduces to a non-spatial probit model. According to LeSage and Pace (2009), it is important to note that, when considering spatial models with and without spatial error component, the SDM can be derived from the spatial error model. Therefore, the SDM should be preferred when facing uncertainty about the true data generating process.

Four main estimators are generally considered when estimating spatial probit models: the expectation-maximization algorithm, Gibbs sampling, recursive importance sampling, and the generalized method of moments. Calabrese and Elkink (2014) simulation results showed that when the sample size increases, the difference between the different estimators becomes smaller. However, when one considers both the estimation of the extent of spatial autocorrelation and the coefficients on the other explanatory variables, the Gibbs estimator (LeSage and Pace, 2009) clearly outperforms the others. In addition, previous studies revealed that Bayesian techniques using Markov Chain Monte Carlo (MCMC) simulations have proven to be a powerful tool in estimating spatial probit models (LeSage, 2000; Holloway *et al.*, 2002; LeSage and Pace, 2009).

The rule in Bayesian estimation framework involves the combination of the data distribution embodied in the likelihood function $p(y|\tau)$ with prior distributions $p(\tau)$ for the parameters $\tau = (\beta, \theta, \rho)$ assigned by the analyst, to produce a posterior distribution $p(\tau|y)$ for the unknown parameters (LeSage and Pace, 2009). This posterior distribution is then specified as: $p(\tau|y) \propto p(y|\tau)\, p(\tau)$.

5.4.2 Estimation technique

LeSage and Pace (2009) suggested the Bayesian Gibbs sampler approach for estimating spatial discrete choice models. For this estimation technique, the latent unobserved observations y^* on the dependent variable are replaced by estimated values. Given values of y^* are therefore used in place of the binary y values, then the same conditional posterior distributions could be derived as in the case of a continuous dependent variable regression model[1]. Once, we derived the conditional distributions, we apply a Markov Chain Monte Carlo procedure (MCMC). For simplicity let's define for this sub-section variable Z and parameter δ such as $Z = (X, WX)$ and $\delta = (\beta, \theta)$. Also note that following LeSage and Pace (2009) for the identification of SDM probit model, we have set $\sigma_\varepsilon^2 = 1$. After choosing some arbitrary starting values for the parameters, we draw each parameter from its conditional distribution ($p(y^*|\delta, \rho, y)$, $p(\delta|y^*, \rho, y)$ and $p(\rho|y^*, \delta, y)$). In total, 2,500 MCMC draws were done with the first 500 draws excluded to account as burn-in. For each draw, we perform the following steps (LeSage and Pace, 2009; Wilhelm and de Matos, 2013):

1. Update $p(y^*|\delta, \rho, y)$ using Gibbs sampling as a truncated multivariate normal distribution subject to $y^* \geq 0$ for $y = 1$ and $y^* < 0$ for $y = 0$:

$$y^* \backsim N\left((I - \rho W)^{-1} Z\delta, \left[(I - \rho W)' (I - \rho W)\right]^{-1}\right)$$

2. Update $p(\delta|y^*, \rho, y)$ using Gibbs sampling from its conditional multivariate

[1]see LeSage and Pace (2009) for more details.

normal distribution and the prior $\delta \backsim N(a, b)$:

$$p(\delta|y^*, \rho, y) \propto N(A, B)$$
$$A = \left(Z'Z + b^{-1}\right)^{-1}\left(Z'(I - \rho W)y^* + b^{-1}a\right)$$
$$B = \left(Z'Z + b^{-1}\right)^{-1}$$

3. Update ρ using the Metropolis-Hastings sampling from its conditional density $p(\rho|y^*, \delta, y)$:

$$p(\rho|y^*, \delta, y) \propto |I - \rho W| exp\left(-\frac{1}{2}\left((I - \rho W)y^* - Z\delta\right)'\left((I - \rho W)y^* - Z\delta\right)\right)$$

The estimation of the Bayesian Spatial Durbin Probit model was conducted in R software using the package *spatialprobit* (Wilhelm and de Matos, 2013).

5.4.3 Direct, indirect and total effects

One advantage in using spatial regression models is to estimate the effects that one unit have on neighbouring units and the effects of neighbours on one unit. LeSage and Pace (2009) demonstrated that for models such as the SDM, the coefficients on the independent variables cannot be interpreted as elasticities. They proposed a method to calculate direct, indirect and total marginal effects, by using the fitted parameters and the expression in matrix as follows:

$$\frac{\partial y}{\partial X_r} = S(W) = (I - \rho W)^{-1}(I\beta_r + W\theta_r). \tag{5.5}$$

This matrix defines the partial derivative of y regarding X, where r represents the r^{th} explanatory variable. An implication of this is that a change in the explanatory variable for a single observation can potentially affect the dependent variable in all other observations. Following LeSage and Pace (2009), three effect estimates are therefore derived from the estimated models: (i) the direct effects or the own–partial derivative ($\partial y_i/\partial X_{ir} = S(W)_{ii}$) , (ii) the indirect effects or the cross–partial derivative ($\partial y_i/\partial X_{jr} = S(W)_{ij}$), and (iii) the total effects, which is the sum of the previous effects. The direct effect expresses the impact of the changes of a characteristic X_r, of a given household located at position i in space, on the adoption response y of this same household. The indirect effect measures the impact of the changes of a characteristic X_r of the household i on the adoption outcome y of a household j

located at position j, where $i \neq j$.

LeSage and Pace (2009) suggested computing scalar summaries of the effects for all observations: the average direct effects, the average indirect effects and the average total effects. The average direct effect is the average of the diagonal elements of the $S(W)$ matrix $((S(W)_{ii} = N^{-1}tr(S(W))))$. The average indirect or spillover effects are the average of the off-diagonal elements of the $S(W)$ matrix, or the difference between the average total effects and the average direct effects. The average total effects to an observation is the sum across the i^{th} row of $S(W)$, the average total effects from an observation is the sum down the j^{th} column of $S(W)$.

5.4.4 Spatial weight matrix

The spatial weight matrix W is a symmetric matrix, where its elements w_{ij} express the proximity of a farm household i with a farm household j. In common practice, W is row standardized so that the sum of the row elements equals one, and the diagonal elements w_{ii} are set to zero. This, therefore, allows the interpretation of model coefficients. Many specifications of weight matrices have been used in the literature and specifying the weight matrix is arbitrary. However, prior knowledge of the study population and economic theory could guide in the specification of these matrices. We consider in our study the inverse distance matrix with several cut-off distances. Neighbours in this specification have different weights, and those with higher weights are closer in distance. The cut-off distances are chosen to determine the distance beyond which spatial effects are no longer relevant. In an inverse distance matrix, elements w_{ij} are defined as $1/d_{ij}$, where d_{ij} is the Euclidean distance between geographic location of household i and j. The cut-off distance approach implies that the weight $w_{ij} = 0$ if the distance between households i and j is beyond a pre-defined distance (i.e. the cut-off distance). To choose the appropriate spatial weights matrix, several models with different thresholds d were first estimated. Then, using the posterior model probability of each model as suggested by LeSage and Pace (2009), we compared these models and chose the one with the highest posterior probability. The threshold values ranged from 1.5 km to 5 km with intervals of 0.5 km. These ranges were chosen based on the characteristics of the population of the study and on previous literature. For rural areas in developing countries, several authors found that the reasonable range for technology spillover is either 1-4 km (Yang and Sharp, 2017), 2–3 km (van Meijl and van Tongeren, 1998; Holloway *et al.*, 2002), 2-4 km (Wollni and Andersson, 2014). Appendix A.3.1 presents a detailed methodology on the model selection approach.

5.5 Data sources and descriptive statistics

5.5.1 Data sources

The data used in this study comes from a cross-sectional survey conducted in Senegal, which randomly sampled 4480 households that mainly produce rain-fed cereals (millet, sorghum, maize, fonio, and rice). The survey was carried out within the framework of the Agricultural Policy Support Project (Projet d'Appui aux Politiques Agricoles, PAPA in French), which was funded by USAID under the "Feed the Future" initiative. The project, implemented by the Senegalese Government, was focused on several commodity value chains (cereals, horticulture), and inputs value chains such as seeds and fertilizers. The Senegalese National Agricultural Research Institute conducted the survey in 2017 with the support of the International Food Policy Research Institute (IFPRI). A two-stage sampling procedure was applied to select survey units. The Enumeration Areas (EAs) defined for the 2013 Senegalese population census were considered as primary units. In total 42 agricultural departments out of 45 were considered. In each agricultural department, between 10 to 36 EAs were randomly selected according to the department size, i.e. the total size of the population. From each enumeration area, 5-10 agricultural households were randomly selected as secondary units. With the data cleaning and after removing observations with no information on the different technology adoption variables (with missing values), a usable sample of 4245 households was prepared for the analysis. However, some observations were found to not have any neighbours for the spatial analysis. Therefore, following Krishnan and Patnam (2013) we dropped these observations which represent 3.9% of the whole sample. This did not affect our analysis as we show later. The final sample is comprised of 4080 households.

Data covered the main agricultural production season of 2016-2017.A structured household questionnaire was used to collect information. This questionnaire included several modules and gathered information on a range of topics such as crop production, membership of farmer-based organizations, household assets, access to rural infrastructure and institutions, and household demographic and socio-economic characteristics. Data collection also included production inputs used, market prices and households use of agricultural technologies such as fertilizers and improved seeds.The survey also covered households located in all six agro-ecological zones.

5.5.2 Productivity enhancing technologies adoption and farmer organizations membership

The dependent variables are binary, indicating the household decision to adopt or not productivity-enhancing technologies (PET), coded as 1 when the household has adopted the PET and 0, otherwise. Two PET technologies are considered in this study, the adoption of improved seeds and inorganic fertilizers. A household is considered as an adopter of improved seeds if it has used any improved cereals seed during the main growing season. Similarly, a farmer is considered an adopter of fertilizers if he/she has used inorganic fertilizers on any cereal crop during the 2016-2017 growing season. Table 5.1 shows descriptive statistics of farmers' adoption of the PET technologies at the household level and farmer organization membership. About 26% of households have adopted improved seeds in cereals farming and 36% have adopted the inorganic fertilizer.

Following the definition of Bernard *et al.* (2015), our variable of interest "organization" is referred to as membership of a rural producer organization that provides farmers with farming and farm-related services including access to inputs, markets and credit, collective sales, and capacities reinforcement. Eight types of farmers organizations were mentioned by the surveyed units: Producer Groups, Economic Interest Groups, Rural Associations, Cooperatives, Women Producer Groups, Federations, Unions, and Networks. The variable organization is binary, coded as 1 if a family member of the household belongs to any of these organizations, and 0 otherwise. About 9% of the households in the sample have at least one person belonging to an organization. The main organizations with the most household members, are the Economic Interest Groups (43.6%), Rural Associations (17.3%), Producer Groups (16.7%), and Cooperatives (15.3%).

Table 5.1: Description of variables of interest

Variables	**Description and measurement**	**Mean (SD)**
	Outcome variable	
Fertilizers	Adoption of fertilizers (1=yes, 0=no)	0.258 (0.438)
Improved seed	Adoption of improved seeds (1=yes, 0=no)	0.361 (0.480)
Organization	Membership in farmer organization (1=yes, 0=no)	0.088 (0.283)
N	Number of Observations	4245

5.5.3 Households characteristics

Following the literature on agricultural technology adoption, several control variables have been included in the models. These includes the household and its head characteristics, households assets and living conditions[2], geographical location, household access to rural institutions, ecological conditions (rainfall and percentages of soil main elements[3]), and various environmental risks faced by the farmer that could affect its adoption decision (e.g. floods, crop diseases, or the break and early stop of rain). Table 5.2 presents the definition and summary statistics of the explanatory variables used in the analysis and their expected signs.

Household socio-economic characteristics variables used include male, age, household size, and education. Male is a dummy variable for the gender of the household head, with value 1 if the household is male-headed and 0 otherwise. The households in our sample are predominantly male-headed, accounting for more than 93% of the sample. The adoption of PET technologies is expected to be positively correlated with gender. Age is a continuous variable and the households head age on average is 53 years. Age is expected to influence the adoption behaviour of farmers, however, the direction can not be defined *a priori*. The relationship between age and adoption of agricultural technology is not straightforward (Adegbola and Gardebroek, 2007). Some authors argue that older farmers are likely to have more dependents, farming experience, and productive resources and information, therefore they are more likely to adopt modern technologies (Sall *et al.*, 2000; Asante *et al.*, 2018). On the other hand, some studies argue that young farmers have a longer-term planning horizon and they can be more willing to take risks and better prone to adopt modern technologies (Zegeye *et al.*, 2001). The average household size of the sample is around 10 indicating the existence of enough family labour for agricultural tasks. We expect that the household size positively influences the adoption of PETs. Education is a

[2]Following the standard approach for calculating a welfare index, we have computed two indexes: a living conditions index and an agricultural implements index. The living conditions index was calculated using various living conditions variables of the household, such as the number of living rooms in the main house, and several dummy transformed variables, such as the type of drinking water, type of cooking water, type of toilet, type of cooking fuel, lightning fuel, type of roofing material for the main house, type of wall material for the main house, type of floor for the main house. Meanwhile, the agricultural implements index was elaborated using dummy variables of the possession of 17 agricultural assets such as donkey carts, horse cart, cattle cart, tractor, sine hoes, ploughs, occidental hoes, sheller, polyculture, arianas, thresher, harvester, sprayer, sower, storage, hangar, atomizer

[3]Using geographical coordinates, the variables rainfall and percentages of clay, silt and sand in soils were retrieved from publicly available databases of the Climate Hazards Center of the University of California (`https://www.chc.ucsb.edu/data`) and of the International Soil Reference and Information Centre (`https://data.isric.org/`), respectively.

binary variable coded as 1 if the household head has not attended any formal school. Most heads in the sample are not educated (more than 60%), they can not read or write. The education dummy is expected to negatively affect PETs adoption, for the simple reason that non educated farmers are less likely to read about the technology and have information about it.

Variables used as proxies for the household wealth are the involvement in off-farm work, the land area owned, and the living conditions and agricultural implements indexes. Off-farm work is a dummy for the involvement of the household in non-farming activities. About 27.1% of households get revenues from activities other than farming. Although off-farm activities might bring additional financial resources to the household that can be invested in productivity-enhancing technologies, the commitment to well-remunerated off-farm works also means less labour available for production, hence discouraging farmers to invest in PETs. Off-farm works can therefore positively or negatively affect the probability of PETs adoption. On average, households in the sample owns about 5.82 hectares of farm land. The area owned and the two indexes are expected to improve the probability of PETs adoption. Variables related to access to infrastructure and institutions include distances to the nearest road and markets, access to extension services, and credit. Only around 11% of the households in our sample have access to extension services and less than 3% have access to credit facilities. Distances to road and markets are expected to have a negative effect on PETs adoption, while access to extension and credit is expected to have a positive effect.

Ecological conditions variables include rainfall[4], the percentages of clay, and silt in soils. Assuming that smaller soil particles constitute good soil conditions for cereals vegetative growth therefore, the expected sign for clay and silt is negative. Farmers who live in areas with good soil fertility, because they use to have good productions, might not invest in PETs. Regarding environmental risks experienced by farmers during the previous five years, we included in the model dummy variables such as floods, rain break, and early rain stop. In terms of expectation, these variables are expected to negatively influence the adoption decision of farmers. Agricultural households that experience such environmental risks are not generally encouraged to use modern technologies.On the contrary, crop disease variable sign is expected to be positive. In addition to the group of spatial explanatory variables (location variables, ecological conditions, environmental risks), two other potential drivers

[4]Besides climate expectation variables, the annual rainfall amounts is also important for the use of some technologies such as fertilizer. In rain-fed agriculture, farmers use fertilizers when the soil has enough moisture to absorb the nutrients in fertilizer.

of households choice in adopting productivity-enhancing technologies are included as explanatory variables: the lagged dependent variable Wy, lagged explanatory variables WX.

Table 5.2: Description of explanatory variables

Variables	**Description and measurement**	**Mean (SD)**	**Expected Signs**	
			Seeds	Fertilizers
Organization	Membership in farmer organization (1=yes, 0=no)	0.088 (0.283)	+	+
	Household and head characteristics			
Male	Gender of household head (1=yes, 0=no)	0.932 (0.251)	+	+
Age	Age of household head (years)	53.074 (13.440)	+/-	+/-
Household size	Number of family members	9.996 (5.438)	+	+
Off-farm	Off-farm work (1=yes, 0=no)	0.271 (0.444)	+/-	+/-
Education	No formal education (1=yes, 0=no)	0.630 (0.483)	-	-
	Assets and living conditions			
Land owned	Total land owned by household (ha)	5.823 (8.367)	+	+
Living conditions	Living conditions index	-0.089 (2.252)	+	+
Implements index	Agricultural implements index	0.013 (1.317)	+	+
	Location variables			
Distance to road	Distance to nearest road (km)	10.154 (14.152)	-	-
Distance to market	Distance to nearest main market (km)	13.560 (11.994)	+	+
	Access to institutions			
Extension	Extension services (1=yes, 0=no)	0.108 (0.310)	+	+
Credit	Credit (1=yes, 0=no)	0.027 (0.162)	+	+
	Ecological conditions			
Rainfall	Annual rainfall (mm)	695.477 (325.058)	+	+
Clay	Percentage of clay (%)	19.996 (7.203)	-	-
Silt	Percentage of silt (%)	13.516 (5.755)	-	-
	Environmental risks			
Floods	Floods (1=yes, 0=no)	0.056 (0.229)	-	-
Crop disease	Crop disease (1=yes, 0=no)	0.107 (0.309)	+	+
Rain break	Breaks in rainfall (1=yes, 0=no)	0.298 (0.457)	-	-
Early Rain Stop	Early stop of rains (1=yes, 0=no)	0.356 (0.479)	-	-
N	Number of Observations	4245		

5.6 Results and discussion

5.6.1 Comparative descriptive analysis

Tables 5.8 and 5.9 (in appendix) present the means of variables included in our models by adoption category and by membership in farmers organizations, with the p-values of computed differences between means. Results show that statistically significant differences are jointly observed for several variables between adopters and non-adopters in the entire sample and in the two sub-groups of organizations members and non-members.

For the adoption of improved seeds, these variables include the living standard index, distance to nearest all-weather road, distance to the nearest market, access to extension services and annual rainfall. Adopters in all groups have a better average living conditions index. For households who are members of farmer organizations, we observe a mean of 0.37 for adopters and -0,17 for non-adopters. Adopters are also significantly distinguishable with their higher access to extension services. However, concerning the annual rainfall, non-adopters in all groups seem to live in areas where the cumulative annual rainfall was higher than that of adopters. Although the differences in means do not allow making causal statement, a particular pattern about the distance of households to the nearest road and markets can be observed. In the sub-sample of non-members, households living closer to roads and markets are more likely to adopt improved seeds, contrarily to the organization members where farmers living farther to roads and markets are the ones who adopt more. This result suggests that members of farmer organization might have some "geographical" ease in the access to improved seeds. This makes sense and it is consistent with previous literature, farmer-based organizations provide their members with seeds and cover the related transaction costs with related membership fees. In addition, taking advantage of social events in these organizations, members could have developed more connections with their neighbouring peers and therefore can provide themselves in seeds with their peers farmers.

For the adoption of mineral fertilizers, significant differences were observed with variables such as gender, household size, area owned, living conditions index, access to extension and credit, and environmental conditions and risks (e.g. annual rainfall, clay, silt and crop diseases). While for some variables, there is no clear pattern between organization membership groups (or sub-samples) about fertilizers adoption, some other variables show distinctively that they might influence the decision of farmers towards adoption. For instance, adopters on average compared to non-adopters, have a higher number of family members (11 vs. 9 for pool sample, 12 vs 10 for members, and 11 vs 9 for non-members), live in better conditions, have better access to extension and credits, and experienced diseases (16% vs. 8% for pool sample, 15% vs. 6% for members, and 16% vs. 8%9 for non-members).

The descriptive analysis suggests that even if a household belongs to producer organizations, there might be other several variables that drive the adoption of improved seeds and inorganic fertilizers. The results also indicate that the use of improved seeds and mineral fertilizers are not necessarily affected by the same factors, as shown by the different distances of the households to institutions such as road and markets that distinguishes adopters of improved seeds from non-adopters but not the adopters of fertilizers from the non-adopters.

5.6.2 Spatial dependence and model selection

Before implementing the Spatial Durbin Probit models, we followed the approach of Läpple and Kelley (2015) by using the predicted values from the non-spatial probit models, and evaluated the strength of spatial dependence in the data through a Moran scatter plots as shown in figure 5.1. These plots show the relationship between the adoption behaviour y^* of a household and the adoption decisions of neighbouring farms. From these plots, black clouds of points can be observed in quadrants I and III indicating that there is a strong and positive spatial dependence between household's decision to adopt both technologies. This visual evidence suggests that farmers with a relatively high propensity to adopt PETs seem to live closer to other farmers with the same level of probability of adoption, and farmers with a relatively low likelihood to adopt tend to live near farmers with low probability to use these technologies. We, therefore, compute several SDM probit models to account for spatial dependence.

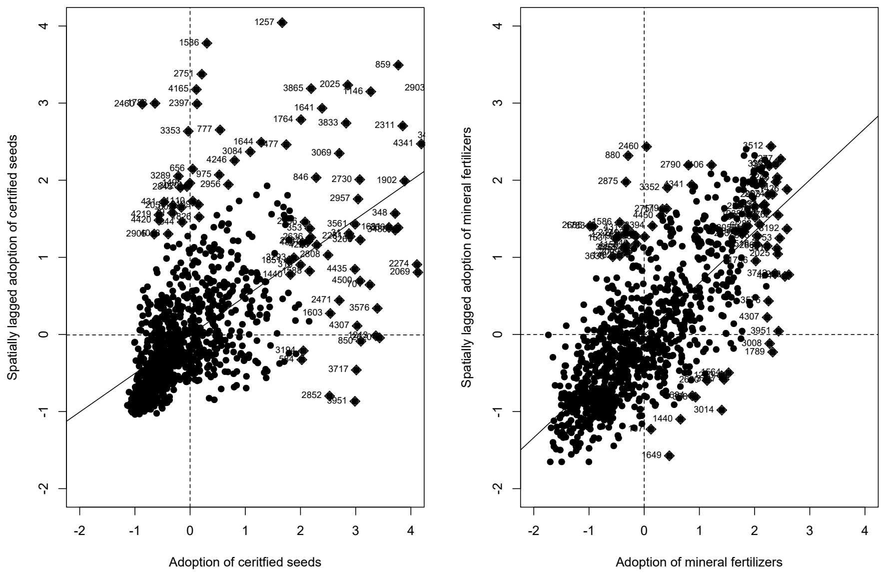


Figure 5.1: Moran Plots

Tables 5.3 and 5.4 present the posterior model probabilities of the alternative Spatial Durbin Probit models with different threshold values ranging from 1 to 5 km, with intervals of 0.5 km. Based on these posterior model probabilities, the preferred models are the ones with the highest value notably the Spatial Durbin Probit model with a 4 km threshold for both technologies. This suggests that spatial spillover effects are assumed to be null beyond this threshold distance. This threshold is higher than those observed in previous studies on technologies adoption in rural areas by Wollni and Andersson (2014) and Yang and Sharp (2017) who both found a threshold of 1.5 km respectively in Honduras for the adoption of organic farming, and in New Zealand for the adoption of best management practices in dairy farming. A high threshold is plausibly due to the extent of our sample, which is of a bigger size and covers a larger area, that is, an entire country. In comparison to other studies, Wollni and Andersson (2014) had a sample size of 241 farms in one state of Honduras and Yang and Sharp (2017) analysis covered 171 dairy farms in one specific producing region. For instance, Läpple *et al.* (2017) had considered a threshold of 45 km when working on 280 dairy farmers in the entire Ireland. The rest of our analysis will focus on these preferred models. Within this radius of 4 km, the households in our sample have on average 8.5 neighbours and an average distance to neighbouring households of 2.7 km.

Table 5.3: SDM Probit posterior model probabilities: Improved seeds

Thresholds (km)	Model Probabilities	ρ values	ρ p-values
d = 1.5	0.0000	0.4395	0.0000
d = 2	0.0014	0.4470	0.0000
d = 2.5	0.0016	0.4526	0.0000
d = 3	0.0040	0.4591	0.0000
d = 3.5	0.0519	0.4619	0.0000
d = 4	0.7132	0.4659	0.0000
d = 4.5	0.0544	0.4664	0.0000
d = 5	0.1733	0.4705	0.0000

Table 5.4: SDM Probit posterior model probabilities: Mineral fertilizers

Thresholds (km)	Model Probabilities	ρ values	ρ p-values
d = 1.5	0.0191	0.5036	0.0000
d = 2	0.0153	0.5052	0.0000
d = 2.5	0.0064	0.5101	0.0000
d = 3	0.1095	0.5175	0.0000
d = 3.5	0.0536	0.5232	0.0000
d = 4	0.2989	0.5290	0.0000
d = 4.5	0.1985	0.5337	0.0000
d = 5	0.2986	0.5362	0.0000

5.6.3 Determinants of productivity-enhancing technologies adoption

Table 5.5 presents the estimates of simple probit models and Spatial Durbin Probit models (SDM Probit). Estimation results using simple probit clearly exhibit statistically significance signs for membership of farmer-based organizations, suggesting that members of these collective action groups are more likely to adopt PETs if we assume that there is no spatial dependence, and therefore no peer neighbours effects in the adoption decisions of farmers. However, estimates from the SDM Probit models show that spatial lag coefficients ρ, are statistically significant at 1% level, indicating the existence of spatial dependence among farmers in the adoption of both PETs, justifying the econometric approach used. In addition, both estimated ρ are positive suggesting that a cereal farmers in Senegal are more likely to adopt a productivity-enhancing technology if their neighbours have adopted it.

Moreover, the estimated coefficients of our variable of interest, i.e., organization membership, are also positive and significant at 1% level implying that membership in a farmer-based organization affects significantly and positively the decision of a farmer to use both PETs. In general, membership in rural organizations helps farmers to develop their social networks and through these networks, members could with certain ease have better access to information about new technologies and about their availability. For example, in Nigeria, Ajah (2015) shows the existence of a positive correlation between agricultural cooperatives and access to farm inputs. Hence, members of collective action groups have a higher probability to adopt new agricultural technologies than non-members. Our results corroborate with studies that have focused on agricultural cooperatives. In Ethiopia Abebaw and Haile (2013) found that membership in cooperative positively and strongly affects the adoption of fertilizer. In China cooperative membership positively and significantly influences the likelihood of farmers to invest in organic soil amendments (Ma *et al.*, 2018). The recent study of Zhang *et al.* (2019) also shows that membership in agricultural cooperatives has a positive effect on the number of technologies adopted by farmers. Furthermore, several specific factors are significantly associated with the adoption of improved seeds and mineral fertilizers.

The number of persons living in a household exerts a positive and statistically significant effect on the adoption of PETs. This suggests that households with a larger number of family members have a higher likelihood to adopt PETs. One explanation is that more family members in a household increase the probability to have a larger social network and therefore more information sources about technologies and its characteristics rendering then their adoption more likely. Bernard and Spielman (2009) and Ma and Abdulai (2016) reported that households who have more family members are more likely to belong to farmer-based organizations which represent good opportunities for farmers to develop their social capital. Our findings are similar to those reported by Hamzakaza *et al.* (2014) in Zambia where household dependency ratio increases the probability to adopt multiple stress-resistant improved common bean varieties. In Tanzania, Letaa *et al.* (2015) also found that the number of household family members (14-65 years) have a positive and significant effect on the adoption of new improved common bean varieties.

Adoption of PETs is also positively affected by the land area owned by the household. Farm households that possess larger areas of land are more likely to adopt PETs. This result corroborates the observations of Khonje *et al.* (2015) in Zambia. Possessed land can generally be used as collateral for credit or converted into cash for the procurement of improved seeds and fertilizers (Asfaw *et al.*, 2012; Khonje *et al.*, 2015). In addition, having large areas of land gives more opportunities for farmers to allocate only a small portion of land to agricultural technologies, spreading, therefore, the risk of technology failure (Mariano *et al.*, 2012). Dummy variable for access to credit is statistically significant in driving the adoption of productivity-enhancing technologies. Farmers who have access to credit are more likely to use improved seeds and inorganic fertilizers. The use of improved technologies such as improved seeds and mineral fertilizers means additional cost for crop production. Sometimes these costs can be very high. In Senegal, a 50kg bag of fertilizer costs on average 12500 FCFA (circa 20 USD) which represents about 81% of the total production cost per hectare in our sample (all crops put together)[5]. Accessing credit might help farmers in covering these additional costs, explaining, therefore, the positive effect that access to credit has on the two PETs adoption. Similar results were observed in Bangladesh, where Ward and Pede (2015) argued that loosening constraints on credit access proves to be beneficial in the adoption of technologies such as hybrid seeds.

Farmers who experienced crop diseases in the last five years are more likely to adopt improved seeds. Crop diseases depending on their importance can lead to great losses in farm production. They are sometimes considered as shocks that affect not only production but also the whole farming household welfare and food security. One interpretation of this result could be that in order to avoid that such shock happens

[5]Figures based on data used.

and ends up with severe losses in productions, farmers might decide to use improved seeds or mineral fertilizers. These technologies are supposed to have inherent abilities to strengthen plants during their vegetative development and to adequately fight any disease that might occur. In the context of Senegal, improved seeds are basically high yielding and disease resistant. Additionally, the use of productivity-enhancing technologies in agriculture are also known to reduce dispersions in outputs hence their importance in managing agricultural risks. Hence the use of improved seeds, mineral or inorganic fertilizers and irrigation are supposed to manage production risks in agriculture.

Besides these explanatory variables, there are some other factors that influence specifically the adoption of each PET. Factors that drive only the adoption of improved seeds are age, distance to market and the dummy variable for breaks in the rain. Age appears to significantly and positively affect the adoption of improved seeds, indicating that older household heads have a higher likelihood to adopt improved seeds than households with younger heads. One explanation is that older farmers might have better access to farming inputs such as land, credit or labour than younger ones. These results are in line with those of Ntshangase *et al.* (2018) who find a positive and significant effect of age on the adoption of no-till conservation agriculture in South-Africa. Results show that farmers who are regularly exposed to breaks in the rain (i.e. during the last five years) are encouraged to use improved seeds. One probable explanation could be that improved seeds are likely to be drought resistant and early maturing. Therefore, farmers are motivated by the idea to harvest in a short period of time, or the possibility to sow even after an early break in rainfall.

In the case of fertilizers adoption, specific drivers are the gender of the household head, living conditions, agricultural implements, and access to extension services. Estimates reveal that variable male significantly affects the adoption of mineral fertilizers, implying that male-headed households are better prone to adopt this PET than female-headed households. This can be explained by the differential access to resources in many sub-Saharan African contexts, that is mostly in favour of men. As demonstrated by Doss and Morris (2001) in Ghana, observed differences in agricultural technologies adoption are due to unequal access between men and women to complementary inputs such as land, labour, and extension services. Proxies and variables used for land, labour and extension services in this study appeared to significantly drive adoption of inorganic fertilizers. The proxy indexes for living conditions assets and possession of agricultural implements of farmers show positive and significant coefficients, suggesting that farmers who possess more assets (for living condition and agricultural production) have higher propensity to adopt mineral fertilizers. This result makes sense, first as stated previously, farmers with high indexes are more wealthy, they can, therefore, afford the additional production costs related to the use of these technologies. Secondly, possessing more agricultural implements (such as oxen or machinery) will also ease difficulties in production works, therefore encouraging farmers to adopt modern technologies.

Results further show that farmers who have access to extension services are more likely to adopt mineral fertilizers. One explanation is that these farmers are better aware of information about fertilizers, their access and therefore are more prone to use them. According to Asfaw *et al.* (2012) adoption of agricultural technologies can also be hampered by a lack of awareness of farmers. Several authors (Wubeneh and Sanders, 2006; Amare *et al.*, 2012; Mariano *et al.*, 2012; Khonje *et al.*, 2015) have observed such a positive relationship between access to extension services and adoption of agricultural technologies.

In regard to the spatially lagged independent variables, some of them show statistically significant coefficients, indicating that a farmer's adoption of PETs is also affected by his/ her neighbours' characteristics. These characteristics include farmer-based organizations membership for both PETs, and extension services and floods for the use of improved seeds. In analysing table 5.5, it worth noting that most expectations on coefficients signs in table 5.2 can be verified. However, these coefficients estimates β, ρ and θ cannot be interpreted as marginal effects, i.e. how changes in the explanatory variables affect the adoption probability of productivity-enhancing technologies. The effects of the explanatory variables on the adoption of PETs are presented and discussed in the next section.

5.6.4 Spatial effects estimation

As stated previously, after the estimation of the models, three types of spatial effects or scalars can be computed: the direct effects, the indirect effects or spatial spillovers and the total effects. The direct effect estimates the effect of a change in an explanatory variable of farmer i on the adoption probability of farmer i, indirect effects or spatial spillovers express the cumulative effect of a change in an explanatory variable of neighbouring farms on the adoption probability of farmer i, and the total effect of an explanatory variable is, therefore, the sum of its direct effect and its indirect effect (LeSage and Pace, 2009). Tables 5.6 and 5.7 present the posterior means of the direct, indirect, and total effects of the coefficients of respective PET adoption model as well as the corresponding 95% credible intervals. The 95% credible intervals were computed using the set of 2000 draws retained from the MCMC estimation.

Results show that for both PETs, the effects of four explanatory variables including organization membership, household size, area owned, crop disease have 95% credible intervals that do not cross zero. In addition, four specific variables (age, distance to market, rainfall, rain break) for the adoption of improved seeds and six specific variables (male, living conditions index, agricultural implements index, extension services, credit, and early rain stop) affect the adoption of mineral fertilizers and have their 95% credible intervals that do not cross zero. Additionally, as expected, the indirect effect estimates are smaller in magnitude than the direct estimates.

In terms of marginal effects, farmer organization membership seems to be the most important factor that determines the adoption of improved seeds. A household's likelihood to adopt improved seeds increases by 20.2% if the household belongs to a farmer-based organization, which includes an 8.2% increase in the probability of adoption of neighbouring households. The likelihood is also very high with fertilizer adoption, where the direct increase of probability is evaluated at 14.3% and the indirect effect at 12%. These results imply that besides the traditional services provided to its members, rural collective action groups might play key roles in agricultural technologies adoption by providing to farmers opportunities to develop social connections, share farming experience, technical knowledge and vital information about technologies.

