ABSTRACT

Title of Dissertation: ADOPTION, IMPACT EVALUATION, AND TECHNICAL EFFICIENCY OF IRON BIOFORTIFIED BEAN PRODUCTION IN RWANDA

José Elías Funes,
Doctor of Philosophy 2020

Dissertation directed by: Professor, Laixiang Sun, Department of Geographical Sciences

Micronutrient malnutrition, also known as hidden hunger, is a public health problem in many developing countries. Hidden hunger limits cognitive and physical development of children and increases both children's and adults' susceptibility to infectious diseases. The most common outcome of iron deficiency is anemia and in Rwanda, iron micronutrient malnutrition is highly pervasive. Thirty seven percent of children under five years of age and nearly 20 percent of women of childbearing age suffer from anemia in the country. Since 2012, HarvestPlus and its partners have been intensively disseminating iron biofortified common beans (*Phaseolus Vulgaris*) (IBB) varieties to help alleviate iron deficiency in Rwanda. On one hand, Rwandan farmers may be uncertain about the economic returns of this new technology owing to insufficient knowledge about the types and costs of inputs needed, the yield distribution, expected market prices, and the demand for the produce. On the other hand, policy makers and donors cannot observe the outcomes that bean farmers

would experience under all treatments of the IBB program. The counterfactual outcomes that a bean farming household would have experienced under other treatments are not observable. In this context, this dissertation uses a multiprong analytical framework to: 1) analyze how peer interactions, households and farm characteristics, as well as regional factors influence smallholder farming households' decisions to grow IBB varieties, 2) evaluate the impact of the IBB program on Rwandan farmer's livelihoods, focusing on the outcomes of yields and incomes for beneficiary households, and 3) estimate the impact of the IBB program on smallholder farming households' technical efficiency. The spatial econometric results indicate spatial interdependence in smallholder farming households' decisions to adopt IBB. In addition to the directly targeted beneficiaries, the spatial parameters from the econometric analysis suggest that the biofortification program affected non-beneficiaries as well. This finding indicates that (1) a household is more likely to grow IBB if the household is near other early IBB adopters who informed them about the nutritional and yield benefits of IBB technology and (2) the propensity of a household to grow IBB varies with the characteristics of neighboring farmers. The impact evaluation analysis supports the hypothesis that IBB growers had significantly higher yields and incomes, as compared to farmers that grew non-biofortified beans, whether improved or traditional. In addition, the impact assessment shows that farmers who grew iron biofortified varieties were relatively more efficient and obtained greater bean production than the control group.

ADOPTION, IMPACT EVALUATION, AND TECHNICAL EFFICIENCY OF IRON BIOFORTIFIED BEAN PRODUCTION IN RWANDA

by

José Elías Funes

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2020

Advisory Committee:
Professor Laixiang Sun, Chair/Advisor
Associate Professor Giovanni Baiocchi
Associate Research Professor Fernando Sedano
Dr. Todd Benson, Senior Research Fellow at IFPRI
Professor Anna Alberini

Preface

Chapters 4, 5, and 6 were presented in peer reviewed AAEA and AAG academic conferences. Chapter 4 and Chapter 5 are under peer-review process for publication in two academic journals.

Dedication

To my beloved wife – Nadya Saber, and to my parents, Mirna and Leo.

Acknowledgements

First and foremost, I would like to thank my wife Nadya for her loving support and patience throughout these past five years.

I would like to thank all of my academic committee members including Professor Laixiang Sun, Dr. Fernando Sedano, Dr. Todd Benson, Dr. Giovanni Baiocchi, and Professor Anna Alberini for their unwavering support throughout my Ph.D. program at the University of Maryland (UMD).

I am especially grateful to Professor Sun, my academic advisor, who has been instrumental in getting this thesis to completion. His superior guidance, unrelentless support, and academic rigor has undeniably helped to direct and shape my research since the start of my candidacy.

My sincere gratitude also goes to Dr. Benson, Senior Research Fellow at IFPRI and Dr. Sedano, Associate Research Professor, for their keen supervision at various stages of the program. I want to also thank Professor Liang for kindly putting me in contact with Professor Sun, as well as the Geography Department's academic faculty.

In addition, I would like to acknowledge Dr. Howard Bouis and Dr. Ekin Birol from HarvestPlus for their strong encouragement to pursue my Ph.D., as well as providing access to the Rwanda Impact Assessment Survey Data.

I also want to thank HarvestPlus for helping to make this dissertation possible by laying the groundwork to conduct this research endeavor. HarvestPlus' principal donors include the UK Government, the Bill & Melinda Gates Foundation, the US government's Feed the Future initiative, the European Commission, and donors for the CGIAR Research Program on Agriculture for Nutrition and Health. HarvestPlus is also supported by the John D. and Catherine T. MacArthur Foundation.

Table of Contents

List of Tables

List of Figures

List of Acronyms

ATT – Average Treatment on the Treated

CIAT – The International Center for Tropical Agriculture

GLM – Generalized Linear Model

IBB – Iron Biofortified Beans

IFPRI – International Food Policy Research Institute

LISA – Local Indicator of Spatial Association

LM test – Lagrange Multiplier Test

LR test – Likelihood-Ratio Test

PS – Propensity Score

RAB – Rwanda Agriculture Board

RWF – Rwandan Franc

SAR – Spatial Autoregressive Model

SDM – Spatial Durbin Model

SFA – C – Stochastic Frontier Analysis Control Group

SFA – T – Stochastic frontier Analysis Treatment Group

SSFA – C – Spatial Stochastic Frontier Analysis Control Group

SSFA – T – Spatial Stochastic Frontier Analysis Treatment Group

TE – Technical Efficiency

TFP - Total Factor Productivity

Chapter 1: Introduction

I. Motivation and Key Topics

Micronutrient malnutrition, also known as hidden hunger, is a public health problem in many developing countries. One-third of the world's population—about 2.5 billion people—are at risk of at least one micronutrient deficiency. Further, as many as 1.3 billion people are at risk of zinc deficiency, about 200 million are at risk of vitamin A deficiency, and 1 billion are at risk of iron deficiency (Saltzman et al., 2017). Hidden hunger limits cognitive and physical development of children and increases children and adults' susceptibility to infectious diseases. These curtail individuals, communities, and countries' abilities to capitalize on economic opportunities—reinforcing the cycle of poverty for generations to come (Alderman, Hoddinott, & Kinsey, 2006; Stein, 2010). Common mechanisms to alleviate this global health problem include direct mineral and vitamin supplementation in health clinics or through outreach programs; food fortification; and most recently, biofortification. The latter involves an agricultural intervention based on the breeding of staple food crops that have a higher micronutrient content together with improved agronomic traits, such as improved yields. Biofortification is considered a cost-effective (Meenakshi et al., 2012), sustainable and scalable intervention to reach rural populations (Saltzman et al., 2013; Bouis, 2017) vulnerable to micronutrient malnutrition.

Using conventional plant breeding techniques, iron-biofortified beans were developed to contain almost twice as much iron as common varieties. These IBB varieties have high iron concentration as well as a good fractional absorption. IBB varieties were developed to reach children < 5 years of age and women of reproductive age who are most vulnerable to iron deficiency in rural populations. The most common outcome of iron deficiency is anemia. Studies have shown that IBB are an efficacious source of iron and regular consumption of iron biofortified beans can not only address iron deficiency (Haas et al., 2016) but also improve cognitive (Murray-Kolb et al., 2017) and physical (Luna, Lung'ago, Gahutu, & Haas, 2015) performance among target populations. Moreover, farmer feedback studies on IBB varieties conducted following early delivery efforts have shown that farmers are willing to expand their production of these beans as well as to share the planting material with others (A Murekezi, Birol, Asare-Marfo, & Ktasvairo, 2013). Consumer acceptance studies found that consumers prefer the IBB varieties over most of local varieties (Oparinde et al., 2016; Abdoul Murekezi, Oparinde, & Birol, 2017). Since 2012, HarvestPlus and its partners have been intensively disseminating IBB varieties to help alleviate iron deficiency in Rwanda.

In Rwanda, micronutrient malnutrition is highly pervasive and adoption rates of improved varieties of staple crops tend to be low (see e.g., (Walker & Alwang, 2015)). Thirty seven percent of children under five years of age and nearly 20 percent of women of childbearing age suffer from anemia (NISR, 2015) in the country. Furthermore, about 25 percent of children and 37 percent of women have iron deficiencies (Petry et al., 2016). At the same time, Rwandans have one of the highest per capita bean consumption rates in the world, with rural households consuming beans on average six days in a given week (Asare-Marfo et al., 2016), and in significant quantities. In terms of bean production, bean farmers in Rwanda have low productivity (FAO, 2020), which translates into low food availability. With less food available, vulnerable populations face an increased risk of malnutrition. To satisfy growing food demand, there are three broad options: 1) increase land under production, 2) boost crop productivity, or 3) food imports. Crop productivity can increase through the adoption of higher yielding varieties and more efficient production techniques.

Following several years of collaborative research between HarvestPlus, the Rwanda Agriculture Board (RAB), and the International Center for Tropical Agriculture (CIAT), four iron biofortified bean (IBB) varieties were officially released for planting in Rwanda in 2010. Another six were released in 2012. Of these, two were bush and eight were climbing types. Climbing beans grow tall, requiring a stake for support, and have a yield potential of 4,000 kilograms per hectare, while bush beans grow between 2 and 3 feet tall, so do not require support. Bush-type IBBs have a yield potential of 3,000 kilograms per hectare.

IBB is a relatively new technology in Rwanda. Since the rollout of the biofortification program, two questions have been raised. The first question relates to the risk a smallholder bean farming household may face in adopting a new agricultural technology. Initially, farmers may be risk-averse to a new technology as they lack information pertaining to the likelihood of possible outcomes (e.g., yield, costs, profitability). As such, a risk-averse attitude would exert a detrimental impact on adoption. In this context, social learning and social interaction often complement and/or act as substitutes in delivering information and facilitating the technology diffusion process. Sources of social influence in the adoption of technology include: (1) endogenous effects, (2) exogenous effects, and (3) correlated effects. In this study, I analyze smallholder farming households' decisions to adopt these newly-released IBB varieties by specifically examining the influence of demand-side factors and the role of peers. To do so, I draw upon several theories from studies on the adoption of agricultural technology, social behavior, and utility maximization to test three hypotheses.

- For the first hypothesis, I test how the adoption behavior of smallholder farming households would be influenced by their neighbor's adoption outcomes, as a result of peer learning about the profitability or the appropriate use of IBB. This phenomenon is known as endogenous effect, which is described as imitation, contagion, bandwagons and social norms.
- For the second hypothesis, I model the effect of contextual factors, wherein the propensity of an IBB grower to behave is correlated with the exogenous characteristics of his/her neighbors.
- For the third hypothesis, I ran a set of regressions with fixed and random effects. The model treats observations from a given village as a cluster and assumes a random effect for each village. In this way, I expect the random effect will produce a weaker spatial relationship. If so, this will confirm the hypothesis that closer neighbors matter more than those farther away.

These analyses all together yield prevalence rates estimates of IBB adoption by district.

The second question I aim to answer is what was the economic impact of the biofortification program on Rwandan smallholder farming households' yields and incomes? From this broader analytical perspective, the impact evaluation investigates the economic impact of the adoption of conventionally-bred IBBs by smallholder farming households. In this analysis, I test the null hypothesis of IBB adoption having absolutely no effect on yields and incomes for any smallholder farming household. Adoption of IBB is expected to improve yields, which may translate into improved iron intake, higher levels of market sales, and subsequent income gains. I also study the heterogeneity of outcome variables as a function of the propensity score (PS) and baseline covariates, key analyses from the perspective of program targeting. The technical efficiency of production analysis, discuss next, is part of the second research question.

In Rwanda, limited access to agricultural technology, such as seed of improved bean varieties, and other complementary inputs, like fertilizers or staking material for climbing beans, can explain low crop productivity in the country. The technical efficiency of production analysis that I undertake investigates the impact of IBB planting material in bringing small farmers closer to their technological frontier for bean production. The analysis provides estimates of a national technological frontier for all bean farmers, a frontier for farmers that grow other improved bean varieties, and a frontier for IBB growers. This analysis provides insights into the IBB program's impact on the production efficiency of bean farming in Rwanda.

II. Background

Common bean (Phaseolus vulgaris) is the main grain legume for direct human consumption for the poorer population of Africa and Latin America. Evidence indicates Mesoamerica as the origin of common bean (Bitocchi et al., 2012). On a global scale, this staple crop is an example of a spillover resulting from a beneficial technology developed in one geographic area—in this case, Latin America—being transferred to another geographic area. An applied example includes the introduced of improved bean varieties in Rwanda, which were developed in collaboration between the International Centre for Tropical Agriculture (CIAT) and the Rwanda Agriculture Board. By 1998, improved bean varieties that originated from CIAT covered about 15 percent of Rwanda's bean area (Larochelle, Alwang, Norton, Katungi, & Labarta, 2014). These new varieties registered an annual incremental production increase of almost 30,000 metric ton with a gross annual value of 8.7 million US dollars (Walker & Alwang, 2015).

Agriculture is an important sector in Rwanda's economy. It accounts for 39 percent of gross domestic product (GDP) and 80 percent of employment (World Bank, 2013). Bean is the most important legume and one of the most vital sources of protein for Rwandan families. Beans are a staple food in Rwanda and as a result, the country ranks first out of 81 countries suitable for investing in iron biofortified beans (IBB) (Asare-Marfo, Gonzalez, Perez, Schwarz, & Zeller, 2013). Rwandans have one of the highest per capita bean consumption rates in the world, with rural households consuming significant quantities of beans on average six days a week (Asare-Marfo et al., 2016; Berti et al., 2012). Beans have the highest share of crop-harvested area in

Rwanda, although there are significant productivity challenges due to limited access to modern agricultural technology. Therefore, there are significant yield gains to be made from the introduction and scaling up of seeds of improved varieties of beans. Iron biofortified beans (IBB) can help improve yields, incomes, nutrition (Haas et al., 2016), and health outcomes (e.g., cognitive and physical functions) (Murray-Kolb et al., 2017) of consuming populations (Luna et al., 2015).

In 2010, the National Agricultural Research System of Rwanda, in collaboration with the CIAT and HarvestPlus, released the first IBB varieties to farmers in Rwanda. These biofortified varieties are micronutrient enriched and demonstrate better yield performance than unimproved bean varieties common in Rwanda. They also tolerate growth-reducing factors, like pests and diseases, and growth-limiting factors, such as droughts. However, little is known about key factors of adoption on the demand side and the role of peer influence. Understanding the factors that drive IBB adoption among bean farmers is critical to help better design policies to increase adoption of improved agricultural technologies to boost food crop productivity in Rwanda.

III. Research objectives

Biofortification has emerged as a cost-effective strategy (Meenakshi et al., 2012) to address micronutrient malnutrition. Since the IBB program rollout in Rwanda in 2012, this agricultural strategy raises two key questions. First, what are the driving factors affecting the adoption of IBB varieties among smallholder farming households? Second, what was the impact of the biofortification program on smallholder farming households' livelihoods? Thus, the objectives of this dissertation are:

- Objective 1: To analyze smallholder farming households’ decisions to adopt IBB in Rwanda, using theories on social interactions and choice behavior.
- Objective 2: To estimate the impact of IBB on smallholder farming households’ yields and incomes in Rwanda
- Objective 3: To estimate the impact of IBB on smallholder farming households’ technical efficiency in bean production in Rwanda.

IV. Significance of this research

It is expected that this empirical research will contribute to the literature in three ways. First, it will provide insight as to what factors drive IBB adoption and what role peer influence plays. Understanding bean farmers’ behavior is critical to design better policies to increase food crop productivity and decrease micronutrient deficiency in Rwanda. This analysis will also provide national and sub-national statistics on adoption rates and on the intensity of adoption of IBB by bean type. With the assistance of spatial econometrics techniques and theories of social interaction and choice behavior, this dissertation will examine how households and farm characteristics, as well as regional factors, influence smallholder farmers' decisions to grow IBB. The dissertation will also assess how important of a role social interaction, proxied by geographic distancing, plays in understanding the interdependence of farmers' decisions to adopt IBB. The robustness of our results will be tested using a simple model of social groups where smallholder farmers are nested within villages. This multilevel mixed model with random effects (village) is expected to produce weaker spatial relationships across villages. This would confirm our hypothesis that for agricultural technology adoption closer neighbors matter more than those farther away.

The second expected contribution to the literature relates to the assessment of the economic impact of improved and biofortified crop varieties on smallholder farming households' outcomes. Notably, I will examine the impact of IBB varieties on Rwandan smallholding farming households' livelihoods, focusing on the yields and incomes of beneficiary households. The findings of this study are expected to support the hypothesis that iron biofortified bean growers had significantly higher yields and potential incomes compared to farmers that grew non-biofortified beans, whether traditional or improved. The empirical analysis will demonstrate the need to control for spatial spillovers, providing evidence that a smallholder farmer's probability of adoption of iron biofortified beans increases if adopting households are located nearby.

The third expected contribution of this research to the literature relates to the estimation of smallholder farming households' unbiased efficiency scores. I estimate a national technological frontier for all bean farmers, a frontier for iron biofortified bean growers (treatment), and a frontier for farmers that grow other improved bean varieties (control). Clustering analysis aims to reveal evidence on where and how this new technology has been effective, thereby providing valuable input into targeting strategies and resource allocation for scaling up of such interventions.

V. Overview of the dissertation

The reminder of this dissertation is organized into six chapters.

Chapter 2 provides a review of the theoretical frameworks commonly used for the analyses presented in chapters 4, 5, and 6. The first part introduces formal definitions of adoption and diffusion of an agricultural innovation, synthetizes commonly cited characteristics and constraints that affect the adoption of such technology, and provides a short summary of the most used theoretical frameworks on agricultural innovation. Economic studies have commonly distinguished between models of technology adoption and models of technology diffusion (path of aggregate demand). The most accepted theoretical frameworks for investigating adoption of technology include (1) economic constraints, (2) adopter perceptions (probit model), (3) innovation-diffusion (epidemic models), and (4) spatial diffusion through social network (spatial econometric models). The fourth framework is employed in chapter 3 to model the role of peer effects on the adoption and diffusion of iron biofortified beans. The second section of this chapter provides a review of the literature related to the evaluation of treatment effects. This section presents a brief review of the studies on the productivity of improved beans in Rwanda, the basis of the framework on potential outcomes, theory on observational studies, and a short discussion on propensity score (PS) matching estimators. This framework is used in chapter 4 for the analysis on impact of IBB on yields and farmer's incomes. The third section of the chapter introduces the total factor productivity (TFP) framework. This framework is used in chapter 5 to model technical efficiency of smallholder farming households.

Chapter 3 presents three conceptual frameworks for use in the analyses that follow. The discussion of each conceptual framework provides a formal definition for each variable and mathematical equations. I use these equations to test the hypotheses presented in Chapter 1. The first conceptual framework introduces two econometric tools: the probit model and the spatial probit model. I used these tools to model latent levels of utility. The second conceptual framework introduces a multivariate matching algorithm to implement Rubin's causal model. The Rubin's causal model tests the null hypothesis of no treatment effect. The third conceptual framework introduces the stochastic frontier analysis (SFA) and stochastic spatial frontier analysis (SSFA) that are applied to estimate bean farmers' technical efficiency.

Chapter 4 presents the analysis of smallholder farming households' adoption of IBB varieties by specifically examining the influence of demand-side factors and the role of peers. I draw upon several theories from studies on the adoption of agricultural technology, social behavior, and spatial econometric methods to build the analytical models. I test for the presence of spatial association among economic agents (farmers), estimate prevalence rates of IBB adoption by district, and examine any potential interactions with contextual factors. I present a disaggregate analysis for climbing and bush bean type. Climbing beans grow tall and need a stake for support with a yield potential of 4,000 kilograms per hectare, while bush beans grow about 2-3 feet tall and do not require support with a yield potential of 3,000 kilograms per hectare. In addition, climbing bean type is mostly grown in the Western and Northern (one spatial regime) regions while bush bean type is mostly grown in the Eastern and Southern regions (second spatial regime). This spatial pattern of adoption suggests a form of heterogeneity that relates to the spatial variability of the parameters that provides evidence for separate analysis. To assess the significance of the two spatial regimes, regions where climbing beans are grown and regions where bush beans are grown, I run a Chow test.

Chapter 5 assesses the economic benefits of IBB delivery efforts in Rwanda. In doing so, I present an analysis of the heterogeneity of outcome variables as a function of PS and baseline covariates. To create the conditions of a natural experiment, this analysis combines quasi-experimental methods and spatial probit methods to address the problem of self-selection bias and spatial spillovers, respectively.

Chapter 6 estimates smallholder bean farming households' efficiency levels. It estimates a national technological frontier for all bean farmers, a frontier for iron biofortified bean growers (treatment group) and a frontier for farmers that grow other improved bean varieties (control group).

Chapter 7 provides an overall conclusion. It provides the outcomes achieved from each objective, as well as research insights and policy implications.

Chapter 2: Literature Review

I. Adoption of Agricultural Technology

This section introduces formal definitions of adoption and diffusion of an agricultural innovation, synthetizes commonly cited characteristics and constraints that affect the adoption of such technology, and provides a short summary of the most used theoretical frameworks on the adoption and diffusion of agricultural innovation.

i. Definitions of Adoption and Diffusion of Agricultural Technology

Technology adoption is defined here as the choice of an individual farmer to acquire and use an agricultural innovation. Many empirical studies measure adoption of an innovation by using one of the two variables: a discrete choice as to whether or not to utilize an innovation, or a continuous variable, such as on the timing or extent of new technology utilization by individual farmers (Sunding & Zilberman, 2001). In this study, I use continuous variables, estimating at national and province levels the total area allocated to bean production with IBB varieties and the number of bean farmers cultivating IBB in season B of 2015. (Season B sowing dates span from around January to March, followed by harvest activities from June to July.)

Established economic literature defines innovation diffusion as an aggregate measure of adoption and thereby analyzes the process through which the innovation penetrates markets and replaces traditional technologies (Sunding & Zilberman, 2001). Measures of innovation diffusion include the percentage of the farming population that adopts new innovations and the area on which the innovation is employed as a share of the total land on which the innovation can be utilized. Namely, innovation diffusion is a process by which an innovation is communicated through certain channels over time among the members of a social system (Rogers, Singhal, & Quinlan, 2019).

A stream of the literature shows the importance of social networks, or peer-effects, as a mechanism for the spatial diffusion of technology (Manski, 1993; Beaman, BenYishay, Magruder, & Mobarak, 2018). This mechanism is especially important when farmers lack information pertaining to the likelihood of the possible outcomes of the new agricultural technology. Farmers may be uncertain about the management practices they should optimally employ, the types and costs of inputs needed, and the economic returns of the new technology. In this context, social learning and social networks often complement or act as substitutes in delivering information and facilitating the technology adoption process.

ii. Commonly Cited Characteristics and Constraints that Affect the Adoption and Diffusion of Innovation Agricultural Technology

The characteristics of a social network—a farmer's social links through which information, goods, money, and services flow—are factors that might induce technology adoption and diffusion (Maertens & Barrett, 2013). Empirical studies have shown the effect of social networks on facilitating the adoption of new agricultural technologies in developing countries. Foster & Rosenzweig (1995) find that farmers with neighbors who have more farming experience have higher profits than those without such neighbors. Krishnan & Patnam (2014) find evidence that social learning was more persistent than learning from extension services for the adoption of new varieties and fertilizer in Ethiopia. Conley & Udry (2010) examine how learning from the experience of others and the flow of information depend on the structure of social networks when there is no access to agricultural extension services. Ward & Pede (2015) find that neighbor effects are a significant determinant of hybrid rice use. They use two specifications to model the endogenous effect of being neighbors—one based on membership in the same village, and the other based on geographical distance.

Other factors driving adoption includes farm size, access to credit, land tenure, human capital, infrastructure, and wealth. Farm size was one of the first factors explored in the empirical literature on adoption. Farm size can have different effects on the rate of adoption depending on the characteristics of the technology. A wide variety of empirical results interpreted in the context of the theoretical literature suggests that farm size is a surrogate for many potentially important factors driving technology adoption, such as access to credit, capacity to bear risk, access to scarce inputs, wealth, and access to information. Chirwa (2005) finds that close to 60 percent of sampled households in Malawi do not use hybrid maize varieties, but that adoption increases with income, education, and farm size.

Many studies have examined the role of access to credit and appropriate financial instruments as a constraint in farmers' adoption decisions (seeds, fertilizers, and pesticides). According to theoretical and empirical research studies, access to capital through either accumulated savings or capital markets is necessary to finance the adoption of many new agricultural technologies, especially for smallholder farmers. Simtowe, Zeller, & Diagne (2009) find higher hybrid maize adoption among households with access to credit in Malawi. Croppenstedt, Demeke, & Meschi (2003) estimating a model of fertilizer use in Ethiopia, find that household cash resources are generally insufficient to cover fertilizer purchases.

There are two methods commonly used in the literature to quantify the role of credit constraints. One is to ask farmers the primary reason why they do not adopt a technology in order to determine if the reasons might be correlated with either wealth or income. However, a problem with this method arises if the returns to the adoption of technology vary by farm scale. The second method is to control for income, scale, and insurance effects to avoid biased adoption estimates (Foster & Rosenzweig, 2010).

A number of empirical and descriptive studies have also considered the effects of land tenure arrangements (which is often considered to be a good proxy for wealth) and the proportion of farms rented on the adoption of new agricultural technology, such as improved, high-yielding varieties. Findings suggest that the form of land tenure (e.g., renters, sharecroppers, landowners) may affect adoption decisions and diffusion rates.

Poor-functioning infrastructure affects the profitability of technology adopted by farmers. The extension and quality of road networks and mobile telephone services rank among the most important infrastructural conditions. In general, transportation limitations tend to reduce competition among input suppliers and middlemen. Empirical evidence shows that travel times between the farm gate and market can be high due in part to underdeveloped road infrastructure. Good transportation is associated with diffusion of technology, better access to inputs, and higher producer prices (Dorosh, Wang, You, & Schmidt, 2012).

Three mechanisms related to human capital have been identified in the literature to explain technology adoption: (1) more educated agents are wealthier, and thus the education–adoption relationship represents an income effect; (2) more educated agents have better access to information; and (3) more educated agents are better able to learn or otherwise internalize new information. The last mechanism has been the principal focus of economists (A. D. Foster & Rosenzweig, 2010). Numerous studies find a significant relationship between education indicators and farm productivity. Since the adoption of innovation generally increases productivity, the importance of education in affecting adoption behavior is implicit. Jamison & Moock (1984) test the effect of schooling and extension contacts on the adoption and diffusion of agriculture technology in Nepal. They find that schooling influences adoptive behavior, but that household income mediates the adoption decision. Weir & Knight (2007) find that the level of education within the household- in Ethiopia is an important factor in adoption, and that early adopters tend to be more educated and to influence their neighbors. Giné & Yang (2009) find that farmers' education, income, and wealth were positively correlated with the take-up of insured loans to adopt a new crop technology in Malawi.

The local and regional geographical setting within villages that can directly and indirectly influence adoption—including geographical variables, such as rainfall, soil type, dominant ethnic group, slope, farmer management practices in a village, population density, road density, and market access. These may vary and, therefore, have an impact on yield differentials across farmers adopting the same technology.

iii. Theoretical Frameworks Applied to the Adoption of Agricultural Technology

Economic studies have commonly distinguished between models of technology adoption and models of technology diffusion (path of aggregate demand). The most accepted theoretical analytical frameworks for investigating the adoption and diffusion of technology include (1) economic constraints, (2) adopter perceptions (probit model), (3) innovation-diffusion (epidemic models), and (4) spatial diffusion through social network (spatial econometric models).

Economic constraint modeling is probably the most extensively used theoretical approach in the literature for examining the adoption of agricultural innovation. Research studies commonly apply the farmer's decision-making model, which is well documented by Feder, Just, & Zilberman (1985). Using this model, it is assumed farmers' decisions result from the maximization of an expected utility against constraints, including – but not limited to – land, access to credit, market access and infrastructure. Farmer profits are defined as a function of the farmer's choice of crops and varieties of these crops to cultivate traditional vs. modern. Therefore, a farmer's income is a function of land allocated to different crops and crop varieties, which can, consequently, be explained by the production function of each crop: yields, inputs (amount and prices), and other associated costs of production. To address the decision over time, Feder, Gershon and T. O'Mara, (1981) and Feder et al., (1985) suggest empirical models that factor in perceived parameters of the production function. These parameters can be updated through learning processes that incorporate prior perception and recent information about yields, input uses, prices and other factors for farmers in any region using Bayesian modeling of learning rules to update farmers' perceptions. Other variables considered in structural equations include extension services and human capital. Other dynamic variables include optimal timing of technology adoption, learning by using, and learning by doing, which require a wealth of data, such as a panel survey, for analysis.