Household size, land area, and crop diseases also appear to have a positive effect on farmers' adoption of agricultural technologies. An increase in household size by one family member increases the likelihood of that household adopting by 0.8% and 1.5% respectively for improved seeds and mineral fertilizers. Regarding the land area owned, PETs adoption probabilities increase, with the expansion of land by one hectare, by 0.6% and 1.3% respectively for improved seeds and mineral fertilizers. In addition, having experienced crop diseases during the last five years increases the probability of adopting mineral fertilizers by 10.9% (direct effect), and also increases the probability of adoption of neighbouring households by 9.2%, adding to a total effect of 20.1% increase in the probability of adopting mineral fertilizers.

The variable age has a positive effect on the probability of adopting improved seeds, however the effects are very small. An increase in age by one year increases the probability of adoption of improved seeds only by 0.1%. An increase in age by one year of the neighbouring household's head increases the given household adoption by only 0.1%, accumulating to a total increase in the probability of adoption of all households of 0.2%. Contrary to the general literature findings, the distance of a household to the nearest market affects positively and significantly the adoption of improved seeds, suggesting that the more, households are far from markets, the likely they are to adopt improved seeds. For example, an increase in one km to the market increases the total probability of adoption of a household by 7.6%. These findings could be explained by the fact that improved seeds are not necessarily found in markets, in contrast to inorganic fertilizes. A farmer can easily obtain improved seeds from the closest farm that grows this kind of seeds, or alternatively with neighbouring households that have beforehand acquired the seeds. Additionally, as argued previously, farmer organizations can mobilize production inputs for its members and therefore reduce overall associated transaction costs.

Two environmental variables are positively and significantly correlated with the adoption of improved seeds. A raise in the level of rainfall by just 1 mm increases the probability of adoption of a household by 0.3%, of nearby farms by 0.2%, with a cumulative total effect on all farms of 0.4%. The effects are more important with farmers that have experienced a break in rainfall during the last five years. Experiencing breaks in the rain increases the likelihood of a given farmer to use improved seeds by 3.2%.

One additional level of the living conditions index of a household increases the probability of adopting fertilizes by over 4%, whereas the spillovers effects of neighbouring households are at 3.5%, resulting in a total effect of around 8% increase in the probability of adoption on all households. We also find that access to extension services and credit has a positive and significant effect on fertilizer adoption. Households who have access to extension services and credit are 8.3% and 17% respectively more likely to adopt fertilizers. Meanwhile, indirect effects from neighbouring farmers are estimated at 7% and 14.2% respectively for extension and credit. These figures denote the importance of information and financial resources in agricultural technologies adoption.

The early stop in rain seem to be the only environmental variable that discourages households from adopting inorganic fertilizers. Farmers who have experienced an early stop in rain in the last five years are in total 7.5% less likely to adopt mineral fertilizers, with neighbouring farmers 3.4% less likely to use mineral fertilizers due to a possible early stop in the rain. This result makes sense if we consider the fact that fertilizers use needs necessary more water to be efficient. Our sample is made of rain-fed cereals farmers and some of them are living in regions where rainfall length is generally short. These farmers do not, therefore, take the risk of investing in fertilizers and finally not getting full benefits from them.

5.6.5 Robustness checks

We conducted two robustness checks by using two other types of weight matrix that help to preserve the whole data sample. Then, we compared the obtained results using those weight matrices specifications to the inverse distance specification. For the first type of weight matrix, we follow Pede *et al.* (2018) and specify a spatial weight matrix W_P of a power form with: $w_{ij} = \exp\left(-d_{ij}^2/s^2\right)$, where d_{ij} is the euclidean distance between households i and j, s is the cut-off distance of 4 km. For the second type of weight matrix, similarly to Läpple *et al.* (2017), we design a weight matrix W_D based on the population density[6].

Specifically, we include population density data in our spatial weight matrix, meaning that the strength of influence is based on population density as well as proximity between farm households. Such approach should help to overcome the missing neighbours problem due to sampling processes. The missing neighbour problem suggests that some households in our sample may have fewer neighbours merely due to sampling issues (Läpple *et al.*, 2017). To design the population density weight matrix, we constructed a population density distance (d_{ij}^{π}) that measures the average population density (π) between households i and j, such that $d_{ij}^{\pi} = (\pi_i + \pi_j)/2$, which is later normalised on the maximum across the sample. Then we superimpose these average population densities on a contiguity matrix based on a cut-off distance of 20 km. This cut-off distance permits us to preserve all the households in the sample. Therefore, the population density weight matrix is computed as a multiple of a contiguity matrix and the density averages values d_{ij}^{π}. Here, households are neighbours if they live within a distance of 20 km, however the strength of their spatial relationship is determined by the average population density value. Both weight matrices were row-normalised.

Results of the robustness analysis are presented in the appendix. Results of SDM

[6]We retrieved the population density data using households coordinates and databases of the Center for International Earth Science Information Network - Columbia University at https://doi.org/10.7927/H49C6VHW (Center for International Earth Science Information Network - CIESIN - Columbia University, 2018)

probit estimates using the different weight matrices specifications are presented in tables 5.10 and 5.13, and the computed spatial effects are reported in tables 5.11, 5.12, 5.14 and 5.15. When comparing these results to the inverse distance weight matrix specification, we can observe that the new specifications provide approximatively the same results. For instance, as in the preferred models, these results show that spatial lag coefficients ρ, are statistically significant and with the same direction, but with higher magnitudes. Regarding the factors that affects the adoption of PETs, most variables that are significant in the preferred model are also significant here, with the same directions and with slight differences in the magnitudes.

5.7 Conclusion

This paper analyses the effects of farmers' neighbourhood and membership in farmer-based organizations on the adoption of two productivity-enhancing technologies - improved seeds and inorganic fertilizers in Senegal. A Bayesian Spatial Durbin Probit model was employed to account for spatial effects in estimations, and empirical analyses were done using a data set of 4080 farm households from five agro-ecological regions of Senegal.

The empirical results show that, after assuming a neighbourhood effect radius of 4 km, high spatial dependence exists among households adoption behaviours. Farmers seem to adopt productivity-enhancing technologies if their neighbours do. In addition, findings reveal that membership in an organization, number of family members, size of land owned, and experiencing crop diseases during the last five years increase the probability to use both productivity-enhancing technologies. These findings suggest a complementary effect of membership in farmer organizations, in facilitating information exchange between farmers. Our results are also in concordance with previous literature regarding the relevance of neighbourhood effects on technology adoption in the developing world.

Therefore, policy interventions in Senegal that aim to increase rates of technologies uptake and agricultural productivity should consider the spatial effects of technologies diffusion. Geography, neighbourhood and social interactions seem to play a considerable role in technology adoption, hence extension services should take such information into account when designing technology diffusion programs.

Table 5.5: Spatial Durbin Probit Models estimates

	Improved seeds			Mineral fertilizers		
	Probit	SDM Probit		Probit	SDM Probit	
		X	WX		X	WX
Intercept	-0.911 (0.154)***	-0.626 (0.221)***		-1.959 (0.155)***	-1.092 (0.214)***	
Organization	0.754 (0.079)***	0.432 (0.097)***	0.430 (0.125)***	0.846 (0.083)***	0.456 (0.106)***	0.532 (0.133)***
Male	0.027 (0.092)	0.027 (0.106)	0.017 (0.143)	0.361 (0.094)***	0.286 (0.113)**	0.106 (0.151)
Age	0.003 (0.002)	0.004 (0.002)**	-0.001 (0.003)	-0.001 (0.002)	0.003 (0.002)	-0.008 (0.003)***
Household size	0.017 (0.004)***	0.016 (0.005)***	-0.005 (0.007)	0.023 (0.004)***	0.026 (0.005)***	-0.015 (0.007)**
Off-farm	-0.067 (0.053)	-0.021 (0.064)	0.009 (0.079)	-0.089 (0.050)*	0.013 (0.066)	-0.089 (0.080)
Education	-0.013 (0.048)	0.010 (0.061)	-0.040 (0.075)	-0.072 (0.047)	-0.072 (0.058)	0.028 (0.072)
Area owned	0.013 (0.003)***	0.013 (0.004)***	-0.005 (0.005)	0.022 (0.004)***	0.023 (0.005)***	-0.012 (0.005)**
Living index	-0.003 (0.011)	0.008 (0.017)	-0.008 (0.020)	0.079 (0.011)***	0.133 (0.018)***	-0.078 (0.021)***
Agricultural index	0.027 (0.020)	0.037 (0.027)	-0.015 (0.031)	0.147 (0.021)***	0.131 (0.027)***	-0.021 (0.032)
Distance to road	-0.008 (0.002)***	-0.060 (0.099)	0.055 (0.099)	0.003 (0.002)	0.108 (0.101)	-0.107 (0.101)
Distance to market	-0.012 (0.002)***	0.162 (0.098)*	-0.170 (0.098)*	-0.003 (0.002)	-0.101 (0.096)	0.099 (0.096)
Extension	0.331 (0.072)***	0.105 (0.094)	0.190 (0.110)*	0.355 (0.072)***	0.263 (0.097)***	-0.090 (0.116)
Credit	0.300 (0.135)**	0.238 (0.143)*	-0.024 (0.205)	0.732 (0.146)***	0.532 (0.165)***	0.217 (0.223)
Rainfall	-0.001 (0.000)***	0.009 (0.006)	-0.010 (0.006)*	0.000 (0.000)***	-0.009 (0.006)	0.009 (0.006)
Clay	-0.005 (0.007)	0.034 (0.023)	-0.043 (0.024)*	0.008 (0.006)	-0.005 (0.023)	0.008 (0.023)
Silt	0.034 (0.008)***	0.019 (0.028)	0.002 (0.029)	0.031 (0.007)***	0.010 (0.028)	0.007 (0.029)
Floods	0.230 (0.093)**	0.022 (0.126)	0.247 (0.142)*	0.015 (0.094)	-0.036 (0.130)	0.047 (0.150)
Crop diseases	0.033 (0.073)	0.183 (0.102)*	-0.299 (0.120)**	0.548 (0.068)***	0.348 (0.097)***	0.169 (0.112)
Rain break	0.188 (0.050)***	0.116 (0.066)*	0.040 (0.079)	-0.044 (0.049)	-0.089 (0.071)	0.100 (0.083)
Early rain stop	0.021 (0.048)	-0.021 (0.065)	0.058 (0.074)	-0.162 (0.047)***	-0.127 (0.068)*	-0.020 (0.078)
ρ		0.466 (0.024)***			0.529 (0.020)***	
Log Likelihood	-2110.226			-2308.522		
Num. obs.	4080	4080	4080	4080	4080	4080

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 5.6: Effect estimates of Spatial Durbin Probit models: Improved seeds

	Direct Effects			Indirect Effects			Total Effects		
	Lower 5%	Mean	Upper 95%	lower 5%	Mean	Upper 95%	lower 5%	Mean	Upper 95%
Organization	0.076	0.119	0.163	0.052	0.082	0.115	0.128	0.202	0.275
Male	-0.037	0.009	0.057	-0.025	0.006	0.040	-0.062	0.015	0.097
Age	0.000	0.001	0.002	0.000	0.001	0.001	0.000	0.002	0.003
Household size	0.002	0.005	0.007	0.002	0.003	0.005	0.004	0.008	0.012
Off-farm	-0.035	-0.005	0.023	-0.024	-0.004	0.016	-0.059	-0.009	0.039
Education	-0.023	0.003	0.030	-0.016	0.002	0.020	-0.040	0.005	0.050
Area owned	0.002	0.004	0.005	0.001	0.003	0.004	0.003	0.006	0.009
Living index	-0.005	0.002	0.010	-0.004	0.002	0.007	-0.009	0.004	0.017
Agricultural index	-0.002	0.010	0.022	-0.002	0.007	0.016	-0.004	0.017	0.037
Distance to road	-0.063	-0.017	0.027	-0.043	-0.012	0.019	-0.106	-0.029	0.046
Distance to market	0.002	0.045	0.092	0.001	0.031	0.064	0.003	0.076	0.155
Extension	-0.014	0.028	0.070	-0.010	0.020	0.048	-0.024	0.048	0.118
Credit	-0.000	0.065	0.129	-0.000	0.045	0.090	-0.001	0.111	0.220
Rainfall	0.000	0.003	0.005	0.000	0.002	0.004	0.000	0.004	0.009
Clay	-0.001	0.010	0.020	-0.001	0.007	0.014	-0.002	0.016	0.034
Silt	-0.007	0.006	0.018	-0.005	0.004	0.012	-0.012	0.009	0.030
Floods	-0.050	0.006	0.063	-0.036	0.004	0.043	-0.086	0.010	0.106
Crop diseases	0.006	0.052	0.100	0.004	0.036	0.070	0.010	0.088	0.169
Rain break	0.002	0.032	0.063	0.002	0.022	0.043	0.004	0.054	0.106
Early rain stop	-0.034	-0.006	0.023	-0.024	-0.004	0.016	-0.058	-0.010	0.040

Table 5.7: Effect estimates of Spatial Durbin Probit models: Mineral fertilizers

	Direct Effects			Indirect Effects			Total Effects		
	Lower 5%	Mean	Upper 95%	lower 5%	Mean	Upper 95%	lower 5%	Mean	Upper 95%
Organization	0.091	0.143	0.198	0.075	0.120	0.170	0.166	0.264	0.367
Male	0.031	0.090	0.150	0.026	0.076	0.128	0.057	0.165	0.279
Age	-0.000	0.001	0.002	-0.000	0.001	0.002	-0.000	0.002	0.004
Household size	0.005	0.008	0.011	0.004	0.007	0.009	0.010	0.015	0.020
Off-farm	-0.030	0.004	0.039	-0.025	0.003	0.033	-0.055	0.007	0.071
Education	-0.053	-0.023	0.007	-0.045	-0.019	0.006	-0.098	-0.042	0.013
Area owned	0.005	0.007	0.010	0.004	0.006	0.008	0.009	0.013	0.018
Living index	0.033	0.042	0.051	0.027	0.035	0.044	0.061	0.078	0.095
Agricultural index	0.028	0.041	0.056	0.023	0.035	0.048	0.051	0.076	0.104
Distance to road	-0.017	0.034	0.088	-0.015	0.028	0.073	-0.032	0.062	0.162
Distance to market	-0.082	-0.032	0.019	-0.068	-0.027	0.016	-0.150	-0.059	0.035
Extension	0.033	0.083	0.133	0.028	0.070	0.111	0.061	0.153	0.243
Credit	0.086	0.170	0.255	0.072	0.142	0.216	0.158	0.312	0.469
Rainfall	-0.006	-0.003	0.000	-0.005	-0.002	0.000	-0.010	-0.005	0.001
Clay	-0.013	-0.002	0.010	-0.011	-0.001	0.008	-0.024	-0.003	0.018
Silt	-0.011	0.003	0.018	-0.010	0.003	0.015	-0.021	0.006	0.033
Floods	-0.077	-0.011	0.058	-0.065	-0.009	0.048	-0.141	-0.021	0.105
Crop diseases	0.060	0.109	0.160	0.050	0.092	0.135	0.110	0.201	0.295
Rain break	-0.066	-0.029	0.008	-0.055	-0.024	0.007	-0.121	-0.053	0.015
Early rain stop	-0.075	-0.041	-0.005	-0.063	-0.034	-0.004	-0.138	-0.075	-0.009

References

Abdoulaye, T., Wossen, T. and Awotide, B. (2018) Impacts of improved maize varieties in Nigeria: ex-post assessment of productivity and welfare outcomes, *Food security*, **10**, 369–379.

Abdulai, A., Monnin, P. and Gerber, J. (2008) Joint estimation of information acquisition and adoption of new technologies under uncertainty, *Journal of International Development*, **20**, 437–451.

Abebaw, D. and Haile, M. G. (2013) The impact of cooperatives on agricultural technology adoption: Empirical evidence from Ethiopia, *Food Policy*, **38**, 82 – 91.

Adegbola, P. and Gardebroek, C. (2007) The effect of information sources on technology adoption and modification decisions, *Agricultural Economics*, **37**, 55–65.

Ahmed, M. H., Geleta, K. M., Tazeze, A. and Andualem, E. (2017) The impact of improved maize varieties on farm productivity and wellbeing: evidence from the East Hararghe zone of Ethiopia, *Development Studies Research*, **4**, 9–21.

Ajah, J. (2015) Comparative analysis of cooperative and non-cooperative farmers' access to farm inputs in Abuja, Nigeria, *European Journal of Sustainable Development*, **4**, 39–50.

Amare, M., Asfaw, S. and Shiferaw, B. (2012) Welfare impacts of maize–pigeonpea intensification in Tanzania, *Agricultural Economics*, **43**, 27–43.

Asante, B. O., Villano, R. A., Patrick, I. W. and Battese, G. E. (2018) Determinants of farm diversification in integrated crop–livestock farming systems in Ghana, *Renewable Agriculture and Food Systems*, **33**, 131–149.

Asfaw, S., Shiferaw, B., Simtowe, F. and Lipper, L. (2012) Impact of modern agricultural technologies on smallholder welfare: Evidence from Tanzania and Ethiopia, *Food Policy*, **37**, 283 – 295.

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

Barham, B. L., Chavas, J.-P., Fitz, D., Salas, V. R. and Schechter, L. (2014) The roles of risk and ambiguity in technology adoption, *Journal of Economic Behavior & Organization*, **97**, 204 – 218.

Barrett, C. B., Christiaensen, L., Sheahan, M. and Shimeles, A. (2017) On the structural transformation of rural Africa, *Journal of African Economies*, **26**, i11–i35.

Bernard, T., Frölich, M., Landmann, A., Unte, P. N., Viceisza, A. and Wouterse, F. (2015) Building trust in rural producer organizations in Senegal: Results from a randomized controlled trial, Tech. rep., IZA Discussion Papers.

Bernard, T. and Spielman, D. J. (2009) Reaching the rural poor through rural producer organizations? A study of agricultural marketing cooperatives in Ethiopia, *Food Policy*, **34**, 60 – 69.

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

Calabrese, R. and Elkink, J. A. (2014) Estimators of binary spatial autoregressive models: A Monte Carlo study, *Journal of Regional Science*, **54**, 664–687.

Case, A. (1992) Neighborhood influence and technological change, *Regional Science and Urban Economics*, **22**, 491–508.

Center for International Earth Science Information Network - CIESIN - Columbia University (2018) Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 11. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC), accessed 08-05-2020.

Conley, T. G. and Udry, C. R. (2010) Learning about a new technology: Pineapple in Ghana, *American Economic Review*, **100**, 35–69.

Cunguara, B. and Darnhofer, I. (2011) Assessing the impact of improved agricultural technologies on household income in rural Mozambique, *Food Policy*, **36**, 378 – 390.

Doss, C. R. and Morris, M. L. (2001) How does gender affect the adoption of agricultural innovations?: The case of improved maize technology in ghana, *Agricultural Economics*, **25**, 27 – 39.

Duflo, E., Kremer, M. and Robinson, J. (2008) How high are rates of return to fertilizer? Evidence from field experiments in Kenya, *American Economic Review*, **98**, 482–88.

Duflo, E., Kremer, M. and Robinson, J. (2011) Nudging farmers to use fertilizer: Theory and experimental evidence from kenya, *American Economic Review*, **101**, 2350–90.

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**, 469–488.

Feder, G. (1980) Farm size, risk aversion and the adoption of new technology under uncertainty, *Oxford Economic Papers*, **32**, 263–283.

Feder, G., Just, R. E. and Zilberman, D. (1985) Adoption of agricultural innovations in developing countries: A survey, *Economic Development and Cultural Change*, **33**, 255–298.

Foster, A. D. and Rosenzweig, M. R. (1995) Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture, *Journal of Political Economy*, **103**, 1176–1209.

Foster, A. D. and Rosenzweig, M. R. (2010) Microeconomics of technology adoption, *Annual Review of Economics*, **2**, 395–424.

Giné, X. and Klonner, S. (2008) *Credit Constraints as a Barrier to Technology Adoption by the Poor: Lessons from South Indian Small-Scale Fisheries*, Palgrave Macmillan UK, London, pp. 221–249.

Hamzakaza, P., Katungi, E., Reyes, B., Maredia, M., Muimui, K. and Ojara, M. (2014) Assessing access and adoption of common bean improved varieties in Zambia, PABRA Research Technical Report, Pan-Africa Bean Research Alliance (PABRA).

Hathie, I., Seydie, B., Samaké, L. and Sakho-Jimbira, S. (2017) Ending Rural Hunger: The Case of Senegal, Tech. rep., The Brookings Institution.

Holloway, G., Shankar, B. and Rahman, S. (2002) Bayesian spatial probit estimation: a primer and an application to HYV rice adoption, *Agricultural Economics*, **27**, 383–402.

Isham, J. (2002) The Effect of Social Capital on Fertiliser Adoption: Evidence from Rural Tanzania, *Journal of African Economies*, **11**, 39–60.

Kabunga, N. S., Dubois, T. and Qaim, M. (2012) Heterogeneous information exposure and technology adoption: the case of tissue culture bananas in Kenya, *Agricultural Economics*, **43**, 473–486.

Kabunga, N. S., Dubois, T. and Qaim, M. (2014) Impact of tissue culture banana technology on farm household income and food security in Kenya, *Food Policy*, **45**, 25 – 34.

Karlan, D., Osei, R., Osei-Akoto, I. and Udry, C. (2014) Agricultural Decisions after Relaxing Credit and Risk Constraints *, *The Quarterly Journal of Economics*, **129**, 597–652.

Kassie, M., Jaleta, M. and Mattei, A. (2014) Evaluating the impact of improved maize varieties on food security in Rural Tanzania: Evidence from a continuous treatment approach, *Food Security*, **6**, 217–230.

Katungi, E. M., Larochelle, C., Mugabo, J. R. and Buruchara, R. (2018) The effect of climbing bean adoption on the welfare of smallholder common bean growers in Rwanda, *Food Security*, **10**, 61–79.

Khonje, M., Manda, J., Alene, A. D. and Kassie, M. (2015) Analysis of adoption and impacts of improved maize varieties in Eastern Zambia, *World Development*, **66**, 695 – 706.

Knight, J., Weir, S. and Woldehanna, T. (2003) The role of education in facilitating risk-taking and innovation in agriculture, *The Journal of Development Studies*, **39**, 1–22.

Koussoubé, E. and Nauges, C. (2017) Returns to fertiliser use: Does it pay enough? Some new evidence from Sub-Saharan Africa, *European Review of Agricultural Economics*, **44**, 183–210.

Krishnan, P. and Patnam, M. (2013) Neighbors and Extension Agents in Ethiopia: Who Matters More for Technology Adoption?, *American Journal of Agricultural Economics*, **96**, 308–327.

Langyintuo, A. S. and Mekuria, M. (2008) Assessing the influence of neighborhood effects on the adoption of improved agricultural technologies in developing agriculture, *African Journal of Agricultural and Resource Economics*, **2**, 1–19.

LeSage, J. and Pace, R. (2009) *Introduction to Spatial Econometrics*, Statistics, textbooks and monographs, CRC Press.

LeSage, J. P. (2000) Bayesian estimation of limited dependent variable spatial autoregressive models, *Geographical Analysis*, **32**, 19–35.

Letaa, E., Kabungo, C., Katungi, E., Ojara, M. and Ndunguru, A. (2015) Farm level adoption and spatial diffusion of improved common bean varieties in southern highlands of Tanzania, *African Crop Science Journal*, **23**, 261–277.

Liu, E. M. (2013) Time to change what to sow: Risk preferences and technology adoption decisions of cotton farmers in china, *The Review of Economics and Statistics*, **95**, 1386–1403.

Läpple, D., Holloway, G., Lacombe, D. J. and O'Donoghue, C. (2017) Sustainable technology adoption: a spatial analysis of the Irish Dairy Sector, *European Review of Agricultural Economics*, **44**, 810–835.

Läpple, D. and Kelley, H. (2015) Spatial dependence in the adoption of organic drystock farming in Ireland, *European Review of Agricultural Economics*, **42**, 315–337.

Ma, W. and Abdulai, A. (2016) Does cooperativemembership improve household welfare? Evidence from apple farmers in China, *Food Policy*, **58**, 94–102.

Ma, W., Abdulai, A. and Goetz, R. (2018) Agricultural Cooperatives and Investment in Organic Soil Amendments and Chemical Fertilizer in China, *American Journal of Agricultural Economics*, **100**, 502–520.

Magnan, N., Spielman, D. J., Lybbert, T. J. and Gulati, K. (2015) Leveling with friends: Social networks and Indian farmers' demand for a technology with heterogeneous benefits, *Journal of Development Economics*, **116**, 223 – 251.

Manski, C. F. (1993) Identification of Endogenous Social Effects: The Reflection Problem, *The Review of Economic Studies*, **60**, 531–542.

Marenya, P. P. and Barrett, C. B. (2009) State-conditional Fertilizer Yield Response on Western Kenyan Farms, *American Journal of Agricultural Economics*, **91**, 991–1006.

Mariano, M. J., Villano, R. and Fleming, E. (2012) Factors influencing farmers' adoption of modern rice technologies and good management practices in the Philippines, *Agricultural Systems*, **110**, 41 – 53.

Matuschke, I. and Qaim, M. (2008) Seed market privatisation and farmers' access to crop technologies: The case of hybrid pearl millet adoption in India, *Journal of Agricultural Economics*, **59**, 498–515.

Matuschke, I. and Qaim, M. (2009) The impact of social networks on hybrid seed adoption in India, *Agricultural Economics*, **40**, 493–505.

McArthur, J. W. and McCord, G. C. (2017) Fertilizing growth: Agricultural inputs and their effects in economic development, *Journal of Development Economics*, **127**, 133 – 152.

Minten, B. and Barrett, C. B. (2008) Agricultural technology, productivity, and poverty in Madagascar, *World Development*, **36**, 797 – 822.

Munshi, K. (2004) Social learning in a heterogeneous population: technology diffusion in the indian green revolution, *Journal of Development Economics*, **73**, 185 – 213.

Ntshangase, N. L., Muroyiwa, B. and Sibanda, M. (2018) Farmers' perceptions and factors influencing the adoption of no-till conservation agriculture by small-scale farmers in Zashuke, KwaZulu-Natal Province, *Sustainability*, **10**, 1–16.

Pede, V. O., Areal, F. J., Singbo, A., McKinley, J. and 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, *Agricultural Economics*, **49**, 301–312.

Sall, S., Norman, D. and Featherstone, A. (2000) Quantitative assessment of improved rice variety adoption: the farmer's perspective, *Agricultural Systems*, **66**, 129 – 144.

Tessema, Y. M., Asafu-Adjaye, J., Kassie, M. and Mallawaarachchi, T. (2016) Do neighbours matter in technology adoption? The case of conservation tillage in northwest Ethiopia, *African Journal of Agricultural and Resource Economics*, **11**, 211–225.

Van den Broeck, K. and Dercon, S. (2011) Information flows and social externalities in a tanzanian banana growing village, *The Journal of Development Studies*, **47**, 231–252, pMID: 21506301.

van Meijl, H. and van Tongeren, F. (1998) Trade, technology spillovers, and food production in China, *Weltwirtschaftliches Archiv*, **134**, 423–449.

Verkaart, S., Munyua, B. G., Mausch, K. and Michler, J. D. (2017) Welfare impacts of improved chickpea adoption: A pathway for rural development in Ethiopia?, *Food Policy*, **66**, 50 – 61.

Ward, P. S. and Pede, V. O. (2015) Capturing social network effects in technology adoption: the spatial diffusion of hybrid rice in Bangladesh, *Australian Journal of Agricultural and Resource Economics*, **59**, 225–241.

Wilhelm, S. and de Matos, M. G. (2013) Estimating Spatial Probit Models in R, *The R Journal*, **5**, 130–143.

Wollni, M. and Andersson, C. (2014) Spatial patterns of organic agriculture adoption: Evidence from Honduras, *Ecological Economics*, **97**, 120 – 128.

Wopereis-Pura, M., Watanabe, H., Moreira, J. and Wopereis, M. (2002) Effect of late nitrogen application on rice yield, grain quality and profitability in the Senegal River valley, *European Journal of Agronomy*, **17**, 191 – 198.

Wubeneh, N. G. and Sanders, J. (2006) Farm-level adoption of sorghum technologies in Tigray, Ethiopia, *Agricultural Systems*, **91**, 122 – 134.

Yang, W. and Sharp, B. (2017) Spatial dependence and determinants of dairy farmers' adoption of best management practices for water protection in New Zealand, *Environmental Management*, **59**, 594–603.

Zegeye, T., Tadesse, B. and Tesfaye, S. (2001) Determinants of adoption of improved maize technologies in major maize growing regions in Ethiopia, in *Enhancing the Contribution of Maize to Food Security in Ethiopia: Proceedings of the Second National Maize Workshop of Ethiopia : 12-16 November 2001, Addis Ababa, Ethiopia* (Eds.) M. Nigussie, D. Tanner, S. Twumasi-Afriyie, E. A. R. Organization, I. Maize and W. I. Center,

Ethiopian Agricultural Research Organization, pp. 125–136.
Zhang, S., Sun, Z., Ma, W. and Valentinov, V. (2019) The effect of cooperative membership on agricultural technology adoption in Sichuan, China, *China Economic Review*, **xxx**, xxx.

Appendix

Table 5.8: Comparative descriptive by organization membership: improved seeds adoption

	All			Members			Non-Members		
	Adopters	Non-Adopters	P-values	Adopters	Non-Adopters	P-values	Adopters	Non-Adopters	P-values
Male	0.94 (0.23)	0.93 (0.26)	0.07	0.97 (0.16)	0.91 (0.29)	0.01	0.94 (0.25)	0.93 (0.26)	0.44
Age	53.85 (13.51)	52.83 (13.39)	0.03	52.66 (12.10)	49.14 (11.55)	0.01	54.13 (13.80)	53.04 (13.45)	0.04
Household size	10.83 (5.76)	9.66 (5.22)	<0.01	11.13 (6.06)	11.54 (6.15)	0.53	10.77 (5.70)	9.56 (5.14)	<0.01
Off-farm	0.24 (0.43)	0.28 (0.45)	0.01	0.19 (0.40)	0.39 (0.49)	<0.01	0.25 (0.44)	0.28 (0.45)	0.18
Education	0.61 (0.49)	0.63 (0.48)	0.20	0.56 (0.50)	0.41 (0.49)	<0.01	0.62 (0.48)	0.65 (0.48)	0.21
Area owned	7.26 (9.42)	5.18 (6.63)	<0.01	5.84 (7.09)	5.18 (4.94)	0.31	7.58 (9.86)	5.18 (6.71)	<0.01
Living conditions	0.15 (2.28)	-0.13 (2.25)	<0.01	0.37 (2.50)	-0.17 (1.99)	0.02	0.10 (2.23)	-0.13 (2.26)	0.01
Agricultural index	0.20 (1.43)	-0.06 (1.27)	<0.01	-0.28 (1.65)	-0.27 (1.34)	0.94	0.31 (1.34)	-0.05 (1.27)	<0.01
Distance to road	8.30 (11.20)	10.77 (15.06)	<0.01	13.90 (16.52)	6.39 (8.77)	<0.01	7.02 (9.11)	11.01 (15.30)	<0.01
Distance to market	11.50 (9.44)	14.23 (12.67)	<0.01	14.89 (12.95)	12.05 (9.81)	0.02	10.72 (8.25)	14.35 (12.80)	<0.01
Extension	0.19 (0.39)	0.08 (0.27)	<0.01	0.52 (0.50)	0.33 (0.47)	<0.01	0.11 (0.31)	0.07 (0.25)	<0.01
Credit	0.05 (0.21)	0.02 (0.13)	<0.01	0.10 (0.30)	0.09 (0.29)	0.66	0.03 (0.18)	0.01 (0.12)	<0.01
Rainfall	642.94 (316.50)	715.44 (328.22)	<0.01	580.88 (378.89)	893.88 (340.13)	<0.01	657.19 (298.79)	705.64 (324.78)	<0.01
Clay	19.58 (7.06)	20.19 (7.22)	0.02	23.24 (5.64)	23.37 (5.73)	0.82	18.74 (7.09)	20.01 (7.25)	<0.01
Silt	13.52 (5.69)	13.53 (5.75)	0.96	15.78 (4.58)	15.20 (5.46)	0.29	13.00 (5.79)	13.44 (5.75)	0.05
Floods	0.08 (0.27)	0.05 (0.22)	<0.01	0.09 (0.28)	0.04 (0.21)	0.10	0.07 (0.26)	0.05 (0.22)	0.01
Crop diseases	0.11 (0.31)	0.11 (0.31)	0.83	0.16 (0.37)	0.08 (0.27)	0.01	0.10 (0.30)	0.11 (0.31)	0.32
Rain break	0.36 (0.48)	0.27 (0.44)	<0.01	0.28 (0.45)	0.25 (0.44)	0.54	0.38 (0.48)	0.27 (0.44)	<0.01
Early rain stop	0.38 (0.49)	0.35 (0.48)	0.03	0.35 (0.48)	0.34 (0.47)	0.79	0.39 (0.49)	0.35 (0.48)	0.02
N	1044	3036	4080	195	158	353	849	2878	3727

Table 5.9: Comparative descriptive by organization membership: inorganic fertilizers adoption

	All			Members			Non-Members		
	Adopters	Non-Adopters	P-values	Adopters	Non-Adopters	P-values	Adopters	Non-Adopters	P-values
Male	0.96 (0.20)	0.92 (0.28)	<0.01	0.97 (0.16)	0.88 (0.33)	0.01	0.96 (0.21)	0.92 (0.28)	<0.01
Age	53.12 (13.30)	53.08 (13.49)	0.93	51.51 (11.87)	49.99 (12.22)	0.29	53.46 (13.56)	53.20 (13.53)	0.58
Household size	11.02 (5.96)	9.37 (4.94)	<0.01	11.74 (6.47)	10.20 (4.85)	0.02	10.87 (5.84)	9.34 (4.94)	<0.01
Off-farm	0.27 (0.44)	0.27 (0.45)	0.89	0.25 (0.44)	0.35 (0.48)	0.10	0.27 (0.45)	0.27 (0.44)	0.79
Education	0.57 (0.50)	0.66 (0.47)	<0.01	0.47 (0.50)	0.56 (0.50)	0.13	0.59 (0.49)	0.67 (0.47)	<0.01
Area owned	7.22 (9.54)	4.87 (5.91)	<0.01	6.03 (6.74)	4.29 (4.36)	<0.01	7.47 (10.02)	4.89 (5.97)	<0.01
Living conditions	0.20 (2.26)	-0.21 (2.25)	<0.01	0.42 (2.40)	-0.64 (1.80)	<0.01	0.15 (2.23)	-0.19 (2.26)	<0.01
Agricultural index	0.16 (1.41)	-0.08 (1.25)	<0.01	-0.20 (1.63)	-0.48 (1.17)	0.07	0.24 (1.35)	-0.07 (1.25)	<0.01
Distance to road	9.85 (12.35)	10.30 (15.16)	0.31	12.44 (15.34)	5.60 (8.42)	<0.01	9.30 (11.55)	10.48 (15.33)	0.01
Distance to market	13.39 (11.28)	13.61 (12.37)	0.56	14.05 (12.57)	12.51 (9.10)	0.20	13.25 (10.99)	13.65 (12.48)	0.32
Extension	0.17 (0.38)	0.07 (0.26)	<0.01	0.50 (0.50)	0.27 (0.44)	<0.01	0.10 (0.30)	0.07 (0.25)	<0.01
Credit	0.05 (0.22)	0.01 (0.10)	<0.01	0.12 (0.33)	0.03 (0.17)	<0.01	0.04 (0.19)	0.01 (0.10)	<0.01
Rainfall	739.82 (314.63)	672.97 (330.97)	<0.01	641.57 (376.60)	927.58 (362.53)	<0.01	760.61 (295.91)	663.07 (325.77)	<0.01
Clay	21.10 (6.60)	19.44 (7.43)	<0.01	22.74 (5.71)	24.74 (5.33)	<0.01	20.75 (6.72)	19.23 (7.42)	<0.01
Silt	14.39 (5.50)	13.05 (5.80)	<0.01	15.16 (5.09)	16.46 (4.62)	0.02	14.22 (5.57)	12.91 (5.81)	<0.01
Floods	0.06 (0.24)	0.05 (0.22)	0.16	0.08 (0.27)	0.04 (0.20)	0.15	0.06 (0.24)	0.05 (0.22)	0.39
Crop diseases	0.16 (0.37)	0.08 (0.27)	<0.01	0.15 (0.36)	0.06 (0.24)	0.01	0.16 (0.37)	0.08 (0.27)	<0.01
Rain break	0.30 (0.46)	0.29 (0.45)	0.45	0.22 (0.42)	0.39 (0.49)	<0.01	0.31 (0.46)	0.28 (0.45)	0.05
Early rain stop	0.34 (0.47)	0.37 (0.48)	0.05	0.30 (0.46)	0.45 (0.50)	0.01	0.34 (0.47)	0.36 (0.48)	0.24
Num. obs.	1460	2620	4080	255	98	353	1205	2522	3727

Table 5.10: Spatial Durbin Probit Models estimates with W_P

	Improved seeds		Mineral fertilizers	
	X	WX	X	WX
Intercept	-0.566 (0.286)**		-1.233 (0.269)***	
Organization	0.434 (0.091)***	0.417 (0.150)***	0.428 (0.103)***	0.596 (0.160)***
Male	0.029 (0.098)	0.113 (0.202)	0.219 (0.106)**	0.580 (0.195)***
Age	0.003 (0.002)	-0.002 (0.004)	0.004 (0.002)**	-0.011 (0.003)***
Household size	0.012 (0.005)**	-0.004 (0.007)	0.022 (0.005)***	-0.011 (0.007)
Off-farm	0.045 (0.064)	-0.083 (0.094)	0.032 (0.062)	-0.130 (0.091)
Education	-0.003 (0.054)	-0.010 (0.084)	-0.114 (0.056)**	0.089 (0.087)
Area owned	0.012 (0.004)***	-0.001 (0.006)	0.017 (0.004)***	-0.013 (0.006)**
Living index	0.010 (0.016)	-0.004 (0.021)	0.127 (0.016)***	-0.081 (0.021)***
Agricultural index	0.050 (0.024)**	-0.038 (0.036)	0.115 (0.026)***	0.006 (0.036)
Distance to road	-0.039 (0.028)	0.035 (0.028)	-0.005 (0.028)	0.006 (0.028)
Distance to market	0.083 (0.024)***	-0.091 (0.025)***	0.010 (0.023)	-0.012 (0.023)
Extension	0.097 (0.094)	0.161 (0.127)	0.326 (0.095)***	-0.229 (0.131)*
Credit	0.252 (0.134)*	-0.295 (0.260)	0.551 (0.155)***	-0.215 (0.277)
Rainfall	0.002 (0.002)	-0.002 (0.002)	-0.001 (0.002)	0.001 (0.002)
Clay	-0.004 (0.015)	-0.001 (0.016)	-0.007 (0.016)	0.013 (0.017)
Silt	0.048 (0.019)**	-0.035 (0.021)*	0.011 (0.019)	0.000 (0.021)
Floods	0.046 (0.121)	0.251 (0.162)	0.086 (0.120)	-0.045 (0.159)
Crop diseases	0.138 (0.098)	-0.229 (0.132)*	0.324 (0.094)***	0.193 (0.127)
Rain break	0.155 (0.064)**	-0.024 (0.089)	-0.044 (0.064)	0.041 (0.087)
Early rain stop	0.040 (0.064)	-0.004 (0.084)	-0.085 (0.065)	-0.059 (0.084)
ρ	0.553 (0.025)***		0.615 (0.021)***	
Num. obs.	4245		4245	

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 5.11: Effect estimates of Spatial Durbin Probit models with W_P: Improved seeds

	Direct Effects			Indirect Effects			Total Effects		
	Lower 5%	Mean	Upper 95%	lower 5%	Mean	Upper 95%	lower 5%	Mean	Upper 95%
Organization	0.079	0.119	0.163	0.083	0.128	0.178	0.164	0.247	0.338
Male	-0.038	0.007	0.051	-0.040	0.008	0.055	-0.078	0.015	0.105
Age	0.000	0.001	0.002	0.000	0.001	0.002	0.000	0.002	0.003
Household size	0.001	0.003	0.006	0.001	0.004	0.006	0.002	0.007	0.011
Off-farm	-0.015	0.013	0.041	-0.017	0.014	0.045	-0.032	0.026	0.086
Education	-0.025	-0.000	0.025	-0.027	-0.000	0.026	-0.052	-0.001	0.050
Area owned	0.002	0.003	0.005	0.002	0.004	0.005	0.003	0.007	0.010
Living index	-0.004	0.003	0.010	-0.004	0.003	0.011	-0.009	0.006	0.021
Agricultural index	0.002	0.014	0.025	0.002	0.015	0.027	0.005	0.028	0.052
Distance to road	-0.023	-0.011	0.002	-0.025	-0.011	0.002	-0.047	-0.022	0.003
Distance to market	0.012	0.023	0.034	0.012	0.024	0.037	0.024	0.047	0.069
Extension	-0.016	0.026	0.071	-0.018	0.028	0.077	-0.033	0.055	0.147
Credit	0.009	0.070	0.130	0.010	0.074	0.141	0.019	0.144	0.267
Rainfall	-0.000	0.000	0.001	-0.000	0.000	0.001	-0.001	0.001	0.003
Clay	-0.008	-0.001	0.006	-0.009	-0.001	0.006	-0.016	-0.003	0.012
Silt	0.005	0.013	0.022	0.005	0.014	0.023	0.010	0.027	0.045
Floods	-0.042	0.013	0.069	-0.044	0.014	0.073	-0.086	0.026	0.142
Crop diseases	-0.006	0.038	0.081	-0.006	0.041	0.089	-0.012	0.079	0.170
Rain Break	0.012	0.042	0.072	0.014	0.046	0.078	0.026	0.088	0.148
Early rain stop	-0.018	0.011	0.040	-0.018	0.012	0.043	-0.035	0.023	0.083

Table 5.12: Effect estimates of Spatial Durbin Probit models with W_P: Mineral fertilizers

	Direct Effects			Indirect Effects			Total Effects		
	Lower 5%	Mean	Upper 95%	lower 5%	Mean	Upper 95%	lower 5%	Mean	Upper 95%
Organization	0.079	0.130	0.181	0.103	0.174	0.248	0.184	0.304	0.426
Male	0.014	0.067	0.121	0.018	0.089	0.164	0.032	0.156	0.285
Age	0.000	0.001	0.002	0.000	0.002	0.003	0.001	0.003	0.005
Household size	0.004	0.007	0.009	0.006	0.009	0.013	0.010	0.016	0.022
Off-farm	-0.021	0.010	0.040	-0.027	0.013	0.054	-0.047	0.022	0.094
Education	-0.062	-0.035	-0.008	-0.084	-0.046	-0.011	-0.145	-0.081	-0.019
Area owned	0.004	0.005	0.007	0.005	0.007	0.010	0.008	0.012	0.017
Living index	0.031	0.039	0.047	0.040	0.052	0.065	0.071	0.091	0.111
Agricultural index	0.022	0.035	0.048	0.029	0.047	0.066	0.051	0.082	0.114
Distance to road	-0.016	-0.002	0.012	-0.021	-0.002	0.016	-0.037	-0.004	0.029
Distance to market	-0.009	0.003	0.015	-0.012	0.004	0.020	-0.021	0.006	0.034
Extension	0.050	0.098	0.144	0.066	0.131	0.197	0.118	0.229	0.340
Credit	0.096	0.170	0.249	0.125	0.227	0.339	0.223	0.398	0.587
Rainfall	-0.001	-0.000	0.000	-0.002	-0.000	0.001	-0.003	-0.001	0.001
Clay	-0.010	-0.002	0.006	-0.013	-0.003	0.008	-0.023	-0.005	0.014
Silt	-0.006	0.004	0.013	-0.008	0.005	0.018	-0.015	0.008	0.031
Floods	-0.032	0.026	0.086	-0.043	0.034	0.115	-0.075	0.060	0.200
Crop diseases	0.053	0.099	0.146	0.067	0.133	0.199	0.120	0.232	0.343
Rain Break	-0.045	-0.013	0.018	-0.060	-0.017	0.024	-0.104	-0.030	0.043
Early rain stop	-0.058	-0.026	0.008	-0.079	-0.034	0.010	-0.136	-0.060	0.018

Table 5.13: Spatial Durbin Probit Models estimates with W_D

	Improved seeds		Mineral fertilizers	
	X	WX	X	WX
Intercept	-1.001 (0.536)*		-0.735 (0.540)	
Organization	0.464 (0.087)***	0.839 (0.267)***	0.494 (0.092)***	0.453 (0.264)*
Male	0.033 (0.099)	0.307 (0.329)	0.195 (0.102)*	0.820 (0.342)**
Age	0.003 (0.002)*	0.001 (0.007)	0.006 (0.002)***	-0.023 (0.007)***
Household size	0.012 (0.005)***	-0.014 (0.010)	0.018 (0.005)***	0.002 (0.010)
Off-farm	0.017 (0.058)	-0.017 (0.149)	0.033 (0.058)	-0.201 (0.146)
Education	-0.049 (0.053)	0.220 (0.146)	-0.163 (0.051)***	0.357 (0.144)**
Area owned	0.013 (0.003)***	0.006 (0.008)	0.017 (0.003)***	-0.024 (0.008)***
Living index	0.006 (0.014)	0.042 (0.025)*	0.117 (0.014)***	-0.061 (0.027)**
Agricultural index	0.044 (0.024)*	-0.030 (0.056)	0.115 (0.025)***	-0.057 (0.057)
Distance to road	-0.007 (0.006)	0.006 (0.007)	0.004 (0.006)	-0.004 (0.007)
Distance to market	-0.001 (0.006)	-0.004 (0.006)	-0.020 (0.005)***	0.020 (0.006)***
Extension	0.153 (0.079)*	-0.193 (0.206)	0.340 (0.088)***	-0.291 (0.215)
Credit	0.263 (0.133)**	-1.548 (0.650)**	0.567 (0.154)***	0.047 (0.658)
Rainfall	-0.000 (0.001)	0.000 (0.001)	0.001 (0.001)	-0.000 (0.001)
Clay	-0.016 (0.010)	0.007 (0.014)	-0.010 (0.010)	0.009 (0.014)
Silt	0.045 (0.012)***	-0.037 (0.016)**	0.010 (0.012)	-0.004 (0.016)
Floods	0.073 (0.112)	0.361 (0.199)*	0.034 (0.115)	0.108 (0.208)
Crop diseases	0.244 (0.085)***	-0.670 (0.187)***	0.389 (0.081)***	0.336 (0.179)*
Rain break	0.210 (0.056)***	-0.054 (0.127)	-0.055 (0.057)	0.059 (0.125)
Early rain stop	0.081 (0.056)	-0.088 (0.102)	-0.107 (0.059)*	-0.025 (0.105)
ρ	0.675 (0.037)***		0.748 (0.027)***	
Num. obs.	4245		4245	

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 5.14: Effect estimates of Spatial Durbin Probit models with W_D: Improved seeds

	Direct Effects			Indirect Effects			Total Effects		
	Lower 5%	Mean	Upper 95%	lower 5%	Mean	Upper 95%	lower 5%	Mean	Upper 95%
Organization	0.087	0.128	0.169	0.165	0.256	0.357	0.258	0.384	0.517
Male	-0.036	0.009	0.053	-0.072	0.018	0.108	-0.108	0.028	0.160
Age	0.000	0.001	0.002	0.000	0.002	0.004	0.000	0.003	0.005
Household size	0.001	0.003	0.005	0.002	0.007	0.011	0.004	0.010	0.017
Off-farm	-0.022	0.005	0.031	-0.043	0.009	0.062	-0.064	0.014	0.093
Education	-0.038	-0.013	0.010	-0.079	-0.027	0.020	-0.116	-0.040	0.030
Area owned	0.002	0.004	0.005	0.004	0.007	0.010	0.006	0.010	0.015
Living index	-0.005	0.002	0.008	-0.010	0.003	0.016	-0.014	0.005	0.025
Agricultural index	0.001	0.012	0.023	0.002	0.024	0.048	0.003	0.036	0.070
Distance to road	-0.005	-0.002	0.001	-0.010	-0.004	0.002	-0.015	-0.006	0.003
Distance to market	-0.003	-0.000	0.002	-0.006	-0.001	0.004	-0.009	-0.001	0.006
Extension	0.006	0.043	0.078	0.011	0.086	0.161	0.017	0.128	0.237
Credit	0.012	0.073	0.135	0.023	0.146	0.277	0.035	0.219	0.405
Rainfall	-0.000	-0.000	0.000	-0.001	-0.000	0.000	-0.001	-0.000	0.001
Clay	-0.009	-0.004	0.000	-0.019	-0.009	0.000	-0.028	-0.013	0.000
Silt	0.007	0.013	0.018	0.014	0.025	0.038	0.021	0.038	0.056
Floods	-0.029	0.020	0.070	-0.058	0.039	0.142	-0.087	0.059	0.209
Crop diseases	0.028	0.068	0.107	0.055	0.136	0.225	0.085	0.203	0.327
Rain Break	0.033	0.058	0.085	0.062	0.116	0.177	0.097	0.174	0.256
Early rain stop	-0.004	0.022	0.047	-0.007	0.043	0.096	-0.011	0.065	0.142

Table 5.15: Effect estimates of Spatial Durbin Probit models with W_D: Mineral fertilizers

	Direct Effects			Indirect Effects			Total Effects		
	Lower 5%	Mean	Upper 95%	lower 5%	Mean	Upper 95%	lower 5%	Mean	Upper 95%
Organization	0.094	0.136	0.178	0.260	0.391	0.539	0.358	0.527	0.708
Male	0.008	0.054	0.100	0.023	0.155	0.294	0.030	0.208	0.389
Age	0.001	0.002	0.002	0.002	0.005	0.007	0.003	0.006	0.010
Household size	0.003	0.005	0.007	0.007	0.014	0.021	0.010	0.019	0.028
Off-farm	-0.018	0.009	0.036	-0.051	0.025	0.103	-0.069	0.034	0.139
Education	-0.068	-0.045	-0.023	-0.205	-0.130	-0.062	-0.272	-0.175	-0.085
Area owned	0.003	0.005	0.006	0.009	0.014	0.019	0.013	0.019	0.025
Living index	0.026	0.032	0.039	0.068	0.093	0.122	0.097	0.126	0.159
Agricultural index	0.021	0.032	0.043	0.057	0.092	0.130	0.079	0.123	0.172
Distance to road	-0.002	0.001	0.004	-0.005	0.003	0.011	-0.007	0.004	0.015
Distance to market	-0.008	-0.006	-0.003	-0.024	-0.016	-0.008	-0.032	-0.021	-0.012
Extension	0.056	0.094	0.136	0.151	0.271	0.403	0.208	0.365	0.532
Credit	0.089	0.155	0.227	0.251	0.446	0.675	0.342	0.601	0.898
Rainfall	-0.000	0.000	0.000	-0.000	0.000	0.001	-0.000	0.001	0.002
Clay	-0.007	-0.003	0.002	-0.021	-0.008	0.005	-0.029	-0.011	0.007
Silt	-0.003	0.003	0.008	-0.008	0.008	0.025	-0.011	0.011	0.033
Floods	-0.045	0.009	0.061	-0.131	0.026	0.175	-0.174	0.035	0.232
Crop diseases	0.070	0.108	0.145	0.194	0.310	0.432	0.267	0.417	0.570
Rain Break	-0.041	-0.015	0.012	-0.120	-0.044	0.033	-0.160	-0.060	0.045
Early rain stop	-0.057	-0.030	-0.003	-0.167	-0.086	-0.009	-0.224	-0.116	-0.011

Chapter 6

Climate variability and farm inefficiency: a spatial stochastic frontier analysis of Senegalese agriculture

K. Christophe Adjin, Christian H. C. A. Henning

Abstract

This paper aimed to analyse Senegalese farmers' technical efficiency in the context of climate variability and spatial heterogeneity. To achieve this, firstly using simulated data, we evaluated the newly developed spatial stochastic frontier estimation technique based on skew-normal distributions. Secondly, using cross-sectional survey data we conducted an empirical analysis for 4423 Senegalese farm households. Simulation results show that the estimation approach used is appropriate and produces consistent results with large sample sizes, although it might suffer from a "starting values" problem. Empirical findings reveal that agricultural production in Senegal mostly depends on the allocated area and it is highly affected by climatic factors such as rainfall and temperature. Moreover, within a radius of 4 km the technical efficiency of farms appears to be significantly affected by unobserved spatial features. Furthermore, this farm technical efficiency can on average be increased by 20%, when accounting for spatial heterogeneity.

Keywords: Climate variability, Farm efficiency, Spatial heterogeneity, Senegal.

JEL Codes: Q54, C21, D24.

6.1 Introduction

In Senegal, agriculture remains an important sector of the economy, and it accounts for approximately 32% of the country's total employment, more than 16% of the national GDP (World Bank, 2019) and 21%[1] of total exports (Republique du Senegal, 2018). However, the sector is challenged by many factors, and among those, climate variability. Senegalese agriculture is mostly rain-fed and less than 1% of agricultural land is under irrigation. The sector is therefore highly vulnerable to rainfall variability (D'Alessandro *et al.*, 2015). During the period 1941-2000, rainfall in the country has been characterized by a great annual irregularity and the shortening of the duration of the rainy season. Also, from year to year, the amount of precipitation substantially varies both temporally and spatially (Ba, 2006). According to Jalloh *et al.* (2013), the country will experience in the coming decades an increase in rainfall variability and droughts due to climate change, especially in the eastern part. The impacts of such erratic weather and climate shifts in Senegal will lead to productivity losses in the agricultural sector (Jalloh *et al.*, 2013). Moreover, such climate conditions also pose serious challenges for the cereals production sub-sector, which constitutes primary sources of food for rural populations. According to Jalloh *et al.* (2013) changing climate conditions combined with population growth, could lead to a 30% reduction in per capita cereals production by 2025.

The presence of erratic weather and climate conditions implies that Senegalese farmers would have to integrate not only conventional production factors but also climatic risks factors (such as precipitation, temperature, etc.) into their production systems. For instance, the study of Ba (2006) has shown that close droughts due to climate change has led to a transformation of not only some productions systems but also the decline or disappearance of some crops. Globally, climate change and its impact on agricultural productivity have gained increased attention during recent decades among research communities (e.g. Hughes *et al.* (2011); Lachaud *et al.* (2017)). As observed by Mulwa and Kabubo-Mariara (2017), despite the increasing number of climate change and variability studies and efficiency studies, there is a lack of literature linking climate change and variability to farm-level efficiency in Africa. Notable exceptions are the papers of Sherlund *et al.* (2002) in Cote d'Ivoire, Mulwa and Kabubo-Mariara (2017) in Kenya and Oyekale (2012) in Nigeria. This study, therefore aims to analyse agricultural technical efficiency in climate variability conditions in the Senegalese context, and it contributes to the literature in several

[1]Calculation is based on values of exportation figures in Republique du Senegal (2018) and includes exports of fishes products, groundnut products, cotton and cotton fabrics.

ways.

First, few studies have analysed the farm efficiencies of Senegalese agriculture. Furthermore, previous studies have focused on groundnut farming (Thiam and Bravo-Ureta, 2003), rice sub-sector (Diagne *et al.*, 2013; Ngom *et al.*, 2016; Seck, 2017), vegetable sub-sector (Dedehouanou, 2014), or some selected crops (Okuyama *et al.*, 2017). Our analysis wants to take the agricultural sector as a whole. Moreover, although several studies have shown the significant effect that climatic factors can have on farmers technical efficiencies, none of these previous studies has really integrated them into their modelling. We, therefore, incorporated in our designated stochastic frontier approach models climatic factors namely rainfall, temperature, and their anomalies. Furthermore, to take into account spatiality and understand how its unobserved features affect generally farm production and technical efficiency, we took advantage of a new approach developed in spatial econometrics and apply it to the Senegalese context. Finally, we extended de Graaff (2020) simulation works in the spatial stochastic frontier field, primarily by conducting simulations for the spatial lag models, and secondly by generating and varying (via spatial correlation coefficients) the spatial weight matrices used in the simulation analysis.

Simulation results confirm that the estimation method designed by de Graaff (2020) based on the skew-normal distribution of errors in the stochastic production frontier is appropriate and can lead to consistent estimates. Although the Maximum Likelihood estimation procedure is sensitive to the starting values, notably the sign of inefficiency dispersion coefficient. Empirical results revealed that farmers technical efficiency in Senegal is affected by both climatic features and spatiality. These results have important policy implications when designing adaptation strategies to climate change.

The paper is divided into six sections, including this introduction. The second section presents a brief literature on spatial stochastic frontier models. The third section describes the approach used and the estimation technique. Section 6.4 presents the simulation results, while section 6.5 presents and discusses the empirical results. Finally, section 6.6 presents the conclusions and explores some policy implications.

6.2 Literature on spatial stochastic frontier

Recently, in the field of productivity and efficiency analysis, analysts have tried to incorporate spatial interactions into frontier models, resulting in some novel spatial frontier models. Based on this literature and assuming some spatial interactions between decision-making units, one could distinguish four main types of models designed to introduce spatial interactions into the stochastic frontier approach. These models are (i) the spatial auto-regressive stochastic frontier (SAR), (ii) the spatial lag on the exogenous variable stochastic frontier (SLX), (iii) the spatial error stochastic frontier (SEM) and (iv) the spatial inefficiency stochastic frontier (SIM).

The SAR model integrates the spatial dependence into the dependent variable, rendering this latter auto-regressive. It has been widely used by modellers in many fields and contexts. Regarding the efficiency analysis field, although few studies are known, this specification is the one of most interest to researchers. Barrios and Lavado (2010) used the SAR-type stochastic frontier model accounting for spatial externalities. They applied it to data sets from the Philippines to demonstrate the importance of spatial components, and to derive unbiased estimations of technical efficiency. To evaluate the impact of programs that aim to improve the productivity of firms, this SAR-type model was also used by Affuso (2010) in Tanzania to analyse the efficiencies of matched subsamples of treated and non-treated farmers. Using this spatial autoregressive model, Glass *et al.* (2013) extended the decomposition of the standard factor productivity growth to incorporate direct and indirect impacts of units, introducing at the same time the concept of efficiency spillovers. The authors applied their specification to a panel data of 40 European countries from 1995 to 2008. Glass *et al.* (2014) applied the same approach in a cost frontier model to highlight the importance of efficiency spillovers in the U.S. manufacturing sector. Han *et al.* (2016) by allowing endogenous interaction effects in the frontier model, used data from 21 OECD countries from 1960 to 2001, to derive the spillover effects of public capital stock. Similarly, Ramajo and Hewings (2018) used a SAR model of stochastic frontier approach to estimate the technical efficiency of 120 European Union regions over the period 1995-2007. Their study revealed a strong geographic pattern of regional efficiency showing productivity convergence of European regions during the same period.

The spatial lag of exogenous variable, like the spatial lag model of X (SLX) defined by LeSage and Pace (2009), characterize a link between an output of a particular decision-making unit and the inputs of its neighbours. This model intends to explain the spatial influence of neighbours' input values (or the indirect flow of resources) on a given unit's output. In the context of efficiency analysis, Adetutu *et al.* (2015) applied such a model to account for local spatial dependence and to shift the production frontier.

The spatial error stochastic frontier is similar to the spatial error model of LeSage and Pace (2009) and unlike the previous models, it is designed to bring out the spatial heterogeneity between production units. Druska and Horrace (2004) used such specification, by including spatial auto-regressive disturbances into the classical SFA. Using generalized moments method on a panel data, they estimated time-invariant inefficiencies and concluded to the existence of spatial correlations into data and this latter affects the magnitude and variability of the production function and the estimated technical efficiencies.

The spatial inefficiency stochastic frontier model is an auto-regressive specification concerning the inefficiency error term of the composite error. This specification illustrates the spatial correlation between the levels of efficiency of neighbouring units. To account for the possible unknown geographical variation of the outputs of farms in Brazil, Schmidt *et al.* (2009) included in their model, a latent spatial structure in the inefficiency error term. Their findings showed that standards models induce significantly different inefficiencies across units. Tonini and Pede (2011) also incorporated spatial dependency in the inefficiency term to measure total factor productivity of European agriculture from 1993 – 2006 in 29 countries. They found that not allowing for spatial dependency underestimates the cumulated technical inefficiency changes. Areal *et al.* (2012) used such specification in a Bayesian setting to show spatial dependence in technical efficiency in a panel data of dairy farms in England and Wales. More recently Pede *et al.* (2018) and Skevas (2020) followed the approach of Areal *et al.* (2012) and applied it respectively to irrigated and rainfed agroecosystem rice farming in the Philippines and to Dutch dairy farms. Fusco and Vidoli (2013) applied the same model to a cross-sectional data of wine industries in Italy and demonstrated the uniform and strong spatial dependence between neighbouring units. Tsionas and Michaelides (2016) also used this same stochastic frontier model with the decomposition of inefficiency into an idiosyncratic and a spatial spillover component. They applied their method to a production data of Italian regions over the period 1970 - 1993. Carvalho (2018) proposed a spatial Bayesian random effects stochastic frontier model that allows for unobserved heterogeneity and spillovers between firms' efficiencies. de Graaff (2020) estimated the SIM model after assuming a skew-normal distribution for the composite error term.

From these four main models, several combinations could be made to form different types of spatial stochastic frontiers. Pavlyuk (2013), for instance, developed a full model with all types of spatial interaction. However, its estimation is challenging due to identification and computation issues. After some restrictions on spatial parameters, Pavlyuk (2011, 2013) used data sets of the regional tourism markets in the Baltic States and the European airports to demonstrate significant spatial dependencies. Orea and Álvarez (2019) developed a stochastic frontier model that allows for cross-sectional spatial correlation in both the noise and inefficiency terms, and that can be estimated by maximum likelihood and non-linear least squares. Glass *et al.* (2016) considered a spatial Durbin stochastic frontier model which combines the SAR and SLX specifications. They applied it respectively to European countries and demonstrated the asymmetry between efficiency spillovers to and from European countries.

6.3 Spatial stochastic production frontier approach

6.3.1 The model

From the seminal works of Aigner *et al.* (1977) and Meeusen and van Den Broeck (1977), the standard stochastic production frontier is defined as:

$$y_i = f\left(X_i, \beta\right) \exp\left(v_i - u_i\right), \tag{6.1}$$

where y_i is the observed output of the farm unit $i\,(i = 1, 2, \ldots.., N)$, X_i is a vector of inputs, $f\left(X_i, \beta\right)$ is the production function, with β as the parameters to be estimated; v_i is a two-sided stochastic term that accounts for statistical noise, u_i is a non-negative stochastic term representing farm inefficiency. In this model, the possible production y_i, is bounded above by the stochastic quantity $f\left(X_i, \beta\right) \exp\left(v_i\right)$ that consists of a deterministic part $f\left(X_i, \beta\right)$ common to all farms and a farm-specific part $\exp\left(v_i\right)$ that captures the effect of random shocks. Errors v_i are assumed to be independently and identically distributed as a normal distribution $N^{+}(0, \sigma_v^2)$, independent of u_i, which are assumed to be non-negative with either a half-normal distribution (Aigner *et al.*, 1977), a truncated normal distribution (Stevenson, 1980), an exponential distribution (Meeusen and van Den Broeck, 1977) or a gamma distribution (Greene, 1990). Assuming a Cobb–Douglas production function for the output production y_i, a logarithmic specification of the function 6.1 gives in vector notation:

$$\ln\left(y\right) = \ln\left(X\right) + v - u. \tag{6.2}$$

Following previous literature, spatial interactions can be incorporated into this standard SFA in two main ways, through the dependent variable resulting in the so-called spatial auto-regressive stochastic frontier model (SAR-SFA), and through the errors giving the spatial error stochastic frontier model (SEM-SFA). The SAR-SFA model can be written as:

$$\ln(y) = \rho W \ln(y) + \ln(X)\beta + v - u, \tag{6.3}$$

where ρ is the spatial lag parameter and W is the spatial weight matrix, and y, X, β, v, and u are as defined previously. Furthermore, the SEM-SFA can be specified as:

$$\ln(y) = \ln(X)\beta + \epsilon, \quad \text{with} \quad \epsilon = v - u = \lambda W \epsilon + \tilde{\epsilon}, \tag{6.4}$$

where λ is the spatial error lag parameter, and W, y, X, β, v, and u are as defined previously.

6.3.2 Estimation of parameters

To estimate parameters of the standard SFA with the maximum likelihood technique, one usually assumes that v follows a normal distribution ($v \sim N(0, \sigma_v^2)$) and u a half-normal distribution ($u \sim N^+(0, \sigma_u^2)$). Therefore, the marginal density function of ϵ is given by:

$$f(\epsilon) = \frac{2}{\sigma}\phi\left(\frac{\epsilon}{\sigma}\right)\mathit{\Phi}\left(-\frac{\epsilon\gamma}{\sigma}\right), \tag{6.5}$$

where $\sigma^2 = \sigma_u^2 + \sigma_v^2$, $\gamma = \frac{\sigma_u}{\sigma_v}$, $\mathit{\Phi}(.)$ is the standard normal cumulative distribution function, and $\phi(.)$ is the standard normal probability density functions. From equation 6.5 the log likelihood function for N farms as proposed by (Aigner *et al.*, 1977) is:

$$\ln L = -\frac{N}{2}\ln\left(\frac{\pi\sigma^2}{2}\right) + \sum_i \ln\Phi\left(-\frac{\epsilon\gamma}{\sigma}\right) - \frac{1}{2\sigma^2}\sum_i \epsilon^2. \tag{6.6}$$

Once the error in the designed stochastic frontier exhibits a more complex structure, estimations of the parameters become cumbersome. Therefore, for the estimation of the spatial models, we adopted a skew-normal distribution approach proposed by de Graaff (2020) which enables us to straightforwardly estimate models 6.3 and 6.4 using the maximum likelihood technique. In the skew-normal approach, the composite error ϵ can be rewritten as a sum of a normal and a truncated normal distributions (de Graaff, 2020) which gives:

$$\epsilon = \delta \mid \mu \mid + \sqrt{1-\delta^2}\nu, \tag{6.7}$$

where μ and ν are independent variables $N(0,1)$ and $\delta \in (-1,1)$. The stochastic variable ϵ is generated by means of convolution, however ϵ can also be obtained by conditioning:

$$\epsilon = (\nu \mid \mu > 0)\,, \tag{6.8}$$

where (μ,ν) follows a bivariate normal distribution with δ as a correlation coefficient. It can be shown that both equations 6.7 and 6.8 lead to the same skew-normal density function:

$$\epsilon Z \sim SN\,(\alpha) = 2\phi\,(x)\,\Phi\,(\alpha x)\,, \tag{6.9}$$

where α is the skewness parameter that determines the shape of the density function[2]. As shown by de Graaff (2020), if $\epsilon \sim SN\,(\alpha)$ and $\ln\,(y) = \ln\,(X)\,\beta$, then the affine transformation of $\ln\,(y) \sim SN\,(\ln\,(X)\,\beta, \sigma^2, \alpha)$ can be expressed as:

$$\epsilon \sim 2\phi\left(\ln\,(y) - \ln\,(X)\,\beta; \sigma^2\right)\Phi\,(\alpha\,(\ln\,(y) - \ln\,(X)\,\beta))\,. \tag{6.10}$$

In this case, $\ln\,(X)\,\beta$, σ^2, and α could be defined as a location parameter, a scale parameter and a skewness parameter respectively. The relation between equations 6.2 and 6.10 can be defined by stating $\ln\,(y) - \ln\,(X)\,\beta = \pi\,(v \mid u > 0) = \epsilon$, where

$$\epsilon = \begin{bmatrix} \mu \\ \nu \end{bmatrix} \sim N\,(0, \Omega^*)\,,\ \Omega^* = \begin{bmatrix} 1 & \delta' \\ \delta' & \sigma^2 \end{bmatrix}, \tag{6.11}$$

and where $\alpha = \delta^2\sqrt{1-\delta^2}$, $\delta = \sigma_u$, and $\sqrt{1-\delta^2}\sigma_\epsilon = \sigma_v$. The latter equality denotes the intrinsic relation between u and v which is implicit in specification 6.2. It important to note that specification 6.2 only holds when $\delta < 0$. From density equation 6.10, the log likelihood for N observations can be specified as (de Graaff, 2020):

$$\ln L = -\frac{N}{2}\ln\pi - \frac{N}{2}\ln\left(\sigma^2\right) - \frac{1}{2}\epsilon'\epsilon + \sum_i \ln\left(2\Phi\,(\alpha\epsilon_i)\right), \tag{6.12}$$

where ϵ_i is the i_{th} observation of the vector ϵ. As argued by de Graaff (2020), a skew-normal distribution allows the use of a single error term instead of a composite one. This process has several advantages when working with multivariate distributions, and the interpretation of the parameters seems as well more intuitive (using scale, location, and skewness parameters). However, a disadvantage is the need to use a re-parametrization of the parameters in order to estimate them properly. The log likelihood function specified in equation 6.12 can be straightforwardly

[2] See de Graaff (2020); Azzalini (1985, 2005); Azzalini and Valle (1996); Azzalini and Capitanio (1999); Arellano-Valle and Azzalini (2006) and Arellano-Valle and Azzalini (2008) for more literature on the skew-normal distributions.

adapted to the spatial lag stochastic frontier and spatial error stochastic frontier model (de Graaff, 2020). The log likelihood for the spatial lag stochastic frontier for N units is expressed as:

$$\ln L = -\frac{N}{2}\ln\pi - \frac{N}{2}\ln\left(\sigma^2\right) + \ln \mid I - \rho W \mid -\frac{1}{2}\epsilon^{'}\epsilon + \sum_i \ln\left(2\Phi\left(\alpha\epsilon_i\right)\right), \tag{6.13}$$

where $\epsilon = \frac{1}{\sigma}\left[\ln\left(y\right) - \ln\left(X\right)\beta\right]$ and I the identity matrix. The log likelihood for a stochastic frontier model with spatial dependence in the error term can be defined as:

$$\ln L = -\frac{N}{2}\ln\pi - \frac{N}{2}\ln\left(\sigma^2\right) + \ln \mid I - \lambda W \mid \frac{1}{2}\epsilon^{'}\epsilon + \sum_i \ln\left(2\Phi\left(\alpha\epsilon_i\right)\right), \tag{6.14}$$

where $\epsilon = \frac{1}{\sigma}\left[I - \lambda W\right]\left[\ln\left(y\right) - \ln\left(X\right)\beta\right]$, and I, N, and w are defined as previously.