Another less cited theoretical approach is a framework based on farmers' perceptions of technology characteristics. Agricultural technology attributes (e.g., yields, drought resistance, pest resistance) and consumers' subjective perceptions (e.g., taste) can be significant in explaining decisions to use a technology. Seminal papers by Adesina (1993) and Adesina (1995) assume farmers' adoption decisions to be based on the non-observable underlying utility function, U (M, A), where M is a vector of farm and farmer-specific attributes of the adopter, and A is a vector of the attributes associated with the technology. Adesina (1993) examines farmer technology adoption conditioned to farmer perceptions of technology-specific characteristics of mangrove rice varieties in Sierra Leone. Adesina (1995) explores farmers' perceptions of modern sorghum and rice varietal technologies in Burkina Faso and Guinea.

An associated theory to adoption of a new technology is the product price treadmill, in which farmers continually seek to improve their incomes by adopting new agricultural innovations. Early adopters make profits for a short time because of their lower unit production costs. As more farmers adopt the technology, production goes up, prices go down, and profits are no longer possible, even with lower production costs. Average farmers are forced by lower product prices to adopt the technology and lower their production costs to remain in the market. The laggards who do not adopt agricultural innovations are lost in the price squeeze and leave room for their more successful neighbors to expand (Sunding & Zilberman, 2001).

The last theoretical framework, spatial diffusion through social interaction, is perhaps the most recent to propagate in the literature (Conley & Udry, 2010). Seminal work by Manski (1993) identifies three sources of social influence: (1) endogenous effects, (2) exogenous network effects, and (3) correlated effects. Most examples in agricultural technology diffusion place an emphasis on modeling the endogenous effects. Common spatial econometric methods used include the spatial error model and the spatial lag model (Holloway, 2002; Ward & Pede, 2015). I extend these methods by modeling the endogenous and exogenous network effects using the spatial Durbin model (Anselin, 1988; LeSage, 2009).

This dissertation explores the spatial structure of the IBB adoption process in Rwanda. IBB is a relatively new technology in Rwanda and spatial contextualization is a key factor for understanding the diffusion of IBB varieties. I assume that the decision of a bean growing household to adopt an IBB variety is spatially correlated. Therefore, the bean farming household's decision to adopt IBB depends not only on her own and her farm level characteristics, but it is also correlated with the decisions of neighboring bean farmers and their characteristics. Thus, a bean growing household that is close to a household that has adopted IBB has a higher probability of being an IBB adopter as well, which is the endogenous effect.

Another condition relates to the social characteristics of a group as the main factor in spatial clustering, which is the likelihood of an individual to behave, on average, in agreement with their social group. Whether or not the diffusion of IBB varieties is geographically driven, the spillover effects will determine a strong spatial relationship, i.e., similar farms will tend to be localized in the same geographical area. Therefore, within a region, it is possible to find similar economic structures, wealth and management practices levels, and family structures.

The theoretical framework applied to the adoption of agricultural technology draws concepts on social interaction from Conley & Udry (2010) and Ward & Pede (2015) and concepts on optimal choice and utility maximization from Abdulai, Monnin, & Gerber (2008). To implement this framework, I make two broad assumptions (1) a smallholder farming household's decision to grow IBB varieties is based on utility maximization theory; and (2) new IBB varieties produce higher yields conditioned on the use of modern inputs and management practices. Another salient uncertainty follows from the fact that the smallholder farming household is less familiar with IBB varieties. The structure of the production function of a smallholder farming household for period *t* and future period *t+1* is specified in Eq. (1),

$$y_{h,t+1=} f(X_{ht}, k_{ht}, \omega_h) + \varepsilon_{h,} \tag{1}$$

where $y_{h,t+1}$is the smallholder farming household's future output, X_{ht} is the quantity of inputs used in the current period, k_{ht} is the smallholder farming households' level of information used in the current period, ω_h environmental conditions, and $\varepsilon_{h,}$is a *i.i.d* disturbance for household *h* with zero mean and σ_ε^2. $\varepsilon_{h,}$is assumed to follow a normal distribution.

The profit function is shown in Eq. (2),

$$\Pi_{h,t+1} = P_{t+1}f(X_{ht}, k_{ht}, \omega) - C_t = max[P_{t+1}f(X_{ht}, k_{ht}, \omega) - \gamma X_{ht} - \zeta k_{ht}] \quad (2)$$

$\Pi_{h,t+1}$ indicates that the value given by the function is the maximum profits obtainable given local market prices. *C* is the cost of production. Only the variables X_{ht} and k_{ht} are under the smallholder farming household's control. The bean farming household chooses levels of these inputs, X and k, in order to maximize profits. The smallholder farming household maximum profits ultimately depend on these three exogenous prices, *P*, γ, and ζ, together with the form of the production function.

The other two sets of assumptions include (1) smallholder farming household's profit expectation depends not only on their own experiences, preferences, etc., but also on their social interaction with other farmers experiences, expectation, and constraints; and (2) social interaction occurs in local places and its strength depends on the relative social geographic distance between IBB adopters and their neighbors. Therefore, smallholder farming households are assumed to maximize expected profits Π_{hi} as stated in Eq. (3),

$$EU\ (\Pi_{hi}) \equiv EU[\Pi_{hi}|f\left(x_{hi}, k_{hi}, \omega, d_{ij}\right), f\left(x_{hj}, k_{hj}\omega\right)] \quad (3)$$

E denotes the expectation operator, *U* is the von-Neuman-Morgenstern utility function; x_{hi} and k_{hi} denote the smallholder farming household inputs decision; x_{hj} and k_{hj} are the inputs decision of neighboring farmers which in turn are a function of the social geographic distance d; and ω denotes environmental conditions.

Smallholder farming households will grow IBB varieties if the expected marginal benefit is greater than the marginal benefit of growing a traditional bean variety. However, the expected marginal benefit is not observable. The smallholder farming household either adopts or does not adopt IBB varieties. In this case, discrete-choice models become instrumental. They are commonly used to investigate a wide range of areas in agricultural economics, including technology adoption and land-use decision-making. I start from the basic empirical model, which is based on farmers' decisions on whether to adopt an IBB variety.

II. Impact Evaluation

In this section, I provide a review of the literature related to the evaluation of treatment effects. I present a brief review of the studies on the productivity of improved beans in Rwanda, the basis of the framework on potential outcomes, theory on observational studies, and a short discussion on PS matching estimators.

i. Productivity of improved bean varieties in Rwanda

Larochelle et al., (2014) examined the impact of improved bean varieties on bean farmers' livelihoods in Rwanda. Their study was based on a comprehensive household survey conducted in 2011/12. Adopters of improved bean varieties were shown to report yield gains of 42 kg/ha compared to households that planted local varieties. Farmers that grew climbing bean varieties had 28 percent higher yields than farmers that grew bush bean varieties. More recently, based on the same nationally representative cross-sectional survey data that is used in this study, Vaiknoras & Larochelle (2018) evaluated the impact of IBB bush variety RWR2245 on productivity, consumption, purchases, and sales. The authors found that RWR2245 growers had on average 49 percent higher yields than traditional bush bean growers, (note in that study the yield was measured as a ratio of quantity of grain harvested to quantity of seed planted).

ii. Potential Outcome Framework

From an econometric point of view, estimating the effects of potential outcomes poses two computational problems: endogeneity and missing counterfactuals (Greene, 2012). The former refers to risks with accurately identifying the causal effect associated with factors that affect both the treatment and the outcome. The latter refers to the fact that just one outcome is being observed. For instance, to measure the impact of IBB adoption on farmers' yields, I would need to run the scenario of bean farmers' production twice for the same farmers, one with IBB adoption and one without IBB adoption. To deal with both econometric problems, I implement the Rubin's causal model (1974, 1978).

In the literature, there are two statistical designs normally used to estimate the causal effect of a treatment or a policy on outcome variables: randomized controlled trials (RCTs) and observational studies. RCTs are considered the gold standard method for causal inference. The basis of an RCT is the random assignment of the treatment to subjects, which uses chance to form comparable groups. When RCTs are not ethical or not feasible, the effect of a treatment can be examined using observational studies.

Observational studies are defined as quasi-experiments. A quasi-experimental design is like an experimental design in that there is a specific investigator-defined intervention for the "treatment" group in the study, but the subjects – in our case, Rwandan bean farmers – are not randomized to receive the treatment (Rosenbaum, 2010). Observational studies are used in a variety of fields from economics to medicine (Austin, 2011, 2014). When properly implemented, such studies can yield results that are almost as reliable and robust as those derived from analysis of RCTs.

A critical requirement to evaluate causal effects on potential outcomes using either statistical design is that the evaluation must be a comparison of $Y_h(1)$ and $Y_h(2)$ for a common set of units, such as bean farmers N. Formally, a causal effect must be a comparison of the ordered sets $\{Y_h(1), h\epsilon N\}$ and $\{Y_h(0), h\epsilon N\}$ (Rubin, 2005). So, assumptions are required to implement potential outcomes scenarios. These assumptions include:

1) The stable unit treatment value assumption, which requires two assumptions. First, it assumes that there is no interference between treatment units $Y_h(1)$ on $Y_h(0)$ or vice versa. Secondly, it assumes that there are no hidden versions of treatments. In addition, I used Kelejian & Prucha (2001) Moran's I to test for the presence of spatial dependence among treatment units.
2) The covariates and the potential outcomes are not affected by how I learn about them, whether by randomized controlled trials or observational studies.

Broadly, there are statistical methods that control for variation in the confounding factors. Common statistical techniques include matched sampling, stratification, model-based adjustments, and combinations of these techniques (Rosenbaum, 2005). In this dissertation, I employ matched sampling methods, i.e. propensity score (PS) matching, which relies on the assumption of “selection on observables”. This means that all variables that affect treatment assignments and outcomes have been measured (Rosenbaum and Rubin 2003).

III. Efficiency analysis

Economic indicators of performance, such as measures of productivity and efficiency, are commonly used to investigate the impact of a new technological-innovation on farmers' outcomes (Duflo, Kremer, & Robinson, 2008). Assessing farmers' efficiency, which is defined as the ability of farmers to utilize the best available technology and to allocate resources productively, together with the impact of an intervention requires the combined application of analytical methods (e.g., for earlier examples (Dinar, Karagiannis, & Tzouvelekas, 2007) (Bravo-Ureta, Greene, & Solís, 2012). Recent literature in productivity analysis and impact evaluation highlights the importance of measuring spillovers to non-beneficiaries e.g. (Gamerman & Moreira, 2004; Schmidt, Moreira, Helfand, & Fonseca, 2009).

When conducting efficiency analysis on a cross-sectional or a panel dataset, a high degree of heterogeneity may lead to biased and inefficient estimates of the efficiency scores. Researchers have approached this problem in different ways. One way is using non-parametric techniques, such as data envelopment analysis (DEA), which ignores the functional form of the production function. Other studies have implemented a two-step approach: the first estimates the frontier, while the second step analyzes the determinants exerting influence over economic agents' efficiencies (Chavas, Petrie, & Roth, 2005; Simar & Wilson, 2007). Greene (2008) proposed the true-fixed effects and the true-random effects models for panel data. When there is spatial heterogeneity, instead of including spatial fixed effects, some authors allow the externalities to spill over throughout the system (Han, Ryu, & Sickles, 2016a). In this research, I implement that latter method as described below.

i. Theoretical Framework

In economics, productivity and efficiency both deal with the economic performance of a production unit. Both refer to the production process in which the economic agent (farmer) transforms a set of inputs X ϵ R_M^+ into a set of outputs Y ϵ R_M^+(Greene, 2008). Efficiency requires the existence of a benchmark (best practices) as it signifies the comparison between observed and optimal values on output or inputs or both. In this study, I evaluate efficiency for bush and climbing bean growers, separately.

To evaluate efficiency levels among bean farmers, I used the total factor productivity (TFP) framework. The *TFP* index is the ratio of total bean production to total inputs employed by the bean farming household, *h*. This index allows for the presence of technical inefficiency in the bean production process. In addition, I measure, through scale efficiency analysis, how close bean farming households are to operating at the optimal scale. The larger the scale efficiency, the closer the farming household is to the optimal scale.

To operationalize the concepts above, I introduce the production function in Eq. (4).

$$Y_h = A_h F(X_h) \tag{4}$$

where Y_h relates to bean farmer's output, X_h the vector of inputs used by the production unit. F(.) represents the body of knowledge available to the producer, and A_i is the index of productivity or the amount of output a given unit can produce from a certain amount of inputs, given the technology level. This is more commonly known as the total factor productivity (TFP), formally Eq. (5),

$$TFP_h \equiv A_h = \frac{Y_h}{F\,(X_h)} \tag{5}$$

The TFP_h index is the ratio of total household bean production to total input employed. I use this framework to evaluate variation in productivity among bean farmers in season B 2015. It is relevant to note that the observed output is equal to the potential level of production, i.e. the frontier output, with no room for technical inefficiency. Allowing the presence of technical inefficiency in production processes lead to Eq. (6),

$$Y_h \leq A_h F(X_h) \tag{6}$$

where the observed level of bean production in household h, Y_h, does not necessarily turn out to be equal to the potential output. Koopmans (1951) introduces the formal definition of technical efficiency and Farrell (1957) operationalized the concept. Here, I build on the case of one output and many inputs. I introduce below the output-oriented measure of technical efficiency.

$$TE_o(X_i, Y_i) = [max\{\emptyset : \emptyset Y \leq_i A_i F(X_i)\}]^{-1} \tag{7}$$

Solving for Y_i,

$$Y_i = TE_{\circ}(X_i Y_i). A_i F(X_i) \tag{8}$$

where $TE_o(X_i, Y_i) \leq 1$. If the framework allows for technical inefficiency, maximum potential output will be equal to the observed output corrected for the output-oriented technical score, which is equal to 1 for fully efficient smallholder bean farming households. From a theoretical point of view, differences in productivity levels are due to differences in factors relating to: (1) technology, (2) scale of production, and (3) externalities. In the first stage, I estimate smallholder farming households' efficiency and scale of efficiency. Scale efficiency expresses how close the bean farmer is to the optimal scale size; the larger the scale efficiency, the closer the firm is to optimal scale. I employed stochastic models to estimate efficiency and DEA analysis to explore smallholder farming households scale of production. In the second stage, I explore factors that exert influence on households' efficiency levels.

ii. Spillovers

I recognize that the biofortification program rollout might have created spatial spillovers in the technical efficiency of non-beneficiary bean farmer growers. Spillovers can be the result of the interactions between economic agents from a local to a global perspective. These interactions may include spillovers of knowledge, technology, and social behavior.

Knowledge spillovers can create economic value for other agents. For instance, knowledge spillovers may occur when information is exchanged between farmers about the benefits of a new agricultural technology (Besley & Case, 1993; A. D. Foster & Rosenzweig, 2010), such as the exchange of information on the nutritional and agronomic benefits of IBB. Conley & Udry, (2010) show that pineapple farmers in Ghana follow the decisions made by other, more experienced farmers, when deciding to adopt a new technology. For example, a farmer would determine the amount of farmland devoted to a crop by considering the amounts allocated by the other farmers in the system.

Technology spillovers refers to the benefits that smallholder farming households receive from research efforts without incurring shared costs. At a broader spatial scale, international spillovers from public agricultural research and development

(R&D) represent a high percentage of agricultural productivity growth (Alston, 2002). Specifically, agricultural R&D and technology spillovers among geographical areas (countries-to-countries, states-to-neighboring states) occur when research conducted by one geographic area transfers benefits to other geographic area(s). An illustrative example includes the 10 varieties of iron biofortified beans, which were released in Rwanda between 2010 and 2012, following years of collaborative research between HarvestPlus, CIAT and RAB. Parental lines of improved bean varieties are bred in CIAT, which is headquartered in Latin America, and distributed to national agricultural research services in Africa, Asia and Latin America for further development, adaptation and release.

The adoption of a new agricultural technology, when released, will depend not only on varying physical geographical variables, like climate, terrain, and soil, but also on other regional and economic factors, such as road infrastructure, accessibility to markets, and institutional setting. To shed light on the spatial patterns of growing IBB varieties and the determinants of farmers' technical efficiency, I conducted a spatial clustering analysis to better understand the geographic concentration of advanced farmers versus less advanced farmers.

This section presented three theoretical frameworks that are combined to study (1) smallholders' decision and the role of peers to adopt IBB, (2) the economic benefits of the iron biofortified bean program, and (3) the efficiency of iron bean production. The next section provides details of the analytical methods employed in the analyses to answer the questions and hypotheses presented in the first chapter.

Chapter 3: Methodology

This chapter presents three conceptual frameworks for use in the analyses presented and discussed in chapters 4, 5, and 6. The first conceptual framework introduces two econometric tools: the probit model and the spatial probit model. I employed these econometric tools to model latent levels of the utility of adopting IBB in smallholder farming households and their peers. I run separate analysis for IBB climbing type and IBB bush type. Climbing bean type is mostly grown in the Western and Northern regions (one spatial regime) while bush bean type is mostly grown in the Eastern and Southern regions (second spatial regime). This spatial pattern of IBB adoption suggests a form of heterogeneity that relates to the spatial variability of the parameters that provide evidence for separate analysis. To assess the significance of these two spatial regimes, I run a Chow test. In my dataset, I classify households based on these two spatial regimes. One spatial regime includes smallholder farming households growing IBB climbing type in the Western and Northern regions. The second spatial regimes include smallholder farming households growing IBB bush type in the Eastern and Southern. The second conceptual framework introduces a multivariate matching algorithm to implement Rubin's causal model. This matching algorithm tests the null hypothesis of no treatment effect. The third conceptual framework introduces the stochastic frontier analysis applied to estimate bean farmers' technical efficiency.

I. Conceptual framework: choice model

i. Logit model

A smallholder farming household's expected profit from adopting a biofortified seed, instead of a local seed or other improved variety, depends on different sets of variables: prices of inputs and outputs; fixed factors, such as farm assets and land holdings; soil characteristics; socioeconomic characteristics, such as education and wealth; neighborhood influences (expected profits to neighbors from adoption); and factors on the supply side, such as seed availability in the market.

I start with a basic latent regression model as shown in equation (9). I analyze the outcome of a discrete choice as a reflection of an underlying regression function. The basic theory is that the farmer makes a marginal benefit or marginal cost estimation based on the utilities achieved (Greene 2012). I model the difference between benefit and cost as an unobserved variable, $y^* = \Pi_{1i} - \Pi_{0i}$, which represents the difference in utility where Π_{1i} represents the utility associated with variety 1, and Π_{0i}, the utility from other varieties, such that

$$y^* = x'\beta + \varepsilon \tag{9}$$

I assume that ε has a mean of zero. Our only observation of the data generation process is

$$\begin{aligned} y &= 1 \text{ if } y^* > 0 \\ y &= 0 \text{ if } y^* < 0. \end{aligned} \tag{10}$$

The smallholder farming household either did adopt (Y = 1) or did not adopt (Y = 0) biofortified beans in season B of 2015. I hypothesized that a set of intrinsic factors, such as farmer characteristics, plot characteristics, and environmental factors, gathered in a vector x, explain farmers' decisions, so that:

$$\text{Prob}\,(Y = 1|x) = \text{F}(x, \beta)$$
$$\text{Prob}(Y = 0|x) = 1 - \text{F}(x, \beta)\,. \tag{11}$$

The set of parameters β reflects the impact of changes in x on the probability. For instance, the marginal effect of household head age on the probability of adoption of IBBs may be a factor of interest.

Normally, the estimation of P(X) = Pr(C = 1 | X) is done by means of a probit or logit model. However, when there are spatial effects, conventional models calculated by maximum likelihood are not adequate. By construction, the errors of a spatial logit model are heteroscedastic, and estimates based on the hypothesis of homoscedasticity in the presence of heteroscedastic errors are inconsistent (Greene, 2012) . Therefore, spatial probit models are used to calculate the probability, P(X)=Pr(D = 1 | X), or propensity, of being an IBB grower for each observation.

ii. Spatial choice model

In this model, I test two hypotheses, whether the propensity of an individual smallholder farming household to grow a new IBB variety depends on (1) the prevalence of IBB adoption of neighboring farmers and 2) on the prevalence of the distribution of the characteristics of neighboring smallholder farming households. In spatial econometrics, social interaction is operationalized by constructing a spatial structure that defines the interdependences among farmers in which preferences, local knowledge, and constraints faced by one farmer are directly influenced by the characteristics and choices of other farmers. I use spatial econometric theory on Bayesian spatial probit modeling presented by LeSage (2008).

In contrast to the standard probit and logit model, where y_h^* represents the latent unobservable utility that depends not only on observable determinants of household *h* represented by X, spatial probit modeling also depends on latent variables y_{hj}^* of neighboring households.

The general model for social-spatial interaction takes the following (matrix) form:

$$y^* = \rho W y^* + X\beta + \gamma W x + u \tag{12}$$

$$u = \alpha + \gamma W u + \varepsilon \tag{13}$$

where the matrix $W(n\times n)$, called the spatial weight matrix, captures the dependence structure between neighboring farmers. The variable Wy^* denotes the endogenous interaction effects among the dependent variables across neighboring farmers, *Wx* the exogenous interaction effects among the independent variables, and *Wu* the interaction effects among the disturbance terms of the different spatial units. ρ is called the spatial autoregressive coefficient, λ the spatial autocorrelation coefficient, **α** represents a $n\times 1$ vector of fixed but unknown parameters to be estimated, and β is a $n\times k$ matrix of unknown parameters to be estimated.

For the first hypothesis, I test endogenous effect which is also described in the literature as imitation, contagion, bandwagons, social norms, and keeping up with the Joneses. Similar to the standard probit and logit model, as presented in Section 3.I.i, where y_h^* represents the latent unobservable utility that depends not only on observable determinants of household *h* represented by *X*, spatial probit modeling also depends on the latent utility of the neighboring household y^*_{hj}. Restriction γ = 0 and λ = 0 give rise to the spatial autoregressive model (SAR).

In more detail, the SAR model, according to LeSage, (2009), is

$$y^*_h = \rho W y^*_h + \beta X + \varepsilon, \varepsilon \sim N(0, I_n) , \tag{14}$$

The data generating process for y is

$$y^*_h = (I_n - \rho W)^{-1} X\beta + (I_n - \rho W)^{-1} \varepsilon$$
$$\varepsilon \sim N(0, I_n). \tag{15}$$

where $(I - \rho W)^{-1}$ is the "Leontief inverse" that links the decision of the smallholder farmer, y_i to all X_1 the system through a so-called spatial multiplier (Wilhelm & de Matos, 2013).

For the second hypothesis, I model the effect of contextual factors on smallholder farming household's decision to adopt IBB planting material. I employ a variation of the SAR model in the analysis of contextual effects—the Bayesian spatial Durbin model (SDM). This model allows variables from neighboring farmers contained in the matrix *X* to exert an influence on the propensity of IBB adoption by household h_i. This is accomplished by adding average-neighbor values of the explanatory variables, created using the matrix product *WX*

LeSage, (2009) provides the data generation process of the SDM as

$$y^*_h = \rho W y^*_h + \alpha\iota + \beta X + \theta WX + \varepsilon; \varepsilon \sim N(0, I_n) . \tag{16}$$

The spatial lag latent dependent variable Wy^*_h involves the *n* x *n* spatial weight matrix *W* that contains elements consisting of either one or zero. All elements of the matrix *W* is row standardized (non-negative and each row sums to 1). The scalar parameter ρ measures the strength of dependence, with a value of zero indicating independence. A non-spatial probit model emerges when ρ = 0.

iii. Average Marginal Effect

In spatial models, a change in some explanatory variable xi for observation *i* will not only affect the observations y_i directly (direct impact), but also affect neighboring observations y_j(indirect impact). These impacts potentially also include feedback loops from observation *i* to observation *j* and back to *i* (Lacombe & LeSage, 2018). The scalar summary measure of indirect effects cumulates the spatial spillovers falling on all other observations, but the magnitude of impact will be greatest for nearby neighbors.

LeSage, et al. (2011) construct a matrix version of the own partial and cross partial derivatives, where d(.) represents the n × 1 vector on the diagonal of a diagonal matrix D(.), where the nondiagonal elements are zeros. By construction, D(.) is symmetric. The n × 1 vector d(f(n)) contains the probability density function (pdf) evaluated at the predictions for each observation and associated n × n diagonal matrix D(f(n)), which has d(f(n)) on the diagonal. Using the matrix of own partial and cross partial derivatives, LeSage et al. show that an n × 1 vector of total effects can be written as:

$$\begin{aligned}
\frac{\partial \Pr(y=1)}{\partial x'_{\mathrm{v}}}\,\iota_\eta &= \left[\, D\,(\big(f(\eta)\big)\iota_\eta + \rho D\big(f(\eta)\big)W\iota_\eta + \ \rho^2 D(\big(f(\eta)\big)W^2\iota_\eta + \cdots \right]\beta_{\mathrm{v}} \\
&= \left[\, D\,(\big(f(\eta)\big)\iota_\eta + \rho D\big(f(\eta)\big)\iota_\eta + \ \rho^2 D(\big(f(\eta)\big)\iota_\eta + \cdots \right]\beta_{\mathrm{v}} \\
&= (D\,(\big(f(\eta)\big)\iota_\eta)(1-\rho)^{-1}\,\beta_{\mathrm{v}} \\
&= (d\,\big(f(\eta)\big))\iota_\eta)(1-\rho)^{-1}\,\beta_{\mathrm{v}}\,. \qquad (17)
\end{aligned}$$

As a scalar summary measure of average total effect, LeSage et al. (2011) uses an average of the vector of total effect,

$$n^{-1}(d\,\big(f(\eta)\big)'\iota_\eta)(1-\rho)^{-1}\,\beta_{\mathrm{v}}\,. \qquad (18)$$

The average direct effect is

$$\frac{1}{n}tr\left(\frac{\partial \Pr(y=1)}{\partial x'_{\mathrm{v}}}\right)$$

$$= \left[\, tr(D\,(f(\eta))) + \rho tr(D(f(\eta))W) + \ \rho^2 tr(D(f(\eta))W^2) + \cdots \right] \frac{\beta_{\mathrm{v}}}{n}. \quad (19)$$

For the average indirect effect, they propose using the difference between the average total effect and the average direct effect.

iv. Multilevel model

As a robustness test, I ran a new set of regressions with fixed and random effects. To do so, our multilevel data structure includes villages in the upper level of the model and smallholder farming households nested within those villages as the lower level. I carried out a multilevel Bayesian analysis of latent Gaussian models using the Integrated Nested Laplace Approximation (INLA) (Rue, Martino, & Chopin, 2009). In a sample of villages, the model with fixed and random effects treats observations from a given village as a cluster and assumes a random effect for each village. I define $\mu_{ij} = (Y_{ij} | U_j)$. Let Y_{ij}be the response of smallholder farmer *i* in cluster *j*, i = i_1,..., i_n. In our case, the responses are adoption of IBB planting material. I implemented *i.i.d* random effect term *U* at the upper village level. The *i.i.d* random effect representation implies 1) strong intra-village dependence between the lower-level observations here smallholder farmers; and 2) weak inter-village dependence. The general mixed model has the form,

$$g(\mu_{ij}) = \ \gamma X \ + \ U_j;\ i = 1, \ldots, n_j; j = 1, \ldots, 81\,; u_j \sim \ N\ (0, \tau_u) \qquad (20)$$

g is the link function, for binary outcomes is the logit link. X_{ij} denotes a vector of explanatory variables such as household head age, years of farming experience, household size, wealth index, and number of bean varieties cultivated, for fixed effect model parameters γ. U_j denotes the vector of random effects for village j. This is common to all observations in the cluster. The random effect vector U_j is assumed to have a multivariate normal distribution $N\ (0, \tau_u)$. The covariance matrix τ_u depends on unknown variance components and correlation parameters. Parameters pertaining to the random effects may be also of interest as a useful summary of the degree of heterogeneity of the population of smallholder farmers. Note that n_j represents the number of smallholder farmers in village *v*. Village *v* is indexed from 1 to 81. I expect that the random effect (villages) will produce a weaker spatial relationship, whether this is true, I would expect to have robust result confirming our hypothesis that closer neighbors matter more than those farther away.

v. Specification

I used the nomenclature M1 and M2 for the two specifications used for each of the non-spatial and spatial probit models as specified in subsections 3.I.i and 3.I.ii, respectively. Specification M1 aims to test how household characteristics, such as wealth (proxied by a household asset index - see Appendix), household composition, iron bean consumption, and years of farming experience, play a role on IBB adoption. In addition, it explores the role of a number of varieties used to manage risk of food insecurity due to crop failure caused by the prevalence of drought. Specification M2, on the other hand, looks at the importance of household technical capacity measured through a management index connected to household education level and household size. The latent variable for the adoption of IBB corresponds to the unobserved profitability. For the construction of the normalized spatial weights matrix, I determine the inverse bilateral distances between all bean farmers in the data.

The specification of the covariates is key, in line with economic theory. To provide further details for each specification and its covariates, in the following section I provide a brief description of the main covariates used through chapters 4, 5, and 6. Table 1 shows relevant factors driving IBB adoption, which include household characteristics, farm characteristics, management practices, and regional geographical variables. I considered spillovers between bean farmers being geographically bound within a radius of 3 kms. For reasons of interpretation, I row sum normalize W.