6.3.3 Measurement of technical efficiency

After estimating the likelihood functions in equations 6.13 and 6.14, the obtained parameters $\hat{\epsilon}$, $\hat{\delta}$ using $\hat{\alpha}$ and $\hat{\sigma}$ can therefore be used to draw simulations from $u \mid \epsilon$ and derive the expectation for each farm. Following de Graaff (2020), we will use the generic formula for $u \mid \epsilon$ defined by Domınguez-Molina *et al.* (2003) as a normal distribution with mean and variance equal to:

$$\text{Mean} = \frac{1}{\frac{\delta^2}{\sigma^2}+1}\frac{\delta}{\sigma^2}\epsilon, \quad \text{Variance} = \frac{1}{\frac{\delta^2}{\sigma^2}+1}.$$

6.4 Monte Carlo simulation

6.4.1 Simulation procedure

Following a similar approach as de Graaff (2020), we set up a simulation procedure with a base cross-section data generating process defined as:

$$\begin{aligned} Y &= 1 + \ln(A) + \epsilon, \\ \text{where } \epsilon &= \delta \mid u \mid + \sqrt{1-\delta^2} v \\ \text{and } \ln(A) &\sim U(5, 14) \\ u &\sim N(0, 0.3) \\ v &\sim N(0, 0.3). \end{aligned} \tag{6.15}$$

Here, v and u are drawn using a normal distribution, the production input is drawn using a uniform distribution with associated coefficient β_1 equals to 1, and the constant term (denoted β_0) is also set to 1. From this base specification we randomly generated data sets by considering alternative scenarios (see table 6.1). The experiments were then performed by varying the number of observations N (250 and 1000), and the values of the coefficient of correlation δ (-0.2, -0.5, and -0.8). However, contrary to de Graaff (2020) we also vary the values of the spatial dependence coefficients λ and ρ (0.2, 0.5, and 0.8), by considering a randomly generated spatial weight matrix of the type, inverse geographic distance, using latitudes and longitudes drawn from a uniform distribution ($U(0, 20)$). Note that the weights matrices are all row-standardized, and their diagonals set to 0. For each scenario, 1,000 replications are made.

Table 6.1: Monte Carlo simulation scenarios

Variables	Description	Base scenario	Alternative scenarios
N	Number of observations	250	1000
λ , ρ	Spatial correlation	0.2	0.5, 0.8
δ	Inefficiency coefficient	-0.2	-0.5, -0.8

For each generated data, we estimate the values of parameters β_0 (constant term), β_1, σ, α, δ, the likelihood value, the mean technical efficiency. We also compute the true values of α using the estimated δ, the true likelihood value, and the true mean technical efficiency. This process is replicated 1,000 times for all the scenarios defined in Table 6.1. For each scenario, we compute the mean value and its standard deviation. We also calculate the bias and root mean squared error (RMSE) for all estimated values using the following formula:

$$\text{Bias}_k = \frac{1}{1000}\sum_{r=1}^{1000}\left(E_{kr} - T_{kr}\right), \quad \text{RMSE}_k = \sqrt{\frac{1}{1000}\sum_{r=1}^{1000}\left(E_{kr} - T_{kr}\right)^2},$$

where Bias_k is the bias in scenario k, RMSE_k is the root mean squared error in scenario k, E_{kr} is the estimated value in replication r of scenario k, and T_{kr} is the true value in replication r of scenario k. Moreover, for a better analysis, we also plot for each scenario box-plots for estimated parameters, which better exhibit the distribution of the parameters (medians, ranges, biases with the true data generating process values).

6.4.2 Simulation results

Tables 6.6, 6.7, and 6.8 present the simulation results for different setups of λ (0.2, 0.5, 0.8) respectively (case of spatial error SFA). Furthermore, figures 6.5, 6.6, and 6.7 in appendix show the distributions of parameters β_0, β_1, δ, σ and λ. These results exhibits some patterns. Estimated coefficients β_0 and β_1 shows relatively very low biases and RSME for all scenarios. For instance, the biases for β_1 are close to 0. Also, β_0 and β_1 distributions show that they converge to their true values when the sample size is high. For parameter λ, with low true values, biases and standard deviations are high. Parameter σ exhibits low biases whether N=250 or N=1000, and when values of δ are -0.2 or -0.5. Biases seems to increase with the increase (absolute term) of values of δ. In some cases, estimated values of σ are totally out of their range as shown by figure 6.5 for δ=0.8. Regarding parameter δ, biases diminishes with the increase in sample size in all scenarios. However, box-plots show relatively skewed distributions for δ with wide range in most scenarios (when δ=-0.2, -0.5), denoting the difficult precision in the estimation of this parameter when the sample size is small and the absolute value of true δ is low. As α and δ are related, the same pattern is observed for α. In conclusion, as also observed by de Graaff (2020), values of the estimated parameters converge to their true values when the sample size becomes larger (as shown by the box-plots, a wider range in the distribution of estimated values is globally observed when N=250, compared to N=1000).

Tables 6.9, 6.10, and 6.11 reports respectively the simulation results for the different setups of ρ (0.2, 0.5, 0.8) (case of spatial auto-regressive SFA), and figures 6.8, 6.9, and 6.10 in the appendix plot in this specific case the distributions of parameters β_0, β_1, δ, σ and ρ respectively. One can observe that parameter β_0 shows high biases and RMSE in most scenarios, especially when true $\rho = 0.8$. But the estimated coefficient β_1 shows very low biases and RMSE in all scenarios. In the case of estimated parameter α, the RMSE are high for all scenarios. In addition, the biases are high when $\delta = -0.2$. Similar pattern can be observed for the estimated values of δ.

These simulation works show that the suggested estimation technique can provide consistent results, mostly in the case of the spatial error stochastic frontier model. However, it is worth noting that for the efficiency parameter, it seems important to use a starting value with a negative sign which is the expected sign of the coefficient. Indeed, during simulations, we observed that when the starting value of the inefficiency parameter δ has a positive value, the estimated coefficient gives values close to zero. We conducted several simulation rounds and observed similar results. Considering the time frame of this PhD work, we did not pursue the simulation work with the positive starting values and neither do we present the results. This is the only issue that we observed in the simulation, and even when using the positive starting values, this did not affect the estimation of other coefficients. We believe this problem might be related to the re-parametrization in the equations: centered versus non-centered parameterizations, which has been widely discussed in literature[3]. The starting values used for the simulation work and empirical analysis are defined following principles of stochastic frontier analysis and they are presented in table 6.2.

Table 6.2: Starting values for maximum likelihood estimation

	Estimated models		
Parameters	SFA	SAR-SFA	SEM-SFA
β_0	β_0^{OLS}	β_0^{SAR}	β_0^{SEM}
β_1	β_1^{OLS}	β_1^{SAR}	β_1^{SEM}
σ	σ^{OLS}	σ^{SAR}	σ_{SEM}
λ		λ^{SEM}	λ^{SEM}
ρ		ρ^{SEM}	ρ^{SEM}
δ	$-\sigma^{OLS}\sqrt{1-\pi/2}$	$-\sigma^{SAR}\sqrt{1-\pi/2}$	$-\sigma^{SEM}\sqrt{1-\pi/2}$

[3]The re-parametrization problem has been mentioned to us via e-mails by de Graaff (2020) himself, however, we did not find any solution to address it in the present work.

6.5 Empirical application

6.5.1 Model specification

This section describes the specifications of the production frontiers and of the weight matrices used. We specify a base production function (model 1), using a Cobb-Douglas functional form between the output and the inputs used[4]. Therefore the estimated parameters of the conventional inputs can be interpreted as partial elasticities of production. From this base model, we specified three variant models: in the first variant (model 2), we incorporated climate variables directly into the non-stochastic component of the production function. Meanwhile in the second variant (model 3), in addition to the climate variables, we included some other environmental variables. Our general model is therefore specified as follows:

$$\ln(y_i) = \theta_0 + \sum_{i=1}^{k} \theta_k \ln(x_{ik}) + \sum_{i=1}^{l} \gamma_l D_{il} + \sum_{i=1}^{m} \Delta_m Z_{im} + \varepsilon_i \tag{6.16}$$

where y_i represents the output of the i^{th} farmer, x_{ik} denote vectors of the production inputs k; D_{il} represent climatic variable l; Z_{im} represent other non-stochastic environmental variables m; θ_0, θ_k, γ_l, and Δ_m are parameters to be estimated and ε_i is the error term. The output here is the household total crop production. The three inputs are land, labour and operating costs. The climatic variables are rainfall, temperature, and rainfall and temperature anomalies. The other non-stochastic environmental variables include the components of soil, the farmer education, the use of improved seeds, and membership in farmers organizations. In this empirical analysis, we estimated this model using a step-wise approach. First, we assumed no spatial interaction and estimated the model as a linear model (OLS). Secondly, after testing for the presence of spatial dependence in the production function (we discussed this later), we incorporated spatial interaction via the spatial weight matrix W and estimated linear spatial models (Spatial lag and Spatial error). The spatial lag is of the form $\ln(y_i) = \rho W \ln(y_i)\, \theta_0 + \sum_{i=1}^{k} \theta_k \ln(x_{ik}) + \sum_{i=1}^{l} \gamma_l D_{il} + \sum_{i=1}^{m} \Delta_m Z_{im} + \varepsilon_i$ and the spatial error is of the form:$\ln(y_i) = \theta_0 + \sum_{i=1}^{k} \theta_k \ln(x_{ik}) + \sum_{i=1}^{l} \gamma_l D_{il} + \sum_{i=1}^{m} \Delta_m Z_{im} + \varepsilon_i$, with $\varepsilon_i = \lambda W \varepsilon_i + \tilde{\varepsilon_i}$, with ρ and λ as parameters to be estimated. Finally, assuming that ε follows a skew-normal distribution, we estimated a standard SFA and then spatial frontier models (spatial SFA lag and Spatial SFA error), to derive farmers technical efficiency scores robust to the presence of spatial

[4]The Cobb-Douglas functional has convenient properties. As argued by O'Donnell (2016), it generally satisfies non-negativity and monotony, unlike Translog, a commonly used alternative.

dependence. The performances of the estimated models are compared to select the "best" fitting model based on various tests.

Prior to estimating the spatial models, we specify the weight matrix for the Senegalese agricultural sector. The spatial weight matrix W is a symmetric matrix, where its elements w_{ij} express proximity of a household i with a household j. In common practice, to enable an interpretation of model coefficients, W is row standardized so that the sum of the row elements equals to one. In addition, the diagonal elements w_{ii} are set to zero, in order to prevent the effect of the i^{th} household from directly predicting itself. Many specifications of weight matrices have been used in the literature, and specifying the weight matrix is arbitrary. However, prior knowledge of the study population and economic theory can help to guide in the specification of these matrices. Following Areal *et al.* (2012) and Pede *et al.* (2018), we specify a spatial weight matrix of a power form with: $w_{ij} = \exp\left(-d_{ij}^2/s^2\right)$, where d_{ij} is the euclidean distance between households i and j, s is the cut-off distance. We choose a cut-off distance of 4 km based on a previous study in this thesis. It is assumed therefore that beyond this cut-off distance there is no spatial effect.

Moreover, before implementing the spatial models, we checked for spatial interaction by computing the Moran index using the dependent variable and the residuals of the OLS of each model[5]. Using the same variables, we plotted Moran scatter plots to visually show the spatial interaction. Furthermore, we conducted a series of tests including the standard Lagrange multiplier (LM)(Anselin, 1988) test and its robust counterparts (Anselin *et al.*, 1996)[6].

6.5.2 Data source and variables

The data used for the empirical analysis was derived from a survey conducted in Senegal, which randomly sampled 4480 households that mainly produce rainfed cereals. The survey was implemented under the Agricultural Policy Support Project

[5]Moran's I statistic for outcome is computed as: $I = (\sum_i \sum_j w_{ij}(Y_i - \bar{Y})(Y_j - \bar{Y}))/(\sum_i (Y_i - \bar{Y})^2)$, where w_{ij} is the spatial weight between households i and j; Y_i is the outcome of household i; and $\bar{Y}$ is the mean of the outcome. The range of Moran's I is $(-1, 1)$, with 1 indicating perfect spatial similarity (or positive spatial correlation), 0 indicating no spatial correlation, and -1 indicating perfect dispersion (or negative correlation). If we observe a significant spatial autocorrelation based on Moran's I statistic, spatial regressions models should be used to correct for the spatial autocorrelation errors.

[6]The standard two LM tests are: $LM_{error} = [e'We/(e'e/N)]^2/[tr(W^2 + W'W)]$ and $LM_{lag} = [e'Wy/(e'e/N)]^2/D$. Robust LM tests are defined as: $RLM_{error} = [e'We/(e'e/N)]^2/[tr(W^2 + W'W)]$ and $RLM_{lag} = [e'Wy/(e'e/N)]^2/D$, where e denotes the estimated residual from the non-spatial model; N is the number of farmers; and W are defined as previously.

(Projet d'Appui aux Politiques Agricoles, PAPA)[7], which is an initiative of the Government of Senegal funded by USAID-Senegal as part of the "Feed The Future" initiative, and implemented for a period of 3 years (2015 - 2018) by the Senegalese Ministry of Agriculture and Rural Facilities with technical support from the International Food Policy Research Institute (IFPRI). The data, which covers the main agricultural season of 2016/2017, contains information on crop production and different inputs used. After the data cleaning and removing observations with no information on crop production, the final sample comprises of 4245 households located in all six Senegalese agro-ecological zones. Climate data were retrieved, using the surveyed households location coordinates, from publicly available databases of the Climate Hazards Center of the University of California (https://www.chc.ucsb.edu/data). The Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS)[8] for the period 1981-2016 was used for rainfall variables, and the monthly Climate Hazards Group InfraRed Temperature with Station (CHIRTSmax)[9] for the period 1983-2016 was used for temperature variables. Table 6.3 presents the definition and summary statistics of the variables used in the analysis.

The dependent variable used in the models is the total crop production, which is expressed in Franc CFA[10], and it includes all farm crop production outputs valued at the market prices. Farmers produce 32 crops which include major crops, namely cereals (rice, maize, sorghum, millet), groundnut, and cotton. The first input is the land, which is the sum of all land area dedicated to crop production during the 2016/2017 growing season. Variable labour, the second input, is the quantity of total labour (in adult equivalent). The total of operating costs in FCFA, which is the last input, includes the costs of seeds, fertilizers, non-family labour, and other costs such as transport, maintenance, etc. Climate variables include rainfall, rainfall anomaly, temperature and temperature anomaly. Based on study of Ba (2006), the climate in Senegal is made up of two seasons: the rainy season which lasts two to four months (exceptionally five months) depending on the region and the very long dry season. Also, these two seasons are more complex to define because they are

[7] Official website of the project is http://www.papa.gouv.sn/.

[8] CHIRPS data is a 35+ years quasi-global rainfall data set. Spanning 50°S-50°N (and all longitudes) and ranging from 1981 to near-present, CHIRPS incorporates the in-house climatology data CHPclim, 0.05° resolution satellite imagery, and in-situ station data to create gridded rainfall time series for trend analysis and seasonal drought monitoring (Funk *et al.*, 2015).

[9] CHIRTSmax is a global 2-m maximum temperature (Tmax) product that directly combines satellite and station-based estimates of Tmax to produce routinely updated data to support the monitoring of temperature extremes. The CHIRTSmax development process integrated a long-term climatology with satellite information and available station data. The result is a monthly estimate of the daily maximum temperature for the 1983-2016 time period (Funk *et al.*, 2019).

[10] Local currency, 1 FCFA = 0,0017 USD as at 8 May 2020.

Table 6.3: Description of variables

Variables	**Description and measurement**	**Mean**	**SD**	**Max**	**Min**
Production	Total crop production (1000 FCFA)	467.684	764.653	14375	0.180
Land	Total area of land cultivated (Ha)	4.411	4.527	50.000	0.001
Family labor	Total family labor (Adult equivalent)	3.228	2.258	20.100	0.150
Operating costs	Total Operating costs (1000 FCFA)	55.970	118.610	3418.500	1.500
Rainfall	Annual rainfall (m)	0.672	0.292	1.484	0.187
Temperature	Average maximum temperature (Celsius degree)	35.942	1.314	38.43	30.90
Rainfall anomaly	Rainfall anomaly 1981-2015	0.033	0.050	0.170	-0.186
Temperature anomaly	Average maximum temperature Anomaly 1983-2015	0.020	0.006	0.038	0.008
Clay	Percentage of clay (%)	19.97	7.235	40.00	3.00
Silt	Percentage of silt (%)	13.47	5.735	31.00	2.00
Improved seed	Use of improved seed (1=yes, 0=no)	0.259	0.4380	1	0
Education	Formal education (1=yes, 0=no)	0.370	0.483	1	0
FBO Membership	Membership in Farmer Based Organization (1=yes, 0=no)	0.088	0.283	1	0
N	Number of Observations	4423	4423	4423	4423

SD: Standard Deviation

difficult to delineate in time and space. The spatial and temporal features of the rain are very variable from one region to another (Ba, 2006). The spatial features of the Senegalese climate justify once again the spatial approach used in this paper. We, therefore, consider rainfall as the annual precipitation received by each household. Temperature is the yearly average maximum temperature of the household location expressed in degree Celsius. We also computed the anomalies of temperature and rainfall. We follow Lachaud *et al.* (2017) and computed anomalies as the deviation of the 2016 annual rainfall and temperature observation from the long-term mean (1981–2015 for rainfall and 1983-2015 for temperature)[11]. As argued by Lachaud *et al.* (2017) and Barrios *et al.* (2010), there are some advantages associated with the use of anomalies. First, factors like the station location and elevation are less critical, and agricultural decisions are based on expected weather behaviour, therefore production might be affected by any weather deviation from the expectations. In the sample, the mean monthly maximum temperature is about 36 degrees Celsius, while the mean annual rainfall received by a household in the year 2016 is 672 mm. The long-term average annual rainfall over the period 1981-2015 is 651 mm and the long-term average maximum temperature over the period 1983-2015 is 35 degree Celsius. The other variables included in the models that can produce

[11] Anomaly = (current year value - long term mean) / long term mean

cross-sectional variation in the production frontier between farmers are their education, use of improved seeds, percentages of clay and silt in soils, and membership in farmers organizations.

6.5.3 Empirical results

Diagnostics for spatial interdependence

Table 6.4 reports the results of the diagnostic tests for spatial interaction. It shows that the Moran's I statistics are all positive and highly significant ($p < 0.01$). For instance, the Moran I using the residuals are 0.377, 0.328, and 0.310 for models 1, 2 and 3 respectively, indicating that there is a strong positive spatial correlation between farmers crop production, as well as a strong spatial dependence. Farmers with relatively high production seem to live close to other farmers with high production, and farmers with relatively low production tend to live near farmers with low production. These results are corroborated by the Moran plots (Figures 6.3 and 6.4 in the appendix), which show how observations outcomes are strongly and positively correlated to their neighbours' outcomes (one can observe the clustering of values in the upper right quadrant and lower left quadrant, suggesting the positive spatial autocorrelation). These results suggest that spatial correlation should be considered in our analysis. After estimating the OLS, we computed the standard Lagrange Multiplier tests (Anselin, 1988) and their robust counterparts (Anselin *et al.*, 1996). The null hypotheses were mostly rejected at $p < 0.01$, indicating that spatial interaction should be incorporated into our models. At this step, tests results are non-conclusive. However, when comparing tests statistics, the results are mostly in favour of the spatial error models. The test statistics of the error models are much higher than that of the lag models, we therefore continue the analysis with the spatial error models.

Production frontier

Tables 6.12, 6.13, and 6.14 present the results of estimated models 1, 2, and 3 respectively, with different estimation techniques. At this stage of the analysis, the preferred models are the spatial error models for the base model (SEM 1) as well as for the variants (SEM 2 and SEM 3). We still need to decide between theses linear spatial error models (SEM 1, SEM 2, and SEM 3) and the spatial

Table 6.4: Moran Index and Lagrange Multiplier tests

	Model 1	Model 2	Model 3
Moran I Statistic (outcome)		0.418***	
Moran I Statistic (residuals)	0.377***	0.328***	0.299***
Standard LM Error	2049.7***	1550.6***	1285.8***
Robust LM Error	1374.4***	1036.3***	735.57***
Standard LM Lag	716.72***	536.69***	550.7***
Robust LM Lag	41.426***	22.414***	0.494

Notes: $^{***}p < 0.01$. LM: Lagrange Multiplier

SFA error models (SEM-SFA 1, SEM-SFA 2, and SEM-SFA 3) which are spatial frontier analysis. Based on the likelihood ratio tests (LR), the spatial SFA error specifications are preferred. The results of the LR tests comparing the various models are presented in table 6.5. Furthermore, based on the same likelihood ratio test, the preferred spatial SFA error model is the one which includes climatic and other variables (SEM-SFA 3). This model suggests that including climatic variables and other variables bring additional information to the model. We, therefore, continue the discussion with this model. The maximum likelihood estimates of the SEM-SFA 3, which includes climatic and other variables is reported in column (6) of table 6.14. Results show as expected[12] that the model produces statistical significant (at 1% level) and positive partial production elasticities. These results suggest that in Senegal, land and labour are the inputs that make the highest contribution to agricultural production. Land contribution to agricultural production is about 86% while labour, which is mainly made of the family labour, contributes about 9%. Similar results have also been reported in other parts of the developing world, for instance in Latin America and Caribbean by Lachaud *et al.* (2017). Meanwhile, the total operating costs which includes the values of fertilizers, seeds, hired labour and other cost such as transport and maintenance only plays a minor role in the farm crop production, with a contribution of about 2%. When comparing these results with the estimates of other models (OLS 1, OLS 2, OLS 3, SFA 1, SFA 2, SFA 3, SEM-SFA 1, etc.), one can observe similar patterns: statistical significant and positive partial production elasticities, with a slight difference in magnitudes. Neither the standard frontier analysis nor the incorporation of spatial interaction changed the results of the frontier estimates, although the likelihood value improves significantly in the spatial models.

[12]Satisfaction of regularity conditions from production economic theory, i.e., partial output elasticities should be non-negative and less than 1(Lachaud *et al.*, 2017).

Table 6.5: Likelihood ratios tests

	SEM 2	SEM 3	SEM-SFA 1	SEM-SFA 2	SEM-SFA 3
SEM 1	88.393***	152.177***	3074.76***		
SEM 2		63.783***		3068.269***	
SEM 3					3072.986***
SEM-SFA 1				81.90***	150.40***
SEM-SFA 2					68.50***

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Climatic and spatial effects

Climatic variables exhibit statistically significant coefficients with the exception of average maximum temperature. The rainfall parameter is positive, and as expected higher precipitation is beneficial for farm crop production. The variables representing rainfall and temperature anomalies are both negative (-3.193 and -16.967) and significant at 1% level. These results would mean that the deviation of 2016 annual rainfall and temperature from the long-term trend has negatively affected farm crop production.

The spatial error parameter λ (with value of 0.525) of the preferred model is positive and statistically different from zero. In comparison, across all alternative models, this parameter ranges from 0.523 and 0.588, and its values are statistically significant. At the same time, the spatial lag parameter ρ, though smaller than λ, also exhibits positive and statistically significant values (range from 0.310 to 0.346). The results suggest the presence of a high spatiality arising from unobservable spatial factors in the agricultural production sector of Senegal. This implies that computing technical efficiency scores using non-spatial production frontiers would have led to biased estimates.

The standard deviation of the error term (σ) for the preferred model is 0.843 and significant at 1% level. This value is very close in terms of magnitude to the estimated ones in all alternative models, which are also significant. The highest σ can be observed for the standard SFA in specification 1 (SFA 1) and the lowest for the spatial lag SFA in specification 2 (SAR-SFA 2). The dispersion parameter for the inefficiency (α) of the preferred model is negative and statistically significant. This result suggests that most of the farmers are producing below the production frontier. However, it easy to notice that this parameter is not always statistically significant in the alternative specifications, and its magnitude also varies significantly. For instance, in the base model, the value of this parameter α is statistically significant for the standard SFA (SFA 1) but in absolute terms, lower than that of the spatial error SFA (SEM-SFA 1). In specifications 2 and 3, the same pattern is observed, however, the magnitude of the α parameter is much lower than that of the spatial error SFA (SEM-SFA 1, 2, and 3). One explanation is that, first the difference in absolute terms in the inefficiency parameter could be due to the effect of climatic variables which may have acted as a "corrective" factor, hence suggesting that not incorporating the climatic variables would simply lead to missing variable bias. Secondly, these differences are due to the existence of spatial features in farm production. Once again, not having included them in our models would have led to biased estimates. When comparing the "instability" of the α parameter to the "stability" of the σ parameter, one can conclude that climate variability and spatial features are very important factors determining the inefficiency of farm households in Senegal. The results also reveal that components of soils (clay, silt), use of improved seed, formal education of household head and membership in farmer organizations are important factors in improving farm production.

Technical Efficiency

Figures 6.1 shows the kernel distributions of technical efficiency estimates for all model specifications (both non-spatial and spatial models). The average value of technical efficiency scores for the preferred model is 0.792. This result suggests that there is still a room to improve farm efficiency by at least 20%. The mean efficiency estimated for all specifications varies between 0.752 and 0.792. The magnitudes of the means of efficiency scores seem consistent through estimations methods and models. However, their distributions do not follow the same patterns. The distribution of the efficiency scores of the spatial SFA error in all specifications have a more flattened shape and is very distinct from the other models. This can be a consequence of the high absolute value of α observed previously. Moreover, we calculated the percentage difference between the average technical efficiency for non-spatial SFA models and for the different models that account for spatial dependence. Results show that accounting for spatial dependence in the analysis leads to slight increases in the estimates of technical efficiency in all cases. For instance, in specifications 3, the increase of technical efficiency is only about 4% in comparison to estimate of a non-spatial SFA. This result would indicate that the effects of incorporating spatial dependence in the analysis is more observable through the distribution of farm technical efficiency scores (shown by the flattened curve), and suggesting that some households might have a certain level of technical efficiency mostly because of their geographic location. Therefore, we mapped the individual change in technical efficiency scores to explore any specificity of the Senegalese regions, and found that there are no peculiarities between localities in terms of efficiency change. Farmers who have high changes in efficiency could be found in any part of the country. Figure 6.2 mapped first the change in technical efficiency due to spatial heterogeneity (change between models SEM-SFA 3 and SFA 3) and secondly the change in efficiency due to spatial heterogeneity and climatic variability (change between models SEM-SFA 3 and SFA 1). These results might suggest that climate variability is persistent and affects Senegalese farmers efficiency irrespective of where they are located, and they should use adequate farming techniques and technologies in order to have an appropriate level of productivity. Finally, we conducted the same analysis using this time the translog functional form and the results are presented in tables 6.15, 6.16, and 6.17 and figure 6.11 in the appendix. Results show that the obtained estimates of parameters α, σ, and ρ and the kernel distributions of technical efficiencies are similar to those of the Cobb-Douglas functional form.

6.6 Conclusion

The main objective of this paper was to combine spatial econometrics techniques and a stochastic frontier approach, to analyse the effects of spatial dependence and climate variability on farmers' technical inefficiency. The paper contributes to the growing literature of efficiency analysis and to the existing knowledge of the Senegalese agricultural sector by incorporating spatial features in the production functions. For this purpose, we use a recently designed estimation technique for cross-sectional data in spatial econometrics, to conduct simulations and empirical analyses.

Simulation results show that the maximum likelihood estimation technique suggested by de Graaff (2020) and based on the skew-normal distribution of errors in the stochastic production frontier provides consistent results. However, it is worth noting that for the efficiency parameter, it seems important to use a stating value with a negative sign which is the expected sign of the coefficient. Empirical results reveal that Senegalese agriculture is more dependent on land area and dedicated labour. Moreover, results show that farms efficiency is highly affected by both climatic features and spatial heterogeneity, and not accounting for them might lead to biased results for the efficiency distribution. Particularly, we found that rainfall and temperature anomalies negatively affect the production frontier, and farm technical efficiency on average increases when estimations control for spatial heterogeneity and climate variability. Also, we did not find conclusive evidence that could have led to the choice of spatial dependence modelling in the form of a spatial autoregressive-type. Furthermore, findings also reveal that the changes in technical efficiency score could be observed in any part of the country, implying that the effects of climate variability and unobserved spatial features are not specific to any particular region, but common in the entire country.

In terms of policy implications, these results imply that farmers need to adapt to climatic effects by using appropriate and very localized technologies or farming practices that take into account the specific characteristics of their locations. Policy-makers should encourage the design and dissemination of agricultural technologies that are very adaptable to specific conditions of farmers. We also observed that membership in farmer organizations as a measure of social capital improves the production frontier. Farmers organizations and other social groups could complement the efforts of extension services, and be a good entry point of introducing climate adaptation and other farming techniques.

From a research perspective, future studies should pursue the simulations work and investigate more the empirical performances of the skew-normal approach. Such studies should explore the problem of starting values that we have encountered, and the "centered versus non-centered parameterizations" issues. These studies are also encouraged to apply the approach to a broad range of sectors (besides agriculture) in developing countries where most studies are needed and where spatiality is mostly neglected in the analysis.