In Rwanda, more than 80 percent of the economically active population are involved in agriculture. In this study, on average, households with more economically active members have a higher propensity to adopt IBB, suggesting that available labor is a consideration in the decision to adopt. The difference in the average household family size is statistically significant between adopters and non-adopters, suggesting an increased need among adopters to meet food demand in the household.

The average education of members in households that grew IBB is statistically greater than for non-adopting households, indicating that higher education influences adoption of the new technology and is positively correlated with wealth. Three mechanisms related to human capital have been identified in the literature to explain technology adoption: (1) more educated agents are wealthier, and thus the education–adoption relationship represents an income effect; (2) more educated agents have better access to information; and (3) more educated agents are better able to learn/ internalize new information. The last mechanism has been the principal focus of economists.

Adopters managed more plots and varieties over larger cropped land areas. These behaviors could be associated with a household's food security strategy where households use mixed bean seeds (local, improved, and iron biofortified) to minimize the risk of food insecurity associated with crop failure or poor crop yield performance of a specific bean variety.

Adopters had larger farmland. Farm size can have different effects on the rate of adoption depending on the characteristics of the technology. A wide variety of empirical results interpreted in the context of the theoretical literature suggests that farm size is a surrogate for many potentially important factors, such as access to credit, capacity to bear risk, access to scarce inputs, wealth, and access to information (Foster & Rosenzweig, 2010).

In this study, I found that land ownership affects the adoption of IBB. A number of empirical and descriptive studies have also considered the effects of land tenure arrangements (which is often considered to be a good proxy for wealth) and the proportion of farms rented on the adoption of new agricultural technology, such as improved, high-yielding varieties. Findings suggest that the form of land tenure (e.g., renters, sharecroppers, landowners) may affect the adoption decisions and diffusion rates.

About one third of IBB adopters received planting material from friends or relatives. As a proxy of household economic well-being and technical capacity, I used the wealth index and management index, respectively (see section on Multiple Correspondence Analysis in the Appendix i). Adopters were wealthier, more technical in their crop management practices, and experienced higher yields.

Households located in the Northern Province on average had the highest management index, followed by the Western and Southern provinces. Management practices refer to the methods bean farmers use to increase productivity. Households in the city of Kigali were on average wealthier than farmers from other regions. The second-wealthiest rural households were located on the Western region, followed by households in the Northern region. Less wealthy households were in the Southern and Eastern regions.

I included regional geographical variables to control for the disparities in quality of road infrastructure and accessibility to extension services. Travel time to extension services represents a geographic accessibility measure. Access is defined as the time needed to travel from a specific household to the nearest location of interest. Good transportation is associated with diffusion of technology, better access to inputs, and lower transportation costs.

For model comparison, I estimated the log-likelihood as a measure of fit adjusted for model complexity. I also reported two information criteria, the Bayes (Schwarz) information criterion (BIC) and the Akaike information criterion (AIC) measures. To compare multilevel models, I estimated measures of complexity and fit such as model's deviance information criterion (DIC). Smaller values of the DIC indicate a better trade-off between complexity and fit of the model. The Watanabe-Akaike information criterion (WAIC), also known as widely applicable Bayesian information criterion, is similar to the DIC but the effective number of parameters is computed in a different way. See Watanabe (2013) and Gelman, Hwang, & Vehtari (2014) for details.

Table 1 Characteristics of adopters and non-adopters of IBB in Rwanda

Variables	Non-adopters	HIB adopters	p-level*
Household characteristics			
Number of women 12 - 49 years old	2.03	2.22	0.00
Number of individuals 0-19 <= >= 65 years old	2.77	2.92	0.12
Number of people per household	4.80	5.14	0.01
Dependency ratio (children)	1.39	1.43	0.51
Number of individuals per household - Economic active population [18-65]	2.68	2.93	0.00
Female household head (proportion of households)	0.27	0.26	0.66
Number of male members per households	2.22	2.54	0.00
Age of household head (years)	46.77	46.83	0.94
Level of education (average number of years in education per household)	2.82	3.46	0.00
Wealth Index	0.42	0.47	0.00
Years of farming experience	8.10	7.13	0.07
Farm characteristics and management practices			
Number of crops	1.78	1.87	0.08
Number of plots	2.97	3.34	0.00
Number of varieties	2.44	4.34	0.00
Percentage rented in land	13.97	11.06	0.08
Percentage own title	70.24	73.45	0.18
Percentage no title	13.17	14.07	0.63
Percentage share cropping	1.81	0.86	0.08
Total farmland (m^2)	2369.91	3092.79	0.00
Management index	0.39	0.45	0.00
Weighted plot slope (percent)	12.87	12.25	0.19
Land labor ratio (m^2/person)	998.82	1153.23	0.08
Time to plot (minutes)	15.45	15.36	0.94
Land terraced (proportion of households)	0.22	0.26	0.13
Plot irrigated (proportion of households)	0.06	0.09	0.08
Hired labor (proportion of households)	0.35	0.49	0.00
Applied fertilizers (proportion of households)	0.20	0.26	0.03
Applied manure (proportion of households)	0.77	0.86	0.00
Applied compost (proportion of households)	0.59	0.66	0.02
Applied pesticide (proportion of households)	0.09	0.10	0.56
Bean area m^2 (proportion of households)	1545.58	1927.93	0.00
Bean consumption (proportion of households)	0.06	0.09	0.00
Weighted average yield (kg/ha)	850.18	870.44	0.54
Access to credit (proportion of households)	0.21	0.20	0.54
Geography			
Kigali region (proportion of households)	0.02	0.02	0.81
Southern region (proportion of households)	0.27	0.28	0.80
Western region (proportion of households)	0.26	0.16	0.00
Northern (proportion of households)	0.21	0.20	0.54
Travel time (minutes) to cities equal or greater than 50,000 inhabitants	248.80	254.65	0.54
DEM (meters)	1734.27	1658.07	0.00
Drought index	-0.03	-0.03	0.70
Number of observations	962.00	432.00	

*We are testing that the mean difference is zero and is a difference t-test p-value.

II. Conceptual framework: multivariate matching algorithm

To estimate the impact of IBB on smallholder farming households' yields and incomes in Rwanda, I used a multivariate matching algorithm that implements Rubin's causal model. The aim of using Rubin's causal model is to test the null hypothesis of absolutely no effect on IBB adoption H_0: $Y_i(0) - Y_i(1)$ for any bean farming household. Basically, the model seeks to measure the impact of the Rwanda IBB delivery efforts on IBB growers' yield and income outcomes. The three key components of Rubin's model are: (1) subjects – in the case here, beans farming households, (2) treatment (Z) – in the case here, a binary treatment in which Z_i equals 1 if bean farming household h receives IBB and 0 otherwise, and (3) potential outcomes – in the case here, yields, Y_h, or incomes, Z_h, for each bean farmer. Our identification strategy estimates the average treatment on the treated (ATT). ATT is defined as E (Y_{h1} - Y_{h0}|X, Z=1) = $E(Y_{h1}|X, Z = 1) - E(Y_{h0}|X, Z = 1)$, where Y_{h0} is bean yield of regular bean varieties, Y_{h1} is yield of IBB varieties (Z=1), and X is a vector of observed covariates. I use 14 covariates, X, related to household, farm, and regional characteristics. Matching methods like PS find a substitute for $E(Y_{h0}|X, Z = 1)$ based on the statistical independence of (Y_{h0}, Y_{h1}) and Z conditional on *X*, i.e., technology is exogenous after conditioning on *X*. This condition is also referred to as "selection on observables".

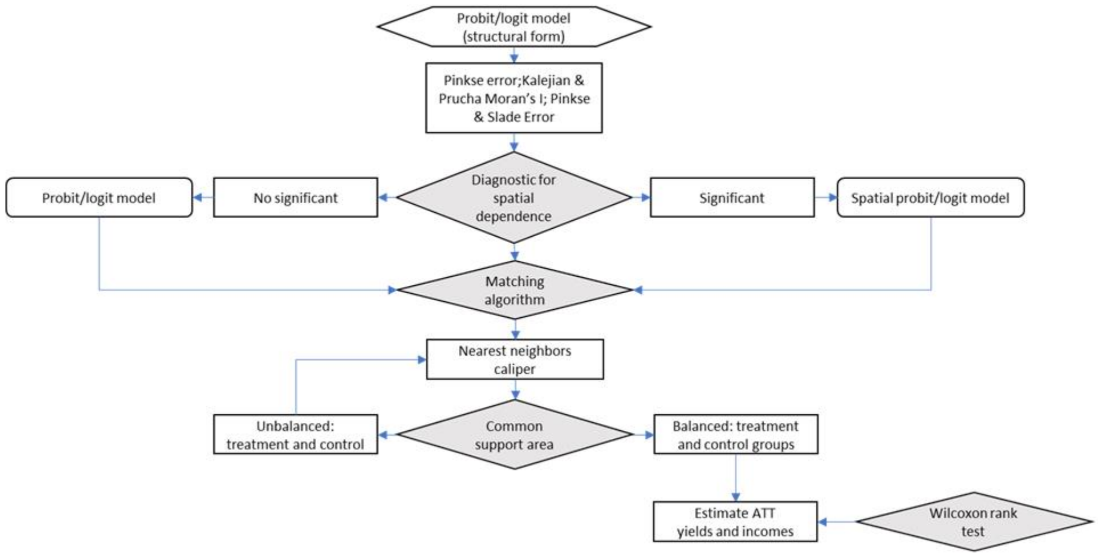
i. Implementation of causal model – multivariate matching algorithm

I use multivariate matching algorithms to implement Rubin's causal model. Matching creates the conditions of a natural experiment in which IBB (bush or climbing) are randomly assigned to two comparable groups: a control group (non-growers) and a treatment group (growers). Existing literature on matching algorithms proposes 5 general steps to estimate the average treatment effect (Caliendo & Kopeinig, 2008; Rosenbaum, 2010). I add an extra step to the estimation, adding statistical tests to check for spatial spillovers that relates to the stable unit treatment value assumption of independence between units. In more detail, these six steps help to identify treatment and control groups by:

1) Defining the structural form and variables selection,
2) Testing for spatial autocorrelation or spatial spillovers,
3) Selecting and defining matching algorithms,
4) Checking overlap/common support areas,
5) Assessing matching qualities, and
6) Running sensitivity tests (Figure 1).

The theoretical properties of the implicit model have been extensively explained elsewhere, so the full development and implementation need not be repeated here.

Figure 1 Implementation steps to estimate average treatment effect



The theoretical motivation for a spatial lag model is based on the literature on peer effects and social interaction. A few studies combine impact evaluation methodology and spatial econometric techniques to control for bias selection and spatial autocorrelation. For example, Chagas, Toneto, & Azzoni, (2012) uses a spatial propensity score-matching algorithm to estimate the effect of cultivating sugarcane on the human development index of sugarcane growing regions. Gonzales, Aranda, & Mendizabal, (2018) used Bayesian spatial-propensity score matching to evaluate the regional effects of microfinance for poverty reduction and women-empowerment at the municipality level in Bolivia.

First, I define the model structural form and pre-treatment covariates. I add a spatial structure to the basic probit model through a distance-neighboring matrix to test for spatial dependence and spillover effects. In contrast to the standard probit/logit model, the spatial probit model y_h^* represents latent unobservable utility that depends not only on observables determinants of household *h*, represented by *X*, but also on other household neighbors' latent variables y_{hj}^*. LeSage (2009) introduces the spatial probit model as the spatial autoregressive model (SAR model), or spatial lag model.

The data generating process for y_h^* is:

$$y_h^* = (I_n - \rho W)^{-1} X\beta + (I_n - \rho W)^{-1} \varepsilon, \varepsilon \sim N(0, I_n) \quad (21)$$

The spatial lag of the latent dependent variable Wy involves the *nxn* spatial weight matrix W(n x n) that contains elements consisting of either 1 or 0. A full description of each parameter is included in Eq 14 and Eq 15. The spatial weight matrix captures the dependence structure between neighboring observations, such as farmers or nearby locations. Wy is a linear combination of neighboring observations. The scalar ε is the dependence parameter and will be assumed to be abs(ε) <1. The k+1 model parameter to be estimated are the parameter vector β and the scalar ρ. I assume that ε follows a multivariate normal distribution, with zero mean and a constant scalar diagonal variance-covariance matrix σ $2I_n$. The fitted values from this model are the estimates of the propensity score. The fitted probabilities from the spatial probit model are extracted and added as an explanatory variable to the matching algorithm. For the estimated probabilities, consistency is necessary, so the first stage needs to be correctly specified and the standard errors need to be adjusted. Therefore, I include the propensity score as a covariate, as it will tend to balance all the observed covariates (Rosenbaum, 2005).

In choosing explanatory variables for the multivariate matching model, some authors suggest including not only statistically significant variables, but also variables known to be theoretically associated with selection and outcome variables. Economic theory and knowledge about the institutional setting should guide the selection of variables (Austin, 2011; Caliendo & Kopeinig, 2008).

I use the household as the unit of analysis. The household unit associates coping strategies, food availability, labor pooling, wealth level, and farm management practices to the outcomes of interest. In addition, using the household as the unit of analysis allows us to explore the link between household and plot management activities, and between household and community activities. For pairing comparable treated and control households, I used 14 covariates for each household. These observed covariates are grouped into three levels: household, farm, and region.

For the household level variables, I use average household education level, number of children under five years of age, age of household head, number of years of farming experience of the household head, the number of economic active males in the household, and household wealth. The average household education level aims to capture human capital. Numerous studies find a significant relationship between education indicators and farm productivity (Weir & Knight, 2000; Foster & Rosenzweig, 2010).

Risk aversion and uncertainty about the benefits of an innovation entails subjective risk, i.e., yield, and objective risks, i.e., pest and weather (Feder et al., 1985). The propensity of farmers to manage risk with the adoption of a new technology is associated with wealth and bigger farms (A. D. Foster & Rosenzweig, 2010). Therefore, as another household level variable, I construct a household wealth index as a proxy for risk aversion. The wealth index includes three groups of assets: household assets, livestock, and other agricultural assets. The number of economically active members in the households was included to explore labor-intensive practices associated with bean production. Bush bean production can be more labor intensive than climbing bean production during the harvesting season. Farming experience is useful in the early stages of adoption, when farmers are still testing the potential agricultural and nutritional benefits of a new technology.

Farm characteristics include management index, plot slope, land labor ratio, and number of bean varieties cultivated. To measure the technical capacity of households and reduce multidimensionality, the management index aims to summarize farmers' management practices, whether farmers have irrigated plots, or used a variety of inputs, such as fertilizers, manure, compost, and or pesticides. Other farmer practices that affect farm output include the number of bean varieties grown and the slope of the plots. Land to labor ratio is a proxy for labor availability. Labor is another factor that influences farmers' decisions to adopt a new technology. A negative correlation between adoption and land-to-labor ratio would support the hypothesis of adoption intensification as population density grows. I observe the positive direct effect associated with the number of bean varieties cultivated. Due to uncertainty in household bean yield in securing food sufficiency in a household and as a coping strategy to secure food security, households manage the risk of crop failure by cultivating multiple bean varieties. The number of varieties could also be because they don't have access to sufficient seed of a single preferred variety.

I include regional geographical variables to control for disparities in road infrastructure and extension services. Travel time to extension services represents a geographic accessibility measure. Access is defined as the time needed to travel from a specific household to the nearest location of interest. Good transportation is associated with diffusion of technology, better access to inputs, and lower transportation costs.

The third step focuses on selecting the proper matching algorithms. To do so, I tested different specifications to reduce either bias or/and variance. I present the results of the most suitable algorithm for our analyses, the nearest neighbor algorithm. This algorithm uses a distance matrix (Euclidean distance) (King & Nielsen, 2019) with calipers which reduces the risk of unfavorable matches if the closest neighbor is far away. A caliper is a non-negative number that measures the degree of similarity of two households in term of their covariates X. A distance matrix is defined as a table with one row for each treated subject and one column for each potential control. For instance, two households with the same value of X would have distance zero. With a caliper of width *w*, if two individuals, say *k* (treatment) and *l* (control), have PS that differ by more than *w*—that is, $|(X_k) - (X_l)|$ > w— then the distance between these individuals is set to ∞, whereas, if $|(X_k) - (X_l)|$ < w, the distance is a measure of proximity of X_k and X_l The caliper width, *w*, is often a multiple of the standard deviation of the PS, (X). A widely use caliper width of 20% of the standard deviation of the PS is a good start (for details, see Appendix).

The raw sample is pre-processed by “trimming” (removing) individual households iteratively, checking for balance on all covariates. The matching algorithm weights on each covariate and the propensity score. As suggested by King & Nielsen (2019), the matching algorithm performs better than the PS alone. Observations that were outside the common support area from either control or treatment groups were discarded. For each value of PS, the common support represents a positive probability of being both treated and control. Estimation of the ATT was based on the region of common support. I implemented the minima and maxima PS criteria. Propensity score values outside this range were removed for the estimation of the ATT. Once the region of common support is defined, households that fall outside this region must be disregarded as their treatment effect cannot be estimated. This exclusion of households may seem as if it would substantially reduce the statistical power of the study to detect effects. However, this is not always the case because of improved covariate balance (see Appendix).

I evaluate each algorithm based on its performance in balancing the measured covariates between treatment and control groups (see Appendix). Inferences about the treatment effect are robust if the matched sample of treated and control bean farmers have similar distributions of measured baseline covariates. The pooled standard deviation for covariate k is $\sqrt{(S_{tk}^2 + S_{ck}^2/2}$. The absolute standardized difference before matching is $sd_{bk} = |\hat{x}_{tk} - \hat{x}_{ck}| / \sqrt{(S_{tk}^2 + S_{ck}^2/2}$ (Rosenbaum, 2010). The absolute standardized difference after matching is $sd_{bk} = |\hat{x}_{tk} - \hat{x}_{cmk}| / \sqrt{(S_{tk}^2 + S_{ck}^2/2}$; where $\hat{x}_{tk}$ and $\hat{x}_{cmk}$ are the means of covariate k for the treated, control, and matched groups, respectively. A covariate is balanced if the standardized bias or the absolute standardized difference mean (ASDM) on each covariate is less than 0.20, ideally below 0.15. The second check is to graphically observe whether the distributions of treated versus control units are even in a kernel density plot (see appendix).

I test the sensitivity of the average treatment effect (ATT) in two ways. First, I employed more than one matching algorithm to conduct the matching procedure. The second method adheres more to the standard definition of sensitivity analysis in an observational study. I ask how the conclusions of the study might change if people who looked comparable were somewhat different (Rosenbaum, 2010).

ii. Heterogenous treatment effect

In the case of considerable bias associated with treatment effect estimate and heterogeneous treatment effect, there are different ways to characterize the heterogeneity. Here, I implement the smoothing-differencing (SD) method. To do so, I fit two separate non-parametric regression models, local polynomial regression, for the outcome variables on the propensity score and covariates; one for each: the treatment group and the control group. The difference between the group-specific regressions gives an estimate of the heterogeneous treatment effects (Zhou & Xie, 2016).

III. Conceptual framework: productivity analysis

In this analysis of technological adoption, I employed economic indicators of performance such as measures of productivity and efficiency. Assessing farmers' efficiency together with the impact of an intervention requires the application of advanced econometric methods. To do so, I employed a two-steps approach. The first step estimates technical efficiency scores on every bean farmer. The second step uses censored statistical methods to identify determinants affecting bean farmers' efficiencies.

i. Technical efficiency analysis

I used stochastic frontier analysis (SFA) and stochastic spatial frontier analysis (SSFA) to estimate bean farmers technical efficiency. In addition, I used an output-oriented DEA model to estimate scale efficiency. Other methods that estimate technical efficiency include DEA solved through linear programming. In broader terms, these frontier analysis techniques for measuring productivity can be frontier and non-frontier techniques modeled either through deterministic or stochastic (Del Gatto, Di Liberto, & Petraglia, 2011).

In contrast to DEA analysis, SFA requires a few more a priori assumptions about the structure of the production function (Greene, 2008). SFA deviations from the frontier is attributed to two factors: a normal error representing randomness and a non-negative error term representing technical inefficiency, the sum of both constitutes the total error.

$$y_i = f(x_i, \beta) + \varepsilon_i \tag{22}$$

$$\varepsilon_i = v_i - u\varepsilon_i \tag{23}$$

where $v_i \sim N(0, \sigma_v^2)$ and $u_i \sim N(0, \sigma_u^2)$

Combining equation 22 and 23,

$$y_i = f(x_i, \beta) + v_i - u_i \tag{24}$$

where *i* indexes cross-section of bean growing households. y_i denotes bean production of household *i*, whereas X_i is a vector (1xK) of N inputs used by household *i*. β is the vector (Kx1) of technology parameters to be estimated, and ε_i is a *i.i.d* disturbance for household *i* with zero mean and variance σ_ε^2 assumes to follow a normal distribution. This term takes care of the stochastic nature of the production process and possible measurement errors of the inputs and output. The composed error (23) consists of a normal error component v_i and a with a non-negative random variable u_i, which represents the technical efficiency term. By assuming that the productivity component follows a non-negative distribution, I am able to estimate the best practice production function rather than the average practice production function.

However, it is to be noted that model (24) does not include any type of spatial dependence between the observations a potentially restrictive specification in empirical applications. Three types of spatial interaction effects can be given on the non-spatial production function (24) (Han et al., 2016a). The first is endogenous effects which explain the dependence between the dependent variable, y_i andy_j. The second is exogenous interaction effects, which explain the dependence between the dependent variable of a specific unit, y_i, and the independent variable of another unit, X. Third, interaction effects among the error terms equation (26). A full model with all types of spatial interaction effects are specified in equation (25).

$$y_i = \rho \sum_{j=1}^{N} W_{ij}\, y_i + \beta_0 + X_i\beta_1 + \gamma \sum_{j=1}^{N} W_{ij}\, x_j + \varepsilon_i \quad (25)$$

$$\varepsilon_i = \lambda \sum_{j=1}^{N} W_{ij}\, \mathcal{E}_j + u_i \quad (26)$$

where, y is a Nx1 vector of observations on the dependent variable, W is an exogenous N x N spatial weight matrix with non-negative elements, ρ is the spatial autoregressive parameter. In this specification, the inclusion of the spatially lagged dependent variable Wy_i on the right-hand side of the equation relates the value of the dependent variable to the values at neighboring locations. X_i is a N x K matrix of observations on explanatory variables with associated K x1 coefficient vector β_i. X_i the spatial lags of the covariates (independent variables) with coefficients γ. ε is a

Nx1 vector of error terms as *i.i.d* disturbance for household h_i with zero mean and variance σ_ε^2. Restriction γ = 0 and λ = 0 give rise to the spatial autoregressive model (SAR) and restriction to ρ = 0 and γ = 0 give rise to the error spatial autoregressive model.

I implemented the first and third methods following the spatial specifications search suggested by Anselin & Rey, (2014). I started with a basic ordinary least squares (OLS) model and ran the Lagrange Multiplier (LM) statistics to decide for either the lag or error specifications. If no spatial autocorrelation evidence is found from the LM-error and the LM-Lag tests, I report the OLS model.

For the efficiency analysis, I transformed the type production function in equation (4) to a frontier model by introducing a non-negative random variable u_i which represents the technical inefficiency of unit *i*. ε_i is divided into two parts: u_i, a non-negative random variable associated with technical inefficiency, and v_i a systematic random error, equation (2). Because I am unable to identify the term -u, I use a relative efficiency measure that accounts for the output of each unit to the output that could be produced by a fully efficiency unit as suggested in (Han et al., 2016a).

The empirical model includes five explanatory variables. The inputs of the production function include total cultivated land area (in square meters), economically active population (number of adults between 15-64 years old), hired labor (whether the household hired labor) as a substitute for household labor, total amount of bean seed (kilograms), and a management index (an index derived from multiple correspondence analysis as a surrogate of technical capacity). The output is measured as the total household bean production (kg). The estimates of technical efficiency are obtained by comparing the input–output bundle of each farm household with the nonparametric and parametric representation of the frontier technology. Reducing the number of variables in the production function would increase the number of inefficient households but would bias efficiency estimates. I shift more importance to the second stage-censored analysis. In all models, I used the multiplicative form and conduct the estimation in log-linear form. The first order coefficients can be interpreted as partial elasticities.

ii. Second stage analysis – factors affecting bean farmers' technical efficiency

Estimated technical efficiency serves as a dependent variable on the post-efficiency analysis. The second stage analysis has two purposes: (1) to explain the variation of relative efficiency and (2) to validate the empirical model from the first stage. A series of control variables are tested using truncated regression. The motivation of this section is to have a better understanding of why some bean farmers are more efficient than others. I explore whether farmers' efficiency is affected by droughts, physical proximity to technical services, crop diversification, or bean growing households' link to local markets to meet the demand for staple food crops. This last variable indicates whether the smallholder farming household sold their crop surpluses to inhabitants residing in the corresponding local village.

Truncated regression has shown robust results. A truncated regression is a distribution that occurs when some values above or below of the variable of interest are omitted. Simar & Wilson (2007) used Monte Carlo experiments to examine the statistical performance of two estimators, namely Tobit and truncated regression, when employing a two-stage approach for non-parametric distance function estimators of technical efficiency. Their experimental results revealed that Tobit regression showed unstable results, whereas the truncated regression was more stable.

iii. Clustering Analysis

Clustering analysis can produce evidence on where this technology has been effective, thereby providing valuable input into targeting strategies and resource allocation for scaling up of such interventions. To identify clusters of farmers with similar and dissimilar technical efficiency scores, I used Local Moran's I statistics. This test identifies five data groups. The first cluster of hotspots is characterized by farmers with high efficiency surrounded by farmers with similar efficiency scores. The opposite of hot spots are cold spots, characterized by bean farmers with low efficiency surrounded by bean farmers with similar low efficiency. The other two data groups are spatial outliers. One set refers to bean farmers with low efficiency scores surrounded by farmers that are more efficient while the second set of outliers reflect the opposite. The firth group are observation without any particular spatial pattern.

Chapter 4: Social Interaction and Technology Adoption: Geographic Diffusion of Iron Biofortified Beans (*Phaseolus Vulgaris*) in Rwanda

I. Introduction

IBB is a relatively new technology. Farmers may be risk-averse when they lack information pertaining to the likelihood of occurrence of possible outcomes (e.g. yield, costs, profitability) from use of the new technology. Such risk-averse attitudes would exert a detrimental impact on adoption. Farmers may be uncertain about the economic returns of the new technology owing to insufficient knowledge about the types and costs of inputs needed, the yield distribution, expected market prices and the demand for the produce (Abadi Ghadim & Pannell, 1999; Tessema, Asafu-Adjaye, Kassie, & Mallawaarachchi, 2016). In this context, social learning and social networks often complement and/or act as substitutes in delivering information and facilitating the technology diffusion process. In the seminal work of Manski, (1993), he identified three sources of social influence in the adoption of technology: (1) endogenous effects, (2) exogenous network effects, and (3) correlated effects. The endogenous effect emphasizes that the adoption behavior of individual farmers would be influenced by their neighbors' adoption outcomes as a result of peer learning about the profitability or the appropriate use of the new technology or due to merely wanting to conform with observed peer behavior. The exogenous network effect highlights contextual interactions, wherein the propensity of an IBB grower to behave is correlated with exogenous characteristics of his/her neighbors. The correlated effect stresses that smallholder farmers in the same group tend to behave similarly because of commonly observed and unobserved characteristics of the group, for example sharing a common institutional or physical environment (Tessema et al., 2016). All these three effects imply a spatial contextualization of the diffusion of IBB varieties, meaning that the decision of a bean growing household to adopt an IBB is spatially correlated. In more detail, farmers' decisions to adopt IBB depend not only on their own farm-level characteristics, but also on the decisions and personal characteristics of neighboring bean farmers. Information about the benefits of growing and consuming IBB varieties is also a factor that affects technology adoption (Abdulai et al., 2008; A. D. Foster & Rosenzweig, 2010). Under this spatial context, a bean growing household that is close to a household who is an IBB grower has a higher probability of being an IBB adopter, which is the endogenous effect. Another condition relates to the social characteristics of a group as the main factor in spatial clustering, which is the likelihood of an individual to behave, on average, in agreement with their social group. Whether or not the diffusion of IBB varieties is geographically driven, the spillover effects will determine a strong spatial relationship, i.e., farms with similar IBB adaptation behavior tend to be localized in the same geographical area. These sources of social influence have differing implications for prediction of policy impacts. Common spatial econometric methods applied to technology adoption include the spatial error model and the spatial lag model. I extend these methods by modeling the endogenous and exogenous network effects using the spatial Durbin model, which allows for an enhanced understanding of IBB adoption as it relates to the characteristics of neighboring smallholder farmers.