Table 6.6: Simulation results for $\lambda = 0.2$

N	Parameters	$\delta=-0.2$					$\delta=-0.5$					$\delta=-0.8$				
		True	Mean	SD	Bias	RMSE	True	Mean	SD	Bias	RMSE	True	Mean	SD	Bias	RMSE
250	β_0	1.000	1.029	0.137	0.029	0.140	1.000	0.946	0.132	-0.054	0.142	1.000	0.917	0.120	-0.083	0.146
	β_1	1.000	1.000	0.007	0.000	0.007	1.000	1.000	0.007	0.000	0.007	1.000	1.000	0.006	0.000	0.006
	σ	0.306	0.319	0.037	0.013	0.039	0.346	0.301	0.039	-0.045	0.059	0.500	0.287	0.040	-0.213	0.217
	α	-0.204	-0.444	0.514	-0.240	0.567	-0.577	-0.523	0.558	0.054	0.560	-1.333	-1.158	0.739	0.175	0.759
	δ	-0.200	-0.312	0.326	-0.112	0.344	-0.500	-0.355	0.335	0.145	0.365	-0.800	-0.635	0.338	0.165	0.376
	λ	0.200	-0.251	1.248	-0.451	1.327	0.200	-0.259	1.365	-0.459	1.440	0.200	-0.241	1.336	-0.441	1.407
	TE	0.871	0.838	0.061	-0.033	0.071	0.869	0.854	0.059	-0.015	0.065	0.782	0.896	0.074	0.114	0.138
	LL	36.349	38.807	11.614	2.457	2.952	49.740	57.270	11.613	7.530	8.244	44.363	102.191	11.595	57.828	58.386
1000	β_0	1.000	1.026	0.104	0.026	0.107	1.000	0.959	0.109	-0.041	0.116	1.000	0.974	0.107	-0.026	0.111
	β_1	1.000	1.000	0.004	0.000	0.004	1.000	1.000	0.003	0.000	0.003	1.000	1.000	0.003	0.000	0.003
	σ	0.306	0.313	0.024	0.007	0.025	0.346	0.298	0.025	-0.048	0.054	0.500	0.296	0.019	-0.204	0.204
	α	-0.204	-0.356	0.360	-0.152	0.390	-0.577	-0.487	0.391	0.090	0.401	-1.333	-1.291	0.336	0.042	0.339
	δ	-0.200	-0.287	0.265	-0.087	0.279	-0.500	-0.380	0.275	0.120	0.300	-0.800	-0.768	0.154	0.032	0.158
	λ	0.200	-0.029	0.903	-0.229	0.931	0.200	0.004	0.799	-0.196	0.822	0.200	0.041	0.752	-0.159	0.768
	TE	0.871	0.849	0.049	-0.022	0.055	0.869	0.871	0.045	0.002	0.047	0.782	0.914	0.048	0.132	0.140
	LL	145.069	147.823	22.293	2.754	3.287	199.056	221.970	22.462	22.914	23.669	177.702	401.685	23.275	223.983	224.536

Notes: TE and LL represent respectively technical efficiency and log likelihood value

Table 6.7: Simulation results for $\lambda = 0.5$

N	Parameters	$\delta = -0.2$					$\delta = -0.5$					$\delta = -0.8$				
		True	Mean	SD	Bias	RMSE	True	Mean	SD	Bias	RMSE	True	Mean	SD	Bias	RMSE
250	β_0	1.000	1.022	0.179	0.022	0.180	1.000	0.874	0.169	-0.126	0.211	1.000	0.784	0.164	-0.216	0.271
	β_1	1.000	1.000	0.007	0.000	0.007	1.000	1.000	0.007	0.000	0.007	1.000	1.000	0.006	0.000	0.006
	σ	0.306	0.319	0.037	0.012	0.039	0.346	0.299	0.039	-0.047	0.061	0.500	0.279	0.044	-0.221	0.225
	α	-0.204	-0.429	0.515	-0.224	0.561	-0.577	-0.475	0.556	0.103	0.565	-1.333	-0.998	0.813	0.336	0.879
	δ	-0.200	-0.300	0.325	-0.100	0.340	-0.500	-0.320	0.337	0.180	0.382	-0.800	-0.541	0.386	0.259	0.464
	λ	0.500	0.399	0.841	-0.101	0.847	0.500	0.400	0.871	-0.100	0.876	0.500	0.377	1.099	-0.123	1.105
	TE	0.871	0.834	0.061	-0.037	0.073	0.869	0.842	0.061	-0.027	0.070	0.782	0.874	0.074	0.092	0.120
	LL	35.913	38.455	11.619	2.542	3.049	49.303	56.897	11.607	7.594	8.308	43.926	101.683	11.609	57.756	58.319
1000	β_0	1.000	1.023	0.146	0.023	0.148	1.000	0.916	0.158	-0.084	0.179	1.000	0.909	0.175	-0.091	0.197
	β_1	1.000	1.000	0.004	0.000	0.004	1.000	1.000	0.003	0.000	0.003	1.000	1.000	0.003	0.000	0.003
	σ	0.306	0.313	0.024	0.007	0.025	0.346	0.299	0.025	-0.048	0.054	0.500	0.295	0.021	-0.205	0.206
	α	-0.204	-0.354	0.364	-0.150	0.393	-0.577	-0.490	0.392	0.087	0.401	-1.333	-1.266	0.380	0.068	0.385
	δ	-0.200	-0.286	0.269	-0.086	0.283	-0.500	-0.382	0.277	0.118	0.301	-0.800	-0.752	0.188	0.048	0.194
	λ	0.500	0.567	0.720	0.067	0.723	0.500	0.587	0.692	0.087	0.697	0.500	0.612	0.681	0.112	0.689
	TE	0.871	0.845	0.052	-0.025	0.059	0.869	0.868	0.048	-0.000	0.050	0.782	0.909	0.051	0.126	0.137
	LL	144.531	147.350	22.290	2.818	3.351	198.519	221.493	22.469	22.974	23.732	177.164	401.120	23.289	223.956	224.510

Notes: TE and LL represent respectively technical efficiency and log likelihood value

Table 6.8: Simulation results for $\lambda = 0.8$

N	Parameters	$\delta = -0.2$					$\delta = -0.5$					$\delta = -0.8$				
		True	Mean	SD	Bias	RMSE	True	Mean	SD	Bias	RMSE	True	Mean	SD	Bias	RMSE
250	β_0	1.000	0.914	0.252	-0.086	0.266	1.000	0.538	0.246	-0.462	0.523	1.000	0.181	0.210	-0.819	0.845
	β_1	1.000	1.000	0.007	0.000	0.007	1.000	1.000	0.007	0.000	0.007	1.000	1.000	0.006	0.000	0.006
	σ	0.306	0.317	0.037	0.011	0.039	0.346	0.295	0.039	-0.051	0.064	0.500	0.263	0.045	-0.237	0.241
	α	-0.204	-0.384	0.516	-0.180	0.547	-0.577	-0.379	0.551	0.199	0.586	-1.333	-0.682	0.820	0.651	1.047
	δ	-0.200	-0.270	0.327	-0.070	0.334	-0.500	-0.255	0.335	0.245	0.415	-0.800	-0.365	0.401	0.435	0.592
	λ	0.800	0.982	0.757	0.182	0.778	0.800	0.987	0.760	0.187	0.782	0.800	0.993	0.789	0.193	0.812
	TE	0.871	0.830	0.076	-0.041	0.087	0.869	0.829	0.081	-0.040	0.092	0.782	0.839	0.074	0.057	0.096
	LL	34.765	37.640	11.635	2.875	3.386	48.156	56.053	11.611	7.897	8.603	42.779	100.466	11.660	57.687	58.259
1000	β_0	1.000	0.935	0.219	-0.065	0.229	1.000	0.637	0.247	-0.363	0.439	1.000	0.470	0.299	-0.530	0.608
	β_1	1.000	1.000	0.004	0.000	0.004	1.000	1.000	0.003	0.000	0.003	1.000	1.000	0.003	0.000	0.003
	σ	0.306	0.313	0.024	0.007	0.025	0.346	0.299	0.025	-0.048	0.054	0.500	0.293	0.025	-0.207	0.209
	α	-0.204	-0.350	0.363	-0.146	0.391	-0.577	-0.486	0.391	0.091	0.401	-1.333	-1.214	0.455	0.120	0.470
	δ	-0.200	-0.283	0.269	-0.083	0.281	-0.500	-0.379	0.276	0.121	0.301	-0.800	-0.719	0.237	0.081	0.251
	λ	0.800	1.158	0.674	0.358	0.763	0.800	1.173	0.666	0.373	0.763	0.800	1.194	0.672	0.394	0.778
	TE	0.871	0.850	0.064	-0.020	0.068	0.869	0.878	0.061	0.009	0.064	0.782	0.918	0.060	0.136	0.149
	LL	143.191	146.310	22.293	3.118	3.646	197.179	220.444	22.467	23.265	24.018	175.824	399.923	23.275	224.098	224.652

Notes: TE and LL represent respectively technical efficiency and log likelihood

Table 6.9: Simulation results for $\rho = 0.2$

N	Parameters	$\delta = -0.2$					$\delta = -0.5$					$\delta = -0.8$				
		True	Mean	SD	Bias	RMSE	True	Mean	SD	Bias	RMSE	True	Mean	SD	Bias	RMSE
250	β_0	1.000	1.161	1.011	0.161	1.023	1.000	1.089	0.935	0.089	0.938	1.000	1.069	0.760	0.069	0.763
	β_1	1.000	1.000	0.007	0.000	0.007	1.000	1.000	0.007	0.000	0.007	1.000	1.000	0.006	0.000	0.006
	σ	0.306	0.320	0.037	0.014	0.040	0.346	0.303	0.038	-0.043	0.058	0.500	0.293	0.036	-0.207	0.210
	α	-0.204	-0.465	0.510	-0.261	0.572	-0.577	-0.557	0.559	0.020	0.559	-1.333	-1.300	0.613	0.034	0.614
	δ	-0.200	-0.329	0.320	-0.129	0.344	-0.500	-0.378	0.335	0.122	0.357	-0.800	-0.718	0.240	0.082	0.254
	ρ	0.200	0.191	0.076	-0.009	0.077	0.200	0.192	0.071	-0.008	0.071	0.200	0.193	0.058	-0.007	0.059
	TE	0.871	0.845	0.054	-0.026	0.061	0.869	0.855	0.057	-0.014	0.062	0.782	0.911	0.032	0.128	0.134
	LL	36.349	38.799	11.596	2.449	2.928	49.740	57.276	11.600	7.536	8.243	44.363	102.302	11.605	57.939	58.497
1000	β_0	1.000	1.100	0.855	0.100	0.860	1.000	1.033	0.789	0.033	0.789	1.000	1.020	0.638	0.020	0.638
	β_1	1.000	1.000	0.004	0.000	0.004	1.000	1.000	0.003	0.000	0.003	1.000	1.000	0.003	0.000	0.003
	σ	0.306	0.313	0.024	0.007	0.025	0.346	0.300	0.025	-0.047	0.053	0.500	0.297	0.018	-0.203	0.204
	α	-0.204	-0.369	0.352	-0.165	0.388	-0.577	-0.517	0.374	0.061	0.378	-1.333	-1.248	0.496	0.085	0.503
	δ	-0.200	-0.300	0.257	-0.100	0.276	-0.500	-0.406	0.260	0.094	0.277	-0.800	-0.739	0.270	0.061	0.276
	ρ	0.200	0.195	0.065	-0.005	0.065	0.200	0.196	0.060	-0.004	0.060	0.200	0.198	0.050	-0.002	0.050
	TE	0.871	0.855	0.047	-0.145	0.153	0.869	0.875	0.045	-0.125	0.133	0.782	0.897	0.123	-0.103	0.161
	LL	145.069	147.837	22.331	2.767	3.335	199.056	221.995	22.530	22.939	23.715	177.702	401.779	23.247	224.077	224.626

Notes: TE and LL represent respectively technical efficiency and log likelihood value

Table 6.10: Simulation results for $\rho = 0.5$

N	Parameters	$\delta = -0.2$					$\delta = -0.5$					$\delta = -0.8$				
		True	Mean	SD	Bias	RMSE	True	Mean	SD	Bias	RMSE	True	Mean	SD	Bias	RMSE
250	β_0	1.000	1.292	1.504	0.292	1.531	1.000	1.207	1.392	0.207	1.407	1.000	1.160	1.137	0.160	1.147
	β_1	1.000	1.000	0.007	0.000	0.007	1.000	1.000	0.007	0.000	0.007	1.000	1.000	0.006	0.000	0.006
	σ	0.306	0.319	0.037	0.012	0.039	0.346	0.303	0.039	-0.044	0.058	0.500	0.293	0.036	-0.207	0.210
	α	-0.204	-0.436	0.518	-0.231	0.567	-0.577	-0.543	0.564	0.035	0.565	-1.333	-1.305	0.605	0.029	0.605
	δ	-0.200	-0.304	0.323	-0.104	0.339	-0.500	-0.366	0.338	0.134	0.363	-0.800	-0.723	0.230	0.077	0.243
	ρ	0.500	0.488	0.071	-0.012	0.072	0.500	0.489	0.066	-0.011	0.067	0.500	0.491	0.055	-0.009	0.055
	TE	0.871	0.839	0.053	-0.032	0.063	0.869	0.853	0.055	-0.016	0.062	0.782	0.912	0.030	0.129	0.134
	LL	35.913	38.361	11.598	2.449	2.930	49.303	56.840	11.604	7.537	8.245	43.926	101.869	11.609	57.943	58.500
1000	β_0	1.000	1.175	1.288	0.175	1.299	1.000	1.096	1.189	0.096	1.193	1.000	1.065	0.966	0.065	0.968
	β_1	1.000	1.000	0.004	0.000	0.004	1.000	1.000	0.003	0.000	0.003	1.000	1.000	0.003	0.000	0.003
	σ	0.306	0.311	0.024	0.005	0.024	0.346	0.298	0.025	-0.049	0.055	0.500	0.298	0.016	-0.202	0.203
	α	-0.204	-0.308	0.371	-0.104	0.385	-0.577	-0.473	0.395	0.104	0.409	-1.333	-1.328	0.242	0.006	0.242
	δ	-0.200	-0.245	0.275	-0.045	0.279	-0.500	-0.368	0.278	0.132	0.307	-0.800	-0.789	0.062	0.011	0.063
	ρ	0.500	0.492	0.061	-0.008	0.062	0.500	0.494	0.057	-0.006	0.057	0.500	0.497	0.047	-0.003	0.047
	TE	0.871	0.831	0.059	-0.039	0.072	0.869	0.868	0.047	-0.001	0.049	0.782	0.919	0.007	0.137	0.137
	LL	144.531	147.283	22.335	2.752	3.321	198.519	221.447	22.526	22.929	23.704	177.164	401.281	23.255	224.117	224.666

Notes: TE and LL represent respectively technical efficiency and log likelihood value

Table 6.11: Simulation results for $\rho = 0.8$

N	Parameters	$\delta = -0.2$					$\delta = -0.5$					$\delta = -0.8$				
		True	Mean	SD	Bias	RMSE	True	Mean	SD	Bias	RMSE	True	Mean	SD	Bias	RMSE
250	β_0	1.000	2.225	3.250	1.225	3.472	1.000	2.026	3.032	1.026	3.200	1.000	1.751	2.513	0.751	2.622
	β_1	1.000	1.000	0.007	0.000	0.007	1.000	1.000	0.007	0.000	0.007	1.000	1.000	0.006	0.000	0.006
	σ	0.306	0.319	0.037	0.013	0.039	0.346	0.303	0.039	-0.044	0.058	0.500	0.293	0.036	-0.207	0.210
	α	-0.204	-0.437	0.518	-0.233	0.568	-0.577	-0.541	0.564	0.037	0.565	-1.333	-1.305	0.604	0.029	0.605
	δ	-0.200	-0.305	0.324	-0.105	0.341	-0.500	-0.365	0.337	0.135	0.363	-0.800	-0.723	0.229	0.077	0.242
	ρ	0.800	0.777	0.062	-0.023	0.066	0.800	0.780	0.058	-0.020	0.062	0.800	0.785	0.049	-0.015	0.051
	TE	0.871	0.839	0.054	-0.032	0.064	0.869	0.853	0.056	-0.016	0.062	0.782	0.912	0.029	0.130	0.134
	LL	34.765	37.224	11.603	2.458	2.945	48.156	55.702	11.610	7.546	8.257	42.779	100.732	11.618	57.953	58.511
1000	β_0	1.000	1.827	2.843	0.827	2.959	1.000	1.645	2.640	0.645	2.716	1.000	1.430	2.178	0.430	2.220
	β_1	1.000	1.000	0.004	0.000	0.004	1.000	1.000	0.003	0.000	0.003	1.000	1.000	0.003	0.000	0.003
	σ	0.306	0.311	0.024	0.005	0.025	0.346	0.298	0.025	-0.049	0.055	0.500	0.298	0.016	-0.202	0.203
	α	-0.204	-0.293	0.383	-0.089	0.393	-0.577	-0.460	0.406	0.117	0.423	-1.333	-1.328	0.242	0.005	0.242
	δ	-0.200	-0.229	0.286	-0.029	0.287	-0.500	-0.356	0.289	0.144	0.323	-0.800	-0.789	0.063	0.011	0.064
	ρ	0.800	0.784	0.054	-0.016	0.056	0.800	0.787	0.051	-0.013	0.052	0.800	0.792	0.042	-0.008	0.043
	TE	0.871	0.818	0.064	-0.052	0.083	0.869	0.856	0.061	-0.012	0.065	0.782	0.919	0.007	0.136	0.137
	LL	143.191	145.948	22.321	2.757	3.323	197.179	220.101	22.528	22.923	23.698	175.824	399.936	23.256	224.112	224.661

Notes: TE and LL represent respectively technical efficiency and log likelihood value

Table 6.12: Production frontier estimates, model without climatic and other environmental variables (Model 1)

	OLS 1	SAR 1	SEM 1	SFA 1	SAR-SFA 1	SEM-SFA 1
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	11.268 (0.029)***	7.151 (0.170)***	11.244 (0.037)***	11.720 (0.093)***	7.486 (0.238)***	12.287 (0.124)***
Land	0.772 (0.012)***	0.627 (0.014)***	0.853 (0.014)***	0.765 (0.013)***	0.624 (0.014)***	0.848 (0.015)***
Family labor	0.073 (0.018)***	0.108 (0.017)***	0.095 (0.017)***	0.073 (0.018)***	0.108 (0.017)***	0.096 (0.017)***
Operating costs	0.032 (0.003)***	0.027 (0.003)***	0.022 (0.003)***	0.033 (0.003)***	0.028 (0.003)***	0.023 (0.003)***
ρ		0.346 (0.014)***			0.345 (0.014)***	
λ			0.587 (0.014)***			0.588 (0.014)***
σ				0.989 (0.043)***	0.883 (0.056)***	0.855 (0.026)***
α				−0.700 (0.163)***	−0.529 (0.261)**	−0.812 (0.109)***
Mean Efficiency				0.784 (0.154)	0.777 (0.122)	0.791 (0.174)
Log Likelihood	-5708.521	-5442.834	-5110.123	-4174.473	-3909.837	-3572.743
N	4423	4423	4423	4423	4423	4423

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors are in parentheses

Table 6.13: Production frontier estimates, model with climatic variables (Model 2)

	OLS 2 (1)	SAR 2 (2)	SEM 2 (3)	SFA 2 (4)	SAR-SFA 2 (5)	SEM-SFA 2 (6)
Constant	10.512 (0.428)***	7.806 (0.428)***	10.822 (0.796)***	10.531 (0.428)***	7.816 (0.301)***	11.928 (0.458)***
Land	0.803 (0.013)***	0.679 (0.015)***	0.860 (0.014)***	0.803 (0.013)***	0.679 (0.013)***	0.858 (0.015)***
Family labor	0.052 (0.018)***	0.082 (0.017)***	0.087 (0.017)***	0.052 (0.018)***	0.082 (0.017)***	0.088 (0.017)***
Operating costs	0.033 (0.003)***	0.028 (0.003)***	0.023 (0.003)***	0.033 (0.003)***	0.028 (0.003)***	0.023 (0.003)***
Rainfall	0.824 (0.060)***	0.652 (0.057)***	0.801 (0.112)***	0.825 (0.045)***	0.652 (0.047)***	0.766 (0.089)***
Temperature	0.020 (0.011)*	−0.009 (0.011)	0.010 (0.021)	0.020 (0.009)**	−0.009 (0.008)	0.002 (0.012)
Rainfall Anomaly	−3.483 (0.262)***	−2.525 (0.252)***	−3.354 (0.478)***	−3.482 (0.320)***	−2.525 (0.151)***	−3.313 (0.424)***
Temperature Anomaly	−20.649 (3.199)***	−14.041 (3.037)***	−18.113 (5.987)***	−20.664 (0.176)***	−14.044 (0.909)***	−17.919 (0.615)***
ρ		0.310 (0.014)***			0.310 (0.014)***	
λ			0.547 (0.015)***			0.550 (0.015)***
σ				0.846 (0.011)***	0.800 (0.009)***	0.821 (0.030)***
α				−0.026 (0.431)	−0.015 (0.464)	−0.668 (0.135)***
Mean Efficiency				0.754 (0.091)	0.752 (0.091)	0.789 (0.166)
Log Likelihood	-5536.755	-5328.830	-5065.926	-4003.861	-3795.935	-3531.792
N.	4423	4423	4423	4423	4423	4423

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors are in parentheses

Table 6.14: Production frontier estimates, model with climatic and other environmental variables (Model 3)

	OLS 3	SAR 3	SEM 3	SFA 3	SAR-SFA 3	SEM-SFA 3
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	10.314 (0.422)***	7.598 (0.419)***	10.449 (0.757)***	10.322 (1.227)***	7.895 (0.298)***	11.737 (0.282)***
Land	0.851 (0.013)***	0.731 (0.015)***	0.867 (0.014)***	0.851 (0.013)***	0.730 (0.015)***	0.863 (0.014)***
Family labor	0.051 (0.018)***	0.080 (0.017)***	0.087 (0.017)***	0.051 (0.018)***	0.080 (0.017)***	0.087 (0.017)***
Operating costs	0.024 (0.003)***	0.019 (0.003)***	0.018 (0.003)***	0.024 (0.003)***	0.019 (0.003)***	0.019 (0.003)***
Rainfall	0.173 (0.084)**	−0.007 (0.080)	0.273 (0.144)*	0.173 (0.063)***	−0.015 (0.062)	0.198 (0.109)*
Temperature	0.013 (0.011)	−0.016 (0.011)	0.011 (0.020)	0.013 (0.013)	−0.017 (0.008)**	−0.000 (0.008)
Rainfall Anomaly	−3.389 (0.260)***	−2.501 (0.249)***	−3.321 (0.456)***	−3.390 (0.201)***	−2.491 (0.207)***	−3.249 (0.360)***
Temperature Anomaly	−6.465 (3.423)*	−1.104 (3.236)	−7.897 (6.033)	−6.460 (0.405)***	−1.047 (0.439)**	−7.088 (0.081)***
Clay soil	0.016 (0.004)***	0.023 (0.003)***	0.013 (0.005)***	0.016 (0.004)***	0.023 (0.004)***	0.014 (0.005)***
Silt soil	0.016 (0.004)***	0.009 (0.004)**	0.013 (0.006)**	0.016 (0.004)***	0.009 (0.004)**	0.014 (0.006)**
Improved seed	0.107 (0.031)***	0.099 (0.030)***	0.109 (0.030)***	0.107 (0.032)***	0.100 (0.030)***	0.114 (0.030)***
Education	0.060 (0.027)**	0.051 (0.025)**	0.066 (0.025)***	0.060 (0.027)**	0.051 (0.025)**	0.068 (0.026)***
FBO Membership	0.324 (0.047)***	0.259 (0.044)***	0.145 (0.045)***	0.324 (0.047)***	0.260 (0.044)***	0.148 (0.045)***
ρ		0.310 (0.014)***			0.310 (0.014)***	
λ			0.523 (0.016)***			0.525 (0.016)***
σ				0.827 (0.013)***	0.822 (0.069)***	0.843 (0.025)***
α				−0.013 (1.433)	−0.425 (0.388)	−0.788 (0.106)***
Mean Efficiency				0.752 (0.089)	0.782 (0.139)	0.792 (0.172)
Log Likelihood	-5436.144	-5222.687	-5034.034	-3903.249	-3689.896	-3497.542
N.	4423	4423	4423	4423	4423	4423

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors are in parentheses

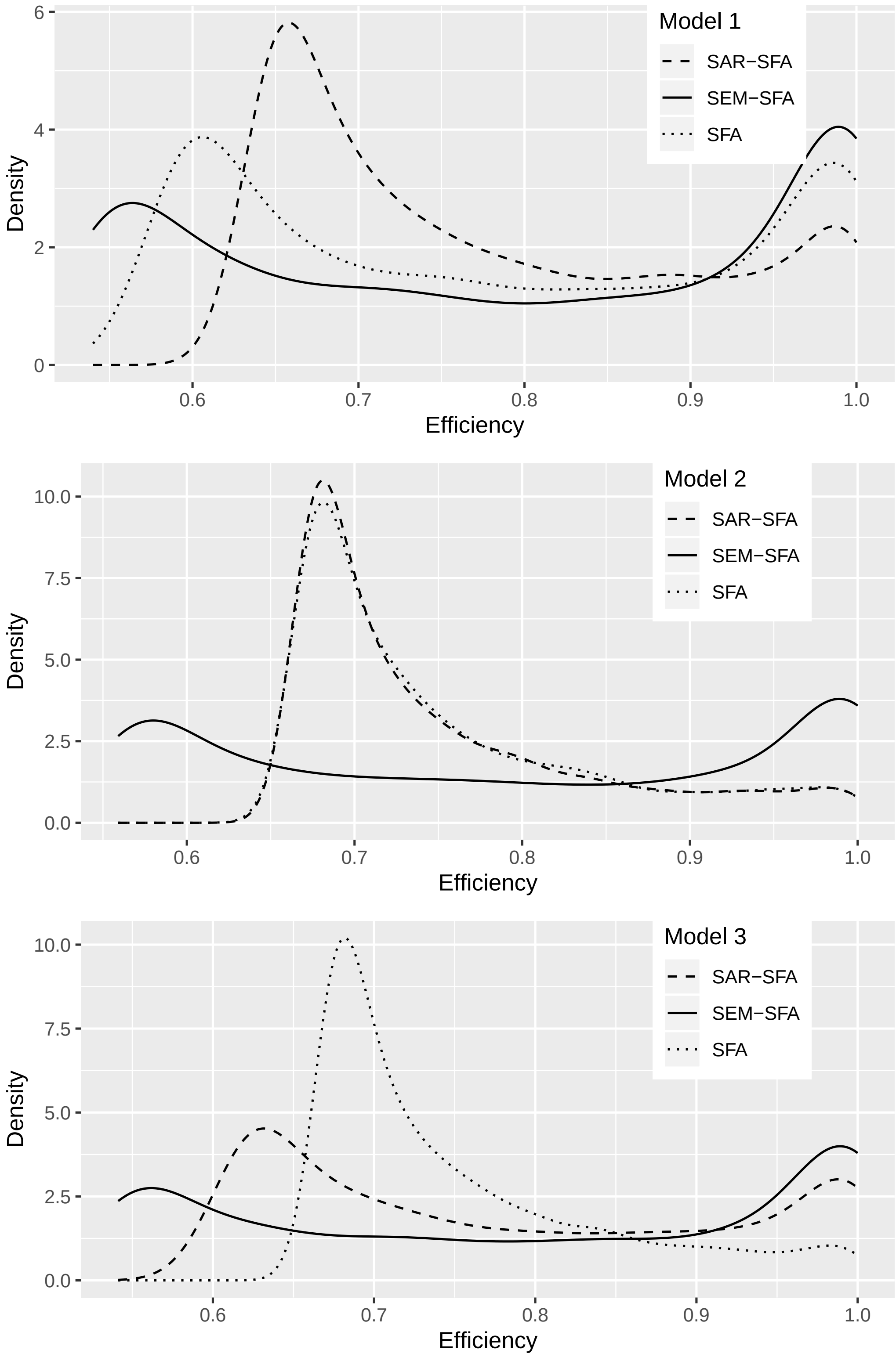
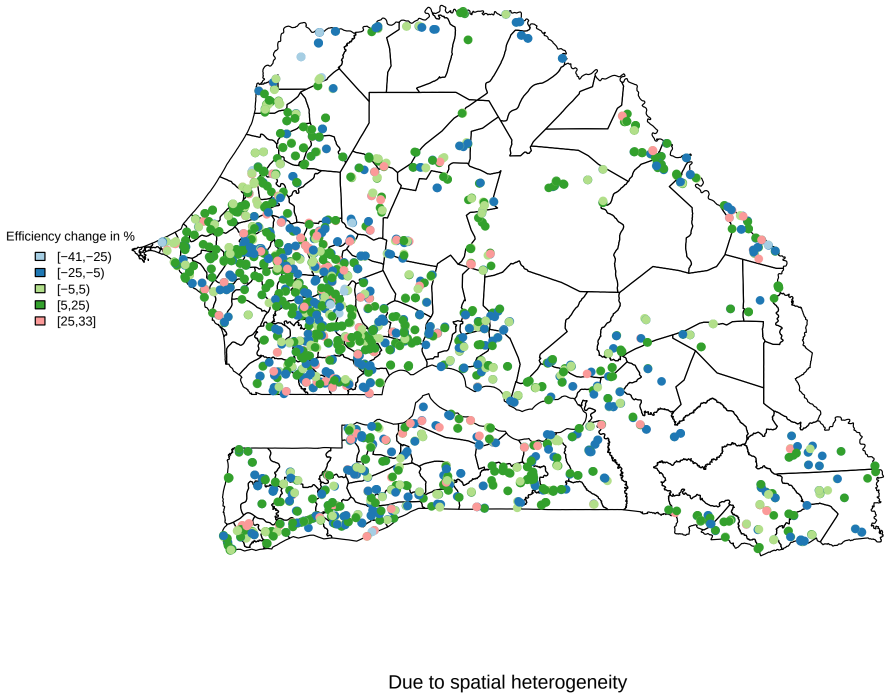


Figure 6.1: Kernel distributions of technical efficiency scores for models 1, 2 and 3



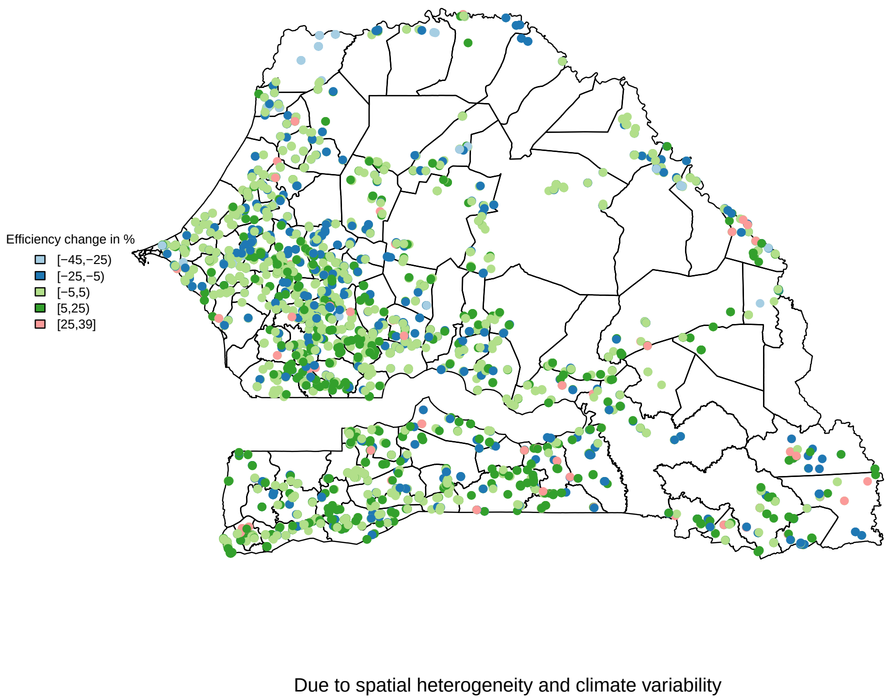


Figure 6.2: Percentage change in individual efficiency scores

References

Adetutu, M., Glass, A. J., Kenjegalieva, K. and Sickles, R. C. (2015) The effects of efficiency and TFP growth on pollution in Europe: a multistage spatial analysis, *Journal of Productivity Analysis*, **43**, 307–326.

Affuso, E. (2010) Spatial autoregressive stochastic frontier analysis: An application to an impact evaluation study, *Available at SSRN 1740382*.

Aigner, D., Lovell, C. and Schmidt, P. (1977) Formulation and estimation of stochastic frontier production function models, *Journal of Econometrics*, **6**, 21 – 37.

Anselin, L. (1988) *Spatial Econometrics: Methods and Models*, NATO Asi Series. Series E, Applied Sciences, Springer Netherlands.

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

Areal, F. J., Balcombe, K. and Tiffin, R. (2012) Integrating spatial dependence into stochastic frontier analysis, *Australian Journal of Agricultural and Resource Economics*, **56**, 521–541.

Arellano-Valle, R. B. and Azzalini, A. (2006) On the unification of families of skew-normal distributions, *Scandinavian Journal of Statistics*, **33**, 561–574.

Arellano-Valle, R. B. and Azzalini, A. (2008) The centred parametrization for the multivariate skew-normal distribution, *Journal of Multivariate Analysis*, **99**, 1362 – 1382, special Issue: Multivariate Distributions, Inference and Applications in Memory of Norman L. Johnson.

Azzalini, A. (1985) A class of distributions which includes the normal ones, *Scandinavian Journal of Statistics*, **12**, 171–178.

Azzalini, A. (2005) The skew-normal distribution and related multivariate mamilies, *Scandinavian Journal of Statistics*, **32**, 159–188.

Azzalini, A. and Capitanio, A. (1999) Statistical applications of the multivariate skew normal distribution, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **61**, 579–602.

Azzalini, A. and Valle, A. D. (1996) The multivariate skew-normal distribution, *Biometrika*, **83**, 715–726.

Ba, B. (2006) *Etude géographique de l'agriculture en Afrique noire: analyse des productions céréalières et des systèmes alimentaires au Sénégal*, Ph.D. thesis, University of Geneva.

Barrios, E. B. and Lavado, R. F. (2010) Spatial stochastic frontier models, Tech. rep., PIDS Discussion Paper Series.

Barrios, S., Bertinelli, L. and Strobl, E. (2010) Trends in rainfall and economic growth in Africa: A neglected cause of the African growth tragedy, *The Review of Economics and Statistics*, **92**, 350–366.

Carvalho, A. (2018) Efficiency spillovers in Bayesian stochastic frontier models: application to electricity distribution in New Zealand, *Spatial Economic Analysis*, **13**, 171–190.

D'Alessandro, S., Fall, A. A., Grey, G., Simpkin, S. and Wane, A. (2015) Senegal. Agricultural sector risk assessment, techreport 96296-SN, World Bank.

de Graaff, T. (2020) On the estimation of spatial stochastic frontier models: an alternative skew-normal approach, *The Annals of Regional Science*, **64**, 267–285.

Dedehouanou, F. (2014) Diversification towards high value export production, farm household productivity and efficiency in Senegal, in *17th Annual Conference on Global Economic Analysis, Dakar, Senegal*, Purdue University, West Lafayette, Global Trade Analysis Project (GTAP).

Diagne, M., Demont, M., Seck, P. A. and Diaw, A. (2013) Self-sufficiency policy and irrigated rice productivity in the Senegal River Valley, *Food Security*, **5**, 55–68.

Domınguez-Molina, J. A., González-Farıas, G. and Ramos-Quiroga, R. (2003) Skew-normality in stochastic frontier analysis.

Druska, V. and Horrace, W. C. (2004) Generalized Moments Estimation for Spatial Panel Data: Indonesian Rice Farming, *American Journal of Agricultural Economics*, **86**, 185–198.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A. *et al.* (2015) The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes, *Scientific data*, **2**, 1–21.

Funk, C., Peterson, P., Peterson, S., Shukla, S., Davenport, F., Michaelsen, J., Knapp, K. R., Landsfeld, M., Husak, G., Harrison, L., Rowland, J., Budde, M., Meiburg, A., Dinku, T., Pedreros, D. and Mata, N. (2019) A high-resolution 1983–2016 tmax climate data record based on infrared temperatures and stations by the climate hazard center, *Journal of Climate*, **32**, 5639–5658.

Fusco, E. and Vidoli, F. (2013) Spatial stochastic frontier models: controlling spatial global and local heterogeneity, *International Review of Applied Economics*, **27**, 679–694.

Glass, A., Kenjegalieva, K. and Paez-Farrell, J. (2013) Productivity growth decomposition using a spatial autoregressive frontier model, *Economics Letters*, **119**, 291 – 295.

Glass, A., Kenjegalieva, K. and Sickles, R. C. (2014) Estimating efficiency spillovers with state level evidence for manufacturing in the us, *Economics Letters*, **123**, 154 – 159.

Glass, A. J., Kenjegalieva, K. and Sickles, R. C. (2016) A spatial autoregressive stochastic frontier model for panel data with asymmetric efficiency spillovers, *Journal of Econometrics*, **190**, 289 – 300, endogeneity Problems in Econometrics.

Greene, W. H. (1990) A gamma-distributed stochastic frontier model, *Journal of Econometrics*, **46**, 141 – 163.