Shaped by local place and constrained by social geographic distance, I model social interactions by setting geographic neighbors' relationships. By social interactions, I refer to interdependence among smallholder farmers in which preferences, tacit knowledge, expectations, and constraints faced by one smallholder farming household are directly influenced by the characteristics and choices of others. I am interested in the importance of tacit knowledge which concerns direct experience and know-how on the use of IBB agricultural technology. In spatial regression analysis, measures of spatial interaction include the spatial autoregressive parameter through different spatial weight structures. The spatial autoregressive parameter represents a way to model structured dependence between observations that arise from peer effects (Case, 1992). The spatial autoregression parameter in technology adoption studies contains important policy information. Mapping interactions of farmers' IBB adaptation behavior can provide guidance to technology delivery programs on how specific initial investments in technology promotion will generate further geographic diffusion.

This chapter analyzes smallholder farming households' adoption of these varieties by specifically examining the influence of demand-side factors and the role of peers. I draw upon several theories from studies on the adoption of agricultural technology, social behavior, and spatial econometric methods to build the models. I test for the presence of spatial association among economic agents (farmers), estimate prevalence rates of IBB adoption by district, and examine any potential interactions with contextual factors. I implement spatial probit models for discrete-choice data using Bayesian modeling. The use of Bayesian modeling to estimate spatial processes allows estimating more realistic models. These methods produce useful measures of direct and spatial spillover impacts from changes in the explanatory variables. By doing all the above, this chapter achieves objective one of this dissertation, which is to analyze smallholder farming households' decisions to adopt iron biofortified beans in Rwanda using theory on social interactions and choice behavior.

II. Data

For this chapter, as well as chapters 4 and 5 of this dissertation, I use nationally representative survey data of rural bean producers in Rwanda that was collected in two stages. In the first stage, as part of a listing exercise, 120 villages were randomly selected. All rural households in the selected villages were interviewed. Rural households were shown a seed sample of ten iron-biofortified bean variety and asked whether they had heard of the variety, had grown it, and, if so, the season they first adopted it and whether they had grown the variety in each subsequent season. The listing exercise was conducted in May and June 2015 (i.e., season 2015B) and included 19,575 households (Asare-Marfo et al., 2016).

In the second stage, 12 households per village were re-interviewed in greater depth in September-October 2015, after the harvest of the same season and included 1,394 households. When possible, six smallholders who grew an IBB in 2015B and six non-adopters were selected randomly in each village. In villages with fewer than six IBB adopters, all adopters were selected and non-adopters were randomly selected to obtain a total of 12 households. Enumerators interviewed the member of the smallholder farming household responsible for bean production decision making during season 2015B. Questions were asked about household demographics and composition, bean farming decision making, asset ownership, bean production and consumption, and iron-biofortified bean adoption history from 2012B to 2015B.

A community survey of 120 villages, conducted along with the main household survey, was administered to key informants, including the village leader, to gather information on village characteristics, services and amenities related to market access, extension, and the presence of formal iron-biofortified bean delivery approaches in the village.

III. Background

Annual food crop production in Rwanda follows the bimodal annual distribution of rainfall that divides crop cultivation into two major seasons: A and B. Season A sowing dates are in September and October, followed by harvest activities from January to March. Season B sowing dates span from January to March, followed by harvest activities in June and July. I analyze adoption of IBB for season B in 2015. At a national level, I estimated 214,130 hectares as having been cultivated with beans in that season, with 58 percent planted with bush and 42 percent with climbing varieties.

The results of the 2015 survey show that local bean varieties still dominate bean production in Rwanda, with 68 percent of cultivated area, followed by improved and IBB varieties, with 21 percent and 11 percent, respectively. The latter figure indicates the intensity of adoption of IBB varieties following only 8 seasons of IBB delivery efforts. IBB bush varieties show a higher intensity of adoption than IBB climbing varieties, at 12 and 10 percent adoption rates, respectively. Of the total population of bean growing households in 2015 Season B, 24 percent had grown IBB varieties.

Four-fifths of the area under bean cultivation comprised plots cultivated with either bush or climbing bean varieties, with the other 20 percent planted with a mixture of both bush and climbing varieties. Plots with only local varieties dominated with 55 percent, composed of bush (33 percent) and climbing (22 percent). The second-largest share was allocated to plots with only improved varieties, at about 16 percent, of which 8 percent were bush varieties, and 8 percent climbing. About 8 percent were plots cultivated with either IBB bush or IBB climbing varieties, at 5 and 3 percent, respectively.

Two out of the ten IBB varieties disseminated throughout Rwanda are IBB bush varieties: RWR2245 and RWR2154. The former had an intensity of adoption of 11 percent of the total land area cultivated with bush bean varieties (157,416 hectares), while the latter represented less than 1 percent. The remaining eight varieties are climbing beans with an intensity of adoption of 10 percent of the total land area cultivated with climbing beans (114,568 hectares). MAC42 was the most popular climbing IBB variety, with almost 5 percent, followed by RWV3316 (~2 percent) and RWV1129 (1 percent). The other five climbing IBB varieties were cultivated in less than 1 percent of the area under bean cultivation.

Beyond bean varieties, I also looked at land ownership and bean consumption. Seventy-one percent of the cultivated land area was owned under a title, 14 percent had no title, and 12 percent was rented land. In descending order, bean utilization breaks down as grain used for home consumption (60 percent), sale in local markets (17 percent), crop given away (6 percent), and saved to be used as seed for next season (6 percent).

At the subnational level, bean production varies by region and by the two bean types, bush and climbing. The geographic distribution of bean production and IBB adoption at district level is shown in Figure 1. The map shows the percentage of total bean area per district normalized by the total national bean area, while the pie charts show the percentage of IBB area cultivated in the district by IBB type. Bean production, about 40%, is concentrated in the Eastern region, gradually decreasing in intensity toward the Central, Southern, and Western regions. Second, the intensity of adoption of IBB varieties mirrors the natural geographic concentration of bush and climbing bean varieties, measured by cultivated area. IBB bush varieties are mostly grown in the Eastern (61 percent) and Southern (32 percent) regions, while IBB climbing varieties dominate in the Western and Northern regions. The Central and Southern regions is a mix of both types, in accordance with the literature. These figures may be partially explained by delivery and marketing efforts to introduce IBB, as illustrated in Figure 3. The high intensity of IBB adoption in the eastern region has been propelled in part by the high density of delivery systems that enhance access to IBB seed to smallholder bean farmers. For practical purposes of this analysis, I re-scaled the density map to squares of 10 km x 10 km as a result of a kernel density function. Therefore, the density values are reported as the number of points or delivery venues per 10 square kilometers. We can observe a high density of delivery mechanisms on the Eastern region with lower density values over the Northern region.

Figure 2 Choropleth map: share of total bean production and pie chart of IBB production by bean type by area, by district

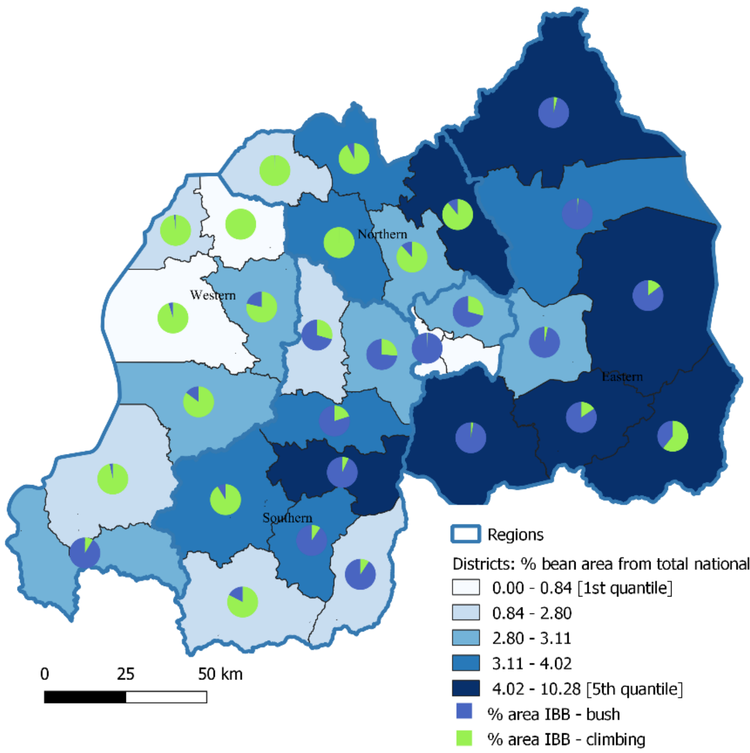
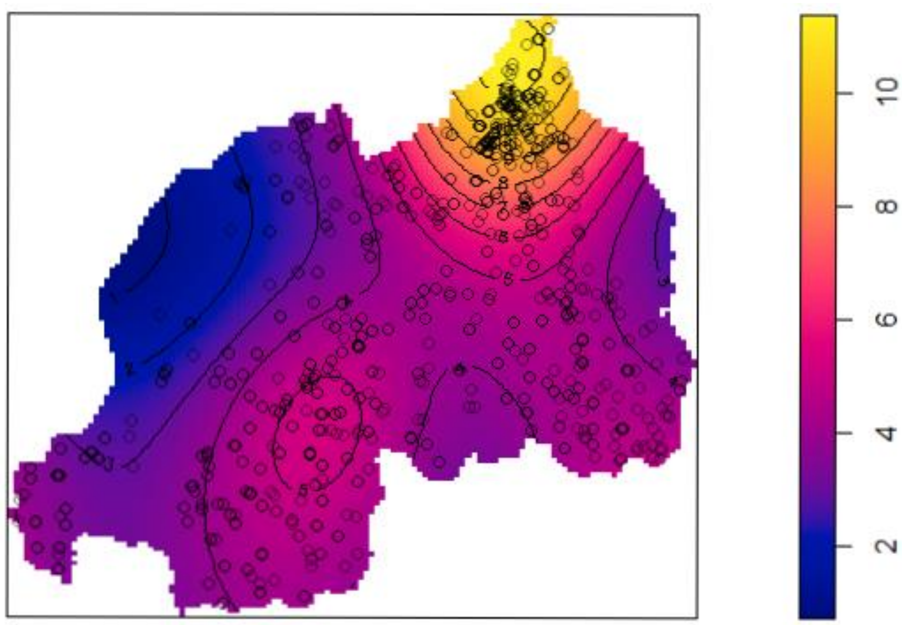


Figure 3 Spatial density of seed delivery



Note: the density values are reported as the number of points or delivery venues per 10 square kilometers.

IV. Results and Discussion

i. Spatial Econometric Analysis: Adoption Model Estimates (non-spatial probit [NSP] vs. spatial probit [SP] models)

The coefficient estimates (posterior means, standard deviations, and Bayesian p-levels) of the two specifications (M1 and M2) for two spatial models (SAR and SDM) and a non-spatial probit (NSP) model are shown in Tables 2 and 7, while Tables 3-6 and 8-11 show the average marginal effects estimates. Tables 4-6 and 8-11 are the basis for inference regarding the effect of changes in the various independent variables on the probabilities that bean farmers will adopt IBB and on the spatial spillover effect on neighboring bean farmers. I also tested the robustness of our results. Table A.1 in the Appendix shows posterior means (standard deviations) of a multilevel spatially structured fixed and random effects model. I described specifications M1 and M2 in chapter 3.

For each scenario, I use a standard generalized linear model (GLM) probit model and two spatial probit models. I describe and compare the average marginal effects for each model. There are three common covariates in both specifications: number of children in the household, age of household head, and accessibility to extension services. The specification M1 aims to test how household characteristics, such as wealth (based on an asset-based wealth index as explained in Appendix A, section I), household composition, and years of farming experience, play a role in IBB adoption. In addition, the M1 scenario explores the role of the number of varieties used to manage the risk of food insecurity due to crop failure caused by drought. The specification M2, on the other hand, looks at the importance of household management technical capacity measured through a management index in connection with education level and household size (see Appendix A, section I). M2 does not include the wealth index because of its positive correlation with the management index and education level.

For all non-spatial probit models, I computed and reported a diagnostics test (Kelejian-Prucha (error)) for spatial dependence. The results of the test for all probit models were positive and significant; therefore, spatial probit models are used to calculate the probability, P(X)=Pr(D = 1 | X), or propensity, of being an IBB grower for each observation. I report the marginal direct and indirect effects just for the spatial probit models. Models are compared using log-likelihood and information criteria, such as AIC (Akaike Information Criterion) and BIC (Schwartz' Bayesian Information Criterion). For comparing models using the log-likelihood value, models with log-likelihood values closer to zero are considered better models. While for comparing models using the information criteria, models with smaller values of these criteria are considered better models. The Chow test provides evidence that the coefficient are not equal across regimes rejecting the null hypothesis. The value of the test statistics is 3.65 which is significant. There is evidence that the model coefficients indeed are not constant across the two spatial regimes, indicating spatial heterogeneity and suggesting the need to carry out separate analysis for IBB bush type and IBB climbing type.

ii. Bush bean analyses

Table 2 shows the results of M1 and M2 on bush bean varieties. I observe that the signs of some covariates are consistent in the spatial probit models and non-spatial probit model. Years of farming experience, planting material from friends and relatives, numbers of varieties cultivated, management index, and the number of economic active male members in the household have a positive influence on the propensity of adoption of IBB bush varieties, while the numbers of children in the household have a negative influence. In the following, I discuss the significant level, magnitude and sign of the average marginal effect of each variable in greater detail with the summary measures of direct, indirect, and total impacts.

Tables 3–6 show the significant level and marginal effect outputs for the non-spatial and spatial probit models for bush bean growers. The first column lists all variables used in each model specification. Columns with the headings direct, indirect effect and total effect show the posterior means and their respective lower (5 percent) and upper (95 percent) bounds confident intervals for the SAR spatial probit model. The last column corresponds to the marginal effect of the standard non-spatial probit model, which is equivalent to the direct effect of the SAR models.

Spatial probit models, such as SAR and SDM, allowed us to disentangle the total marginal effect into direct and indirect impacts. The direct effect measures how a change in an explanatory variable in household h_i affects the dependent variable in household h_i, plus any feedback effects. The indirect effects measure how changes in the explanatory variables associated with household h_i cumulatively impact the dependent variable in all other households. These effects are referred to as spatial spillovers. The statistically significant spatial autocorrelation coefficient ρ on the lagged dependent variable in the spatial Durbin model suggests spatial autocorrelation, which means that the SDM specification helps correct biases of estimated coefficients and improve the efficiency of the estimations.

In general and comparing the absolute values of the coefficients, the SAR model's indirect effects are smaller than the direct effects. The SAR model's direct impacts were statistically indifferent from the direct effects of the non-spatial probit model's direct impacts in terms of both sign and magnitude. Below, I provide a discussion of the average marginal effect for the SAR models.

Table 3 presents own partial derivatives (direct effect), $Z_i/X_v i$ and cross-partials derivatives (indirect effect), $Z_i/X_v j$, or spatial spillover effects in the case of spatial dependence. Of the household characteristics evaluated at the sample means in Table 3, farming experience had a positive and significant effect of 1 percent and a spatial spillover effect of about 0.3 percent for every additional year of farming experience, resulting in a total effect of 1 percent, ceteris paribus. Indirect effects are accumulated across all neighboring farmers, so the impact on individual farmers is likely smaller than the direct effects. Farming experience is useful in early stages of adoption, when farmers are still testing the potential agricultural benefits of IBB. Having economically active male household members had a positive effect, with one additional male member increasing the adoption rate by 4 percent, and a positive indirect effect of 2 percent. Households are averse to adopting new varieties given the risk and uncertainty of their future performance. As a coping strategy to minimize the chances of food insecurity, households manage the risk of crop failure by cultivating multiple varieties. Growing of an additional variety increases the probability of IBB adoption by 7 percent and a spatial spillover of 3 percent.

Specification M2 aims to check the consistency of the estimated casual effect (Tables 2 and 5). I use four new covariates: household access to title land, household size, education level, and management index or technical capacity. The management index and education level variables were precluded in specification M1 because of collinearity with the wealth index variable. The scalar parameter ρ in Table 2 measures the strength of spatial dependence of the IBB adoption propensity, with a value of zero indicating independence. The estimated ρ in each M2 model is statistically significant at 1 percent level and its magnitude varies from 0.38 to 0.42, suggesting significant positive spatial autocorrelation in bean farmers' decisions regarding adoption of new iron biofortified bean varieties. In other words, there is a global social multiplier in the system that indirectly affect non-beneficiaries. Of the 7 variables in Table 5, the effect of household management practices is statistically significant and positive. When the management index increases from the lowest value zero to the largest value 1, the probability of the adoption will increase by 23 percent, with a spillover effect of around 17 percent, and a total effect of 40 percent. As with any new technology, IBB varieties would be more frequently used by farmers who already use other agricultural inputs, such as fertilizers or manure, to increase farm productivity. Household education level, household size, and other variables were not significant.

Tables 4 and 6 summarize the SDM's marginal effects of bush bean growers. Of the nine explanatory variables included in specification M1 (table 4), six are statistically significant at the 5 percent level for the estimates of the direct effect, which are "number of economic active males in HH", "farming experience", "number of bean varieties cultivated", "IBB planting material from friends or relatives", and "travel time to extension services". For specification M2 (table 6), out of the seven included covariates, two variables—management index and access to extension services—are statistically significant at the 5 percent level for both the direct and the estimates of the indirect effects. The latter effect confirms the presence of spatial spillovers or peer effects. Proximity to extension centers has a positive direct effect of 4 percent and a spatial spillover effect of 2 percent, which seems consistent with the notion of accessibility of farmers to agricultural extension agents. Extension services can help guide farmers, particularly on the agricultural superiority of improved varieties such as iron biofortified ones, strengthening farmers' knowledge and experience in agricultural best practices. Poor-functioning infrastructure affects the profitability of technology adopted by farmers, and road networks (extension and quality) and mobile services rank among the most important infrastructural conditions. In general, transportation limitations tend to reduce competition among input suppliers and middlemen. Empirical evidence shows that travel times between the farm gate and market can be high due in part to underdeveloped road infrastructure. Good transportation is associated with a diffusion of technology, better access to inputs, and higher producer prices (Dorosh & Wang, 2010).

The management index also exerts a positive direct and indirect impact on the propensity of iron bean adoption, suggesting that I would observe increased adoption rates on bean farmers that already use other agricultural practices like irrigation, terracing, fertilizer, pesticides, manure, and/or compost. The indirect impact from management practices on nearby farmers is almost half the magnitude of the direct impact, suggesting a very significant spatial spillover impact in adoption rates, in comparison with other variables.

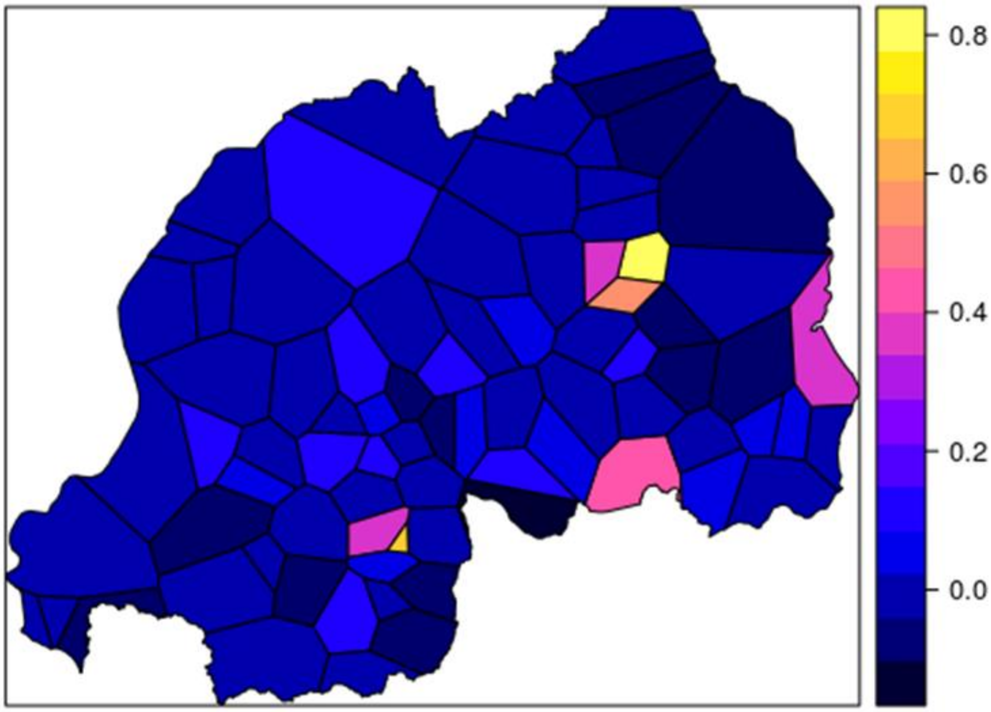
Figure 4 shows point estimates of village-level random effect. The values of the point estimates change from one village to its neighbors ranging from 0.1 to 0.8 with a higher prevalence of villages with negative point estimates. These villages have a lower probability to adopt IBB bush varieties. This analysis provides evidence that geographic distance significantly slows the spread of new agricultural technology. However, I can observe clusters of villages in the Eastern, Kigali, and Southern regions have positive point estimates which increase their probability to adopt IBB bush varieties. Random effects is a useful measure of the degree of spatial heterogeneity of the smallholder farming households in Rwanda.

Table 2 SAR, SDM, and GLM probit model estimate for bush bean farmers

Variables	M1-SAR	M1-SDM	M1-NSP	M2-SAR	M2-SDM	M2-NSP
Rho	0.309***	0.297***		0.423***	0.385**	
	(0.091)	(0.079)		(0.091)	(0.101)	
Constant	-5.180	-7.663	-3.998	-5.135	-4.404	-2.919
	(8.918)	(9.389)	(8.941)	(8.063)	(8.091)	(7.967)
Household size				0.005	0.001	0.011
				(0.038)	(0.038)	(0.037)
Number of economic active males in HH	0.171**	0.171***	0.161**			
	(0.063)	(0.063)	(0.062)			
Number of children under 5 years old in HH	-0.166	-0.167**	-0.166**	-0.142	-0.138	-0.165
	(0.109)	(0.099)	(0.103)	(0.101)	(0.101)	(0.100)
Log (HH head age)	2.646	3.986	1.935	2.679	2.769	1.365
	(4.769)	(5.062)	(4.81)	(4.314)	(4.315)	(4.263)
$(\text{Log(HH head age)})^2$	-0.479	-0.666	-0.377	-0.401	-0.412	
	(0.634)	(0.678)	(0.644)	(0.569)	(0.569)	
HH average years of schooling				0.024	0.026	0.015
				(0.039)	(0.041)	(0.039)
Wealth Index	0.408	0.377	0.310			
	(0.513)	(0.527)	(0.495)			
Number of varieties cultivated	0.309***	0.311***	0.293***			
	(0.042)	(0.042)	(0.041)			
Farming experience (years)	0.038***	0.039***	0.035***			
	(0.010)	(0.011)	(0.010)			
Share of land area with legal title				0.002	0.001	0.002
				(0.001)	(0.002)	(0.001)
Management Index				0.717**	0.818**	0.830**
				(0.298)	(0.293)	(0.285)
Land labor ratio (m^2/person)	-0.000	-0.000	-0.000			
	(0.000)	(0.000)	(0.000)			
Travel time to technical services	-0.008	0.094	-0.002	0.016	0.124	0.020
	(0.022)	(0.085)	(0.024)	(0.017)	(0.071)	(0.020)
Planting material from friends or relatives	10.1***	17.16***	26.11			
	(3.02)	(4.273)	(803.8)			
W-HH size					0.020	
					(0.070)	
W-Number of economic active males in HH		0.081				
		(0.135)				
W-Number of children under 5 years old in HH		-0.122			-0.329*	
		(0.187)			(0.187)	
W-Log (HH head age)		-0.008			-0.015	
		(0.013)			(0.009)	
W-HH average years of schooling					-0.093	
					(0.076)	
W-Wealth Index		0.671				
		(1.012)				
W-Number of varieties cultivated		0.108				
		(0.076)				
W-Farming experience (years)		0.011				
		(0.021)				
W-Share of land area with legal title					0.002	
					(0.002)	
W-Management Index					-0.194	
					(0.459)	
W-land labor ratio (m2/person)		-0.0002**				
		(0.0001)				
W-Travel time to technical services		-0.106			-0.106	
		(0.085)			(0.071)	
W-Planting mat. from friends or relatives		-1.844*				
		(1.002)				
Kelejian-Prucha (error)				3.987***		4.618***
Log likelihood	-205.480	-197.387	-289.821	-211.031	-295.813	-294.343
BIC	485.679	525.532	685.493	490.553	653.892	644.724
AIC	434.960	436.775	613.642	444.060	611.626	606.686
Number of observations	506	506	506	506	506	506

Notes: () are standardized errors. *, ** and *** denote significance at the 10%, 5%, and 1% level, respectively.

Figure 4 Point estimates of village-level random effect: IBB bush M1.IID



Note: Random effect values range from -1 to 1. Values around zero mean spatial randomness, values close to 1 suggests clustering and values close to -1 suggests a negative spatial association.

Table 3 SAR and GLM probit model effects estimates for bush bean growers (M1)

	Direct effect			Indirect effect			Total effect		
Variables	Lower 0.05	Coefficient	Upper 0.95	Lower 0.05	Coefficient	Upper 0.95	Lower 0.05	Coefficient	Upper 0.95
Num of economic active males in HH	0.016	0.038	0.061	0.004	0.016	0.034	0.022	0.0547	0.089
Num of children under 5 yrs. old in HH	-0.077	-0.037	0.005	-0.038	-0.016	0.002	-0.109	-0.0532	0.006
Log(HH head age)	-1.111	0.596	2.364	-0.459	0.255	1.147	-1.577	0.8518	3.368
(Log(HH head age))²	-0.345	-0.108	0.12	-0.167	-0.046	0.045	-0.488	-0.1541	0.173
Wealth Index	-0.091	0.091	0.279	-0.040	0.037	0.128	-0.132	0.1283	0.393
Number of varieties cultivated	0.056	0.069	0.083	0.011	0.029	0.054	0.075	0.0988	0.129
Farming experience (years)	0.005	0.008	0.012	0.001	0.004	0.007	0.006	0.0122	0.018
Land labor ratio (m²/person)	0.000	0.000	0.000	0.000	0.000	0.000	-0.000	0.000	0.000
Travel time to technical services	-0.009	-0.002	0.006	-0.005	-0.001	0.002	-0.015	-0.0029	0.008
Planting material from friends or relatives	1.269	2.263	3.465	0.433	0.907	1.450	1.973	3.170	4.555

Table 4 SDM probit model effects estimates for bush bean growers (M1)

	Direct effect			Indirect effect			Total effect		
Variables	Lower 0.05	Coefficient	Upper 0.95	Lower 0.05	Coefficient	Upper 0.95	Lower 0.05	Coefficient	Upper 0.95
Numb of economic active males in HH	0.014	0.036	0.058	0.005	0.016	0.031	0.020	0.051	0.085
Numb of children under 5 years old in HH	-0.069	-0.035	-0.001	-0.035	-0.016	-0.001	-0.102	-0.051	-0.002
Log(Household head age)	-0.881	0.836	2.521	-0.363	0.371	1.263	-1.251	1.207	3.849
(Log(Household head age))²	-0.371	-0.140	0.088	-0.185	-0.062	0.036	-0.552	-0.202	0.127
Wealth Index	-0.095	0.078	0.256	-0.038	0.035	0.129	-0.133	0.113	0.372
Number of varieties cultivated	0.052	0.065	0.078	0.013	0.029	0.047	0.070	0.094	0.118
Farming experience (years)	0.005	0.008	0.012	0.001	0.004	0.007	0.006	0.012	0.018
Land labor ratio (m²/person)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Travel time to technical services	-0.009	0.020	0.050	-0.003	0.009	0.024	-0.013	0.029	0.071
Planting material from friends or relatives	2.279	3.601	5.229	0.765	1.521	2.369	3.314	5.122	7.099

Table 5 SAR and GLM probit model effects estimates for bush bean growers (M2)

Variables	Direct effect			Indirect effect			Total effect		
	Lower 0.05	Coefficient	Upper 0.95	Lower 0.05	Coefficient	Upper 0.95	Lower 0.05	Coefficient	Upper 0.95
Household size	-0.019	0.002	0.022	-0.015	0.001	0.018	-0.034	0.003	0.039
Number of children under 5 years old in HH	-0.101	-0.046	0.010	-0.083	-0.034	0.006	-0.177	-0.079	0.016
Log(Household head age)	-1.388	0.860	3.193	-1.001	0.662	2.619	-2.440	1.522	5.696
(Log(Household head age))^2	-0.432	-0.129	0.167	-0.359	-0.099	0.123	-0.766	-0.228	0.296
HH average years of schooling	-0.014	0.008	0.028	-0.010	0.006	0.022	-0.024	0.013	0.049
Share of land area with legal title	0.000	0.001	0.001	0.000	0.000	0.001	0.000	0.001	0.002
Management Index	0.073	0.230	0.386	0.043	0.167	0.311	0.124	0.397	0.659
Travel time to extension services	-0.004	0.005	0.015	-0.003	0.004	0.012	-0.006	0.009	0.026

Table 6 SDM probit model effects estimates for bush bean growers (M2)

Variables	Direct effect			Indirect effect			Total effect		
	Lower 0.05	Coefficient	Upper 0.95	Lower 0.05	Coefficient	Upper 0.95	Lower 0.05	Coefficient	Upper 0.95
Household size	-0.019	0.000	0.020	-0.013	0.000	0.013	-0.032	0.000	0.033
Number of children under 5 years old in HH	-0.095	-0.043	0.007	-0.069	-0.027	0.004	-0.159	-0.070	0.011
Log(HH head age)	-1.290	0.868	3.064	-0.770	0.579	2.308	-2.007	1.447	5.181
(Log(HH head age))^2	-0.420	-0.129	0.159	-0.311	-0.085	0.089	-0.711	-0.215	0.246
HH average years of schooling	-0.013	0.008	0.030	-0.008	0.005	0.020	-0.021	0.014	0.048
Share of land area with legal title	0.000	0.000	0.001	0.000	0.000	0.001	-0.001	0.001	0.002
Management Index	0.095	0.256	0.400	0.036	0.162	0.323	0.147	0.418	0.701
Travel time to extension services	0.003	0.039	0.075	0.001	0.025	0.058	0.005	0.064	0.128

iii. Climbing bean analyses

Scenario 1 for climbing bean adopters shows a different spatial pattern (Table 7). Contrary to IBB bush adopters, the propensity of adoption of IBB climbing varieties increases with household's wealth and risk-taking households or households with less farming experience. In the SDM, the direct and direct effects turned statistically significant at the 5 percent level (Table 9) for four covariates: household's wealth, number of bean varieties cultivated, household head farming experience, and IBB planting material received from friends or relatives. The largest total marginal effect was associated with IBB planting material from friends or relatives followed by household's wealth, which increased the probability of adoption by 25 percent, the number of bean varieties grown by the household, which increased the propensity of adoption by 5 percent. To cope with the risk associated with crop failure and food insecurity, bean farmers cultivate more than a single bean variety.

In scenario M2, three variables turned statistically significant at the 5 percent level: management practices, household size, and household education level (table 7). Household size positively affects the propensity of IBB adoption, with a positive direct impact of increasing adoption by 2 percent for an additional member in the household. Larger households have the capacity to increase the labor availability required with the adoption of a new variety, such as IBB, while household education level had a direct impact of increasing the probability of adoption by 3 percent. Most notably, the results suggest that the average education level of household members (rather than the education level of the head of household) influence the adoption of new technology, and it is positively correlated with wealth. Farming households that are more educated are wealthier, and thus the education–adoption relationship may represent an income effect (Jolliffe, 2002). Also, as it was reported in the descriptive statistics, wealth may be correlated with the scale of operation, as adopters tend to manage more and larger plots.

Tables 8 to 10 summarize the observed values of the estimates of the marginal effects for specifications M1 and M2 of the SAR and SDM models for climbing bean growers. The SDM model effect for specification M1 (Table 9), out of the nine covariates, just four covariates are statistically significant at the 5 percent level: farming experience, number of bean varieties cultivated, wealth index, and IBB planting material from friends or relatives. The positive direct effect of the number of bean varieties cultivated and planting material from friends or relatives suggest that higher values of these variables for bean growing household h_i lead to an increase in the propensity of adoption of IBB climbing varieties. Farming experience indicates a negative direct effect suggesting that household heads with less farming experience are more likely to adopt IBB climbing varieties. Socioeconomic characteristics of neighboring bean farmers, such as their household size and level of education exert positive spatial spillovers on IBB adoption rates. Higher magnitude of the estimated parameters of these covariates increases the propensity of the adoption of IBB climbing varieties. For specification M2 (Table 11), out of the seven included covariates, three variables—management index, household size, and household average years of schooling—are statistically significant at the 5% level for both the direct and the estimates of the indirect effects. Contrary to IBB bush varieties, where the last two covariates were not statically significant.

The national adoption rate of IBB varieties was 28 percent. To better understand patterns of IBB adoption at sub-national level, Figure 5 contains a choropleth map of the prevalence rates of IBB adoption by bean types at the district level. From this map, I highlight two spatial patterns. First, from the choropleth map, it is clear that the rates of adoption for IBB bush varieties are higher in the Eastern region and gradually decreasing toward the Central, Southern, and Western regions. IBB climbing varieties, on the other hand, have higher rates of IBB adoption in the Western and Northern regions. IBB bush varieties have higher probabilities of adoption in the Central and Southern regions.

Figure 6 shows point estimates of village-level random effect for IBB climbing growers. The values of the point estimates range from -0.8 to 0.6 with a higher prevalence of villages with point estimates that range between -0.4 and 0. These villages have a lower probability to adopt IBB climbing varieties and do not follow a particular spatial pattern. On the other hand, the spatial footprint of villages with positive point estimates is less frequent. These villages have higher probabilities to adopt IBB climbing varieties and tend to form clusters. Geographic diffusion of iron biofortified planting material occurs among these neighboring villages that exhibit: interdependent decision-making patterns, as well as similar characteristics relative to the group.

Table 7 SDM probit model effects estimates for climbing bean growers (M1)

Variables	M1-SAR	M1-SDM	M1-NSP	M2-SAR	M2-SDM	M2-NSP
Rho	0.3128***	0.293*		0.419***	0.405**	
	(0.08145)	(0.093)		(0.093)	(0.096)	
Constant	-1.446	-2.266	-1.907	-4.699	-3.867	-4.947***
	(8.491)	(8.651)	(8.394)	(7.728)	(7.756)	(7.649)
Household size				0.075***	0.075***	0.070***
				(0.034)	(0.035)	(0.033)
Number of economic active males in HH	0.059	0.047	0.039			
	(0.058)	(0.061)	(0.057)			
Number of children under 5 years old in HH	-0.006	0.008	-0.001	0.005	0.046	0.017***
	(0.097)	(0.100)	(0.094)	(0.102)	(0.099)	(0.095)
Log (HH head age)	-0.803	-0.299	-0.548	1.794	1.527	1.764
	(4.532)	(4.619)	(4.475)	(4.081)	(4.113)	(4.063)
(Log(HH head age))²	0.201	0.128	0.158	-0.227	-0.187	-0.221
	(0.603)	(0.612)	(0.595)	(0.533)	(0.538)	(0.533)
Level education				0.094***	0.101***	0.0961***
				(0.033)	(0.033)	(0.032)
Wealth Index	1.023**	1.17**	1.072**			
	(0.471)	(0.464)	(0.465)			
Number of varieties cultivated	0.227***	0.236***	0.219***			
	(0.037)	(0.039)	(0.038)			
Farming experience (years)	-0.036**	-0.038***	-0.035**			
	(0.017)	(0.016)	(0.016)			
Share of land area with legal title				-0.001	-0.001	-0.001
				(0.001)	(0.001)	(0.001)
Management Index				0.502*	0.517*	0.566**
				(0.266)	(0.303)	(0.279)
Land labor ratio (m²/person)	-0.000	-0.000	-0.000			
	(0.000)	(0.000)	(0.000)			
Travel time to technical services	0.008	0.022	0.002	-0.008	0.003	-0.012
	(0.020)	(0.077)	(0.023)	(0.017)	(0.067)	(0.021)
Planting material from friends or relatives	9.584***	11.73**	27.75			
	(1.412)	(2.968)	(669.1)			
W-Household size					-0.041	
					(0.061)	
W-Number of economic active males in HH		0.108				
		(0.113)				
W-Number of children under 5 years old in HH		-0.329*			-0.203	
		(0.194)			(0.175)	
W-Log (HH head age)		0.009			0.002	
		(0.012)			(0.009)	
W-Level education					-0.045	
					(0.055)	
W-Wealth Index		-0.959				
		(0.919)				
W-Number of varieties cultivated		-0.034				
		(0.066)				
W-Farming experience (years)		-0.019				
		(0.019)				
W-Share of land area with legal title					-0.004	
					(0.003)	
W-Management Index					0.517	
					(0.303)	
W-Land labor ratio (m2/person)		0.000				
		(0.000)				
W-Travel time to technical services		-0.011			-0.014	
		(0.078)			(0.068)	
W-Planting mat. From friends or relatives.		0.054				
		(0.867)				
Kelejian-Prucha (error)			4.013***			4.582***
Log likelihood	-249.459	-244.114	-252.6889	-326.712	-322.465	-326.984
BIC	569.522	616.663	575.9799	717.607	754.042	711.733
AIC	520.919	528.296	527.3779	673.423	678.930	671.968
Number of observations	613	613	613	613	613	613

Notes: () are standardized errors. *, ** and *** denote significance at the 10%, 5%, and 1% level, respectively

Figure 5 Choropleth map: prevalence rate of IBB adoption, by district

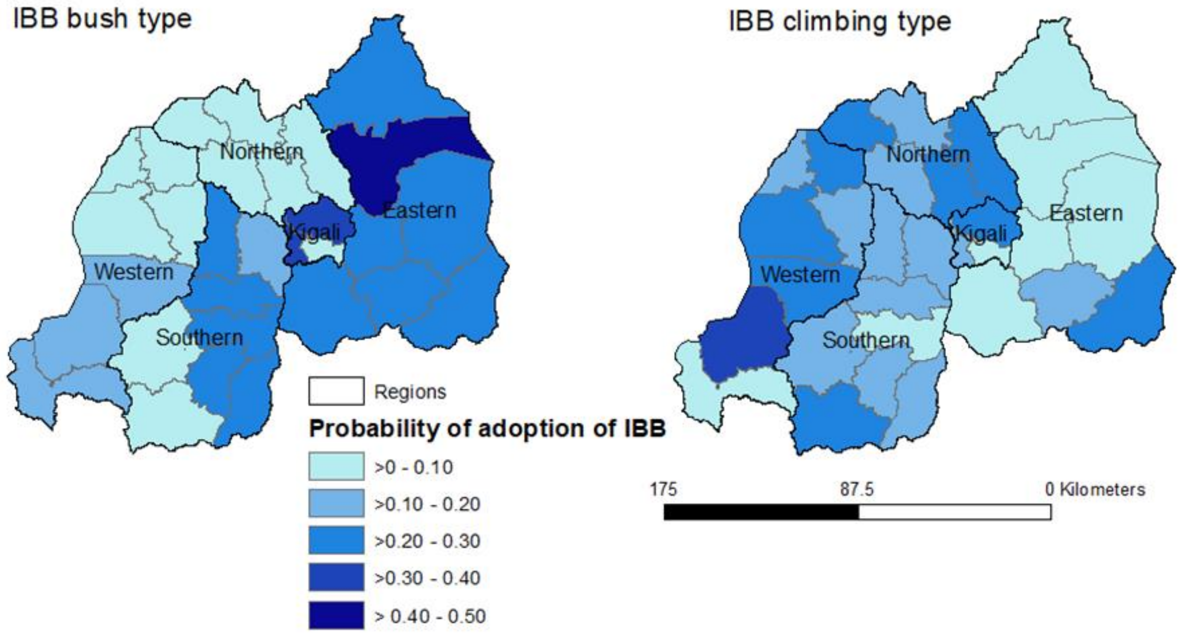
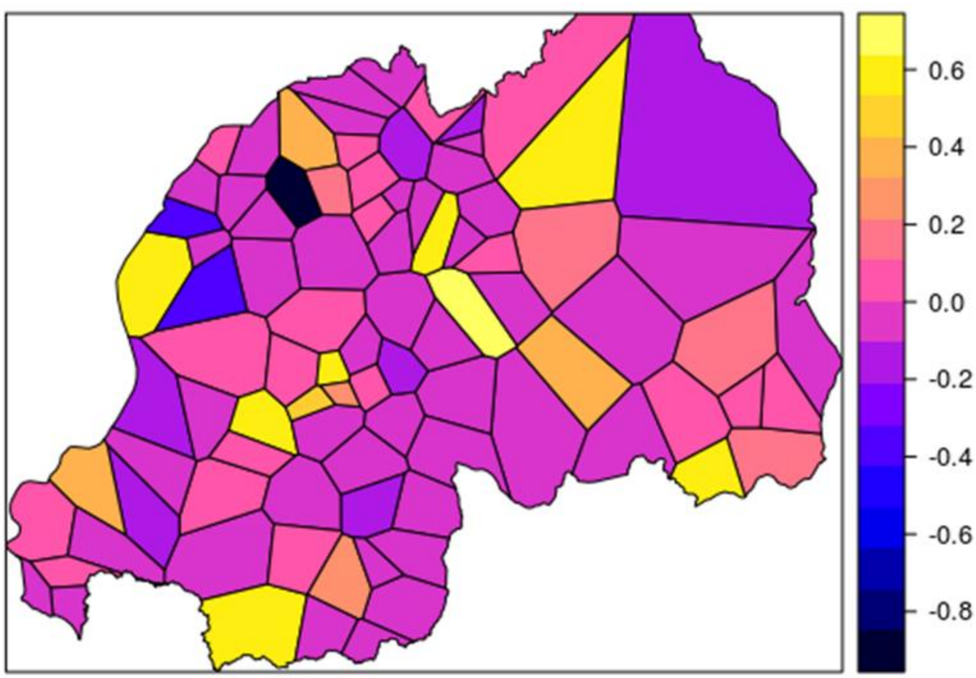


Figure 6 Point estimates of village-level random effect: IBB climbing M1.IID



Note: Random effect values range from -1 to 1. Values around zero suggest spatial randomness, residual values close to 1 suggest spatial clustering and residual values close to -1 suggest a negative spatial association.

Table 8 SAR and GLM probit model effects estimates for climbing bean growers (M1)

	Direct effect			Indirect effect			Total effect		
Variables	Lower 0.05	Coefficient	Upper 0.95	Lower 0.05	Coefficient	Upper 0.95	Lower 0.05	Coefficient	Upper 0.95
Number of economic active males in HH	-0.008	0.013	0.034	-0.003	0.006	0.017	-0.011	0.019	0.052
Number of children under 5 years old in HH	-0.037	-0.001	0.033	-0.019	-0.001	0.015	-0.055	-0.002	0.047
Log(Household head age)	-1.788	-0.171	1.490	-0.843	-0.086	0.680	-2.579	-0.257	2.101
(Log(Household head age))²	-0.175	0.043	0.260	-0.081	0.021	0.127	-0.259	0.064	0.373
Wealth Index	0.055	0.222	0.389	0.022	0.104	0.210	0.080	0.326	0.582
Number of varieties cultivated	0.036	0.049	0.061	0.011	0.023	0.037	0.052	0.073	0.093
Farming experience (years)	-0.014	-0.008	-0.002	-0.008	-0.004	-0.001	-0.022	-0.012	-0.003
Land labor ratio (m²/person)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Travel time to technical services	-0.005	0.002	0.009	-0.002	0.001	0.005	-0.007	0.003	0.014
Planting material from friends or relatives	1.506	2.090	2.648	0.491	0.959	1.427	2.349	3.049	3.796

Table 9 SDM probit model effects estimates for climbing bean growers (M1)

	Direct effect			Indirect effect			Total effect		
Variables	Lower 0.05	Coefficient	Upper 0.95	Lower 0.05	Coefficient	Upper 0.95	Lower 0.05	Coefficient	Upper 0.95
Number of economic active males in HH	-0.013	0.010	0.032	-0.005	0.004	0.016	-0.018	0.014	0.046
Number of children under 5 years old in HH	-0.036	0.002	0.036	-0.014	0.001	0.016	-0.048	0.002	0.052
Log(Household head age)	-1.692	-0.069	1.587	-0.754	-0.015	0.729	-2.327	-0.083	2.303
(Log(Household head age))²	-0.186	0.028	0.243	-0.087	0.010	0.105	-0.273	0.038	0.340
Wealth Index	0.097	0.251	0.414	0.022	0.106	0.231	0.130	0.357	0.619
Number of varieties cultivated	0.037	0.051	0.064	0.008	0.021	0.039	0.049	0.072	0.096
Farming experience (years)	-0.014	-0.008	-0.003	-0.007	-0.003	-0.001	-0.020	-0.012	-0.004
Land labor ratio (m²/person)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Travel time to technical services	-0.023	0.005	0.031	-0.009	0.002	0.015	-0.031	0.007	0.045
Planting material from friends or relatives	1.514	2.524	3.660	0.423	0.999	1.609	2.240	3.523	4.733

Table 10 SAR and GLM probit model effects estimates for climbing bean growers (M2)

Variables	Direct effect			Indirect effect			Total effect		
	Lower 0.05	Coefficient	Upper 0.95	Lower 0.05	Coefficient	Upper 0.95	Lower 0.05	Coefficient	Upper 0.95
Household size	-0.019	0.002	0.022	-0.015	0.001	0.018	-0.034	0.003	0.039
Number of children under 5 years old in HH	-0.101	-0.046	0.010	-0.083	-0.034	0.006	-0.177	-0.079	0.016
Log(HH head age)	-1.388	0.860	3.193	-1.001	0.662	2.619	-2.440	1.522	5.696
(Log(HH head age))²	-0.432	-0.129	0.167	-0.359	-0.099	0.123	-0.766	-0.228	0.296
HH average years of schooling	-0.014	0.008	0.028	-0.010	0.006	0.022	-0.024	0.013	0.049
Share of land area with legal title	0.000	0.001	0.001	0.000	0.000	0.001	0.000	0.001	0.002
Management Index	0.073	0.230	0.386	0.043	0.167	0.311	0.124	0.397	0.659
Travel time to extension services	-0.004	0.005	0.015	-0.003	0.004	0.012	-0.006	0.009	0.026

Table 11 SDM probit model effects estimates for climbing bean growers (M2)

	Direct effect			Indirect effect			Total effect		
Variables	Lower 0.05	Coefficient	Upper 0.95	Lower 0.05	Coefficient	Upper 0.95	Lower 0.05	Coefficient	Upper 0.95
Household size	0.004	0.022	0.039	0.002	0.014	0.030	0.007	0.036	0.066
Number of children under 5 years old in HH	-0.035	0.013	0.060	-0.024	0.009	0.044	-0.060	0.022	0.100
Log(HH head age)	-1.489	0.431	2.356	-0.958	0.319	1.820	-2.400	0.751	4.149
(Log(HH head age))²	-0.307	-0.053	0.199	-0.236	-0.039	0.127	-0.537	-0.093	0.326
HH average years of schooling	0.013	0.029	0.044	0.006	0.019	0.037	0.022	0.048	0.076
Share of land area with legal title	-0.001	0.000	0.000	-0.001	0.000	0.000	-0.002	0.000	0.001
Management Index	0.012	0.148	0.294	0.006	0.100	0.226	0.019	0.247	0.495
Travel time to extension services	-0.031	0.001	0.034	-0.021	0.001	0.024	-0.050	0.002	0.055

Table 12 Multilevel spatial regression models for bush and climbing bean growers

Variables	Bush		Climbing	
	M1.IID	M2.IID	M1.IID	M2.IID
Constant	8.139	5.941	4.436	7.835
	(15.502)	(13.414)	(14.933)	(13.353)
Household size		0.020		0.128***
		(0.064)		(0.059)
Number of economic active males in HH	0.293***		0.056	
	(0.102)		(0.099)	
Number of children under 5 years old in HH	-0.318**	-0.280*	-0.014	0.039
	(0.184)	(0.170)	(0.168)	(0.167)
Log (HH head age)	4.115	2.850	-0.401	2.641
	(8.352)	(7.178)	(7.957)	(7.091)
$(\text{Log(HH head age)})^2$	-0.773	-0.462	0.210	-0.330
	(1.121)	(0.949)	(1.057)	(0.929)
HH average years of schooling		0.030		0.170***
		(0.066)		(0.056)
Wealth Index	0.545		2.002**	
	(0.882)		(0.847)	
Number of varieties cultivated	0.513***		0.397***	
	(0.073)		(0.071)	
Farming experience (years)	0.065***		-0.067**	
	(0.018)		(0.031)	
Share of land area with legal title		0.003		-0.002
		(0.002)		(0.003)
Management Index		1.387***		0.970**
		(0.491)		(0.506)
Land labor ratio (m^2/person)	0.000		0.000	
	(0.000)		(0.000)	
Travel time to technical services	-0.003	0.038	0.006	-0.024
	(0.045)	(0.037)	(0.044)	(0.042)
Planting material from friends or relatives	20.481***		22.064**	
	(12.315)		(12.818)	
Log Likelihood	-269.670	-342.630	-310.600	-375.700
DIC	453.540	605.390	535.290	664.980
WAIC	454.890	605.160	537.400	663.220
Number of observations	506	506	613	613

Notes: () are standardized errors. *, ** and *** denote significance at the 10%, 5%, and 1% level, respectively

V. Conclusions

This chapter contributes to the research literature in two ways. First, it provides a national and a sub-national analysis on the intensity of adoption and adoption rates of iron biofortified beans (IBB) by bean type. Second, the chapter examines, with the assistance of spatial econometrics techniques and theories of social interaction, and choice behavior, how households and farms characteristics, as well as regional factors, influence smallholder farmers' decisions to grow IBB. I used a cross-sectional nationally representative survey of bean producing households on bean varieties grown in 2015 season B in Rwanda. I employed two spatial probit specifications, spatial autoregressive model (SAR) and spatial Durbin model (SDM), to empirically assess the interdependence of farmers' decisions to adopt IBB. Robustness of the results was tested by setting a simple social grouping where smallholder farmers are nested within villages. This multilevel fixed model with random effect (village) produced weaker spatial relationship across villages. This confirms our hypothesis that closer neighbors matter more than those farther away.

The tabular analyses of the data indicate that local bean varieties still dominate the area under bean cultivation, followed by improved and IBB varieties, respectively. Given that IBB varieties were only released 3-5 years prior to the time of the survey, 11 percent coverage figure indicates the intensity of adoption of IBB, suggesting early stages within a long-run S-shaped adoption curve. The spatial econometric results indicate interdependence on farmers' decisions to adopt IBB. In addition to the directly targeted beneficiaries, the parameter ρ suggests that the biofortification program affected non-beneficiaries as well. This finding indicates that

(1) a household is more likely to grow IBB if the household is near other early IBB adopters which communicate about the nutritional and yield benefits of IBB technology and (2) the propensity of a household to grow IBB varies with the characteristics of neighboring farmers. A non-spatial probit model would not measure spatial association as an indicator of interaction of farmers in a social network.

Structural factors are the main direct and indirect determinants for predicting the likelihood of adoption of IBB varieties. For IBB bush growers, these factors include the number of economically active male members in a household, and management practices. In absolute terms, the largest total marginal effect is management practices. For IBB climbing growers, household size and education level were most effective to exert direct and indirect effect in the adoption of IBB. Common factors that influence the adoption of IBB varieties for both bush and climbing bean growers include the number of years of farming experience and the number of varieties cultivated. Farming experience has a negative direct impact, as well as a negative spatial spillover on the household's propensity to adopt IBB climbing varieties. In contrast to the adoption of IBB bush varieties, years of farming experience have a positive direct impact on the adoption of IBB varieties and a positive spatial spillover. The second common factor that influences the adoption of IBB is the number of varieties cultivated. I observe a positive direct effect associated with the number of varieties cultivated, suggesting that a higher value of this variable leads to an increase in the propensity to adopt IBB varieties by the household. I considered the variable planting material from friends or relatives of a smallholder farmer as covariate. The coefficient for this covariate for both IBB bush and climbing growers was positive and significant, which further supports the positive role of social interaction in technology diffusion.

Some general policy implications can be drawn from the above results. First, drafting a differentiated geographical targeting strategy for bush and climbing bean varieties as a function of farmer and farm characteristics might increase adoption rates on the most vulnerable groups in rural areas. Second, if schooling augments the propensity to adopt climbing IBB varieties, increasing educational levels on the nutritional and agronomic benefits of IBB might be an effective policy to stimulate technological diffusion. Third, strengthening partnerships with extension services will stimulate adoption. Fourth, considering spatial econometric techniques in program evaluation helps to assess the impact of policies on indirect beneficiaries. Finally, considering progressive farmers and strengthening social group activities when designing technology-promotion programs increases the spread of information and geographical diffusion of IBB. In sum, social interaction works as an invisible cost-effective delivery strategy of IBB in rural population that are likely at risk of iron deficiency. These activities might continue to support scaling up of delivery activities for IBB varieties in rural areas of Rwanda.

Chapter 5: Impact of iron biofortified beans on yields and farmers' incomes: The case of Rwanda

I. Introduction

In Rwanda, micronutrient malnutrition is highly pervasive. Thirty seven percent of children under five years of age and nearly 20 percent of women of childbearing age suffer from anemia (NISR, 2015) in the country, and about 25 percent of children and 37 percent of women have iron deficiencies (Petry et al., 2016). At the same time, Rwandans have one of the highest per capita bean consumption rates in the world, with rural households consuming beans on average six days a week (Asare-Marfo et al., 2016), and in significant quantities (Berti et al., 2012). However, in terms of bean production, bean farmers in Rwanda have low productivity (FAO, 2020) translating into low food availability. With less food available, vulnerable populations face an increased risk of malnutrition. To satisfy growing food demand, there are two broad options: 1) increase land under production, 2) food imports, or/and 3) boost crop productivity. Crop productivity can increase through the adoption of higher yielding varieties and more efficient production techniques.

Following several years of collaborative research between HarvestPlus, the Rwanda Agriculture Board (RAB), and the International Center for Tropical Agriculture (CIAT), four iron biofortified bean (IBB) varieties were officially released for planting in Rwanda in 2010. Another six were released in 2012. Of the IBB varieties released, two were bush and eight were climbing types. Farmer feedback studies on the IBB varieties conducted following early delivery efforts have shown that farmers are willing to grow these beans in increasing areas as well as to share the planting material with others. Consumer acceptance studies found that consumers prefer the IBB varieties over most of their local varieties (Oparinde et al., 2016).

In addition to their nutritional benefits, biofortified crop varieties are bred to yield as well as the best yielding current varieties. However, these hypotheses haven't been tested. Therefore, in Rwanda, adoption of IBB is expected to improve yields, which may translate into one or a combination of the following: (1) higher household consumption of IBB, that will translate into improved iron intakes, and (2) higher levels of market sales and hence subsequent income gains. Assessment of income gains is particularly important in Rwanda. The share of agriculture in GDP (32 percent) is higher than the average for low-income countries (27 percent) and higher than the average in Sub-Saharan Africa (12 percent). Currently, there is very little empirical evidence on these hypotheses and also a lack of evidence on the productivity levels across the two bean types, especially for IBB varieties. By filling this important niche in the literature, this chapter achieves objective two of this dissertation:

Objective 2: to estimate the potential impact of iron biofortified beans on smallholder farming households' outcomes: yields and incomes in Rwanda

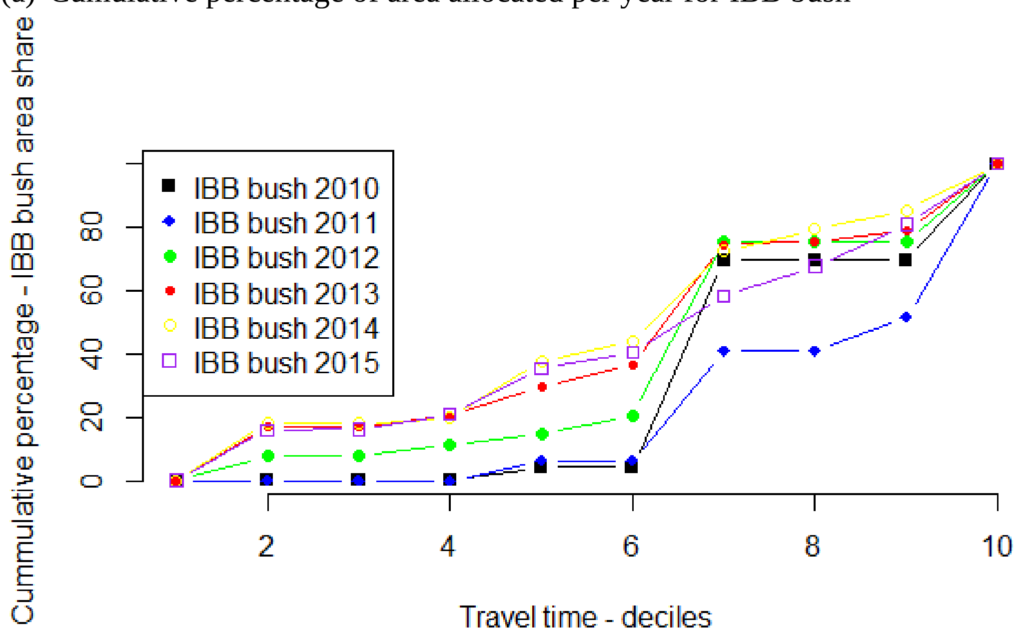
II. Background

i. IBB adoption

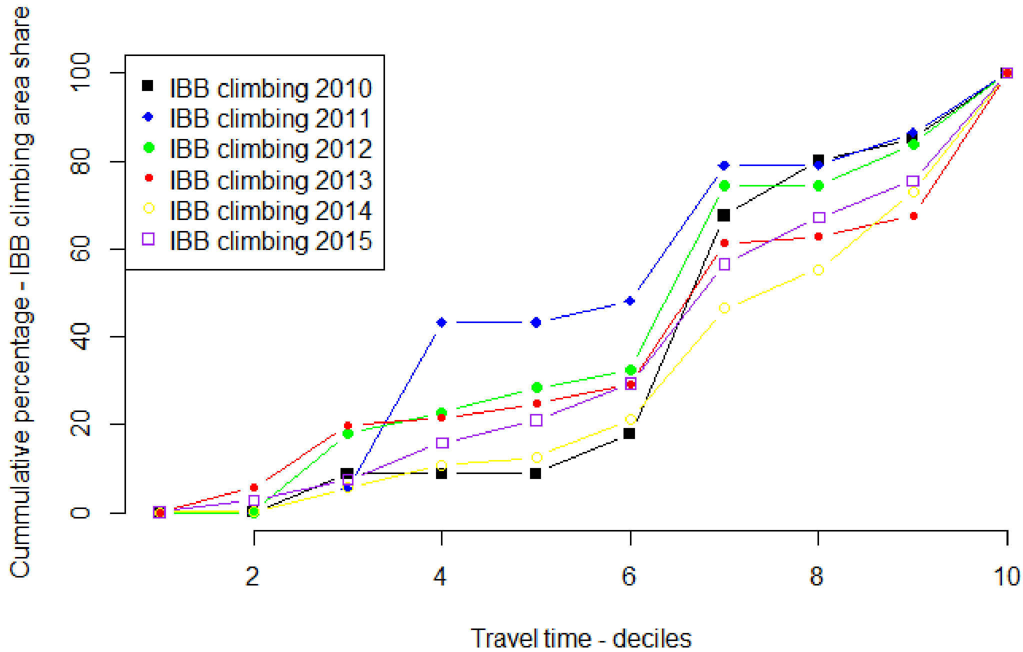
IBB are grown in every district of Rwanda. By the end of 2015 Season B, ten biofortified IBB varieties had been disseminated to smallholder farming households in Rwanda, out of which two are iron biofortified bush bean varieties: RWR2245 and RWR2154 and eight are iron biofortified climbing bean varieties. Figure 7 shows percentage of IBB area from 2010 to 2015 plotted against travel time to markets (see appendix A). Bush IBB adoption was highest in rural areas from onset and gradually increased over time to areas closer to markets of 50,000 inhabitants or more. On the other hand, climbing IBB adoption was highest in areas closer to markets and over time expanded to smallholder farming households living in rural areas.