Han, J., Ryu, D., Sickles, R. *et al.* (2016) How to measure spillover effects of public capital stock: a spatial autoregressive stochastic frontier model, Tech. rep.

Hughes, N., Lawson, K., Davidson, A., Jackson, T. and Sheng, Y. (2011) Productivity pathways: climate-adjusted production frontiers for the australian broadacre cropping industry, in *Australian Agricultural and Resource Economics Society Conference, 2011 Conference (55th), February 8–11, 2011, Melbourne, Australia*, 2011, 422-2016-26937, p. 33.

Jalloh, A., Nelson, G., Thomas, T., Zougmoré, R. and Roy-Macauley, H. (2013) *West African Agriculture and Climate Change: A Comprehensive Analysis*, Climate change in Africa, International Food Policy Research Institute.

Lachaud, M. A., Bravo-Ureta, B. E. and Ludena, C. E. (2017) Agricultural productivity in latin america and the caribbean in the presence of unobserved heterogeneity and climatic effects, *Climatic Change*, **143**, 445–460.

LeSage, J. and Pace, R. (2009) *Introduction to Spatial Econometrics*, Statistics, textbooks and monographs, CRC Press.

Meeusen, W. and van Den Broeck, J. (1977) Efficiency estimation from cobb-douglas production functions with composed error, *International Economic Review*, **18**, 435–444.

Mulwa, R. and Kabubo-Mariara, J. (2017) Productive efficiency and its determinants in a changing climate: A monotonic translog stochastic frontier analysis, Tech. rep., Environment for Development Initiative.

Ngom, C. A. B., Sarr, F. and Fall, A. A. (2016) Mesure de l'efficacité technique des riziculteurs du bassin du fleuve sénégal, *Économie rurale. Agricultures, alimentations,* *territoires*, pp. 91–105.

O'Donnell, C. (2016) Using information about technologies, markets and firm behaviour to decompose a proper productivity index, *Journal of Econometrics*, **190**, 328 – 340.

Okuyama, Y., Maruyama, A., Takagaki, M. and Kikuchi, M. (2017) Technical efficiency and production potential of selected cereal crops in senegal, *Journal of Agriculture and Rural Development in the Tropics and Subtropics (JARTS)*, **118**, 187–197.

Orea, L. and Álvarez, I. C. (2019) A new stochastic frontier model with cross-sectional effects in both noise and inefficiency terms, *Journal of Econometrics*, **213**, 556 – 577.

Oyekale, A. S. (2012) Impact of climate change on cocoa agriculture and technical efficiency of cocoa farmers in south-west nigeria, *Journal of Human Ecology*, **40**, 143–148.

Pavlyuk, D. (2011) Application of the spatial stochastic frontier model for analysis of a regional tourism sector, *Transport and Telecommunication*, **12**, 28–38.

Pavlyuk, D. (2013) Distinguishing between spatial heterogeneity and inefficiency: Spatial stochastic frontier analysis of european airports, *Transport and Telecommunication*, **14**, 29–38.

Pede, V. O., Areal, F. J., Singbo, A., McKinley, J. and 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, *Agricultural Economics*, **49**, 301–312.

Ramajo, J. and Hewings, G. J. (2018) Modeling regional productive performance using a spatial stochastic frontier approach: New evidence for europe (1995-2007), *Regional Studies*, **52**, 1372–1387.

Republique du Senegal (2018) Note d'Analyse du Commerce Exterieur, Edition 2017, Tech. rep., Agence Nationale de la Statistique et de la Démographie, ANSD.

Schmidt, A., Moreira, A., Helfand, S. and Fonseca, T. (2009) Spatial stochastic frontier models: accounting for unobserved local determinants of inefficiency, *Journal of Productivity Analysis*, **31**, 101–112.

Seck, A. (2017) Fertiliser subsidy and agricultural productivity in senegal, *The World Economy*, **40**, 1989–2006.

Sherlund, S. M., Barrett, C. B. and Adesina, A. A. (2002) Smallholder technical efficiency controlling for environmental production conditions, *Journal of Development Economics*, **69**, 85 – 101.

Skevas, I. (2020) Inference in the spatial autoregressive efficiency model with an application to dutch dairy farms, *European Journal of Operational Research*, **283**, 356 – 364.

Stevenson, R. E. (1980) Likelihood functions for generalized stochastic frontier estimation, *Journal of Econometrics*, **13**, 57 – 66.

Thiam, A. and Bravo-Ureta, B. (2003) Technical efficiency measures for a sample of senegalese groundnut producers using pooled cross-section time series data., *International Arachis Newsletter*, pp. 36–39.

Tonini, A. and Pede, V. (2011) A generalized maximum entropy stochastic frontier measuring productivity accounting for spatial dependency, *Entropy*, **13**, 1916–1927.

Tsionas, E. G. and Michaelides, P. G. (2016) A spatial stochastic frontier model with spillovers: Evidence for italian regions, *Scottish Journal of Political Economy*, **63**, 243–257.

World Bank (2019) World bank data, world Bank Data Base.

Appendix

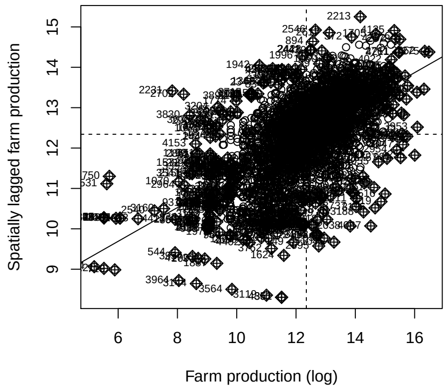


Figure 6.3: Moran plots using dependent variable

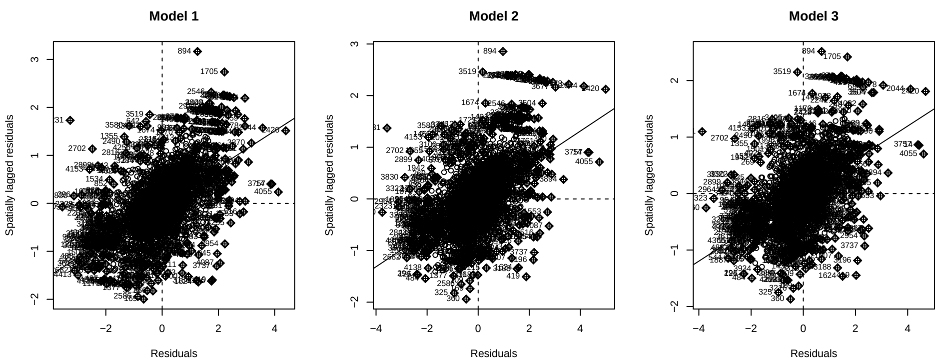


Figure 6.4: Moran plots using OLS residuals

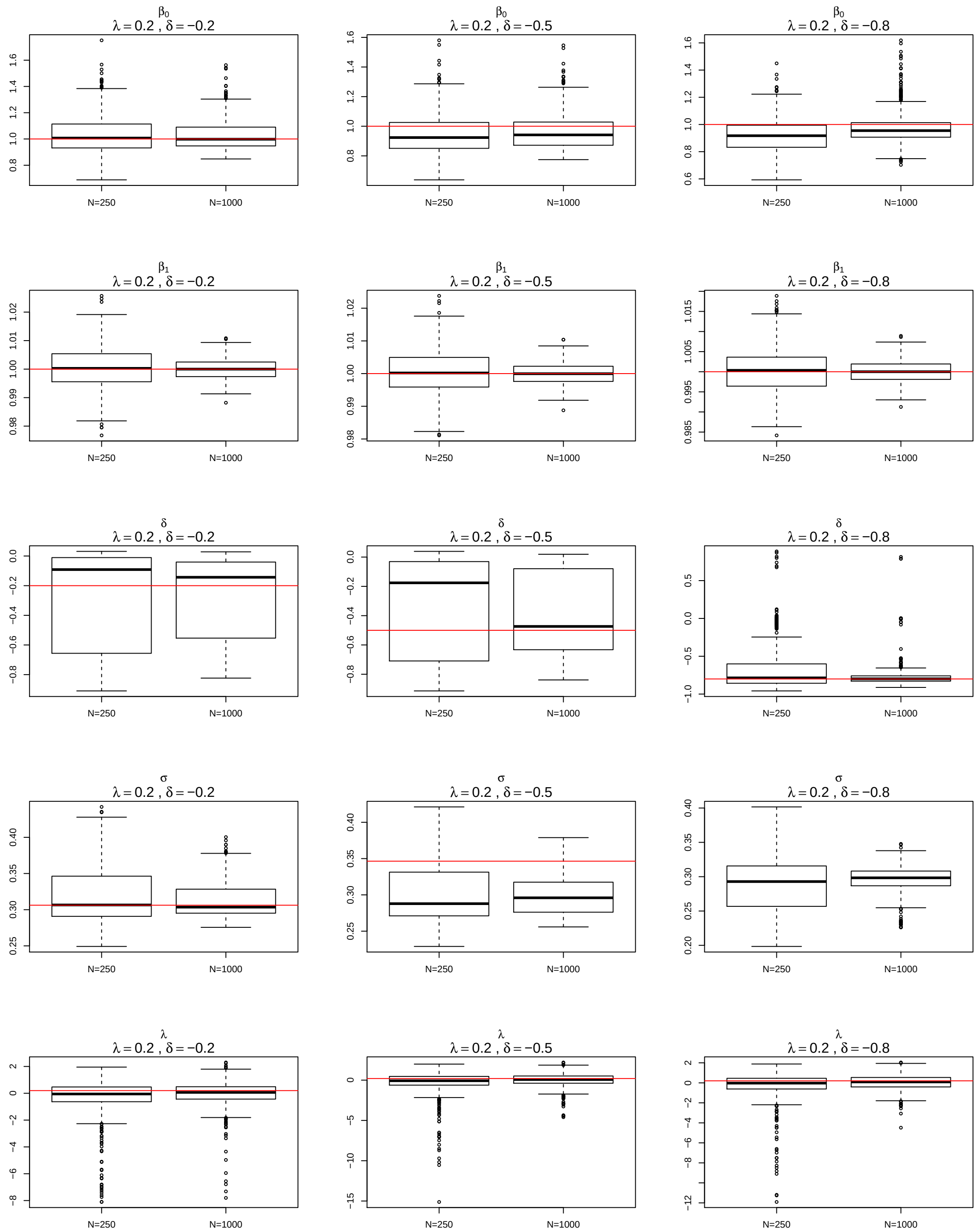


Figure 6.5: Box plots of β_0, β_1, δ, σ and λ for true λ=0.2

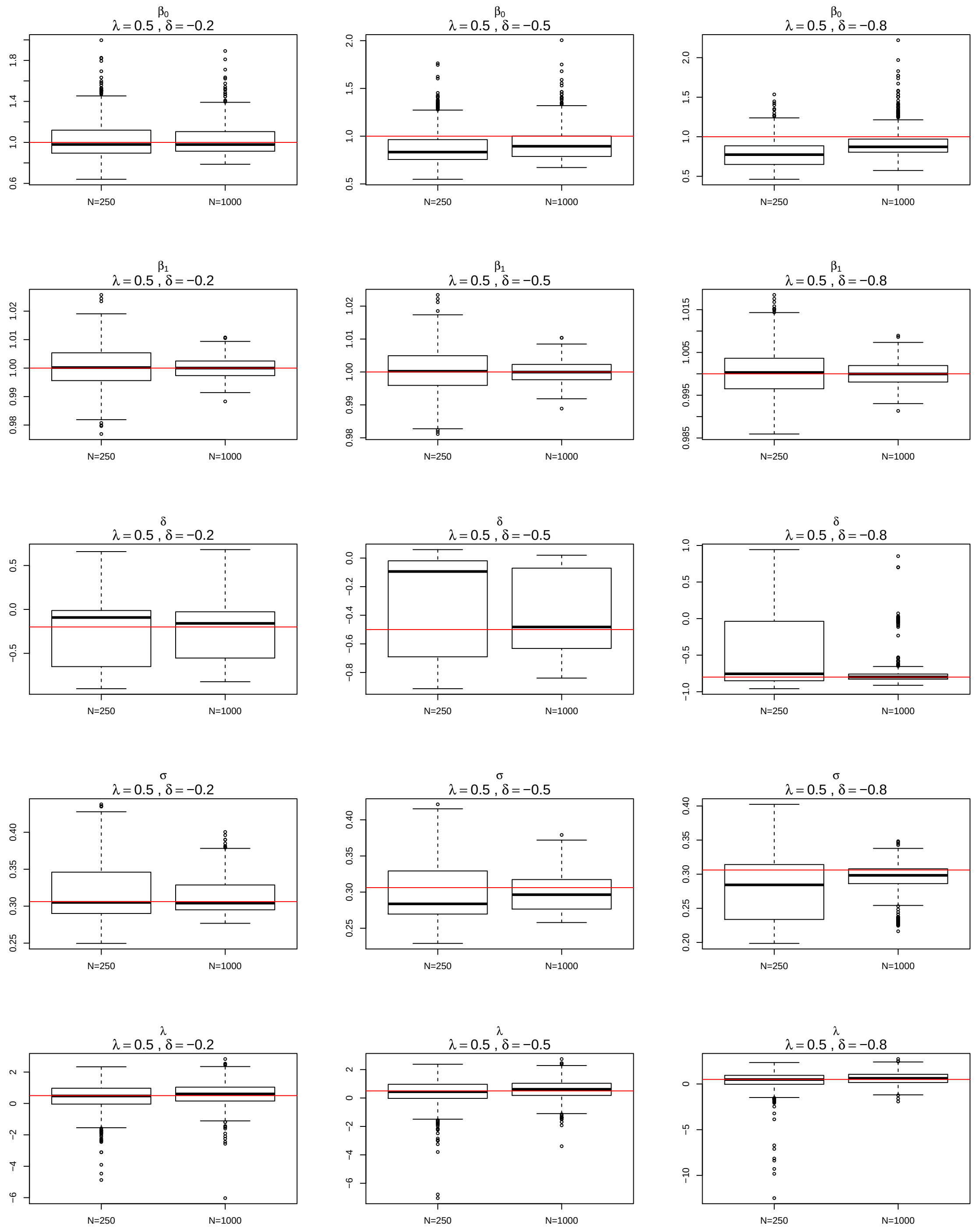


Figure 6.6: Box plots of β_0, β_1, δ, σ and λ for true λ=0.5

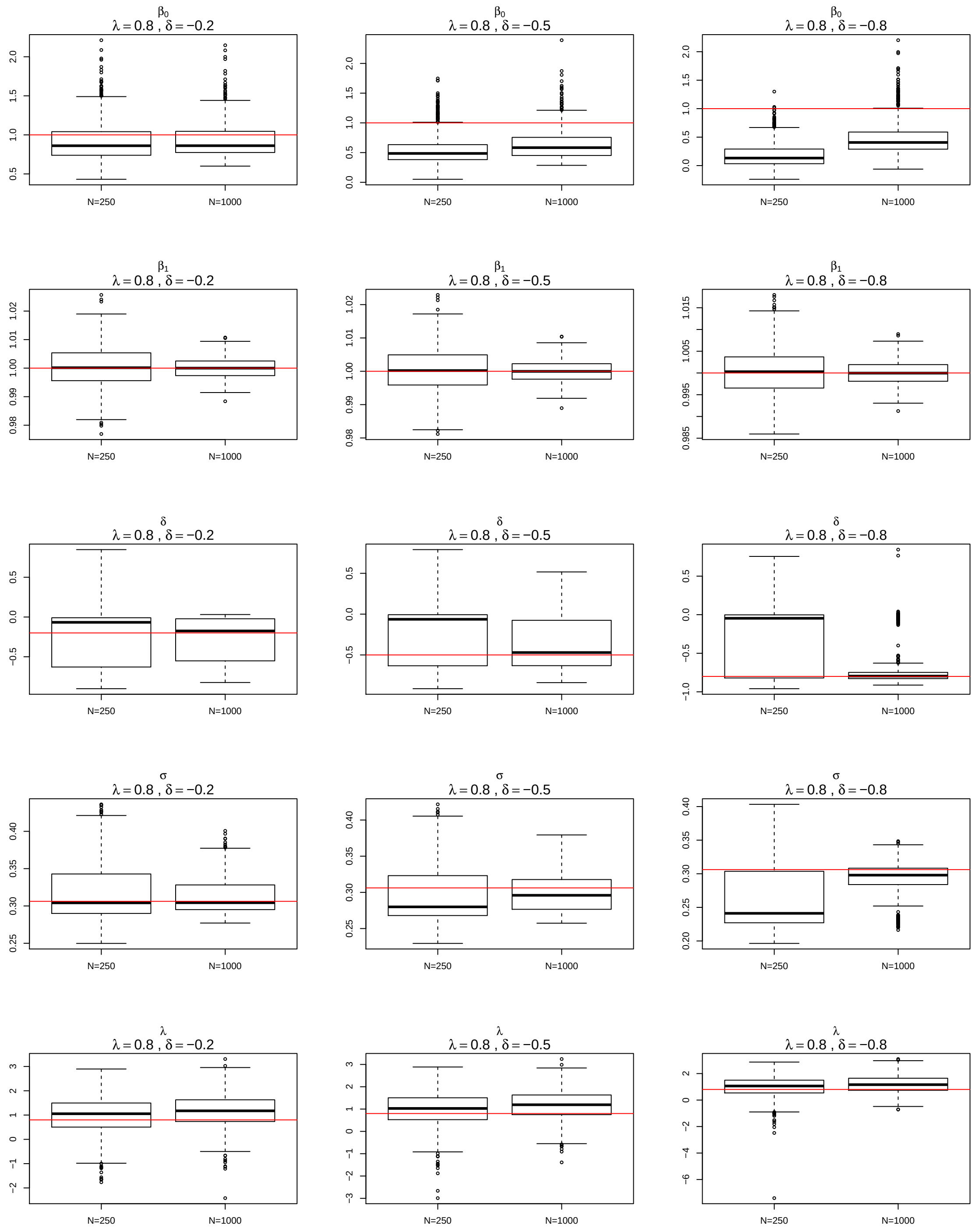


Figure 6.7: Box plots of β_0, β_1, δ, σ and λ for true λ=0.8

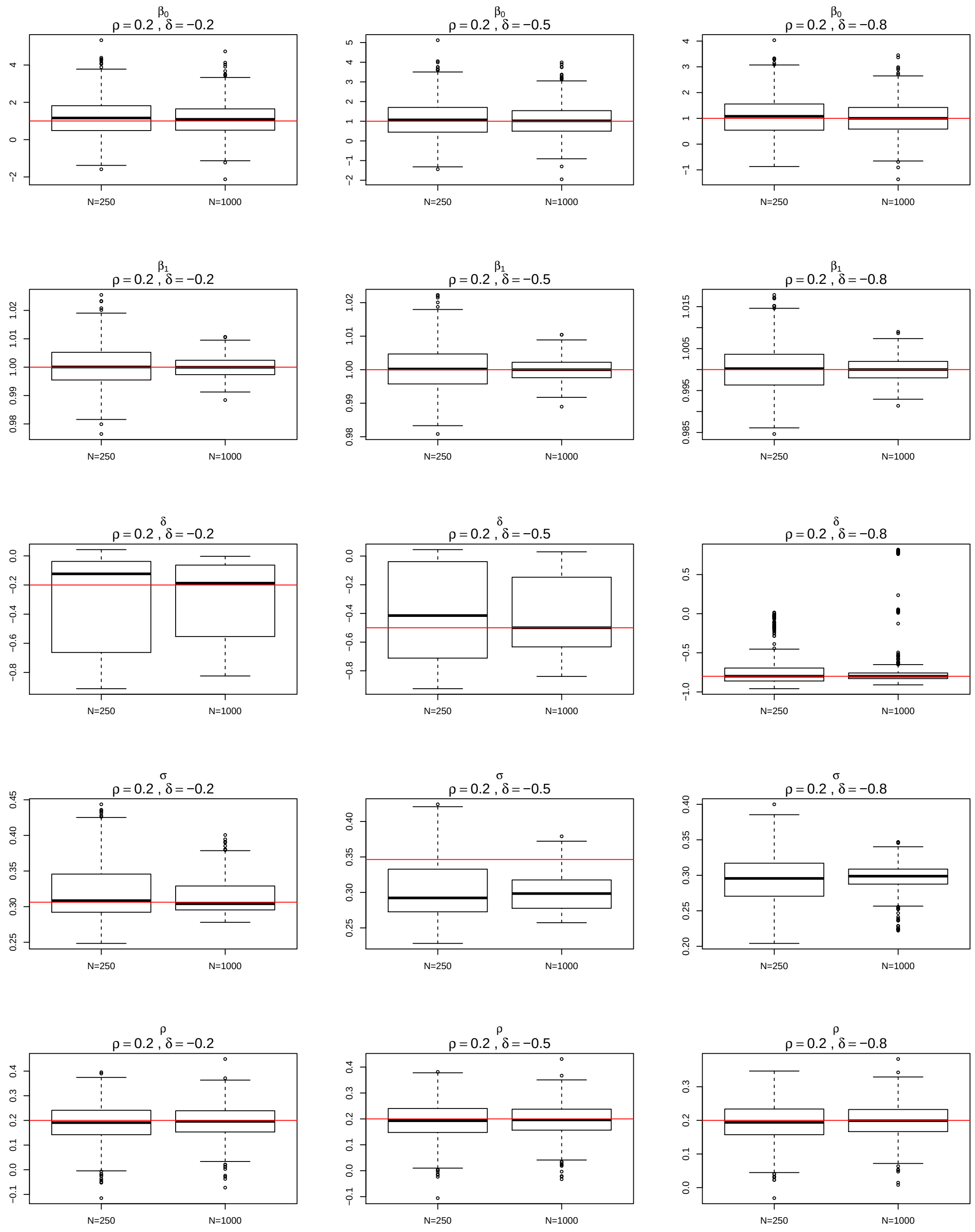


Figure 6.8: Box plots of β_0, β_1, δ, σ and ρ for true ρ=0.2

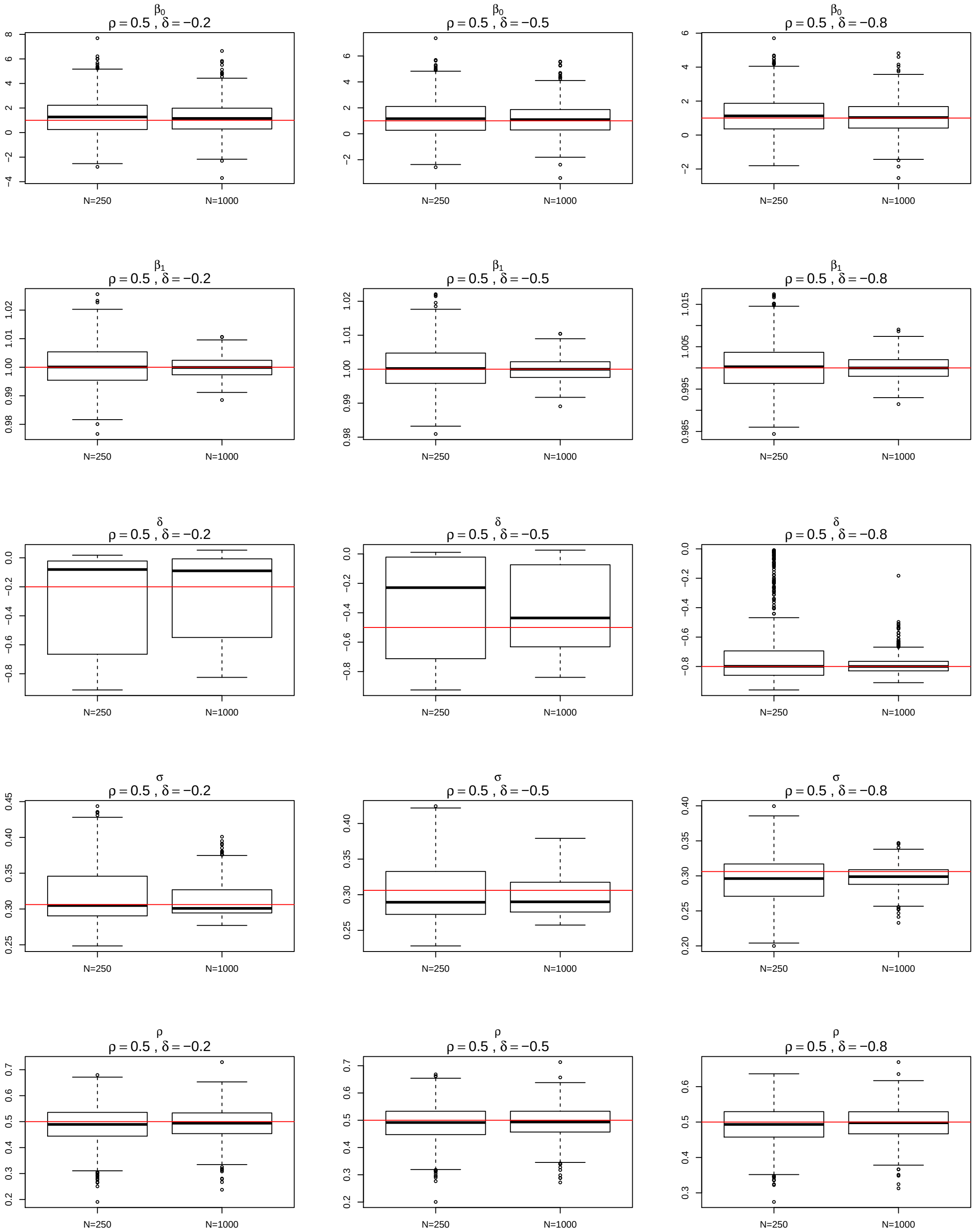


Figure 6.9: Box plots of β_0, β_1, δ, σ and ρ for true ρ=0.5

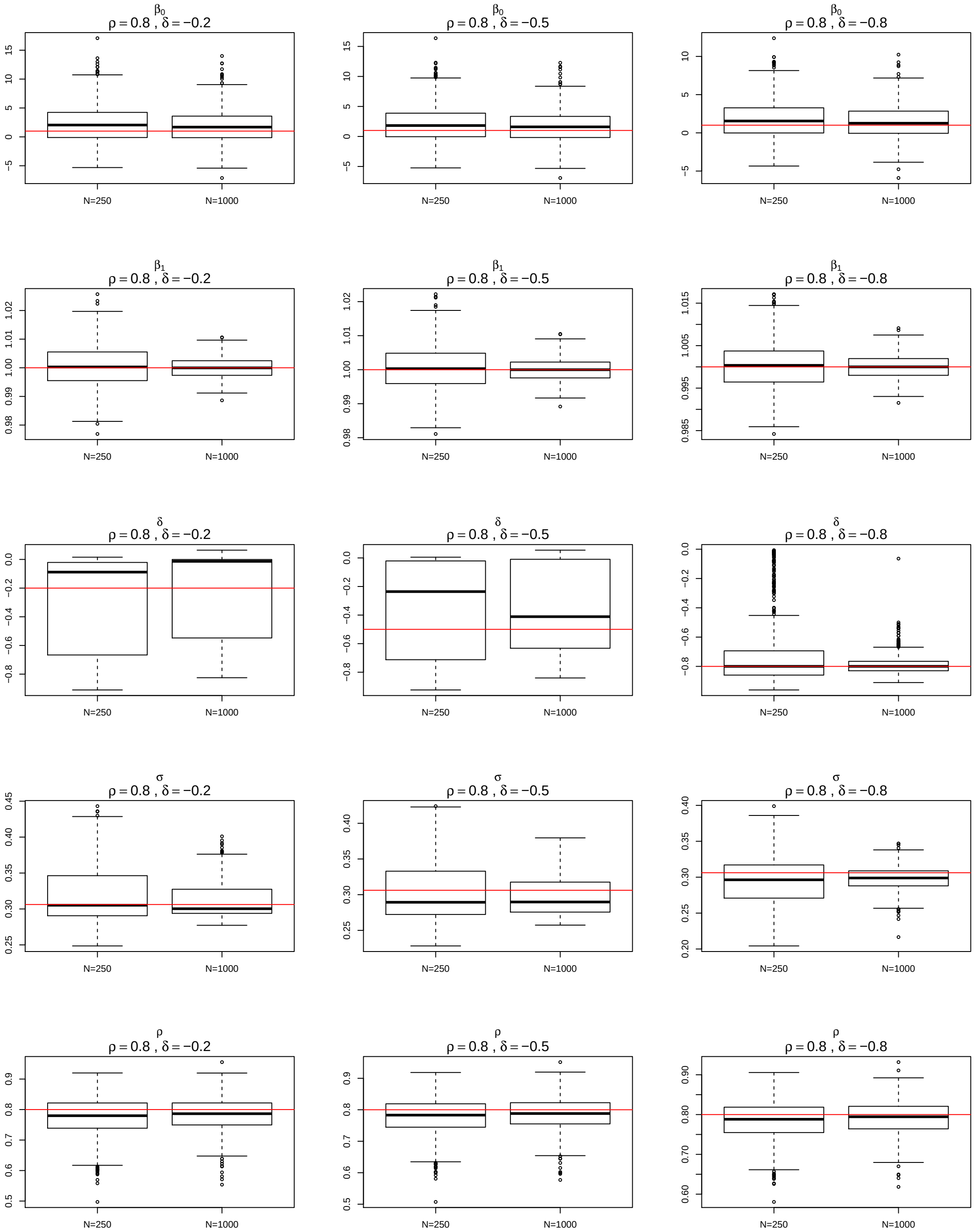


Figure 6.10: Box plots of β_0, β_1, δ, σ and ρ for true ρ=0.8

Table 6.15: Translog production frontier estimates, model without climatic and other environmental variables (Model 1)

	OLS	SAR	SEM	SFA	SAR-SFA	SEM-SFA
Constant	11.335 (0.040)***	7.435 (0.171)***	11.313 (0.044)***	11.918 (0.059)***	7.934 (0.202)***	11.313 (0.479)***
Land	0.774 (0.027)***	0.626 (0.027)***	0.786 (0.026)***	0.767 (0.027)***	0.623 (0.026)***	0.786 (0.026)***
Family labor	−0.024 (0.039)	0.002 (0.037)	0.008 (0.034)	−0.031 (0.039)	−0.003 (0.037)	0.008 (0.035)
Operating costs	−0.167 (0.014)***	−0.146 (0.013)***	−0.105 (0.012)***	−0.170 (0.014)***	−0.148 (0.013)***	−0.105 (0.013)***
Land x Land	−0.040 (0.010)***	−0.044 (0.009)***	−0.015 (0.009)	−0.043 (0.010)***	−0.044 (0.009)***	−0.015 (0.009)
Family labor x Family labor	0.110 (0.034)***	0.131 (0.031)***	0.101 (0.029)***	0.111 (0.033)***	0.131 (0.031)***	0.101 (0.029)***
Operating costs x Operating costs	0.040 (0.002)***	0.035 (0.002)***	0.025 (0.002)***	0.041 (0.002)***	0.035 (0.002)***	0.025 (0.002)***
Land x Family labor	0.068 (0.016)***	0.054 (0.015)***	0.042 (0.014)***	0.070 (0.016)***	0.055 (0.015)***	0.042 (0.014)***
Land x Operating costs	−0.018 (0.003)***	−0.013 (0.003)***	−0.005 (0.003)*	−0.019 (0.003)***	−0.014 (0.003)***	−0.005 (0.003)*
Family labor x Operating costs	−0.006 (0.004)	−0.005 (0.004)	−0.004 (0.003)	−0.005 (0.004)	−0.004 (0.004)	−0.004 (0.003)
ρ		0.328 (0.014)***			0.325 (0.015)***	
λ			0.570 (0.014)***			0.570 (0.014)***
σ				1.029 (0.028)***	0.919 (0.030)***	0.729 (0.008)***
α				−1.007 (0.104)***	−0.807 (0.120)***	−0.000 (0.355)
Mean Efficiency				0.788 (0.167)	0.790 (0.167)	0.748 (0.088)
Log Likelihood	-5553.057	-5306.002	-5037.265	-4010.806	-3769.499	-3504.370
Num. obs.	4423	4423	4423	4423	4423	4423

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors are in parentheses

Table 6.16: Translog production frontier estimates, model with climatic variables (Model 2)

	OLS	SAR	SEM	SFA	SAR-SFA	SEM-SFA
Constant	10.485 (0.417)***	7.938 (0.418)***	10.757 (0.755)***	10.512 (0.350)***	7.948	11.978 (0.188)***
Land	0.821 (0.027)***	0.692 (0.027)***	0.802 (0.026)***	0.821 (0.027)***	0.692 (0.026)***	0.799 (0.026)***
Family labor	-0.011 (0.038)	-0.002 (0.036)	0.006 (0.034)	-0.011 (0.038)	-0.002 (0.036)	0.005 (0.034)
Operating costs	-0.153 (0.013)***	-0.140 (0.013)***	-0.104 (0.012)***	-0.153 (0.013)***	-0.140 (0.013)***	-0.106 (0.012)***
Land x Land	-0.030 (0.009)***	-0.034 (0.009)***	-0.013 (0.009)	-0.030 (0.009)***	-0.034 (0.009)***	-0.014 (0.009)
Family labor x Family labor	0.084 (0.032)***	0.111 (0.031)***	0.096 (0.029)***	0.085 (0.032)***	0.111 (0.030)***	0.097 (0.029)***
Operating costs x Operating costs	0.039 (0.002)***	0.034 (0.002)***	0.025 (0.002)***	0.039 (0.002)***	0.034 (0.002)***	0.026 (0.002)***
Land x Family labor	0.065 (0.015)***	0.054 (0.015)***	0.043 (0.014)***	0.065 (0.015)***	0.054 (0.015)***	0.043 (0.014)***
Land x Operating costs	-0.019 (0.003)***	-0.015 (0.003)***	-0.006 (0.003)**	-0.019 (0.003)***	-0.015 (0.003)***	-0.006 (0.003)**
Family labor x Operating costs	-0.008 (0.004)**	-0.006 (0.003)*	-0.004 (0.003)	-0.008 (0.004)**	-0.006 (0.003)*	-0.004 (0.003)
Rainfall	0.851 (0.058)***	0.677 (0.056)***	0.809 (0.106)***	0.852 (0.048)***	0.677 (0.042)***	0.761 (0.080)***
Temperature	0.020 (0.011)*	-0.007 (0.010)	0.012 (0.020)	0.020 (0.010)**	-0.007 (0.007)	0.003 (0.006)
Rainfall Anomaly	-2.938 (0.256)***	-2.083 (0.247)***	-2.983 (0.455)***	-2.937 (0.290)***	-2.082	-2.930 (0.200)***
Temperature Anomaly	-19.511 (3.124)***	-13.217 (2.978)***	-17.773 (5.682)***	-19.572 (0.853)***	-13.223 (0.279)***	-17.238
ρ		0.293 (0.014)***			0.293 (0.014)***	
λ			0.527 (0.015)***			0.530 (0.016)***
σ				0.818 (0.012)***	0.777 (0.009)***	0.828 (0.025)***
α				-0.036 (0.459)	-0.017 (0.451)	-0.755 (0.110)***
Mean Efficiency				0.755 (0.092)	0.752 (0.090)	0.791 (0.173)
Log Likelihood	-5388.054	-5194.900	-4992.608	-3855.159	-3662.006	-3456.784
Num. obs.	4423	4423	4423	4423	4423	4423