Figure 7 Cumulative percentage of area allocated per year for (a) IBB bush and (b) IBB climbing varieties

(a) Cumulative percentage of area allocated per year for IBB bush



(b) Cumulative percentage of area allocated per year for IBB climbing



ii. Bean prices

In addition to the household survey, a community survey was implemented. Data on bean prices in 120 local markets during periods of both low and high seasonal availability of beans was collected in the community survey. To assign bean prices from local markets to the location of the 1,397 bean growing households, I employed spatial interpolation methods to create bean price surfaces across our study area. In this study, I used Thiessen polygons. Thiessen polygons (also known as Voronoi diagrams) are obtained by assigning to each bean farmer the bean price reported at the nearest local market. Basically, each Thiessen polygon represents a local market, which influences farmers' prices much more so than any other sample price point or local market. This process results in the partitioning of space into a tessellation of many local markets, which corresponds to the notion of spatial local market areas where only transportation costs matter (Anselin & Le Gallo, 2006).

Figure 8 Thiessen polygon interpolation of bean prices (RWF/kg) in local markets in Rwanda during periods of low availability of beans (left) and high availability of beans (right) in season B 2015

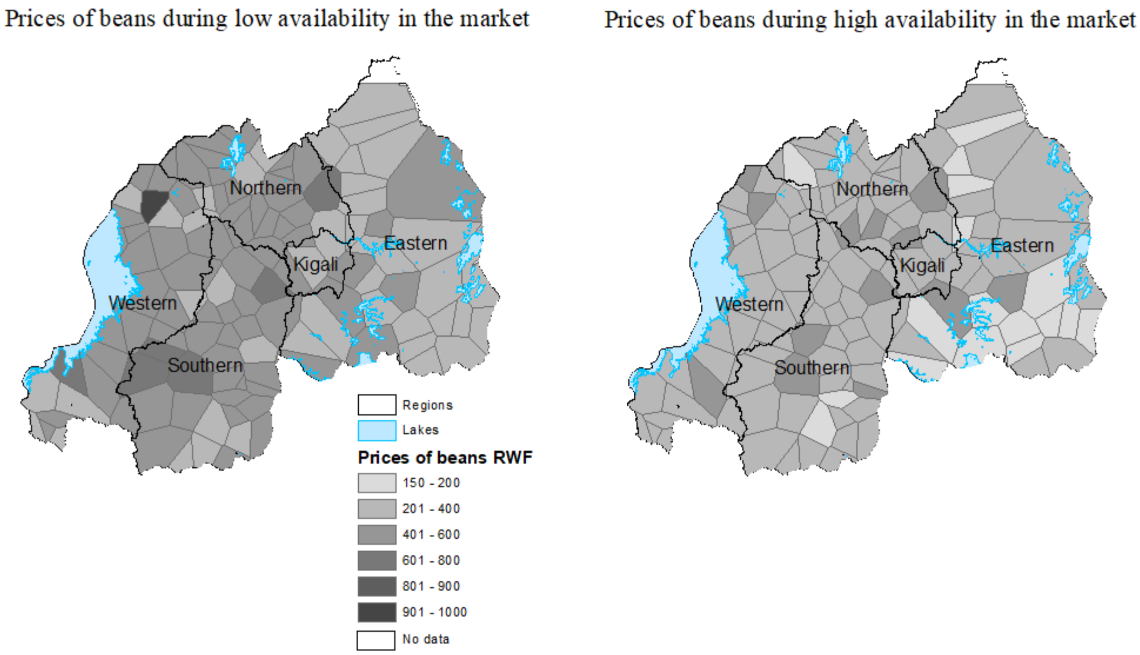


Figure 8 shows prices of regular beans in local markets during periods of low and high availability of beans in season B 2015. During low availability periods, prices ranged from 290 RWF/kg to 1000 RWF/kg, with an overall mean of 470 RWF/kg. This overall mean hides considerable variability across provinces. For instance, the average bean prices for local markets in districts in Eastern Province was 398 RWF/kg, while it was 425 RWF/kg in Kigali city, 485 RWF/kg in districts in Northern Province, and 500 RWF/kg in the Southern and Western provinces. During high availability periods, bean prices dropped by 29%. Bean prices ranged from 150 RWF/kg to 600 RWF/kg with a mean of 320 RWF/kg. A similar pattern was observed with lower prices in the eastern part of the country. To estimate production values for the analysis here of the impact of IBB on smallholder farming household income, I used the midpoint of prices during the low and the high availability periods for beans in season B 2015 in local markets.

III. Results and Discussion

i. Impact of iron biofortified beans on yields and smallholder farming households' incomes

Tables 19 and 21 provide the results of the probit models for IBB bush and IBB climbing varieties, respectively. The bottom of the tables shows the diagnostics tests of spatial dependence, all of which are significant, hence, rejecting the null hypothesis that IBB adopters are randomly distributed. I re-estimated the probability of adoption and added it as a lagged variable in the multivariate matching algorithm. The Appendix provides a discussion of the results of the multivariate matching algorithm. Chapter 4 provides a full discussion of the marginal effects of the propensity to adopt IBB varieties. For the purpose of this chapter, the discussion is focused on the impact of the treatment effect and on the heterogeneity of impacts on outcome variables.

With respect to the impact, tables 13 and 14 provide summary statistics of the treatment effect on bean yields and agricultural incomes for both IBB bush growers and IBB climbing growers, respectively. As shown in table 1, IBB bush growers increased their bean yields by 153 kg/ha, which was a 23 percent increase over prior yields. The 95 percent confidence interval around this estimate ranged between 73 kg/ha and 240 kg/ha. For IBB climbing growers as shown in table 2, their bean yields increased by 20 percent (equivalent to 182 kg/ha).

I also examined IBB bean impact on smallholder farming households' incomes. The conversion rate (PPP) from dollars to Rwandan francs is based on the 2015 exchange rate of RWF 750=US $1.00. IBB bush growers increased their potential incomes by 27 percent (equivalent to $84/ha). IBB climbing growers increased their potential incomes by 23 percent (equivalent to $110/ha. The total national economic value added for both IBB bean types in season B 2015 equaled 2.5 million US dollars. Extrapolated from the onset of the iron biofortification bean program, the total economic value added equaled 9 million US dollars.

Table 13 Models for potential outcomes (ATT) for IBB bush growers – estimates and 95% confidence intervals (CI)

	Summary		
Treatment effect	Mean	P-value	CI
Yield gain (kg/ha)	153.54	0.04	73 - 240
	(72.36)		
Income gain ($/ha)	84.41	0.02	39 - 127
	(37.19)		
Total added value ($)	1,524,106		

ATT = average treatment on the treated, based on 1000 simulations. The model reported the average yields of bush type of IBB: 805 kg/ha and the average yields of regular bush beans: 652 kg/ha. N=206

Table 14 Models for potential outcomes (ATT) for IBB climbing growers – estimates and 95% confidence intervals (CI)

	Summary		
Treatment effect	Mean	P-value	CI
Yield gain (kg/ha)	181.53 (90.97)	0.04	78 - 292
Income gain ($/ha)	109.74 (47.97)	0.02	56 - 162
Total added value ($)	1,236,169		

ATT = average treatment on the treated, based on 1000 simulations. The model reported the average yields of climbing type of IBB: 1081 kg/ha and the average yields of regular climbing beans: 899 kg/ha. N=206

The results of the sensitivity analysis by IBB type are presented in Table 23 in the Appendix. This analysis checks for the presence of hidden bias due to unobserved covariates that are related to the treatment assignment mechanism. The larger gamma value indicates the group difference is more resistant to hidden bias. For IBB bush adopters, a change of 0.04 or 0.08 on the odds of treatment assignment will change the treatment effect from significant to non-significant at the 5 and 10 percent significance levels, respectively.

This indicates that the treatment effect could be quite easily altered by accounting for some unobserved covariates. For IBB climbing adopters, a change of 0.09 or 0.15 on the odds of treatment assignment will change the treatment effect from significant to non-significant at 5 and 10 percent significance levels, respectively. This indicates that the conclusion of treatment effect is less sensitive to being altered by accounting for some presently unobserved covariates.

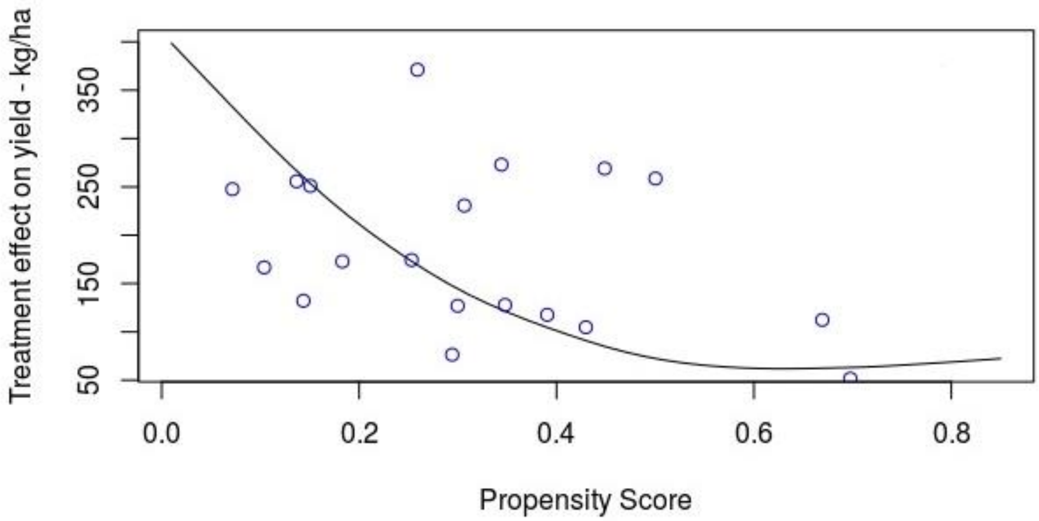
While our estimates of impacts are significantly positive, impacts on outcomes are heterogeneous across farming households. I invoke the ignorability assumption that after I control for a set of observed covariates, there are no additional confounders between farming households that adopted and did not adopt IBB.

ii. Heterogeneous treatment effect over PS, land-labor ratios, and travel time to markets

Figures 9 and 10 show how the treatment effect on yields and incomes varied over PS for bush and climbing bean growers, respectively. For the bush bean growing smallholder farming households, the impact on yields was highest and positive for farming households less likely to grow IBB. The impact declines to a threshold before the impact on yields starts to increase again. Figure 10 (b) shows the case of climbing bean growers. The results suggest that households with the lowest PS had the greatest yields and incomes gains. Farming households with PS higher than 0.6 tended to have lower treatment effect on yields and incomes. The impact on incomes was similar to yields, but it declined to a positive value and remained relatively even for farming households with PS greater than 0.6 for both growing mechanisms. These findings were unexpected and suggests the presence of negative and positive selections.

Figure 9 Heterogeneity of treatment effect of PS over (a) yields and (b) incomes - IBB bush

(a) Yields



(b) Incomes

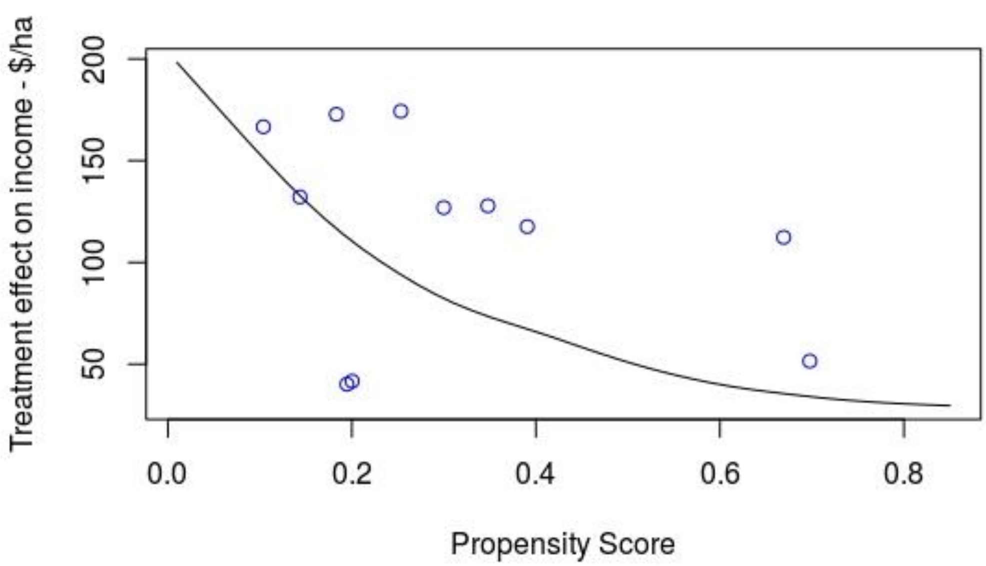
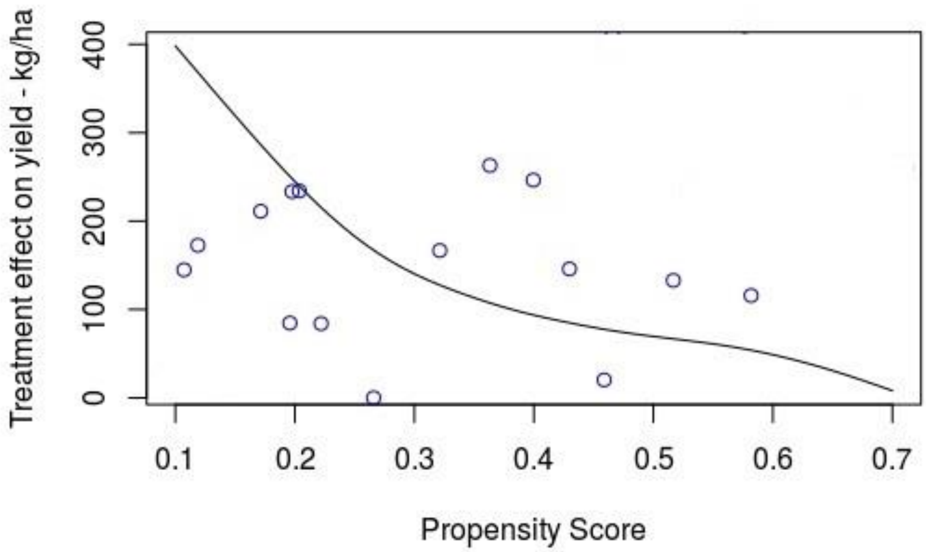
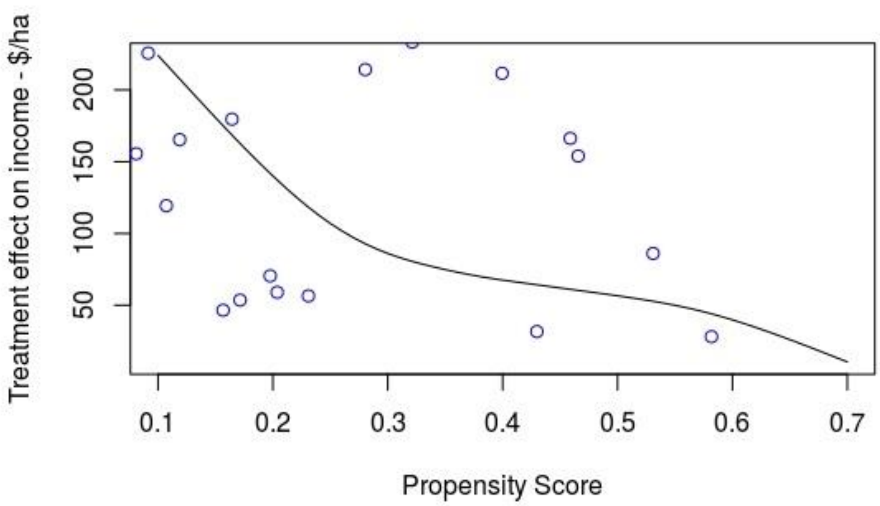


Figure 10 Heterogeneity of treatment effect of PS over (a) yields and (b) incomes - IBB climbing

(a) Yields



(b) Incomes

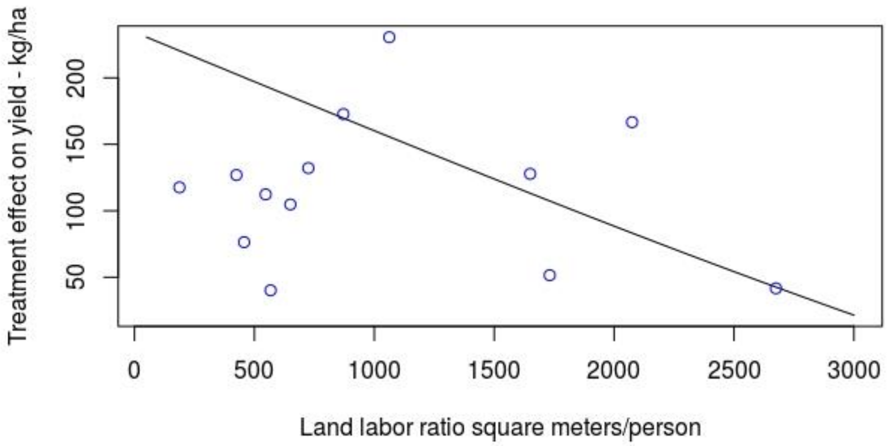


These findings suggest that the delivery of IBB seed was most effective in increasing the yields and incomes of farmers who were less likely to adopt IBB. This outcome is partially explained by the fact that extension agencies have been tailored by design to target smallholder farming households with less access to agricultural inputs, less technologically advanced, and less wealthy. The aim of the biofortification program is to reduce micronutrient deficiency, rather than maximizing adoption. A possible explanation for the low propensity of IBB adopters with relatively large yields and incomes might be related to the proximity of farming households with respect to markets and land-labor ratios. In Rwanda, IBB adoption has led to an increase in yields and an increase in incomes, thus creating a win-win situation in smallholding farmers. Our results of negative selection are similar to previous research that highlights the problem of negative selection. Verhofstadt & Maertens (2015) reported negative selection on cooperative memberships; they found the largest impacts among smallholder farming households with the lowest probability to become a cooperative member.

Another significant finding of this study relates to how treatment effect on yields and incomes varied over land-labor ratios. Figures 11 and 12 show these relationships for IBB bush growers and IBB climbing growers, respectively. I observed a negative relationship between land-labor ratios and outcomes. The treatment impact on yields and incomes was the highest for farming households with small land-labor ratios in both groups of IBB growers. Most evidence from within low-income countries suggests that agricultural productivity and scale are inversely related (Hazell, Poulton, Wiggins, & Dorward, 2010).

Figure 11 Heterogeneity of treatment effect of land-labor ratios over (a) yields and (b) farmers' incomes - IBB bush

(a) Yields



(b) Incomes

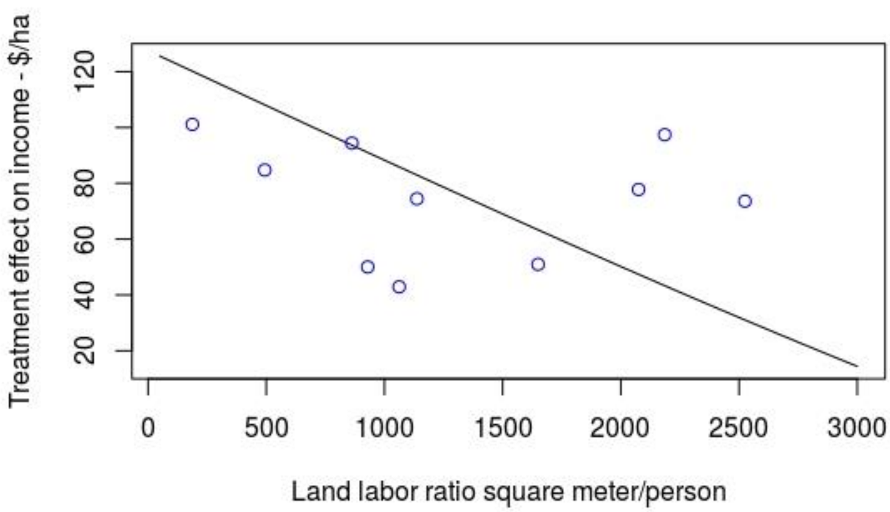
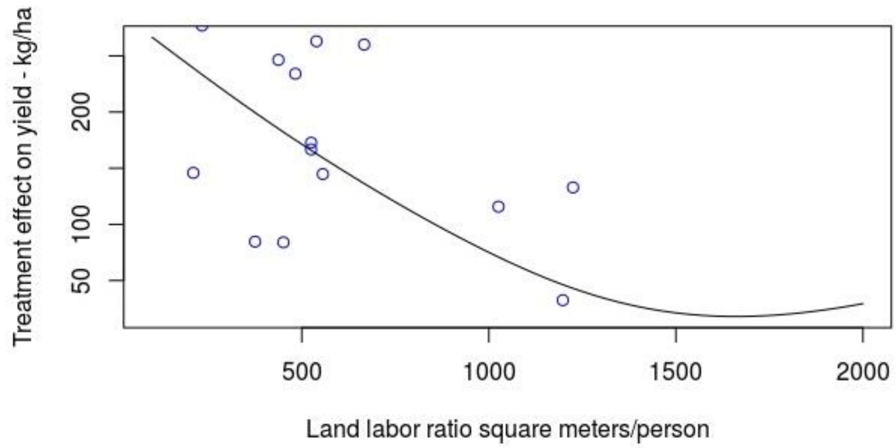
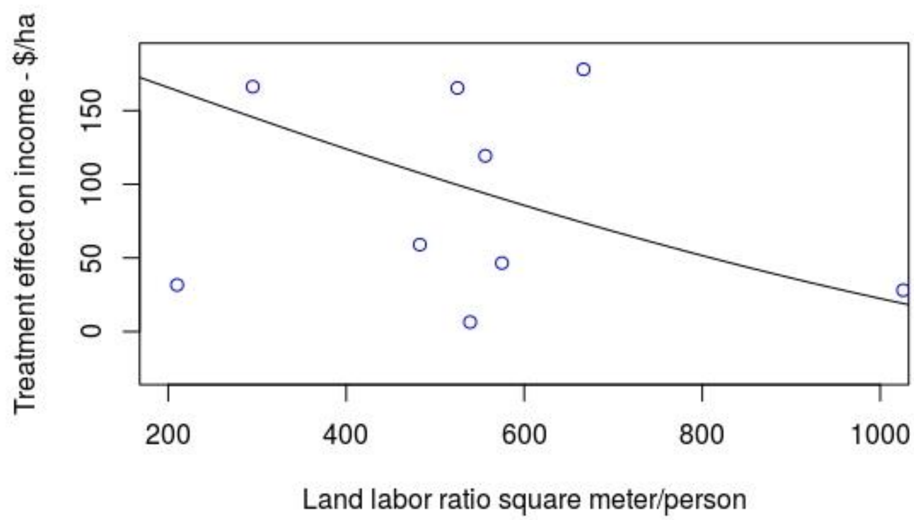


Figure 12 Heterogeneity of treatment effect of land-labor ratios over (a) yields and (b) farmers' incomes - IBB climbing

(a) Yields



(b) Incomes



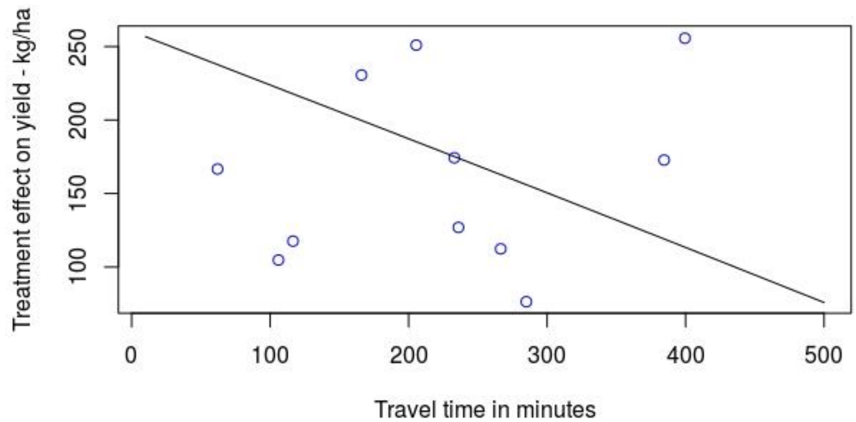
The literature has highlighted several explanations for the inverse relationship, including risk, labor market, and omitted variable bias. Ali & Deininger (2015) reported evidence of a strong inverse relationship between land-labor/farm size and productivity in Rwanda. They argued that labor market imperfections drive the inverse relationship, rather than other unobserved factors. Foster & Rosenzweig

(2017) reported that smallholder farms that are exclusively managed and run by family members have lower operational costs. Larson, Otsuka, Matsumoto, & Kilic (2014) underlying premise is that small farms are productive in the African context and they do not experience economies of scale. Given the consensus that smaller farms have a lower land-labor ratio than large farms, I could argue that small farms enjoy higher land productivity in the short-run. However, over the long-run, land productivity would tend to decrease given heavy cultivation of the land. This study provides evidence of this inverse relationship based on smallholder farming households in a similar environment, markets, and technology frontier.

In addition, it is important to ask how treatment effects on yields and agricultural incomes vary over travel time to markets. Travel time to the nearest market is a key factor for increasing agricultural productivity. This is because it determines farmers' physical accessibility to agricultural inputs and influences smallholder farming households' ability to sell crop surpluses. By way of illustration, in Sub-Saharan Africa, Dorosh et al., (2012) reported that agricultural production and adoption of technology were highly correlated with the proximity to urban markets. Figures 13 and 14 show how ATT varies over travel time to midsize towns. I found a negative relationship between travel time to markets and treatment effects over yields and incomes. These findings imply that the delivery of IBB seeds was most effective in increasing the yields of farming households located closer to midsize markets. Similarly, treatment impact over income was greater for farming households closer to markets than farming households in most rural areas.

Figure 13 Heterogeneity of treatment effect of travel time to markets over (a) yields and (b) farmers' incomes - IBB bush

(a) Yields



(b) Incomes

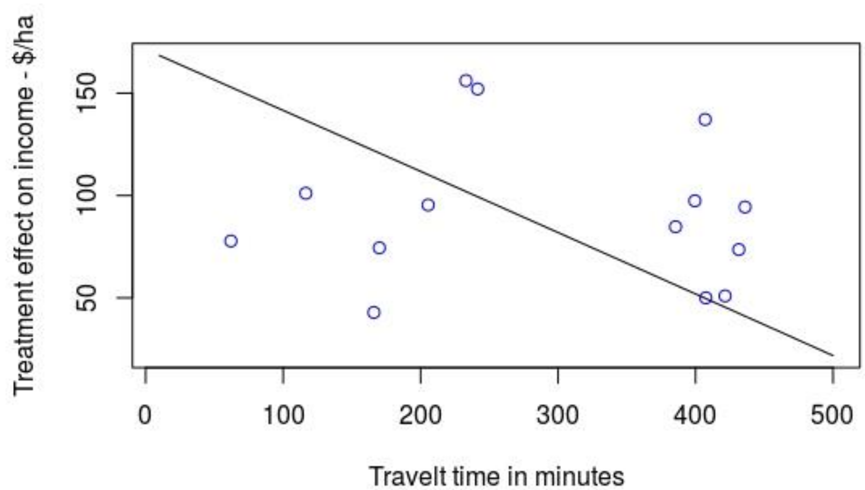
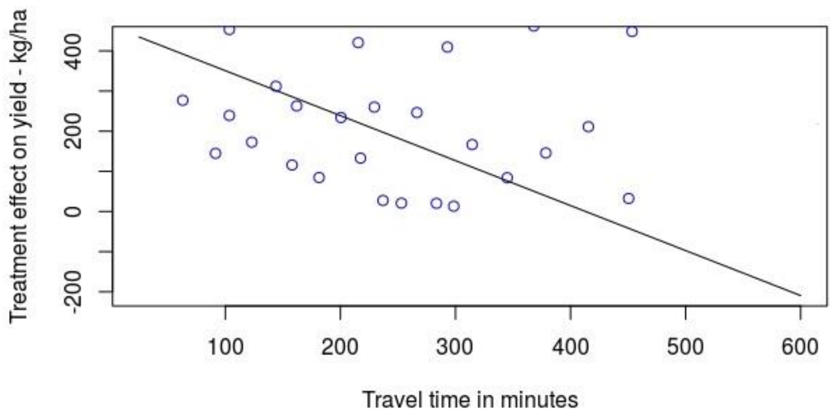
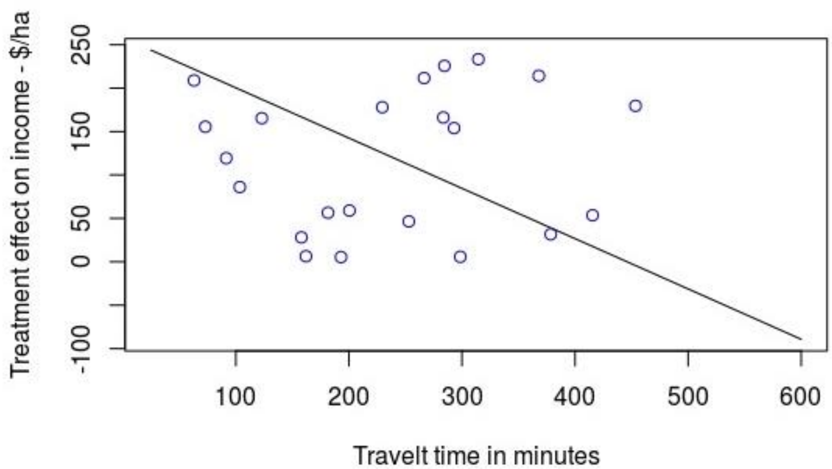


Figure 14 Heterogeneity of treatment effect of travel time to markets over (a) yields and (b) farmers' incomes - IBB climbing

(a) Yields



(b) Incomes



This combination of findings provides some support for the conceptual premise that: (1) farming households closer to markets might be able to trade beans at higher prices, allowing for higher income gains and lower transaction costs; and (2) farming households in remote areas might face higher transportation costs in order to trade their crop surpluses in larger markets. On the first premise, providing nutrition information about the IBB varieties, these varieties might receive a premium from consumers. The farming households that benefit from proximity to midsize towns tend to have smaller land-labor ratios.

Together, these results provide important insights into geographic targeting for introducing iron biofortified staple food crops. I found evidence that impact on smallholders' outcomes varied by their propensity to adopt IBB, land-labor ratios, and travel time to markets. The incidence of higher income along with bean yields increases as a farmer's propensity to adopt IBB decreases (negative selection), land-labor ratios decrease, and travel time to markets decreases.

iii. Policy Implications

On a policy level, this study provides positive empirical evidence that demonstrates the superior yield and potential income effects of growing biofortified beans as it supports the hypothesis that IBB growers have significantly higher yields than farmers that grew non-biofortified (improved or traditional) beans. This increases the food availability to vulnerable populations by boosting the adoption of higher yielding varieties. These findings provide evidence that supports continuing investment in agricultural research, such as the biofortification of staple food crops, as a cost-effective strategy to reduce micronutrient deficiency and to mitigate rural poverty in Rwanda.

I found evidence of negative selection in the incidence of higher yields and incomes increasing as farmers' propensity to adopt IBB decreases (negative selection). This is partially explained by the early delivery strategy of reaching the most vulnerable populations in remote and rural areas, which have little to no access to agricultural inputs, are less technically advanced, and comparatively considered to be less wealthy farming households.

I also examined the heterogeneous treatment effects on outcomes by propensity score and a set of covariates, key for program targeting and for policy formulation. Understanding the pattern of heterogenous treatment effects across the targeted population can help policymakers to effectively classify target geographic areas for investment. In addition to production, consumption, and micronutrient indicators, a subnational geographic index could factor in covariates. These could include travel time to urban centers as a proxy of delivery costs and consumer's propensity to adopt as a proxy to market segmentation. In summary, such an index could help inform future biofortification programs on how to tailor and assign different treatments to smallholder farmers with various characteristics by geography. Overall, the potential outcomes from tailored extension services include an increase in adoption rates, a strategy to maximize average outcomes (yields and incomes), and a strategy to reduce delivery costs.

Given the mounting evidence of, positive and significant nutrition, cognitive and physical performance impacts of IBB, the policies and investments for their scale up are expected to be beneficial for nutrition, for health, and for income outcomes.

IV. Conclusions

In this chapter, I contributed to the empirical literature on assessing the impact of improved and biofortified crop varieties. More notably, I examined the impact of iron biofortified bean (IBB) varieties on Rwandan smallholder farming households' livelihoods, focusing on the outcomes of yields and incomes for beneficiary households. It supports the hypothesis that IBB growers had significantly higher yields (23 percent for bush and 22 percent for climbing bean types), and potential incomes (24 percent for bush and 25 percent for climbing bean types) than farmers that grew non-biofortified (improved or traditional) beans. Our empirical analysis demonstrated the need to control for spatial spillovers which provide evidence that adopting households nearby increases a smallholder farming household's probability of adoption of iron biofortified beans.

Chapter 6: Efficiency analysis of smallholder bean farming households

I. Introduction

Agriculture is an important sector in Rwanda's economy. It accounts for 39 percent of gross domestic product (GDP) and 80 percent of employment (World Bank, 2013). Rwandans have one of the highest per capita bean consumption rates in the world, with rural households consuming beans on average six days in a given week (Asare-Marfo et al., 2016), and in significant quantities (Berti et al., 2012). Despite beans being naturally high in iron content, a significant proportion of the Rwandan population is at risk of iron deficiency, resulting in loss of economic development and growth. Beans have the highest share of the crop-harvested in Rwanda, though, there are significant productivity challenges. Limited access to agricultural technology, in particular seeds of improved varieties of beans and other complementary inputs such as fertilizers or staking material for higher yielding varieties like climbing beans can partially explain the country's low productivity issue. Therefore, there are significant yield gains to be made from the introduction and scaling up of seeds of improved varieties of beans. Iron biofortified beans (IBB) can not only help improve yields and hence incomes, but also nutrition and health outcomes (e.g., cognitive) and physical functions of consuming populations.

Economic indicators of performance such as measures of productivity and efficiency are commonly used to investigate the impact of a new technological-innovation on farmers' outcomes (Duflo et al., 2008). Assessing farmers' efficiency defined as the ability of farmers to utilize the best available technology and to allocate resources productively, together with the impact of an intervention requires the combines application of analytical methods (e.g., for earlier examples (Bravo-Ureta et al., 2012; Dinar et al., 2007).

When conducting efficiency analysis on a cross-sectional or a panel dataset, a high degree of heterogeneity may lead to biased and inefficient estimates of the efficiency scores. Recent literature has approached this problem in different ways. One way is using non-parametric techniques, such as data envelopment analysis (DEA), which ignore the functional form of the production function. Other studies have implemented a two-step approach: the first step estimates the frontier and the second step analyzes the determinants exerting influence over economic agents' efficiencies (Chavas et al., 2005; Simar & Wilson, 2007). Greene (2008) proposed the true-fixed effects and the true-random effects models for panel data. When there is spatial heterogeneity, instead of including spatial fixed effects, some authors allow the externalities to spill over throughout the system, (Han, Ryu, & Sickles, 2016b). In this chapter, I implement the latter method through a three-pronged approach.

The first prong assesses the importance of social networks in the adoption and diffusion of the technology in question (i.e., IBB). The second prong uses a multivariate matching algorithm method to measure the impact of growing IBB on farmers' bean yield and bean income. This second prong also produces a mechanism for controlling for observable heterogeneity and for producing an unbiased subsample for the third prong, namely the technical efficiency analysis. In this latter prong, spatial stochastic frontier models are fit to each of the data groups (control, treated, and pool).

This chapter contributes to the research literature by presenting a case study on an intervention that delivered a new technology, IBB, in Rwanda. I developed a multi-pronged approach and applied it to a cross-sectional, nationally representative data of bean farmers in Rwanda to estimate farmers' unbiased efficiency scores. I combine spatial econometrics and quasi-experimental methods to estimate a national technological frontier for all bean farmers, a frontier for IBB growers, and a frontier for farmers that grow other improved or traditional bean varieties. I compare standard stochastic frontier models to spatial stochastic frontier models, which help us estimate efficiency spillovers among IBB growers (treatment) and others (control). Clustering analysis produces evidence on where and how this new technology has been effective, thereby providing valuable input into targeting strategies and resource allocation for scaling up of such interventions. By completing the above three-pronged procedure, I achieve objective three of this dissertation:

Objective 3: to estimate the impact of IBB on smallholder farming households' efficiency.

II. Results and discussion

i. Bush bean farmers

Table 15 presents the results of the fitted models for the pool, treated (bush IBB growers), and control (other bush bean growers) data groups. The starting point of the model is the production function estimated with OLS. To this end, I test whether the technical efficiency of farming households is stochastic. I also report the LM test statistics to determine which spatial terms are appropriate. For the former, the highly significant (p=0.001) likelihood test confirms the presence of technical inefficiency LR=$-2 * [lnL_{ols} - lnL_{sfa}]$, while the LM tests suggested different spatial specifications.

As Table 15 shows, in the OLS-pool model the lack of significance of the IBB parameter suggests that there is no significant difference between the two groups of bean farmers: treated and control. However, the LR test (p=0.05), estimated as $LR = 2 * [lnL_{pool} - (lnL_{IBB} + lnL_{control})]$ rejects the null hypothesis of equality of the parameters across the treatment and control groups. Therefore, I estimated separate technology frontiers for each data group. The bottom of Table 15 shows the results of the spatial specification of the LM test statistics. The LM lag test favors the spatial autoregressive model for the pool and for the control data groups while the LM error test favors the spatial autoregressive error model for the treated group. The rho and lambda parameters were significant, justifying the need to use spatial econometrics. It is worth highlighting that the parameter rho (0.115) or global spatial multiplier is significant for the spatial stochastic frontier model - pool (SSFA-pool). This parameter reveals the link or spillover effect in the system between treated and control groups. Bean farmers who did not grow IBB got an indirect benefit of 12 percent in terms of total bean production because they interact with neighbors who grew IBB. In addition, the IBB adoption parameter in both the stochastic frontier analysis - pool (SFA-pool) and the SSFA-pool models is positive and significant (p<0.10).

With respect to the pool specifications of the SSFA; bean area, hired labor, and economic active population positively contributed the most to bean production. For the control group (SSFA-C), bean area and hired labor positively contributed the most, whereas for the treatment group (SSFA-T), hired labor, seed, and bean area had a positive significance and the most influence on bean production. Interpretation of the coefficient requires the estimation of direct and indirect impact effects, which I do not discuss in this chapter to maintain its analytical focus. The sum of all partial production elasticities is larger than 1, suggesting bean farmers are not operating at an efficient scale. Scale efficiency (SE) analysis indicates farms of sizes smaller than SE are too small as they exhibit increasing returns to scale.

In the last two rows of table 15, I include two summary measures of economic performance: relative efficiency and relative efficiency weighted by bean production (kg). In the SSFA-pool model, the relative efficiency of bean farmers estimated was 0.18 while the relative efficiency weighted by bean production was about 0.361. In addition to having a higher relative efficiency, IBB bush growers had a 3 percent higher output than bean farmers with a similar level of inputs growing other bean varieties.

Table 15 Estimation results of an OLS, a non-spatial frontier analysis (SFA) model and a spatial frontier model (SSFA): pooled, control (C) and treatment (T) bush bean growers

Variable	OLS - pool (1)	SFA-pool (2)	SSFA-pool (3)	OLS-C (4)	SFA-C (5)	SSFA-C (6)	OLS-T (7)	SFA-T (8)	SSFA-T (9)
Log(Bean area)	0.514***	0.641***	0.522***	0.582***	0.747***	0.578***	0.447***	0.483***	0.466***
	(0.080)	(0.073)	(0.078)	(0.121)	(0.110)	(0.117)	(0.104)	(0.088)	(0.101)
Log(Management index)	0.801	0.317	0.818	0.993	0.567	1.010	0.381	-0.179	0.179
	(0.537)	(0.462)	(0.524)	(0.779)	(0.515)	(0.748)	(0.752)	(0.612)	(0.732)
Log(Economic active population)	0.541**	0.554**	0.507**	0.509	0.243	0.394	0.674*	1.038***	0.714**
	(0.241)	(0.203)	(0.236)	(0.363)	(0.293)	(0.358)	(0.322)	(0.334)	(0.304)
Hired labor (dummy)	0.627***	0.441***	0.608***	0.445	0.443**	0.466	0.830***	0.642***	0.802***
	(0.181)	(0.150)	(0.176)	(0.304)	(0.207)	(0.292)	(0.222)	(0.192)	(0.205)
Log(seed)	0.150	0.278	0.152	-0.180	-0.069	-0.138	0.768**	0.797***	0.781***
	(0.206)	(0.193)	(0.201)	(0.289)	(0.276)	(0.277)	(0.302)	(0.283)	(0.282)
IBB adoption	0.278	0.238*	0.282*						
	(0.172)	(-0.132)	(0.168)						
Constant	2.300	3.597***	-2.770	3.034***	4.700***	2.449***	1.259*	2.194***	0.040
	(0.534)	(0.509)	(1.740)	(0.749)	(0.684)	(0.769)	(0.719)	(0.677)	(0.897)
Observations	220	220	220	110	110	110	110	110	110
Rho			0.115*			0.185*			0.27
Lmerr	1.385			1.456			6.123***		
Lmlag	4.351			3.321*			1.101		
Log Likelihood	-355.2678	-325.977	-353.676	-184.889	-166.765	-183.189	-165.042	-153.444	-162.920
LR Test			3.183*			3.401*			4.243**
TE			0.179			0.169			0.212
Average TE			0.361			0.333			0.363

Notes: () are standardized errors. *, ** and *** denote significance at the 10%, 5%, and 1% level, respectively

ii. Climbing bean farmers

As table 16 shows, the two pooled models, SSFA and OLS, suggest that there are no significant differences in output between IBB adopters and non-adopters. However, the LR test (p=0.01) provides strong evidence for the estimation of separate technologies for each data group. The LMlag test statistics are significant, suggesting the presence of spatial spillovers. The rho (0.09) parameter for the pool (SSFA-pool) was significant, revealing the presence of spatial spillovers between the treated and control groups, i.e. climbing bean growers who did not grow iron biofortified varieties got an indirect benefit of 9 percent in terms of total bean production. The sum of all partial production elasticities is larger than 1 which indicates increasing returns to scale in all models. Given increasing returns to scale, climbing bean growers are not operating at an efficient scale.

The most influential parameters for both the, SFA-pool and the SSFA-pool models were the management index, bean area, and seed. For the SFA control (SFA-C) model, bean area and economic active household members were the most influential parameters. The most influential parameters for the SSFA control (SSFA-C) model were the management index and the bean area. The SSFA treatment (SSFA-T) model shows something new compared to the control model. After the management index, the seed parameter shows a positive and large influence in households' bean production. The management index indicates whether the farmers have irrigation systems, terraced plots and whether they apply pesticides, compost, manure, and fertilizers on their plots. Farmers that practice better management activities on their farms (i.e., those that have higher management index scores) exhibit higher efficiency. The results indicate that, when controlling for the total bean area, there is a statistically significant and positive relationship between plot size and output of bean production. The quantity of seed used was positive and significantly ($p=0.001$) correlated with the output level. This suggests that the treatment group had access to new technologies, such as IBB planting material, and coupled with other agricultural inputs used by bean farmers within this group, managed to have higher bean production.

In table 16, I observe the relative efficiency of bean farmers estimated with spatial pool stochastic production frontier was 0.262 while the relative efficiency weighted by bean production was 0.395. IBB climbing adopters reported higher relative efficiency than the control group of bean farmers. In addition to having higher relative efficiency, IBB climbing growers had 13 percent higher output than bean farmers with similar level of inputs who grew other bean varieties.

Table 16 Estimation results of an OLS, a non-spatial frontier analysis (SFA) model and a spatial frontier model (SSFA): pooled, control and treatment climbing bean growers

Variable	OLS-pool (1)	SFA-pool (2)	SSFA-pool (3)	OLS-C (4)	SFA-C (5)	SSFA-C (6)	OLS-T (7)	SFA-T (8)	SSFA-T (9)
Log(Bean area)	0.557***	0.611***	0.549***	0.702***	0.729***	0.674***	0.380***	0.484***	0.371***
	(0.051)	(0.052)	(0.050)	(0.065)	(0.066)	(0.062)	(0.081)	(0.080)	(0.074)
Log(Management index)	1.131**	1.044***	1.132***	0.706	0.923	0.593	1.652	1.385***	2.012***
	(0.418)	(0.396)	(0.409)	(0.662)	(0.605)	(0.664)	(0.559)	(0.497)	(0.506)
Log(Economic active population)	0.217	0.257*	0.237	0.414*	0.433**	0.445**	0.046	0.080	-0.001
	(0.175)	(0.150)	(0.172)	(0.232)	(0.211)	(0.214)	(0.257)	(0.191)	(0.233)
Hired labor (dummy)	0.081	0.048	0.084	-0.060	-0.070	-0.069	0.119	0.092	0.151
	(0.117)	(0.104)	(0.114)	(0.167)	(0.151)	(0.161)	(0.170)	(0.138)	(0.161)
Log(seed)	0.564***	0.500	0.549***	0.400*	0.357*	0.403**	0.900***	0.856***	0.931***
	(0.152)	(0.143)	(0.149)	(0.188)	(0.182)	(0.234)	(0.253)	(0.240)	(0.234)
IBB adoption	0.174	0.202**	0.170						
	(0.116)	(0.132)	(0.114)						
Constant	2.396	3.497***	2.013***	2.746***	3.544***	2.728***	1.908**	3.189***	1.768***
	(0.437)	(0.436)	(0.472)	(0.543)	(0.556)	(0.513)	(0.643)	(0.568)	(0.592)
Observations	242	242	242	121	121	121	121	121	121
Rho			0.091*			0.196**			0.236*
Lmerr	3.094*			4.217			0.943***		
Lmlag	2.339			8.025*			3.482*		
Log Likelihood	-303.537	-291.302	-302.04	-140.365	-137.273	-138.019	-155.545	-147.84	-153.425
LR Test			2.983*			4.692**			4.240**
TE			0.262			0.239			0.299
Average TE			0.395			0.340			0.472

Notes: () are standardized errors. *, ** and *** denote significance at the 10%, 5%, and 1% level, respectively

iii. Truncated – second stage analysis

In contrast to the input variables used in the first stage to estimate bean farmers' relative efficiency, the covariates listed in table 17 are sources of inefficiency. Proximity to extension services is an indicator of regional endowment. Bean farmers closer to extension services had higher efficiency than those farmers farther away. The more distant farmers may incur higher transportation costs to access basic agricultural inputs and advice from extension services. Bean farmers with lower transportation costs are more efficient at connecting to markets to meet the demand for staple food crops in the local market. On this last point, proximity to towns with a population greater than 50,000 people was not significant. However, univariate statistical analysis shows a positive correlation for bean farmers' efficiency near local markets.

IBB adopters that sold their IBB surpluses in local markets show a strong positive effect in their relative efficiency. This suggests that bean farmers are likely to increase their participation, as sellers of staple food crops, in functioning markets that give them appropriate incentives to increase their agricultural income. Overall, access to IBB varieties brings about a significant change in crop production efficiency, which in turn improves farm households' income.

I found crop specialization is associated with higher efficiency. The parameter of crop diversification (i.e., the number of crops grown) was negative and significant. This evidence suggests that bean farmers' efficiency increases as they cultivate fewer crops. However, in Rwanda bean farmers grow up to six crops. The fact that most farms are dual or multi-crop farms suggests that the benefits of diversification are significant in Rwandan agriculture, which I do not test in this study. These benefits could manifest in two ways: the presence of economies of scope, reflecting the reduced costs associated with producing multiple outputs; and the risk-reducing effects of diversification (Chavas et al., 2005). From a nutritional security and suitability (transaction cost) point of view, diversification can be efficient and sustainable in agriculture, but might have an adverse impact on yields of any one crop.

The drought index (-1 to 1) parameter is also significant and has the most influence on farmers' efficiency (magnitude of the coefficient). Negative values indicate areas affected by droughts and positive values indicate normal level of rainfall. The drought index helps to evaluate farmers' ability to manage weather shocks. Households in areas less prone to droughts witnessed higher efficiency. This might be associated with farmers' efforts to adapt to climate change by adjusting their farming management plans.

Table 17 Truncated analysis of bean farmers' TE

Variables	Estimate
Travel time to agrotechnical services	-0.013
	(0.007)
Crop count index	-0.076*
	(0.039)
Drought index	0.388**
	(0.155)
Bean farmers link to markets	0.149***
	(0.051)
Intercept	0.072
	(0.051)
Log-Likelihood	289.10
AIC	-566.00

Notes: () are standardized errors. *, ** and *** denote significance at the 10%, 5%, and 1% level, respectively.

iv. Clustering analysis

Figure 15 shows that the Moran's I test (0.2839) confirms the presence of global spatial autocorrelation (pseudo-p < 0.001 randomized with 999 permutations). The scatter plot shows the values of a given location (x-axis) against the values of its neighbors (y-axis). The units are in standard deviations. Figure 16 splits the global spatial association into five observation groups in three categories:

- no significant spatial association;
- two sets of clusters, (2) hot spots (high-high) and (3) cold spots (low-low); and
- two sets of spatial outliers: (4) low-high and (5) high-low.

Figure 15 Moran's I bean farmers' technical efficiency

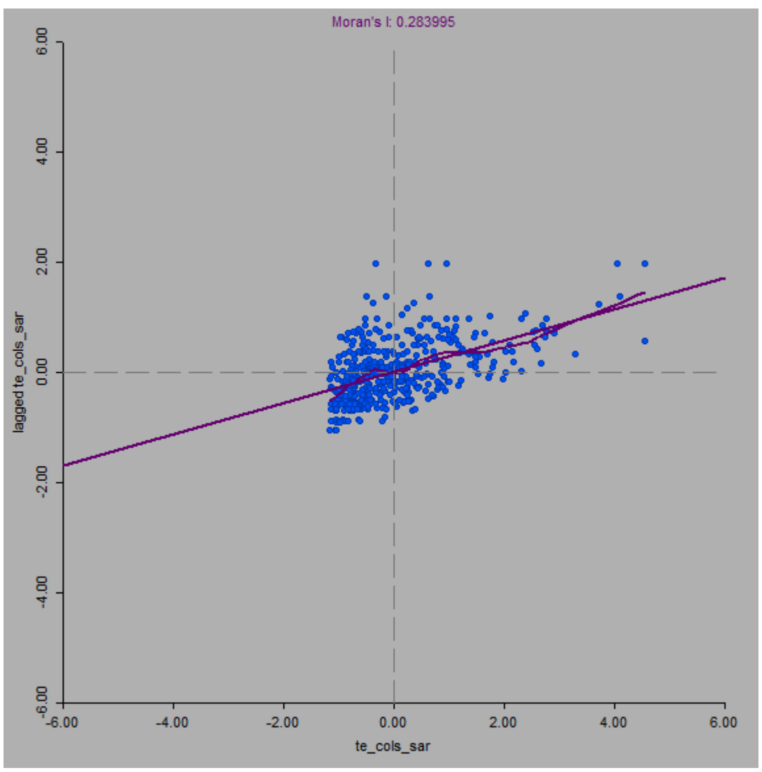
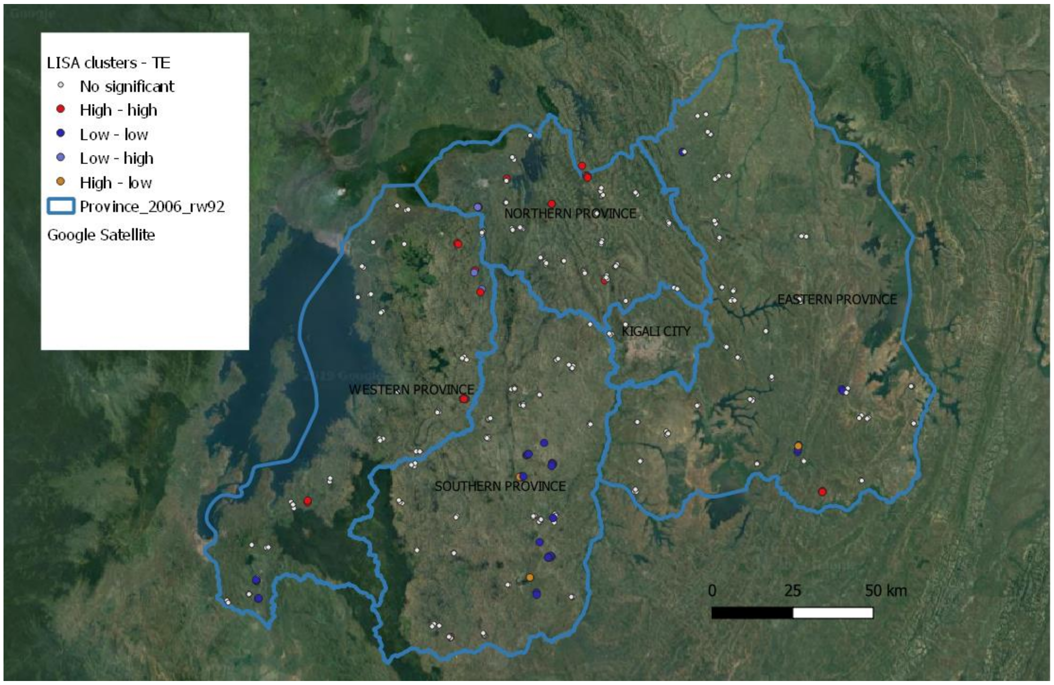


Figure 16 LISA cluster map of farmers' technical efficiency



The second cluster category in the legend in figure 15 are hotspots of bean farmers with high efficiency scores surrounded by farmers with similar efficiency scores. These hotspots, depicted in red, are found in the neighboring districts of Burera, Gakenke, and Musanze in the Northern region and in the districts of Nyabihu and Ngororero in the Western region. More hot spots are observed in the neighboring districts of Rwamagana and Bugeresa in the Eastern region. Hotspots of smallholder bean farming households are associated with attributes, such as households with a higher wealth index, a higher level of education, and a higher management index. In addition, these households tend to be in less vulnerable areas to droughts and sell their crop surpluses to local markets. The opposite hold for cold spots. Cold spots, the third cluster category (depicted in dark blue) of bean farmers with low efficiency scores are observed in the neighboring districts of Ruhango, Nyanza, and Huye in the Southern region and in the districts of Gatsibo and Nyagatare in the Eastern region.

The last two cluster categories are spatial outliers. The first set refers to bean farmers with low efficiency scores surrounded by farmers that are more efficient, depicted in light blue, while the second set of outliers reflect the opposite-farmers with high efficiency surrounded by farmers that are lower efficiency scores, depicted in yellow. Identification and training of farmers with high technical efficiency scores who are in closer proximity to those with lower technical efficiency scores could be a cost-effective strategy for the diffusion of new technologies, such as IBB varieties.