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors are in parentheses

Table 6.17: Translog production frontier estimates, model with climatic and other environmental variables (Model 3)

	OLS	SAR	SEM	SFA	SAR-SFA	SEM-SFA
Constant	10.337 (0.414)***	7.757 (0.412)***	10.435 (0.725)***	10.940 (0.294)***	8.263 (0.272)***	11.780 (0.402)***
Land	0.850 (0.027)***	0.726 (0.026)***	0.810 (0.026)***	0.848 (0.027)***	0.724 (0.026)***	0.807 (0.026)***
Family labor	−0.004 (0.038)	0.006 (0.036)	0.009 (0.034)	−0.007 (0.038)	0.003 (0.036)	0.007 (0.034)
Operating costs	−0.145 (0.013)***	−0.131 (0.013)***	−0.105 (0.012)***	−0.146 (0.013)***	−0.132 (0.013)***	−0.107 (0.012)***
Land x Land	−0.025 (0.009)***	−0.029 (0.009)***	−0.013 (0.009)	−0.025 (0.009)***	−0.029 (0.009)***	−0.014 (0.009)
Family labor x Family labor	0.080 (0.032)**	0.106 (0.030)***	0.094 (0.029)***	0.082 (0.032)**	0.107 (0.030)***	0.095 (0.029)***
Operating costs x Operating costs	0.035 (0.002)***	0.031 (0.002)***	0.025 (0.002)***	0.035 (0.002)***	0.031 (0.002)***	0.025 (0.002)***
Land x Family labor	0.055 (0.015)***	0.044 (0.014)***	0.041 (0.014)***	0.055 (0.015)***	0.044 (0.014)***	0.041 (0.014)***
Land x Operating costs	−0.016 (0.003)***	−0.012 (0.003)***	−0.005 (0.003)**	−0.016 (0.003)***	−0.012 (0.003)***	−0.006 (0.003)**
Family labor x Operating costs	−0.007 (0.004)**	−0.006 (0.003)*	−0.004 (0.003)	−0.007 (0.004)*	−0.005 (0.003)	−0.004 (0.003)
Rainfall	0.295 (0.083)***	0.101 (0.079)	0.335 (0.138)**	0.261 (0.063)***	0.076 (0.062)	0.253 (0.107)**
Temperature	0.014 (0.011)	−0.014 (0.010)	0.013 (0.019)	0.009 (0.008)	−0.017 (0.008)**	0.000 (0.011)
Rainfall Anomaly	−2.946 (0.255)***	−2.142 (0.245)***	−2.968 (0.438)***	−2.905 (0.229)***	−2.110 (0.317)***	−2.889 (0.274)***
Temperature Anomaly	−7.383 (3.352)**	−2.032 (3.181)	−8.240 (5.785)	−7.082 (0.339)***	−1.753 (0.277)***	−7.221
Clay soil	0.012 (0.004)***	0.019 (0.003)***	0.011 (0.005)**	0.013 (0.004)***	0.019 (0.003)***	0.011 (0.005)**
Silt soil	0.015 (0.004)***	0.009 (0.004)**	0.013 (0.006)**	0.015 (0.004)***	0.009 (0.004)**	0.014 (0.006)**
Improved seed	0.122 (0.031)***	0.111 (0.029)***	0.117 (0.029)***	0.125 (0.031)***	0.115 (0.029)***	0.122 (0.029)***
Education	0.051 (0.026)*	0.042 (0.025)*	0.058 (0.025)**	0.051 (0.026)**	0.043 (0.024)*	0.060 (0.025)**
FBO Membership	0.224 (0.046)***	0.177 (0.043)***	0.120 (0.045)***	0.225 (0.046)***	0.178 (0.044)***	0.121 (0.045)***
ρ		0.296 (0.014)***			0.296 (0.014)***	
λ			0.507 (0.016)***			0.508 (0.016)***
σ				0.906 (0.032)***	0.850 (0.032)***	0.842 (0.023)***
α				−0.703 (0.133)***	−0.661 (0.140)***	−0.840 (0.099)***
Mean Efficiency				0.788 (0.160)	0.788 (0.162)	0.794 (0.176)
Log Likelihood	-5316.782	-5116.775	-4963.964	-3782.425	-3582.761	-3425.641
Num. obs.	4423	4423	4423	4423	4423	4423

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors are in parentheses

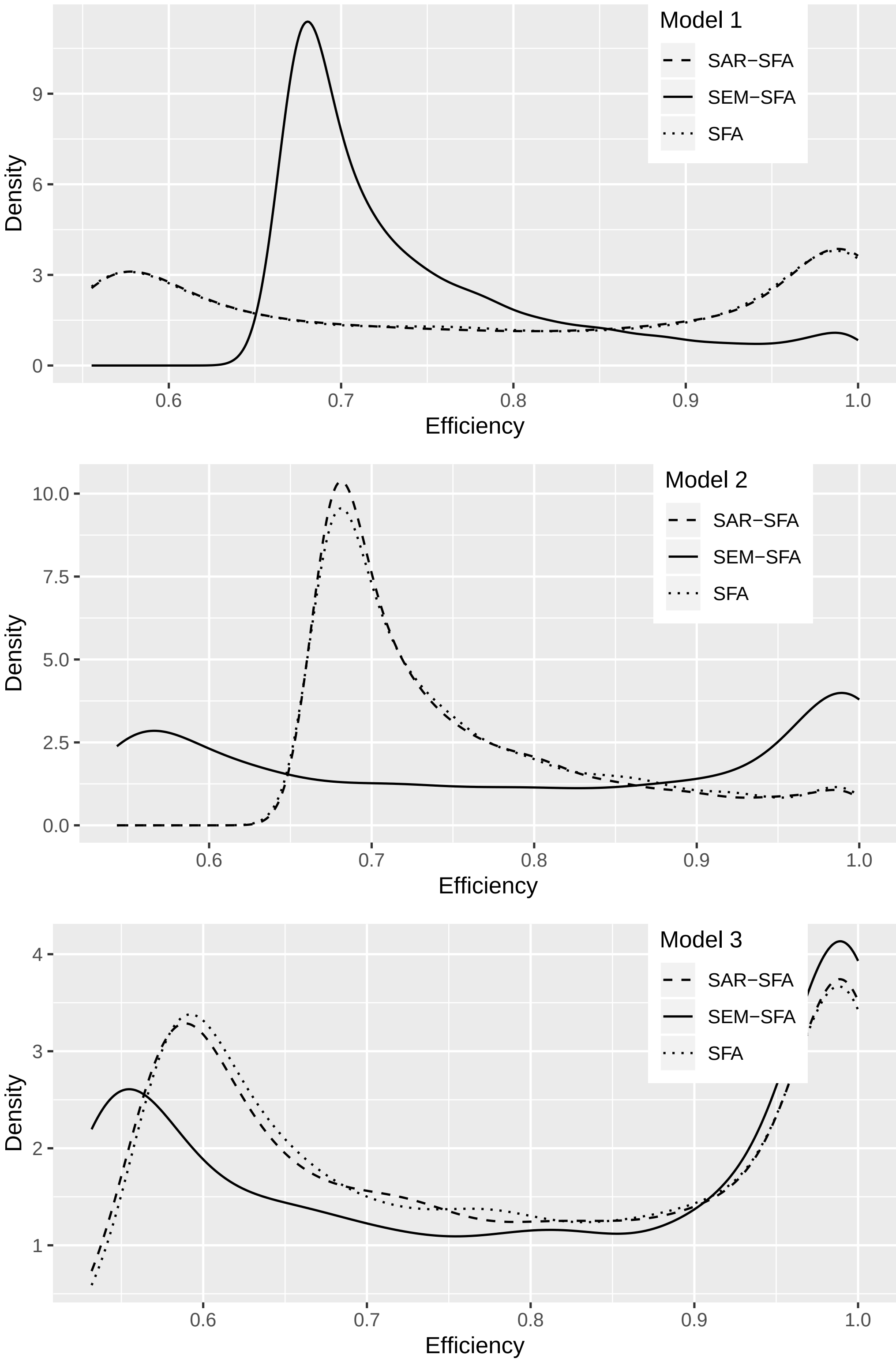


Figure 6.11: Kernel distributions of technical efficiency scores for models 1, 2 and 3 with Translog

Chapter 7
General conclusion

The recent renaissance of the Senegalese cooperative movement and the revival of the country's agricultural sector motivated this research. Agriculture in Senegal remains an important economic sector. However, its productivity is still low due to several binding constraints such as limited access to inputs and improved technologies, mainly faced by smallholder farmers. Meanwhile, agricultural cooperatives or farmer organizations have been promoted for decades as institutions to address such issues. Although in most developing countries, there is a growing and interesting literature on the contribution of these institutions on farm households' performances, less attention has been given to some specific issues related to cooperatives. This research, therefore, provided more comprehensive insights on the role of farmer organizations and of the complementary effects of spatial features on farm households performances in Senegal. To achieve this objective, the study conducted several econometric analyses using primary data, collected in 2017, and which comprises a sample of 4480 Senegalese cereal growers. The main econometric frameworks used include the stochastic frontier approach, spatial econometrics, and impact evaluation methods. These frameworks were used either separately or in combined forms. This chapter summarizes the main findings of the study and suggests some policy implications based on the findings. It also briefly presents the limitations of the research and provides some guidance for future research.

7.1 Main findings

In chapter 2, we illustrated the causal relationship between membership in farmer organizations and the ability of farm households to produce food. Food insecurity is still a serious issue in rural Senegal, and geography plays an important role. We, therefore, applied a generalised spatial two stages least squares technique that control for selection biases and spatial heterogeneity, to derive non-biased and consistent estimates. Results revealed that membership in farmer organizations affects significantly and strongly farm household's food production. Belonging to a farmer organization significantly improves cereals production by at least 19% and the daily food calories available for the household by 13%. In addition, the results show that households' food availability indicators are also positively and significantly correlated with the characteristics of the household (gender of the head, active and dependents members, the possession of agricultural assets), access to extension services and to fertilizers subsidies, and the early stop of rainfall. Furthermore, farmers' food production is also driven by spatial features. Thus, being food insecure also depends on where a farm household lives. Hence, the general conclusion that can be drawn from this chapter is that farmer organizations can help in fighting food insecurity in rural areas by providing the necessary conditions and social networks for access to technologies, knowledge and production inputs.

Chapter 3 complements chapter 2 and its objective was to determine the effectiveness of membership in farmers organizations on household land productivity and net income. The propensity score matching method and the endogenous switching regression models were used to derive treatment effects of membership in farmer organizations. Findings in this chapter first showed that factors influencing households decision to belong to farmers organizations are the education of household head, family size, distance to the nearest road, access to extension and to information on sales, and the location in various agro-ecological zones. Moreover, the study also found out that membership in farmer organizations is a key component of farm households' land productivity and income, and the estimated results appear to be consistent throughout estimation methods. Results show that being a member of a farmer organization helps to increase land productivity by almost twenty percentage points and household income by at least fourteen percentage points. Furthermore, membership in farmer organizations exhibits heterogeneous effects over the propensity score, household characteristics, and types of organization. The estimated treatment effects are negatively correlated with households' likelihood to belong to a farmer-based organization, implying that the effect of membership is stronger for households with the lowest propensity to become members, meanwhile this also suggest possible barriers of entry for these households.

In this dissertation, we also narrowed the analyses and examine the effects of membership in farmer organizations on the technical efficiency of rice farming households. We used an econometric framework that combines a propensity score matching method with the selection corrected stochastic production frontier model and a meta-frontier approach, and we derived for the two groups of farmers (members and non-members of organizations) their group-specific technical efficiency scores, the meta-technology ratios and the meta-technology technical efficiency. Results presented in chapter 4 mainly proved that belonging to farmers organizations affects positively and significantly the production of rice in Senegal. Moreover, members have higher meta-technology ratios, implying that they are operating much closer to the meta-frontier than non-members. However, non-members of farmer organizations appeared to be more technically efficient than members, contrasting with the findings of recent studies.

The objective of chapter 5 was to analyse the complementary roles of neighbourhood and membership in farmer organizations on the adoption of two productivity-enhancing technologies. Social learning constitutes a relevant component of agricultural technology adoption, and recent techniques to incorporate social interactions into the analysis of farmers behaviour include spatial econometrics. Therefore, a Bayesian Spatial Durbin Probit model was employed. Results showed the existence of high spatial dependence among households' adoption behaviours, suggesting that farmers adopt technologies when their neighbours do. Additionally, membership in a farmer organizations, number of family members, size of land owned, and experiencing a crop disease during the last five years are factors that increase the probability of adoption of productivity-enhancing technologies. This chapter findings suggest a complementary effect of neighbourhood and membership in farmer organizations, in facilitating information exchange between farmers.

Chapter 6 presented the results of an assessment analysis of the technical efficiency of Senegalese farmers in the context of climate variability and spatial heterogeneity. In this chapter, we first show with simulation results that the newly developed maximum likelihood technique estimation based on the skew-normal distribution of errors in the spatial stochastic production frontier provides consistent results. However, this estimation technique requires that the starting values for the inefficiency term are designed in an appropriate manner. Empirical results reveal that Senegalese agriculture is more dependent on the land area and dedicated labour. Land's contribution to agricultural production is about 86%, followed by family labour which is about 9%. Moreover, results show that farms efficiency is highly affected by both climatic features and spatial heterogeneity, and not accounting for them might lead to biased results for the efficiency distribution. Farm efficiency can on average be expanded by 21% when accounting for spatial heterogeneity. Furthermore, findings also revealed that the changes in technical efficiency score could be observed in any part of the country, implying that the effects of climate variability and unobserved spatial features are not specific to a region, but common to the entire country.

7.2 Policy implications

The findings of this dissertation have shown that farmer organizations still constitute effective policy instruments to enhance farm production, land productivity, household food security and income, and technology adoption, although the technical efficiency of organizations members is questionable. The findings also revealed that the impact membership in farmer organizations although positive are heterogeneous across the probability of membership, households' characteristics, and the legal type of organizations. Moreover, findings have established the presence and importance of spatial heterogeneity and neighbourhood effects in the analyses. We observed a clustering effect of membership in farmer organizations and the spatial effects in the outcomes of interest. Based on these results, some specific policy recommendations can be made.

Firstly, the Senegalese government and international donors should continue supporting farmer organizations. Such supports can be an effective way to fight rural poverty and food insecurity. However, supporting efforts need to take into account the spatial distribution of households. The approach would at first imply the targeting of areas where the most vulnerable farmers are located. Secondly, we have demonstrated that within farmer organizations, not all farmers seem to benefit from membership. We have also shown the importance of spatial spillovers effects. Therefore, policymakers when designing programmes that are aimed at increasing the rates of technologies uptake, agricultural productivity and efficiency, should consider the social networks from both farmer organizations and neighbourhoods. Such an approach would enable a large range of farmers (not just members of farmer organizations but also non-members) to have access to technical knowledge. Lastly, we also observed spatial heterogeneity in the distribution of farmers technical efficiency scores. Thus, policymakers should encourage the design and dissemination of agricultural technologies that are very adaptable to specific spatial conditions of farmers (e.g. climate adaptation or other farming techniques).

7.3 Future research

This dissertation has shed light on several research issues related to farmer organizations. Nevertheless, there is one key limitation that should be considered for future research. Our study is based on cross-sectional data. Therefore, we were not able to control for unobserved heterogeneity due to time-invariant factors. Future research should use panel data. Such a strategy would also help in bringing better insight of the long-run impact of cooperatives in the developing world.

In chapter 6, due to actual limitations in estimation techniques, the designed spatial stochastic frontier approach could not include an inefficiency model (i.e. a one step model estimation with incorporation of variables in the inefficiency term). It prevented us to conduct robustness checks regarding our spatial efficiency analysis. Therefore, future research should focus on designing estimation techniques that permit the incorporation of environmental variables in a one-step approach framework for the spatial stochastic frontier analysis. Moreover, future studies should also pursue the simulation works done in the same chapter and investigate more the empirical performances of the skew-normal approach. Such studies should explore the problem of starting values that we encountered, and the potential "centered versus non-centered parameterizations" problem. Furthermore, powerful econometrics tools such as the expectation-maximization (EM) algorithm, the weighted method of moments and the Bayesian approach are also promising ways that could be explored to estimate the spatial stochastic frontier model.

This research has considered agriculture as one sector and the various farmer organizations as one entity. However, in Senegal, agriculture is multi-sectoral and farmer organizations are of several legal and organizational forms. Thus, future research could focus on specific sub-sector (such as maize, groundnut, etc) and specific type of organizations (e.g. Economic Interest Group). Such a research approach would be helpful in comparing the different forms of organizations across sub-sectors, and in determining which legal forms of horizontal coordinated institutions should be promoted in the Senegalese context.

Appendix A

Complement to methodology

A.1 Background on methodology

A.1.1 Spatial effects

Three chapters (2, 5, and 6) of this dissertation deal with the existence of spatial effects in the estimations. Spatial effects can be defined as spatial interactions contained in the data of analysis, and ignoring such characteristics of the data can lead to ineffective or biased estimates and misleading inferences (Anselin, 1988b, 2001; Holloway, 2007; LeSage and Pace, 2009). In agricultural economics, past empirical studies mostly ignored the spatial features of the data in their analysis (Skevas *et al.*, 2018). Two types of spatial effects are analysed in this study: the spatial dependence (or neighbourhood) and spatial heterogeneity. Following Anselin (1988b), spatial dependence occurs when there is a functional relationship between observed values at one location and observations at nearby locations (e.g. neighbourhood effects), whereas spatial heterogeneity refers to the variation with locations, of the functional relationship between the observations.

To deal with spatial effects in economic analysis, standard econometrics techniques are not always applicable, therefore analysts resort to spatial econometrics techniques[1]. In spatial econometrics, two categories of regression models have been suggested in the literature. The first category intends to explain the spatial dependence in variables of surveyed units. The spatial autoregressive model (SAR) (LeSage and Pace, 2009) is an example of this category. The second category of models incorporates spatial dependence in the error terms, and are designed to bring out the spatial heterogeneity between surveyed units. The spatial error model (SEM) (Cliff and Ord, 1973) is an example of this category. The spatial autoregressive model can

[1] See (LeSage and Pace, 2009) for detailed information on spatial econometrics models and estimation techniques.

be specified as:

$$y = \rho W y + \beta X + \epsilon,$$

and the spatial error model can be expressed as:

$$y = \beta X + \epsilon \quad \text{with} \quad \epsilon = \lambda W \epsilon + \tilde{\epsilon},$$

where y is a $N \times 1$ vector of dependent variable, X the $N \times K$ matrix of covariates, ϵ and $\tilde{\epsilon}$ are $N \times 1$ errors terms. β is a $N \times K$ parameters to be estimated, and ρ and λ are scalar parameters to be determined. N and K represent the sample size and the number of covariates respectively. W is a pre-specified weight matrix. These models have also been adapted to the field of productivity and efficiency analysis particularly the stochastic frontier approach.

One of the issues when using spatial econometrics techniques is to design a spatial weight matrix prior to estimations. A spatial weight matrix is a square matrix that captures the spatial relations between all surveyed units. Several forms of weight matrices are proposed in literature from the earlier binary contiguity matrix to the general weight matrices based on distances (either geographic, economic, or social, etc.) between spatial units[2]. The specification of the weight matrix constitutes an important step because differences in weight matrices are sources of differences in obtained results, and any misspecification can lead to biased estimates of the spatial effects (Páez *et al.*, 2008; Plumper and Neumayer, 2010; LeSage and Pace, 2014). However, according to LeSage and Pace (2014), the sensitivity of results to weight matrices arises only when models estimates are incorrectly interpreted or the models are misspecified.

When analysing spatial effects, it is necessary to first show the existence of spatial features in the data. The Moran test statistics can be used for that purpose (Moran, 1950; Anselin, 1988b). Secondly, to select the appropriate model, theoretical aspects should be taken into account, with the support of statistical tests such as the Lagrange Multiplier test (Anselin, 1988a) and their robust counterparts (Anselin *et al.*, 1996). These tests can be complemented by the likelihood ratio tests[3].

Estimation of spatial econometrics models requires non-standard estimation methods. Ordinary Least Squares estimator generally produces non-consistent estimates

[2]Literature and discussions on the specifications of spatial weight matrices can be found in Anselin (1988b); Anselin and Bera (1998); Páez *et al.* (2008); Getis (2009), and Plumper and Neumayer (2010)

[3]See Elhorst (2010) for more details on suggested testing procedures.

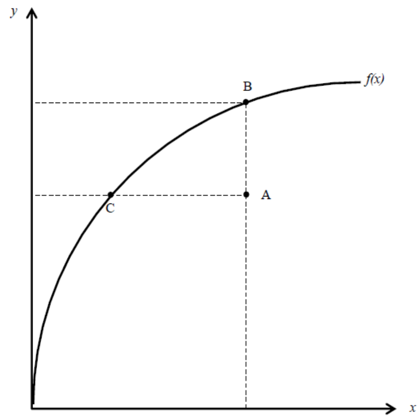
for spatial models (LeSage, 2008). Therefore, several alternatives estimation methods have been suggested in the literature. These estimators include the maximum likelihood technique (Ord, 1975; Anselin, 1988b; Anselin and Bera, 1998); the instrumental variables / generalized method of moments estimators (Kelejian and Prucha, 1998, 1999; Pinkse *et al.*, 2002; Fingleton and Le Gallo, 2008; Lee, 2007; Lin and fei Lee, 2010; Arnold and Wied, 2010); and the Bayesian approach (LeSage and Pace, 2009).

A.1.2 Efficiency analysis

Chapters 4 and 6 present our work on Senegalese farms technical efficiency. Efficiency is a relative concept that expresses how available initial resources are transformed with the aim to obtain a utilizable final result. It is measured with respect to the objective such as maximizing output or revenues or minimizing inputs or costs. Farrell (1957) defined the overall efficiency (or total economic efficiency) of a decision-making unit as a combination of two components: the technical efficiency and the price efficiency (or allocative efficiency). Technical efficiency measures the ability of the decision-making unit to produce the maximum output from given inputs, and the price efficiency measures the ability to use inputs in optimal proportions given their respective prices and the production technology. Measures of efficiencies focus more on technical efficiency, which we investigate in this dissertation.

Estimation of efficiency at the decision-making level requires the use of two widely known frontiers approaches: the non-parametric or mathematical programming approach such as the data envelopment analysis (DEA) and the parametric approach or an econometric approach like the stochastic frontier analysis (SFA), where one imposes a functional form on the production function. As argued by Coelli *et al.* (2005), despite its advantages, the non-parametric approach presents a major issue in agricultural economics, it does not take into account measurement's errors and other statistical noises that might influence the shape and the position of the frontier, which therefore affects the results and their interpretations. Since the seminal works of Aigner *et al.* (1977) and Meeusen and van Den Broeck (1977), efficiency analysis using SFA approach has seen incredible advancements.

Figure A.1 shows the SFA representation of a production function. It plots a single-input single-output case where y is the output and x the input. The curve *f(x)* represents the technology production frontier where any output level is the maximum attainable with the level of input and vice versa. Here, point A represents a technically inefficient farm. Farm A can produce more output with the same level of input and distance $\overline{AB}$ shows the output lost due to technical inefficiency. Farm A could also produce the same quantity of output using less inputs, then distance $\overline{AC}$ would represent the reduction of input without reducing the level of output.

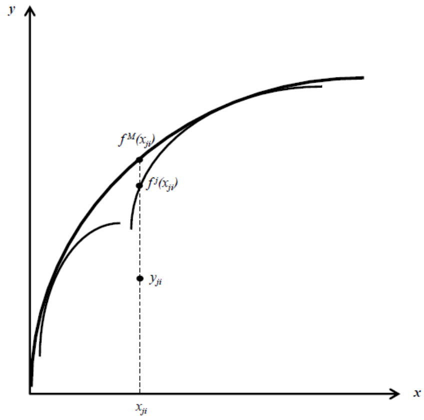


Source:(Kumbhakar and Wang, 2010)

Figure A.1: Technical Inefficiency, single-output single-input

In this thesis, we use the SFA approach for the mentioned advantages. However, as argued by Kumbhakar and Wang (2010) and shown by figure A.1 estimates of inefficiency are conditional on the technology in use (production frontier). Therefore, when estimating the technical efficiencies of farmers, it is essential that the estimates were done with respect to the common technology. Thus, in chapter 4 we used a meta-frontier approach that envelopes the different production frontiers of two groups of farmers (members of farmer organizations and non-members). Figure A.2 shows a representation of the meta-frontier production function. As stated by Huang *et al.* (2014), at any level of input, an associated farm output y_{ji} with respect to the meta-frontier $f^M(x_{ji})$ has the following components: the meta technology ratio $f^j(x_{ji})/f^M(x_{ji})$, the group specific technical efficiency of each farm $y_{ji}/(f^j(x_{ji})\exp(v_{ji}))$, and the technical efficiency of each farmer regarding the meta-frontier $y_{ji}/(f^M(x_{ji})\exp(v_{ji}))$, with v_{ji} being the error term.

Furthermore, when it comes to the specificities of the location of each farm, this approach presents also some limitations that have been early noted by Farrell (1957). Therefore, in chapter 6 we designed a spatial stochastic frontier and analysed farm efficiency by accounting for spatial effects. Reasons that motivate the integration of spatial effects when analysing farm efficiency are of twofold. First, the structural specific characteristics of the location of each farmer (e.g. soil fertility, soil topography, rainfall) differ from one farmer to another. Second, one of the most



Source:(Huang *et al.*, 2014)

Figure A.2: Metafrontier production model

important aspects of agricultural production is the possibility for farmers to have access to productive factors such as improved technologies and the related technical and necessary knowledge. In the African context, the level of rural infrastructures development (such as routes, extension services, fertilizers vending points, credit facilities) generally varies from one location to another. As argued by LeSage and Pace (2009), amenities and characteristics of each location usually constitute unobservable factors that might affect the performance of farmers, and it is difficult to find explanatory variables that capture easily and completely all of these types of latent effects. The integration of spatial interaction to the stochastic frontier approach has resulted in literature in four main models:

- the spatial auto-regressive stochastic frontier (SAR):

$$y = \rho W y + \beta X + v - u,$$

- the spatial lag on the exogenous variable stochastic frontier (SLX):

$$y = \beta X + \theta W X + v - u,$$

- the spatial error stochastic frontier (SEM) and

$$y = \beta X + v - u \quad \text{with} \quad v = \delta W v + \tilde{v},$$

- the spatial inefficiency stochastic frontier (SIM)

$$y = \beta X + v - u \quad \text{with} \quad u = \lambda W u + \tilde{u}, \quad u \geq 0,$$

. where y is the $N \times 1$ vector of output, X the $N \times K$ matrix of inputs, v, u, $\tilde{v}$, and $\tilde{u}$ are $N \times 1$ errors terms. β and θ are $N \times K$ parameters to be estimated, and ρ, δ, λ are scalar parameters to be determined. N and K represent the sample size and the number of inputs respectively. W is a pre-specified weight matrix. Distributions of errors terms are to be chosen depending on the estimation strategy. Chapter 6 offers a detailed review on empirical applications of these models.

A.2 Complement to chapter 2

A.2.1 The Generalized Spatial Two-Stage Least Squares

The Generalized Spatial Two-Stage Least Squares (GS2SLS) is a three-step procedure suggested by Kelejian and Prucha (1998, 1999). In the first step the regression model is estimated by two-stage least squares (2SLS) using the instruments H_n. In the second step the autoregressive parameter ρ is estimated in terms of the residuals obtained via the first step and the generalized moments procedure suggested in Kelejian and Prucha (1999). Finally, in the third step, the model is re-estimated by 2SLS after transforming the model via a Cochrane-Orcutt type transformation to account for the spatial correlation. Consider the following compacted model:

$$y_n = Z_n\delta + u_n, \quad u_n = \rho M_n u_n + \epsilon_n,$$

Applying a Cochrane-Orcutt type transformation to this model yields furthermore:

$$y_{n*} = Z_{n*}\delta + \epsilon_n,$$

where $y_{n*} = y_n - \rho M_n y_n$ and $Z_{n*} = Z_n - \rho M_n Z_n$. In the following y_{n*} and Z_{n*} could also be referred as $y_{n*}(\rho)$ and $Z_{n*}(\rho)$ to indicate the dependence of the transformed variables on ρ. In the first step, the 2SLS estimator is considered as:

$$\tilde{\delta}_n = \left(\hat{Z}_n^{'}\hat{Z}_n\right)^{-1}\hat{Z}_n^{'}y_n$$

where $\hat{Z}_n^{'} = P_{H_n}\hat{Z}_n$ and $P_{H_n} = H_n\left(H_n^{'}H_n\right)^{-1}H_n^{'}$. Although $\tilde{\delta}_n$is consistent, it does not utilize information relating to the spatial correlation of the error term. Therefore the procedure continues with a second step.

For this second step, let define $u_{i,n}$, $\bar{u}_{i,n}$, $\bar{\bar{u}}_{i,n}$ as respectively, the i_{th} elements of u_n,

$\bar{u}_n = M_n u_n$, $\bar{\bar{u}}_{i,n} = M_n^2 u_n$. Similarly, let $\epsilon_{i,n}$, and $\bar{\epsilon}_{i,n}$, be in the i_{th} elements of ϵ_n, $\bar{\epsilon}_n = M_n \epsilon_n$. Then, the spatial correlation model implies that $u_{i,n} - \rho \bar{u}_{i,n} = \epsilon_{i,n}$ and $\bar{u}_{i,n} - \rho \bar{\bar{u}}_{i,n} = \bar{\epsilon}_{i,n}$. From these equations, the following three-equations can be obtained (Kelejian and Prucha, 1998):

$$\begin{aligned} 2\rho n^{-1} \sum u_{i,n} \bar{u}_{i,n} - \rho^2 n^{-1} \sum \bar{u}_{i,n}^2 + n^{-1} \sum \epsilon_{i,n}^2 &= n^{-1} \sum u_{i,n}^2 \\ 2\rho n^{-1} \sum \bar{u}_{i,n} \bar{\bar{u}}_{i,n} - \rho^2 n^{-1} \sum \bar{\bar{u}}_{i,n}^2 + n^{-1} \sum \bar{\epsilon}_{i,n}^2 &= n^{-1} \sum \bar{u}_{i,n}^2 \\ \rho n^{-1} \sum \left[u_{i,n} \bar{\bar{u}}_{i,n} + \bar{u}_{i,n}^2 \right] - \rho^2 n^{-1} \sum \bar{u}_{i,n} \bar{\bar{u}}_{i,n} + n^{-1} \sum \epsilon_{i,n} \bar{\epsilon}_{i,n} &= n^{-1} \sum u_{i,n} \bar{u}_{i,n}. \end{aligned}$$

Kelejian and Prucha (1998) assumptions implied that:

$$\begin{aligned} E\left(n^{-1} \sum \bar{\epsilon}_{i,n}^2\right) &= n^{-1} E\left[Tr\left(\epsilon_n^{'} M_n^{'} M_n \epsilon_n\right)\right] \\ &= n^{-1} Tr\left(E\left(\epsilon_n \epsilon_n^{'} M_n^{'} M_n\right)\right) \\ &= \sigma_\epsilon^2 n^{-1} Tr\left(M_n^{'} M_n\right) \end{aligned}$$

where $Tr(.)$ is the trace operator. Let define $\alpha = (\rho, \rho^2, \sigma_\epsilon^2)$ and $\gamma_n = n^{-1}(E(u_n^{'} u_n)$, $E(\bar{u}_n^{'} \bar{u}_n), E(u_n^{'} \bar{u}_n))^{'}$ then the system of three equations becomes:

$$\Gamma_n \alpha = \gamma_n$$

where,

$$\Gamma_n = \frac{1}{n} \begin{bmatrix} 2E(u^{'}\bar{u}) & -\frac{1}{N} E\bar{u}^{'}\bar{u}) & 1 \\ 2E(\bar{\bar{u}}^{'} u) & -E(\bar{\bar{u}}^{'}\bar{\bar{u}}) & \mathrm{Tr}(M^{'}M) \\ E(u^{'}\bar{\bar{u}} + \bar{u}^{'}\bar{u}) & -E(\bar{u}^{'}\bar{\bar{u}}) & 0 \end{bmatrix}$$

Kelejian and Prucha (1999) suggested two estimators of ρ and σ_ϵ^2 based on estimated values of Γ_n and γ_n. Let $\tilde{u} = y_n - Z_n \tilde{\delta}_n$, $\tilde{\bar{u}}_n = M_n \tilde{u}_n$, and $\tilde{\bar{\bar{u}}}_n = M_n^2 \tilde{u}_n$, where $\tilde{\delta}$ is the 2SLS estimator of the first step. Considering the following estimator for Γ_n and γ_n:

$$G_n = \frac{1}{n} \begin{bmatrix} 2\sum \tilde{u}_{i,n} \tilde{\bar{u}}_{i,n} & -\sum \tilde{\bar{u}}_{i,n}^2 & 1 \\ 2\sum \tilde{\bar{u}}_{i,n} \tilde{\bar{\bar{u}}}_{i,n} & -\sum \tilde{\bar{\bar{u}}}_{i,n}^2 & Tr(M_n^{'} M_n) \\ \sum \left(\tilde{u}_{i,n} \tilde{\bar{\bar{u}}}_{i,n} + \tilde{\bar{u}}_{i,n}^2\right) & -\sum \tilde{\bar{u}}_{i,n} \tilde{\bar{\bar{u}}}_{i,n} & 0 \end{bmatrix}, \; g_n = \frac{1}{n} \begin{bmatrix} \sum \tilde{u}_{i,n}^2 \\ \sum \tilde{\bar{u}}_{i,n}^2 \\ \sum \tilde{u}_{i,n} \tilde{\bar{u}}_{i,n} \end{bmatrix}$$

Then, the empirical form of the relationship $\gamma_n = \Gamma_n \alpha$ is $g_n = G_n \alpha + v_n$ where

v_n can be viewed as a vector of regression residuals. The first set of estimators is based on the first and third elements of the ordinary least squares estimator $\tilde{\alpha_n}$ for α obtained from regressing g_n against G_n. In the second set of estimators (nonlinear least squares estimators), $\tilde{\rho}_n$ and $\tilde{\tilde{\sigma}}^2_{\epsilon,n}$ are defined as the minimizers of:

$$\left[g_n - G_n \begin{bmatrix} \rho \\ \rho^2 \\ \sigma^2_\epsilon \end{bmatrix}\right]' \left[g_n - G_n \begin{bmatrix} \rho \\ \rho^2 \\ \sigma^2_\epsilon \end{bmatrix}\right]$$

Finally, in the third stage of the procedure, the generalized spatial 2SLS estimator is given by:

$$\hat{\delta} = \left[\hat{Z}_{n*}(\rho)\prime \hat{Z}_{n*}(\rho)\right]^{-1} \hat{Z}_{n*}(\rho)\prime y_{n*}(\rho)$$

A.3 Complement to chapter 5

A.3.1 Bayesian approach of model selection

In chapter 5, several weight matrices based on inverse distance matrices were defined. Therefore, to select the best model, Bayesian methods were used to compare the different models resulting from the different weight matrices. LeSage and Pace (2009) argues that a Bayesian approach to model comparison has specific advantages over likelihood-based methods.

The Bayesian approach of model comparison involves considering a set of models based on M alternative weight matrices $W_k, k = 1, ..., M$, while maintaining the other aspects of the model constant. For each model, a likelihood function and priors are specified for the parameters of interest $\tau = (\beta, \theta, \rho)$. Using Bayes' rule, the posterior density of the parameters τ conditional on a specific W_k can be derived by combining the likelihood $p(y^*|\tau_k, W_k)$ with the prior beliefs about the parameters (LeSage and Pace, 2009):

$$p(\tau_k|y^*, W_k) = \frac{p(y^*|\tau_k, W_k)\, p(\tau_k|W_k)}{p(y^*|W_k)},$$

where the term $p(y^*|W_k)$ represents the likelihood of the data given W_k and is referred to as the marginal likelihood for this model comparison situation and is expressed as (LeSage and Pace, 2009):

$$p\left(y^{*}|W_{k}\right)=\int p\left(y^{*}|\tau_{k},W_{k}\right)p\left(\tau_{k}|W_{k}\right)d\tau_{k}.$$

By applying Bayes' theorem to this marginal distribution, one can compute posterior model probabilities for each of the different model based on W. The posterior model probability PMP for a specific weight matrix W_k can be specified as (Juhl, 2020):

$$PMP\left(W_{k},y^{*}\right)=\frac{p\left(y^{*}|,W_{k}\right)p\left(W_{k}\right)}{\sum_{j=1}^{M}p\left(y^{*}|,W_{j}\right)p\left(W_{j}\right)},$$

where $p\left(W_k\right)$ and $p\left(W_j\right)$ are prior model probabilities and equal to 1/M, making each weight matrix equally likely a priori. To select the best fitting model, the weight matrix with the highest posterior model probability is selected.

A.4 Complement to chapter 6

A.4.1 R codes for simulation works

```r
# Packages
library(igraph)
library(spdep)
library(maxLik)
library(GMCM)

# Simulations functions
simul_spsfa <- function ( repetitions = NULL, seed = NULL,
                          nobs = NULL, Wmatr = NULL,
                          swdim = NULL, swneig = NULL, swp = NULL,
                          geolat = NULL, geolon = NULL,
                          minX = NULL, maxX = NULL, dgp = NULL, bcoef = NULL,
                          fake_rho = NULL, fake_lambda = NULL,
                          fake_sdv = NULL, fake_sdu = NULL, fake_delta = NULL,
                          approach = NULL, estim = NULL, efficiency = NULL,
                          method = NULL, control=NULL) {

  # Set the seed
  set.seed ( seed )

  # Matrices to simulations results

  Nobs <<- matrix ( NA, nrow = repetitions, ncol=1)
  Wmatrice <<- matrix ( NA, nrow = repetitions, ncol = 1)
  minXvar <<- matrix ( NA, nrow = repetitions, ncol = 1)
  maxXvar <<- matrix ( NA, nrow = repetitions, ncol = 1)
  dataGP <<- matrix ( NA, nrow = repetitions, ncol = 1)
  rho_true <<- matrix ( NA, nrow = repetitions, ncol = 1)
  lambda_true <<- matrix ( NA, nrow = repetitions, ncol = 1)
  splag_true <<- matrix ( NA, nrow = repetitions, ncol = 1)

  beta_true <<- matrix ( NA, nrow = repetitions, ncol = 2)
  sdv_true <<- matrix ( NA, nrow = repetitions, ncol = 1)
  sdu_true <<- matrix ( NA, nrow = repetitions, ncol = 1)
  alpha_true <<- matrix ( NA, nrow = repetitions, ncol = 1)
  sigma_true <<- matrix ( NA, nrow = repetitions, ncol = 1)
  delta_true <<- matrix ( NA, nrow = repetitions, ncol = 1)
  ditribution <<- matrix ( NA, nrow = repetitions, ncol = 1)
  estimation <<- matrix ( NA, nrow = repetitions, ncol = 1)
  U_true <<- matrix ( NA, nrow = repetitions, ncol = nobs)
  mTE_true <<- matrix ( NA, nrow = repetitions, ncol = 1)
  Loglik_true <<- matrix ( NA, nrow = repetitions, ncol = 1)

  beta_est <<- matrix ( NA, nrow = repetitions, ncol = 2)
  alpha_est <<- matrix ( NA, nrow = repetitions, ncol = 1)
  sigma_est <<- matrix ( NA, nrow = repetitions, ncol = 1)
  delta_est <<- matrix ( NA, nrow = repetitions, ncol = 1)
  splag_est <<- matrix ( NA, nrow = repetitions, ncol = 1)
  U_est <<- matrix ( NA, nrow = repetitions, ncol = nobs)
  mTE_est <<- matrix ( NA, nrow = repetitions, ncol = 1)
  Loglik_est <<- matrix ( NA, nrow = repetitions, ncol = 1)

  timespent <<- matrix ( NA, nrow = repetitions, ncol = 1)
```

```r
# Data generating process function
func_dgp_est <- function ( repetitions = NULL, seed = NULL,
                           nobs = NULL, Wmatr = NULL,
                           swdim = NULL, swneig = NULL, swp = NULL,
                           geolat = NULL, geolon = NULL,
                           minX = NULL, maxX = NULL, dgp = NULL, bcoef = NULL,
                           fake_rho = NULL, fake_lambda = NULL,
                           fake_sdv = NULL, fake_sdu = NULL, fake_delta = NULL,
                           approach = NULL, estim = NULL, efficiency = NULL,
                           method = NULL, control=NULL) {

  rm ( list=ls () )

  if  (Wmatr == "sw"){ # Small World graph weight matrix
    Wmat <- sample_smallworld ( dim = swdim, size = nobs,
                                nei = swneig, p = swp)
    Wmatw <- as.matrix ( as_adjacency_matrix ( Wmat, type = "both") )

  }  else { # Geographic distance weight matrix
    Dgeo <- data.frame (lat = runif(nobs, 0, geolat),
                        lon = runif ( nobs, 0, geolon))
    Dmat <- dist ( cbind ( Dgeo$lon, Dgeo$lat ),
                   method = 'euclidean', upper=TRUE )
    Wmatw <- ( as.matrix ( (1 / Dmat), upper=TRUE, diag = TRUE ) )
    Wmatw[!is.finite(Wmatw)] <- 0
    diag ( Wmatw ) <- 0
  }
  # Row standardized
  Wmatws <- apply ( Wmatw, 2 , function(x) x/rowSums(Wmatw))
  Wmatlw <- mat2listw (Wmatw, style="W")

  # Inverse of matrix (I-rho*W)
  I_rhow <- invIrW ( Wmatws, fake_rho, method = "solve", feasible = NULL)
  I_lambdaw <- invIrW ( Wmatws, fake_lambda, method = "solve", feasible = NULL)

  # Creation of variables
  # Data is constant plus a uniform covariate
  X <- cbind ( c (1), runif ( nobs, minX, maxX ) )

  # Inefficiency term,  # u ~ Normal(0, 0.3)
  u <- rnorm (nobs, mean=0, sd = fake_sdu )

  # Noise # normal idiosyncratic error term
  v <- rnorm (nobs, mean=0, sd = fake_sdv)

      # True values
  fake_alpha <- ( fake_delta / (sqrt( 1 - fake_delta^2 )) )
  fake_sigma <- fake_sdv / (sqrt( 1 - fake_delta^2 ) )

  if ((dgp == "sem")||( dgp == "sde")){
    fake_splag <- fake_lambda
  } else { fake_splag <- fake_rho }
```

```r
if (approach == "halfn") { # normal - half-normal approach
  e <- v - abs(u)
  } else { # skew-normal approach
  e <- fake_delta * abs(u) + sqrt ( 1 - fake_delta^2 ) * v }

# Dependent variable
if ( dgp == "linear") { # No spatial
  y <- X%*%bcoef + e

} else if ( dgp == "sar" ) { # SAR: y = rho*W*y + Xbeta + e
  y <- I_rhow %*% ( X%*%bcoef + e )
  I <- diag(nobs)
  fake_Bmat <- I - fake_splag*Wmatws
  fake_detBmat <- determinant(fake_Bmat, logarithm = TRUE)$modulus
  fake_epsilon <- (1/fake_sigma)*(fake_Bmat%*%y - X%*%bcoef)

  fake_Loglik <- - (nobs / 2) * log (pi) -
    (nobs/2) * log( fake_sigma^2 ) +
    fake_detBmat - (1/2) * (t ( fake_epsilon) %*% fake_epsilon ) +
    sum ( log ( 2*pnorm ( fake_alpha * fake_epsilon ) ) )

} else if ( dgp == "sdb" ) { # SDB: y = rho*W*y + X*beta + W*X*theta + e
  y <- I_rhow %*% ( X[,1] * bcoef[1] + X[,2] * (bcoef[2]/2) +
                      Wmatws %*% X[,2] * (bcoef[2]/2) + e )

} else if ( dgp == "sem" ) { # SEM: y = X*beta + [I-W*lambda]^(-1)*e
  y <- X%*%bcoef + I_lambdaw %*% e
  I <- diag(nobs)
  fake_Bmat <- I - fake_splag*Wmatws

  fake_epsilon <- (1/fake_sigma)*fake_Bmat %*% (y - X%*%bcoef)
  fake_detBmat <- determinant(fake_Bmat, logarithm = TRUE)$modulus

  fake_Loglik <- (-1)*( nobs/2 )*log ( pi ) -
    (nobs/2)*log ( fake_sigma^2 ) +
    fake_detBmat -  ( 1/2 ) * t ( fake_epsilon ) %*% fake_epsilon +
    sum ( log ( 2 * pnorm ( fake_alpha * fake_epsilon ) ) )

  } else  { # dgp == "sde" # SDE:  y = X*beta+W*X*teta+[I-W*lambda]^(-1)*e
  y <- X[,1] * bcoef[1] + X[,2] * (bcoef[2]/2) +
    Wmatws %*% X[,2] * (bcoef[2]/2) + I_lambdaw %*% e
}

#Estimating specific inefficiency

if (dgp == "sem") {
  fake_residuals <- ( 1 / fake_sigma ) * fake_Bmat %*% ( y - X %*% bcoef )
} else if ( dgp == "sar" ) {
  fake_residuals <- ( 1 / fake_sigma ) * (fake_Bmat %*% y - X %*% bcoef )
} else {
  fake_residuals <- ( 1 / fake_sigma ) * (y - X %*% bcoef )
}
```

```
fake_means_u <- ( 1 / ( ( fake_delta^2 / fake_sigma^2 ) + 1 ) ) * (
  fake_delta / fake_sigma^2 ) * fake_residuals
fake_var_u <- (1) / ( (fake_delta^2 / fake_sigma^2 ) + 1)
fake_U <- dnorm(fake_residuals, mean = fake_means_u, sd = fake_var_u )
fake_mTE <- mean ( exp ( -fake_U ) )

start.time <<- Sys.time()

#Estimations
if ( estim == "skewn_sem") {
  #SEM, log likelihood function (deGraaf, 2019)

  SpatialFrontierErrorFun <- function(pars, X, Y, W){
    p <- length(pars)
    lambda <- pars[p]
    alpha <- pars[(p-1)]
    sigma <- pars[(p-2)]
    beta <- pars[(1:(p-3))]

    sigma <- sqrt(sigma^2)
    lambda <- 2*(exp(lambda)/(1+exp(lambda)))-1
    nObs <- dim(X)[1]
    B  <- (diag(nObs) - lambda * W)
    Xb <-  X%*%beta
    alpha <- -sqrt(alpha^2)
    delta <- alpha/(sqrt(1+alpha^2))
    z <- (1/sigma)*(B%*%(Y - Xb))
    term1 <- -(nObs/2)*log(pi) -(nObs/2)*log(sigma^2) + log(det(B)) -
      (1/2)*(t(z)%*%z)
    term2 <- colSums(log(2*pnorm(alpha*z)))

    return(term1 + term2)
  }


  # Starting values (spdep package)
  q <- dim(X)[2]
  erlag <- errorsarlm (y ~ X[,2] , listw = Wmatlw)
  beta0 <- erlag$coefficients[1:(q)]
  delta0 <- -sqrt(erlag$s2*(1-2/pi))
  alpha0 <- delta0/(sqrt(1 - delta0^2))
  sigma0 <- sqrt(erlag$s2)
  splag0 <- erlag$lambda[1]
  startv <- as.vector (c(beta0, sigma0, alpha0, splag0))
  estim_max <- maxLik ( SpatialFrontierErrorFun,
                        X=X, Y=y, W=Wmatws,
                        start = startv,
                        method = method, control= control)

}  else if ( estim == "skewn_sar") {
  #SAR, log likelihood function (deGraaf, 2019)

  SpatialFrontierLagFun <- function(pars, Y, X, W){
```

```
    p <- length(pars)
    rho <- pars[p]
    alpha <- pars[(p-1)]
    sigma <- pars[(p-2)]
    beta <- pars[(1:(p-3))]
    sigma <- sqrt(sigma^2)
    nObs <- length(Y)
    A  <- (diag(nObs) - rho * W)
    Xb <-  X%*%beta
    alpha <- -sqrt(alpha^2)
    delta <- alpha/(sqrt(1+alpha^2))
    z <- (1/sigma)*(A%*%Y - Xb)

    term1 <- -(nObs/2)*log(pi) -(nObs/2)*log(sigma^2) +
      log(det(A)) - (1/2)*(t(z)%*%z)
    term2 <- colSums(log(2*pnorm(alpha*z)))

    return(term1 + term2)
  }

  ## Spatial SAR Degraaf
  q <- dim(X)[2]
  sp_lag_a <- lagsarlm (y ~ X[,2], listw = Wmatlw)
  beta0 <- sp_lag_a$coefficients[1:(q)]
  delta0 <- -sqrt(sp_lag_a$s2*(1-2/pi))
  alpha0 <- delta0/(sqrt(1 - delta0^2))
  sigma0 <- sqrt(sp_lag_a$s2)
  rho0 <- sp_lag_a$rho[1]
  startv <- c(beta0, sigma0, alpha0, rho0)
  estim_max <- maxLik ( SpatialFrontierLagFun,
                      X=X, Y=y, W=Wmatws,
                      start = startv,
                      method = method, control= control)

} else {
  estim_max <- NA   }

beta_hat <- estim_max$estimate[1:q]
sigma_hat <- estim_max$estimate[q+1]
alpha_hat <- estim_max$estimate[q+2]
splag_hat <- estim_max$estimate[q+3]
Loglik_hat <- estim_max$maximum

# Estimating specific inefficiency

delta_hat <- alpha_hat / ( sqrt (1 + alpha_hat^2 ) )
I <- diag ( nobs )
Bmat_hat <- I - splag_hat*Wmatws

if (estim == "skewn_sem") {
  residuals <- ( 1 / sigma_hat ) * Bmat_hat %*% ( y - X %*% beta_hat )
} else if ( estim == "skewn_sar" ) {
  residuals <- ( 1 / sigma_hat ) * (Bmat_hat %*% y - X %*% beta_hat )
```

```
  } else {
    residuals <- ( 1 / sigma_hat ) * (y - X %*% beta_hat )
  }

  means_u <- ( 1 / ( ( delta_hat^2 / sigma_hat^2 ) + 1 ) ) * (
    delta_hat / sigma_hat^2 ) * residuals
  var_u <- (1) / ( (delta_hat^2 / sigma_hat^2 ) + 1)
  U_hat <- dnorm(residuals, mean = means_u, sd = var_u )
  mTE_hat <- mean ( exp ( -U_hat ) )

  # Return of results
  return ( list (nobs=nobs, Wmatr=Wmatr,
                 minX=minX, maxX=maxX, dgp=dgp,
                 fake_rho=fake_rho, fake_lambda=fake_lambda,
                 fake_splag=fake_splag, bcoef=bcoef,
                 fake_sdv=fake_sdv, fake_sdu=fake_sdu,
                 fake_sigma = fake_sigma,
                 fake_delta = fake_delta, fake_alpha = fake_alpha,
                 fake_Loglik = fake_Loglik,
                 fake_U = fake_U, fake_mTE = fake_mTE,
                 Loglik_hat= Loglik_hat,
                 approach=approach, estim=estim,
                 beta_hat=beta_hat, alpha_hat=alpha_hat,
                 sigma_hat=sigma_hat,
                 splag_hat=splag_hat, delta_hat=delta_hat,
                 Bmat_hat=Bmat_hat, residuals=residuals,
                 means_u=means_u,
                 var_u=var_u, U_hat=U_hat, mTE_hat=mTE_hat ) )
}


for (k in 1:repetitions){

  # Run the function
  Results <- func_dgp_est ( repetitions = NULL, seed = NULL,
                            nobs = NULL, Wmatr = NULL,
                            swdim = NULL, swneig = NULL, swp = NULL,
                            geolat = NULL, geolon = NULL,
                            minX = NULL, maxX = NULL, dgp = NULL, bcoef = NULL,
                            fake_rho = NULL, fake_lambda = NULL,
                            fake_sdv = NULL, fake_sdu = NULL, fake_delta = NULL,
                            approach = NULL, estim = NULL, efficiency = NULL,
                            method = NULL, control=NULL)

  # Results
  Nobs[k] <<- (Results$nobs)
  Wmatrice[k] <<- (Results$Wmat)
  minXvar[k] <<- (Results$minX)
  maxXvar[k] <<- (Results$maxX)
  dataGP[k] <<- (Results$dgp)
  rho_true[k] <<- (Results$fake_rho)
  lambda_true[k] <<- (Results$fake_lambda)
  splag_true[k] <<- (Results$fake_splag)
```

```
    beta_true[k,] <<- (Results$bcoef)
    sdv_true[k] <<- (Results$fake_sdv)
    sdu_true[k] <<- (Results$fake_sdu)
    alpha_true[k] <<- (Results$fake_alpha)
    sigma_true[k] <<- (Results$fake_sigma)
    delta_true[k] <<- (Results$fake_delta)
    Loglik_true[k] <<- (Results$fake_Loglik)
    U_true[k,] <<- (Results$fake_U)
    mTE_true[k] <<- (Results$fake_mTE)

    ditribution[k] <<- (Results$approach)
    estimation[k] <<- (Results$estim)

    beta_est[k,] <<- (Results$beta_hat)
    alpha_est[k] <<- (Results$alpha_hat)
    sigma_est[k] <<- (Results$sigma_hat)
    splag_est[k] <<- (Results$splag_hat)
    delta_est[k] <<- (Results$delta_hat)
    U_est[k,] <<- (Results$U_hat)
    mTE_est[k] <<- (Results$mTE_hat)
    Loglik_est[k] <<- (Results$Loglik_hat)

    end.time <- Sys.time()
    time.taken <- difftime (end.time, start.time, units = "secs")
    timespent[k] <<- time.taken
    print (sprintf ("Simulation %s of %s. Processing took %f seconds.",
                    k, repetitions, time.taken ) )

  }

  simul_results <- list ( Nobs = Nobs, Wmatrice = Wmatrice,
                          minXvar = minXvar, maxXvar = maxXvar,
                          dataGP = dataGP, rho_true = rho_true,
                          lambda_true = lambda_true, splag_true = splag_true,
                          beta_true = beta_true, sdv_true = sdv_true,
                          sigma_true = sigma_true,
                          sdu_true = sdu_true, alpha_true=alpha_true,
                          delta_true = delta_true, Loglik_true = Loglik_true,
                          U_true = U_true,  mTE_true = mTE_true,
                          Loglik_est = Loglik_est,
                          ditribution = ditribution,
                          estimation = estimation, beta_est = beta_est,
                          alpha_est = alpha_est, sigma_est = sigma_est,
                          splag_est = splag_est, delta_est = delta_est,
                          U_est = U_est, mTE_est = mTE_est,
                          Loglik_est = Loglik_est, repetitions = repetitions,
                          timespent = timespent  )

  return ( simul_results )
  #saveRDS (simul_results, "simul_results.rds" )
}
```

```r
restable_fun <- function (replications=NULL ) {

  # True values
  NobsF <- replications$Nobs[1]
  WmatriceF <- replications$Wmatrice[1]
  dataGPF <-  dataGP[1]
  beta_trueF <- replications$beta_true[1,]
  rho_trueF <- replications$rho_true[1]
  lambda_trueF <- replications$lambda_true[1]
  sigma_trueF <- replications$sigma_true[1]
  alpha_trueF <- replications$alpha_true[1]
  delta_trueF <- replications$delta_true[1]
  ditributionF <- replications$ditribution[1]
  estimationF <- replications$estimation[1]
  mTE_true_F <- colMeans ( replications$mTE_true)
  Loglik_trueF <-  colMeans (replications$Loglik_true)

  # Means of parameters
  beta_mean <<- colMeans ( replications$beta_est )
  alpha_mean <<- colMeans ( replications$alpha_est )
  sigma_mean <<- colMeans ( replications$sigma_est )
  splag_mean <<- colMeans ( replications$splag_est )
  delta_mean <<- colMeans ( replications$delta_est )
  mTE_est_mean <<- colMeans ( replications$mTE_est )
  Loglik_est_mean <<-  colMeans (replications$Loglik_est)

  # standard deviations of parameters
  beta_sd <- GMCM:::colSds(replications$beta_est )
  alpha_sd <<- sd ( replications$alpha_est )
  sigma_sd <<- sd ( replications$sigma_est )
  splag_sd <<- sd ( replications$splag_est )
  delta_sd <<- sd ( replications$delta_est )
  mTE_est_sd <<- sd ( replications$mTE_est )
  Loglik_est_sd <<-  sd (replications$Loglik_est)

  beta_dif <<- replications$beta_est - replications$beta_true
  alpha_dif <<- replications$alpha_est - replications$alpha_true
  sigma_dif <<- replications$sigma_est - replications$sigma_true
  splag_dif <<- replications$splag_est - replications$splag_true
  delta_dif <<- replications$delta_est - replications$delta_true
  U_dif <<- replications$U_est - replications$U_true
  mTE_dif <<- replications$mTE_est - replications$mTE_true
  Loglik_dif <<- replications$Loglik_est - replications$Loglik_true


  # Biaises
  biais <- function (nrepliq, pardif ) {
    ( 1 / nrepliq ) * sum ( pardif )  }

  # Root Mean square error mse
  rmse <- function (nrepliq, pardif ) {
    sqrt ( ( 1 / nrepliq ) * sum ( ( pardif )^2 ))   }
```

```
beta_biais <- c(biais ( replications$repetitions, beta_dif[,1]),
                biais ( replications$repetitions, beta_dif[,2]))
alpha_biais <- biais ( replications$repetitions, alpha_dif)
sigma_biais <- biais ( replications$repetitions, sigma_dif)
splag_biais <- biais ( replications$repetitions, splag_dif)
delta_biais <- biais ( replications$repetitions, delta_dif)
Loglik_biais <- biais ( replications$repetitions,  Loglik_dif)
mTE_biais <- biais ( replications$repetitions, mTE_dif)

beta_rmse <- c(rmse ( replications$repetitions, beta_dif[,1]),
               rmse ( replications$repetitions, beta_dif[,2]))
alpha_rmse <- rmse ( replications$repetitions, alpha_dif)
sigma_rmse <- rmse ( replications$repetitions, sigma_dif)
splag_rmse <- rmse ( replications$repetitions, splag_dif)
delta_rmse <- rmse ( replications$repetitions, delta_dif)
Loglik_rmse <- rmse ( replications$repetitions,  Loglik_dif)
mTE_rmse <- rmse ( replications$repetitions, mTE_dif)

Infos <- c(ditributionF, estimationF,
           dataGPF, WmatriceF, NobsF)
names(Infos) <- c("ditribution", "estimation",
                  "dataGP", "Wmatrice", "Nobs")
pars_true <- c ( beta_trueF, sigma_trueF, alpha_trueF,
                 delta_trueF, lambda_trueF,
                 mTE_true_F, Loglik_trueF)
pars_est_mean <- c ( beta_mean, sigma_mean, alpha_mean,
                     delta_mean, splag_mean,
                     mTE_est_mean, Loglik_est_mean)
pars_est_sd <- c ( beta_sd, sigma_sd, alpha_sd,
                   delta_sd, splag_sd,
                   mTE_est_sd, Loglik_est_sd)
pars_biais <- c ( beta_biais, sigma_biais, alpha_biais,
                  delta_biais, splag_biais,
                  mTE_biais, Loglik_biais)
pars_rmse <- c ( beta_rmse, sigma_rmse, alpha_rmse,
                 delta_rmse, splag_rmse,
                 mTE_rmse, Loglik_rmse)

result_table <- matrix ( c ( pars_true, pars_est_mean, pars_est_sd,
                             pars_biais, pars_rmse), ncol=5)
result_table <- round (result_table, 3)
colnames (result_table) <- c('True', 'Estimates M', 'Estimates SD',
                             "Bias", "RMSE")
rownames(result_table) <- c("beta_1", "beta_2", "sigma", "alpha",
                            "delta", "splag",
                            "TE", "Loglik")

all_result <- list (replications = replications,
                    result_table = result_table,
                    Infos = Infos)

return(all_result)
```

```
}

applysim <- function(param, simul_spsfa, restable_fun) {

  repetitions <- param$repetitions
  seed <- param$seed
  nobs <- param$nobs
  Wmatr <- param$Wmatr
  swdim <- param$swdim
  swneig <- param$swneig
  swp <- param$swp
  geolat <- param$geolat
  geolon <- param$geolon
  minX <- param$minX
  maxX <- param$maxX
  dgp <- param$dgp
  bcoef <- param$bcoef
  fake_rho <- param$fake_rho
  fake_lambda <- param$fake_lambda
  fake_sdv <- param$fake_sdv
  fake_sdu <- param$fake_sdu
  fake_delta <- param$fake_delta
  approach <- param$approach
  estim <- param$estim
  efficiency <- param$efficiency
  method <- param$method
  control <- param$control

  rep_result <- simul_spsfa ( repetitions = repetitions, seed = seed,
                              nobs = nobs, Wmatr = Wmatr,
                              swdim = swdim, swneig = swneig, swp =swp,
                              geolat = geolat, geolon = geolon,
                              minX = minX, maxX = maxX, dgp = dgp,
                              bcoef = bcoef,
                              fake_rho = fake_rho, fake_lambda = fake_lambda,
                              fake_sdv = fake_sdv, fake_sdu = fake_sdu,
                              fake_delta = fake_delta,
                              approach = approach, estim = estim,
                              efficiency = efficiency,
                              method = method, control=control)

  finalres <- restable_fun (replications=rep_result)
  return(finalres)
}
# Running the simulations

# Parameters setting
param_base <- list( repetitions = 10, seed = 2000,
                    swdim = 1, swneig = 6, swp =0.1,
                    geolat = 20, geolon = 20,
                    minX = 5, maxX = 14,
                    bcoef = c(1, 1),
                    fake_sdv = 0.3, fake_sdu = 0.3,
```

```r
                    efficiency = NULL,
                    method = "BFGS", control=NULL)

# Parameters for nobs=250, Wmatr = "geo", dgp = "sem", ...
param_250_geo_02_sem_08 <- c(param_base,
                            list(nobs = 250, Wmatr = "geo", dgp = "sem",
                                 fake_rho = 0.2, fake_lambda = 0.2,
                                 fake_delta = -0.8,
                                 approach = "skewn", estim = "skewn_sem",
                                 name="resim_250_geo_02_sem_08"))

# Simulations for repetitions=10, nobs=250, Wmatr = "geo", dgp = "sem", ....
resim_250_geo_05_sem_08 <- applysim (param = param_250_geo_02_sem_08,
                                     simul_spsfa = simul_spsfa,
                                     restable_fun = restable_fun)
```

```
## [1] "Simulation 1 of 10. Processing took 2.392616 seconds."
## [1] "Simulation 2 of 10. Processing took 3.653912 seconds."
## [1] "Simulation 3 of 10. Processing took 1.799989 seconds."
## [1] "Simulation 4 of 10. Processing took 2.280699 seconds."
## [1] "Simulation 5 of 10. Processing took 1.758994 seconds."
## [1] "Simulation 6 of 10. Processing took 1.799990 seconds."
## [1] "Simulation 7 of 10. Processing took 2.116800 seconds."
## [1] "Simulation 8 of 10. Processing took 1.772988 seconds."
## [1] "Simulation 9 of 10. Processing took 1.804969 seconds."
## [1] "Simulation 10 of 10. Processing took 1.976870 seconds."
```

References

Aigner, D., Lovell, C. and Schmidt, P. (1977) Formulation and estimation of stochastic frontier production function models, *Journal of Econometrics*, **6**, 21 – 37.

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

Anselin, L. (1988b) *Spatial Econometrics: Methods and Models*, NATO Asi Series. Series E, Applied Sciences, Springer Netherlands.

Anselin, L. (2001) Spatial effects in econometric practice in environmental and resource economics, *American Journal of Agricultural Economics*, **83**, 705–710.

Anselin, L. and Bera, A. (1998) Spatial dependence in linear regression models with an introduction to spatial econometrics, in *Handbook of Applied Economics Statistics* (Eds.) A. Ullah and D. Giles, New York: Marcel Dekker, vol. 237, p. 289.

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

Arnold, M. and Wied, D. (2010) Improved GMM estimation of the spatial autoregressive error model, *Economics Letters*, **108**, 65 – 68.

Cliff, A. and Ord, J. (1973) *Spatial autocorrelation*, Monographs in spatial and environmental systems analysis, Pion.

Coelli, T., Rao, D., O'Donnell, C. and Battese, G. (2005) *An introduction to efficiency and productivity analysis*, Springer US.

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

Farrell, M. J. (1957) The measurement of productive efficiency, *Journal of the Royal Statistical Society: Series A (General)*, **120**, 253–281.

Fingleton, B. and Le Gallo, J. (2008) Estimating spatial models with endogenous variables, a spatial lag and spatially dependent disturbances: Finite sample properties*, *Papers in Regional Science*, **87**, 319–339.

Getis, A. (2009) Spatial weights matrices, *Geographical Analysis*, **41**, 404–410.

Holloway, G. (2007) Special feature: Spatial issues in agricultural economics, *Journal of Agricultural Economics*, **58**, 484–486.

Huang, C. J., Huang, T.-H. and Liu, N.-H. (2014) A new approach to estimating the metafrontier production function based on a stochastic frontier framework, *Journal of Productivity Analysis*, **42**, 241–254.

Juhl, S. (2020) The sensitivity of spatial regression models to network misspecification, *Political Analysis*, **28**, 1–19.

Kelejian, H. H. and Prucha, I. R. (1998) A generalized spatial two stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances, *Journal of Real Estate Finance and Economics*, **17**, 99–121.

Kelejian, H. H. and Prucha, I. R. (1999) A generalized moments estimator for the autoregressive parameter in a spatial model, *International Economic Review*, **40**, 509–533.

Kumbhakar, S. C. and Wang, H.-J. (2010) Estimation of technical inefficiency in production frontier models using cross-sectional data, *Indian Economic Review*, **45**, 7–77.

Lee, L.-F. (2007) GMM and 2SLS estimation of mixed regressive, spatial autoregressive models, *Journal of Econometrics*, **137**, 489 – 514.

LeSage, J. and Pace, R. (2009) *Introduction to Spatial Econometrics*, Statistics, textbooks and monographs, CRC Press.

LeSage, J. P. (2008) An Introduction to Spatial Econometrics, *Revue d'économie industrielle*, **123**, 19–44.

LeSage, J. P. and Pace, R. K. (2014) The biggest myth in spatial econometrics, *Econometrics*, **2**, 217–249.

Lin, X. and fei Lee, L. (2010) GMM estimation of spatial autoregressive models with unknown heteroskedasticity, *Journal of Econometrics*, **157**, 34 – 52.

Meeusen, W. and van Den Broeck, J. (1977) Efficiency estimation from cobb-douglas production functions with composed error, *International Economic Review*, **18**, 435–444.

Moran, P. A. P. (1950) Notes on continuous stochastic phenomena, *Biometrika*, **37**, 17–23.

Ord, K. (1975) Estimation methods for models of spatial interaction, *Journal of the American Statistical Association*, **70**, 120–126.

Páez, A., Scott, D. M. and Volz, E. (2008) Weight matrices for social influence analysis: An investigation of measurement errors and their effect on model identification and estimation quality, *Social Networks*, **30**, 309 – 317.

Pinkse, J., Slade, M. E. and Brett, C. (2002) Spatial price competition: A semiparametric approach, *Econometrica*, **70**, 1111–1153.

Plumper, T. and Neumayer, E. (2010) Model specification in the analysis of spatial dependence, *European Journal of Political Research*, **49**, 418–442.

Skevas, T., Skevas, I. and Swinton, S. M. (2018) Does spatial dependence affect the intention to make land available for bioenergy crops?, *Journal of Agricultural Economics*, **69**, 393–412.