III. Conclusions

In this chapter, I used a multi-pronged approach to estimate the relative efficiency of bean farmers in Rwanda to provide key policy recommendations, as well as support country-program implementation. To control for self-selection, a multivariate matching approach was used to create two comparable matched groups: a treatment group, which included farmers who grew iron biofortified bean in 2015 Season B and a control group who grew other bean varieties. Then, a non-spatial stochastic frontier model and a spatial stochastic frontier model were fit for each of the three data groups: pool, control, and treated. In a second stage analysis, I fit a truncated model to validate and explore the sources of inefficiency. This chapter contributes to the literature by controlling for self-selection bias, missing counter-factual, and spatial spillovers. These analyses were conducted separately for climbing bean farmers and bush bean farmers. Results reveal that farming households who grew IBB varieties were more technically efficient and generated higher total bean production. Overall, the analysis shows a (positive) technology spillover is indeed an improvement in the use of the existing technology.

IBB bush and climbing growers had higher relative technical efficiency and higher total bean production than their counterpart in the control group of bean farmers. However, most bush and climbing bean farmers are not operating at an efficient scale. The sum of all partial production elasticities is larger than 1 which indicates increasing returns to scale in all models. In the second stage analysis, I discovered IBB growers that had market linkages were in areas less vulnerable to weather shocks and had better access to extension services. Based on these factors, they were more likely to exhibit higher relative efficiency scores. Clustering analysis helped to identify hotspots of bean farmers with high and low efficiency. Farmers' efficiency could be increased given the current state of technology and it is possible to reach out to them through targeted, tailored-made interventions. These analyses provide valuable input into targeting strategies and resource allocation for scaling up of such interventions.

Chapter 7: Conclusion

Summary

Micronutrient malnutrition affects key development outcomes, including physical and mental development in children, vulnerability to disease, blindness, and general losses in productivity. In Rwanda, micronutrient malnutrition is highly pervasive, adoption rates of improved varieties of staple crops are low, and bean farmers' productivity levels are low. Since 2012, HarvestPlus and its partners have been intensively disseminating common iron biofortified bean (IBB) varieties to help alleviate iron deficiency in Rwanda. In this dissertation, I analyzed smallholder farming households' decisions to adopt these newly-released IBB varieties by specifically examining the influence of demand-side factors and the role of peers. To do so, I drew upon several theories from studies on the adoption of agricultural technology, social behavior, and utility maximization to test three hypotheses. For the first hypothesis, I tested how the adoption behavior of smallholder farming households would be influenced by their neighbor's adoption outcomes, as a result of peer learning about the profitability or the appropriate use of IBB. This phenomenon is known as endogenous effect, which is described as imitation, contagion, bandwagons, and social norms. For the second hypothesis, I modeled the effect of contextual factors, wherein the propensity of an IBB grower to behave is correlated with the exogenous characteristics of his/her neighbors. For the third hypothesis, I ran a new set of regressions with fixed and random effects. The model treats observations from a given village as a cluster and assumes a random effect for each village. As expected, the random effect produced a weaker spatial relationship. This result confirms the hypothesis that closer neighbors matter more than those farther away. These analyses all together estimate the prevalence of IBB adoption by district.

From a broader analytical perspective, the second part of this dissertation aimed at answering what was the economic impact of the biofortification program in Rwandan smallholder farming households' outcomes: yields and incomes – a question that links the program delivery efforts and its economic impact on beneficiaries. In this analysis, I tested the null hypothesis of absolutely no effect on IBB adoption for any smallholder farming household. The results indicate that the adoption of improved IBB increased smallholder farming households' yields and incomes, which may translate into improved iron intakes, and higher levels of market sales and hence subsequent income gains.

The third part of this dissertation compares technical efficiency across IBB growers and non-IBB growers. After controlling for biases from observed and unobserved variables, IBB bush growers had a 3 percent higher output than non-IBB growers with a similar level of inputs. IBB climbing growers had a 13 percent higher output than bean farmers with similar levels of inputs who grew other bean varieties. In addition to having higher relative efficiency, the empirical results suggest that the frontier for IBB bush growers and the frontier for IBB climbing growers are located above the ones for the non-IBB grower groups.

i. Research Objectives and Outcomes

The first objective of this dissertation aimed to answer a question that relates to the risk a smallholder bean farming household faces in adopting IBB. To do so, I used the von-Neuman-Morgenstern utility framework. I approached this theoretical framework by applying spatial econometric techniques to estimate neighborhood influence and to determine the factors driving the adoption of IBB. To estimate the strength of social interactions and the robustness of the results, I used different techniques including a spatial autoregressive probit model and a spatial Durbin model. The results show that the adoption of both bush and climbing IBB varieties exhibit positive spatial autocorrelation and spatial spillovers. This indicates that geographic diffusion of iron bean planting material occurs among neighboring farmers that exhibit interdependent decision-making patterns, as well as similar characteristics relative to the group.

This dissertation contributes to the research literature in two ways. First, it provides spatially granular statistics on the intensity of adoption and adoption rates of IBB by bean type. Second, the paper examines, with the assistance of spatial econometrics techniques, theories of social interaction, and choice behavior, how household and farm characteristics, the role of peer influence, as well as regional factors, influence smallholder farming households' decisions to grow IBB.

The second objective of this dissertation aimed to inform policy makers of the economic impact of IBB on smallholder farming households’ yields and incomes in Rwanda. To do so, I analyzed the impact of IBB on smallholder farming households’ yields and incomes. Using observational studies and spatial econometrics methods, this dissertation estimated the treatment effect and heterogeneous impact of IBB production on farmers’ yields and potential incomes.

This dissertation adds to the empirical literature on assessing the impact of improved and biofortified crop varieties. I examined the impact of IBB varieties on

Rwandan farmers' livelihoods, focusing on the outcomes of yields and incomes for beneficiary households. It supports the hypothesis that IBB growers had significantly higher yields (23 percent for bush and 22 percent for climbing bean types), and potential incomes (24 percent for bush and 25 percent for climbing bean types) compared with farmers that grew non-biofortified (improved or traditional) beans.

The third objective complements the second objective at estimating the impact of IBB on smallholder farming household's technical efficiency in bean production in Rwanda. This analysis provides key policy recommendations, as well as supports country-program implementation. To control for self-selection, a multivariate matching algorithm was used to create two comparable matched groups: a treatment group, which included farmers who grew iron biofortified bean in 2015 Season B and a control group who grew other bean varieties. Then, a non-spatial stochastic frontier model and a spatial stochastic frontier model were fit for each of the three data groups: pool, control, and treated. In a second stage analysis, I fit a truncated model to validate and explore the sources of inefficiency.

This analysis contributes to the literature by controlling for self-selection bias, missing counter-factual, and spatial spillovers. These results provide less biased technological frontiers including a national technological frontier for all bean farmers, a frontier for IBB growers, and a frontier for farmers that grow other improved or traditional bean varieties. Results reveal that farming households who grew IBB varieties were more technically efficient and generated higher total bean production. Overall, the analysis shows a (positive) technology spillover, which is indeed an improvement in the use of the existing technology.

ii. Research Insights

This dissertation provides recommendations for policy, research, and practice. Research and policies need to be adapted in order to increase the uptake of IBB varieties and to optimize the delivery of agricultural technology to the most vulnerable rural farming households in Rwanda.

On a policy level, this study provides positive empirical evidence demonstrating the superior yield and potential income effects of growing biofortified beans. As such, this study supports the hypothesis that IBB growers have significantly higher yields than farmers that grew non-biofortified (improved or traditional) beans. This strategy proved to increase food availability in vulnerable populations by boosting the adoption of higher yielding varieties. These findings provide evidence to support continuing investment in agricultural research—such as biofortification of staple food crops—as a cost-effective strategy to increase food availability, as well as to mitigate rural poverty in Rwanda.

In terms of research and practice, this dissertation examined the heterogeneous treatment effects on outcomes by the probability of IBB adoption and a set of control covariates—key for program targeting and for policy formulation. Understanding the pattern of heterogeneous treatment effects across the targeted population can help to effectively classify target geographic areas for investment. In addition to production, consumption, and micronutrient indicators a subnational geographic index could factor in covariates—such as travel time to urban centers as a proxy of delivery costs and consumers' propensity to adopt as a proxy of market segmentation. This composite summary geographic index could help inform biofortification programs on how to tailor and assign different treatments to smallholder farming households with various characteristics by geographies. Potential outcomes from tailored extension services include a strategy to increase IBB adoption rates, a strategy to maximize average outcomes (yields and incomes), and a strategy to reduce delivery costs.

Priority should be given in drafting a differentiated geographical targeting strategy for bush and climbing bean varieties as a function of farmer characteristics, farm characteristics, the role of peers, and regional endowment of resources. This might help to increase adoption rates for the most vulnerable groups in rural areas. If schooling augments the propensity to adopt climbing IBB varieties, increasing educational levels on the nutritional and agronomic benefits of IBB might be an effective policy to stimulate knowledge and technological diffusion. Considering progressive farmers and strengthening social group activities when designing technology-promotion programs increases the spread of information and geographical diffusion of IBB. In sum, social interaction works as an invisible cost-effective delivery strategy of IBB in rural population that are likely at risk of iron deficiency. These activities might continue to support scaling up of delivery activities for IBB varieties in rural areas of Rwanda.

Given the mounting evidence pointing to positive and significant nutritional benefits, as well as enhanced cognitive and physical performance resulting from consumption of iron biofortified beans, policy makers should support greater investment and scaling up of biofortification technologies—as it has proven to be a powerful tool to improve lives and livelihoods of smallholder farming households.

Appendices

I. Construction of the wealth and management indices

Multiple correspondence analysis (MCA) was used to create two composite indices: the wealth index and the management index. The first composite index aims to measure household wealth in the absence of data on household income. The second aims to summarize farmers' management practices. The former includes household, livestock, and agricultural assets. The later includes management practice activities, including whether the farmers have irrigation systems and apply pesticides, compost, manure, and fertilizers on their plots. Construction of the management index helped to control for multi-collinearity.

By construction, the wealth and management indices range from 0 to 1: for the wealth index, it is equal to 0 in case of no assets and for the management index, it is equal to 0 in case of no modern agricultural practices. It is equal to 1 when a household own all assets considered in constructing the wealth index or equal to 1 when a household wholly uses modern agricultural practices.

In a nutshell, MCA is the application of a correspondence analysis algorithm to multivariate categorical data coded in the form of an indicator matrix (binary coding of the factors) that consists of an individual × variables matrix, where the rows represent individuals and the columns represent categories of the variables (Asselin 2009). For instance, the wealth index ranks households from poorest to wealthiest. Each household is given an individual wealth index, summarized below.

The functional form to build a composite indicator is as follows:

$$\text{CWIi} = \frac{1}{k}\sum_{k=1}^{k}\sum_{Jk=1}^{Jk} W_{Jk}^{k}\, I_{j_{k}i}^{k}$$

$$W_{j_k}^{k} = \frac{s^k}{\sqrt{\lambda_1}},$$

where k is the number of dimensions (variables) with k = (1, 2, …, K), j is the number of modalities of each dimension with j = (1, 2, …, Jk) and I is the binary (0/1) indicator of each modality. W is the weight determined with MCA (the factor score on the first axe normalized by the eigenvalue λ with s = factor score), and I is the index number indicating household.

The composite indicator is the simple average across dimensions (variables) of the weighted sum of each binary modality of each dimension.

I*p* = binary indicator 0/1; 1 indicates household *h* has the modality, otherwise 0.

W_i = the average global welfare of household *h*.

There are three categories included in the wealth index: household assets, livestock assets, and agricultural assets. Household assets include land, houses, motorcycles, bicycles, cells, radios, TVs, saving accounts, and savings in informal groups. Livestock assets are sheep, goats, cattle, pigs, chicken, rabbits, etc. Agricultural assets include ploughs, wheelbarrows, machetes, shovels, picks, and sprayers.

Variables included in the management index are whether farmers have a terraced plot or irrigated plot, and whether or not they apply fertilizers, manure, compost, and pesticides.

II. Travel time

Travel time is estimated using an algorithm that factors in road quality, speed, slope, and biophysical characteristics. Thirty percent of the population of Rwanda lives within 2 hours of travel time to midsize cities or cities equal or greater than 50,000 inhabitants while 70 percent lives at least 3.5 hours travel time. The latter remote areas are where the most vulnerable population lives. According to the World Bank's poverty assessment report in Rwanda, poor households tend to be more isolated, living at greater distances from markets, public transport facilities, and health centers. Based on the travel time index, the World Bank report finds that the most isolated households are twice as likely to be poor relative to the most connected households.

Table 18 Summary statistics travel time in hours, population (thousands), bean production (thousands), and bean availability in 2015, by travel time decile

TT deciles	Travel time, hrs.	Population	Share population	Bean production, tons	Bean availability, kg/person
1	1	842	7	0	0
2	2	1222	10	10	8
3	3	992	8	7	7
4	4	1318	11	13	10
5	4	1010	9	13	13
6	5	1105	9	18	16
7	5	1550	13	23	15
8	5	1043	9	22	21
9	6	1134	10	22	20
10	6	1486	13	53	35
Average/total	4	11703	100	181	15

III. Bush bean growers, matching members of treatment and control groups

Table 19 shows the results of the probit model for IBB bush growers.

Table 19 Probit model IBB bush

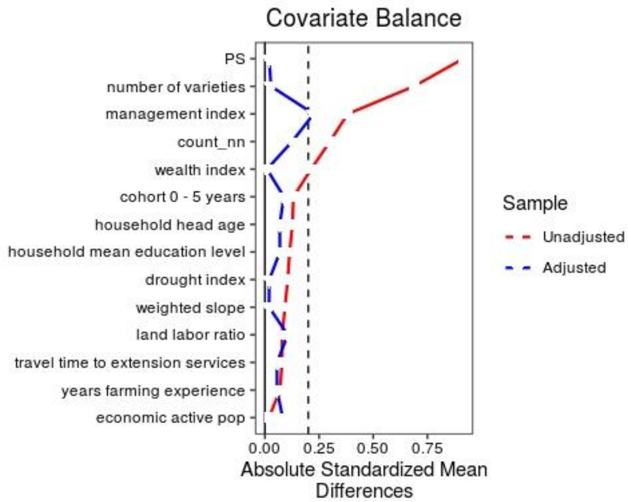
Variable	Coefficient
Constant	-1.02***
	(0.36)
Number of varieties cultivated	0.29***
	(0.04)
Management Index (0-1)	0.78**
	(0.31)
Wealth Index	0.03
	(0.48)
Number of children under 5 years old in household	-0.13
	(0.10)
HH average years of schooling	-0.02
	(0.04)
Drought index	-0.48
	(0.40)
Land labor ratio	-0.00
	(0.00)
HH head age	-0.02**
	(0.01)
Farming experience (years)	0.02**
	(0.01)
Travel time to extension services	0.01
	(0.02)
Weighted slope	-0.00
	(0.01)
Number of individuals per household - economically active members [18-65]	0.09
	(0.07)
% corrected predictes	75.89
Log-likelihood	-258.78
LR test	88.68***
Kelejian-Prucha (error)	2.73***
Pinske (error)	6.96***
Pinske -Slade (error)	5.71**
Number of observations	506

Multivariate matching

Figures 17 and 18 display the comparability of the two groups – treatment and control – with respect to the fourteen baseline covariates for bush type IBB growers. Careful examination of figure 17 allows one to visually assess the performance of the nearest neighbor algorithm to balance each of the baseline covariates, while examination of figure 18 allows us to see the distribution of PS for the treatment and the control groups.

The y-axis lists the name of the fourteen baseline covariates and the x-axis shows the absolute standard difference of means (ADM). The black dashed line shows an ADM benchmark set to 0.20 as suggested by literature (Rosenbaum, 2005). The red line shows the ADM for each of the baseline covariates of the unadjusted dataset, while the blue line shows the ADM for each of the baseline covariates of the adjusted dataset. In the red line, I observe that five out of the fourteen baseline covariates have standardized differences that exceed the 0.20 threshold, in descending order: PS, number of bean varieties, management index, number of neighboring households, and wealth index. This imbalance suggests that making causal inferences on the raw dataset would result in biased and spurious estimates. The blue line, adjusted by the nearest neighbor algorithm, minimized the ADM of the 14 covariates below the 0.20 cutoff value and created a balanced match of 103 pairs of treated and control subjects.

Figure 17 IBB bush - balance of baseline covariates before (unadjusted) and after nearest neighbor matching algorithm measured by the absolute standardized difference of means (ASDM)



The matching algorithm helps to correct the imbalances on most covariates and to create similar PS distributions for the treatment and control groups. To illustrate this fact, the kernel density plot in figure 19 shows a full overlap over the common support area for both distributions, treated vs. control. For the treated-group, PS range from 0.070 to 0.893 with a median equal to 0.308, while for the control-group PS range from 0.076 to 0.896 with a median equal to 0.294.

Figure 18 Nearest neighbor matching algorithm – Kernel density balancing plot – region of common support between treatment and control groups – bush bean growers

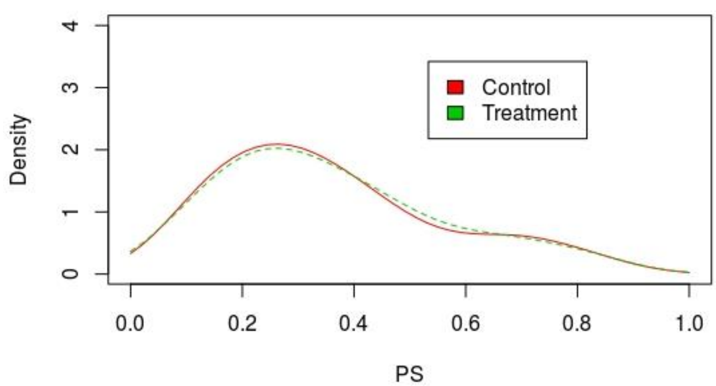


Table 20 Probit model IBB bush nearest neighbor matching algorithm - balancing properties of baseline covariates in treated and control groups

Baseline covariates	Mean treated	Means control	SD control	Mean diff
Propensity score (unbalanced)	0.36	0.21	0.14	0.15
Balanced	0.31	0.30	0.15	0.00
Number of bean varieties cultivated (unbalanced)	3.17	2.01	1.41	1.17
Balanced	2.80	2.70	1.78	0.10
Mean education of HH members, years (unbalanced)	3.66	2.88	1.89	0.78
Balanced	3.51	3.5	2.00	0.00
Household head age, years (unbalanced)	47.80	47.04	16.03	0.77
Balanced	46.78	47.4	15.52	-0.61
Wealth index (0-1) (unbalanced)	0.48	0.41	0.14	0.07
Balanced	0.47	0.48	0.13	-0.01
Cohort 0-5 years old (unbalanced)	0.73	0.83	0.84	-0.10
Balanced	0.68	0.75	0.81	-0.07
Economically active members [18-65] (unbalanced)	2.64	2.61	1.32	0.03
Balanced	2.69	2.81	1.44	-0.12
HH head - years of farming experience (unbalanced)	8.58	7.71	8.35	0.86
Balanced	8.88	8.18	10.39	0.69
Land labor ratio, m/person (unbalanced)	1414.25	1270.53	1814.88	144.00
Balanced	1319.96	1510.18	2115.74	-190.21
Drought index (from -1 to 1) (unbalanced)	-0.04	-0.03	0.17	-0.01
Balanced	-0.05	-0.04	0.18	-0.00
Accessibility to extension services (unbalanced)	5.63	5.41	3.18	0.22
Balanced	5.75	5.90	3.22	-0.15
Plot slope % (unbalanced)	11.52	10.80	7.33	0.72
Balanced	10.98	11.13	6.80	-0.15
Management Index (0-1) (unbalanced)	0.40	0.32	0.23	0.07
Balanced	0.35	0.39	0.19	-0.04
Number of neighbors (unbalanced)	13.88	11.19	7.18	2.69
Balanced	12.83	11.70	7.63	1.13
Number of observations	103	103		

IV. Climbing bean growers, matching members of treatment and control groups

Table 21 shows the results of the probit model for IBB bush growers.

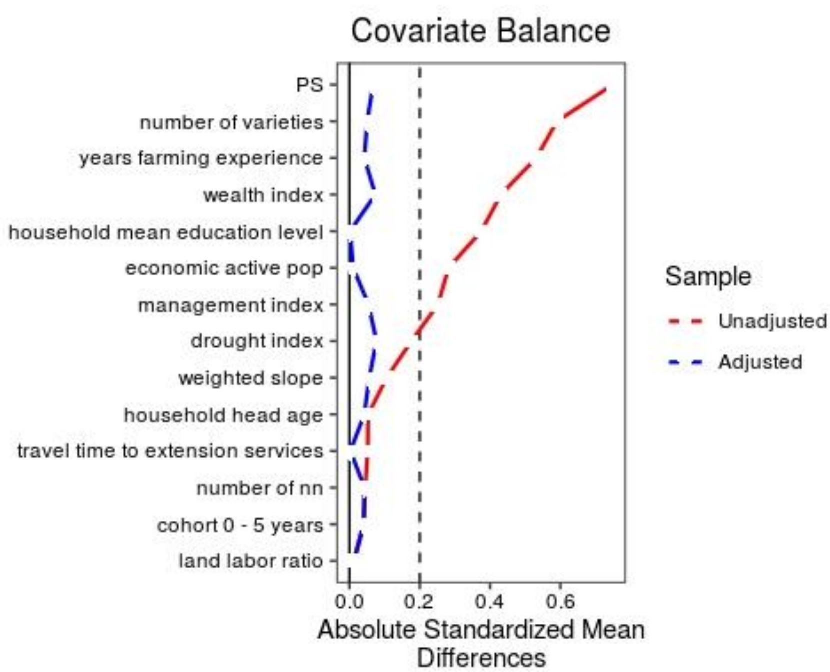
Table 21 Probit model IBB climbing

Variable	Coefficient
Constant	-1.93***
	(0.38)
Number of varieties cultivated	0.22
	(0.03)
Management Index (0-1)	0.21
	(0.30)
Wealth Index	1.24***
	(0.46)
Number of children under 5 years old in household	-0.02
	(0.10)
HH average years of schooling	0.05
	(0.04)
Drought index	1.03***
	(0.45)
Land labor ratio	-0.00
	(0.00)
HH head age	0.02**
	(0.01)
Farming experience (years)	-0.04**
	(0.02)
Travel time to extension services	-0.02
	(0.02)
Weighted slope	-0.01
	(0.01)
Number of individuals per household - economically active members [18-65]	-0.09*
	(0.06)
% corrected predicted	77.81
Log-likelihood	-297.23
LR test	92.20***
Kelejian-Prucha (error)	4.36***
Pinske (error)	19.46***
Pinske -Slade (error)	17.70***
Number of observations	613

Multivariate matching

Figure 19 reports the absolute standard difference of means (ADM) of the unadjusted sample and of the fitted algorithm (adjusted). In the unadjusted sample (red line), I observe differences that exceed the 0.20 benchmark in 7 of the 14 covariates. In descending order, the covariates with the largest difference: PS, the number of bean varieties cultivated, years of farming experience, wealth index, average household education level, number of EA males in household, and management index. The blue line, adjusted dataset fitted by the matching algorithm, minimized these differences below the 0.20 cutoff value in all baseline covariates, in most cases below 0.15. The matching algorithm created 130 balanced pairs of treated subjects to control subjects.

Figure 19 IBB climbing - balance of baseline covariates before (unadjusted) and after nearest neighbor matching algorithm measured by the absolute standardized difference of means (ASDM)



The matching algorithm helps to correct the imbalances and to create similar PS distributions. The kernel density plot (figure 20) shows, over the common support area, a full overlap of the PS distributions of the treatment and control groups. For the treated-group PS range from 0.086 to 0.782 with a median value equal to 0.282, while for the control-group PS range from 0.086 to 0.779 with a median value equal to 0.291.

Figure 20 Nearest neighbor matching algorithm – Kernel density balancing plot – region of common support between treatment and control groups – climbing bean growers

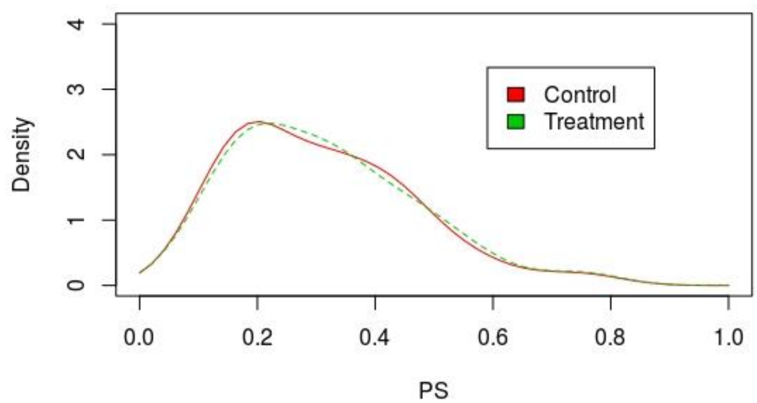


Table 22 IBB climbing - nearest neighbor matching algorithm - balancing properties of baseline covariates in treated and control groups

Baseline covariates	Mean treated	Means control	SD control	Mean diff
Propensity score (unbalanced)	0.36	0.21	0.14	0.15
Balanced	0.31	0.30	0.15	0.00
Number of bean varieties cultivated (unbalanced)	3.17	2.01	1.41	1.17
Balanced	2.80	2.70	1.78	0.1
Mean education of HH members, years (unbalanced)	3.66	2.88	1.89	0.78
Balanced	3.51	3.50	2.00	0.00
Household head age, years (unbalanced)	47.80	47.04	16.03	0.77
Balanced	46.78	47.4	15.52	-0.61
Wealth index (0-1) (unbalanced)	0.48	0.41	0.14	0.07
Balanced	0.47	0.48	0.13	-0.01
Cohort 0-5 years old (unbalanced)	0.69	0.72	0.80	-0.02
Balanced	0.72	0.75	0.78	-0.03
Economically active members [18-65] (unbalanced)	3.15	2.74	1.38	0.41
Balanced	3.03	3.04	1.44	-0.01
HH head - years of farming experience (unbalanced)	5.72	8.54	10.68	-2.82
Balanced	5.78	6.00	5.18	-0.22
Land labor ratio, m/person (unbalanced)	669.66	678.61	1047.11	-8.94
Balanced	687.48	697.43	815.37	-9.94
Drought index (from -1 to 1) (unbalanced)	-0.01	-0.03	0.14	0.02
Balanced	-0.02	-0.03	0.13	0.01
Accessibility to extension services (unbalanced)	5.43	5.58	2.73	-0.15
Balanced	5.43	5.43	2.71	0.00
Plot slope % (unbalanced)	13.08	14.00	8.40	-0.91
Balanced	13.67	13.18	7.74	0.48
Management Index (0-1) (unbalanced)	0.50	0.45	0.21	0.06
Balanced	0.48	0.49	0.17	-0.01
Number of neighbors (unbalanced)	5.81	5.88	1.96	-0.08
Balanced	5.75	5.67	1.75	0.08
Number of observations	130	130		

Table 23 Sensitivity analysis using one-tailed P-value for departures from random assignment.

Gamma Value	IBB bush p-value	IBB climbing p-value
1	0.040	0.025
1.01	0.045	0.028
1.02	0.051	0.031
1.03	0.057	0.035
1.04	0.063	0.039
1.05	0.070	0.043
1.06	0.077	0.047
1.07	0.084	0.052
1.08	0.092	0.057
1.09	0.100	0.062
1.1	0.108	0.068
1.11	0.117	0.074
1.12	0.126	0.081
1.13	0.135	0.088
1.14	0.144	0.095
1.15	0.154	0.102
1.16	0.164	0.110

Note: Gamma values range from 1 to 1.16 with an increment of 0.01 in Gamma - Wilcoxon's statistic

Bibliography

Abadi Ghadim, A. K., & Pannell, D. J. (1999). A conceptual framework of adoption of an agricultural innovation. *Agricultural Economics*. https://doi.org/10.1016/S0169-5150(99)00023-7

Abdulai, A., Monnin, P., & Gerber, J. (2008). Joint estimation of information acquisition and adoption of new technologies under uncertainty. *Journal of International Development*. https://doi.org/10.1002/jid.1422

Adesina, A. (1993). Technology characteristics, farmers' perceptions and adoption decisions: A Tobit model application in Sierra Leone. *Agricultural Economics*, *9*(4), 297–311. https://doi.org/10.1016/0169-5150(93)90019-9

Adesina, A. (1995). Farmers' perceptions and adoption of new agricultural technology: evidence from analysis in Burkina Faso and Guinea, West Africa. *Agricultural Economics*, *13*(1), 1–9. https://doi.org/10.1016/0169-5150(95)01142-8

Alderman, H., Hoddinott, J., & Kinsey, B. (2006). Long term consequences of early childhood malnutrition. *Oxford Economic Papers*. https://doi.org/10.1093/oep/gpl008

Ali, D. A., & Deininger, K. (2015). Is there a farm size-productivity relationship in African agriculture? Evidence from Rwanda. *Land Economics*. https://doi.org/10.3368/le.91.2.317

Alston, J. M. (2002). Spillovers. *Australian Journal of Agricultural and Resource Economics*, *46*(3), 315–346. https://doi.org/10.1111/1467-8489.00030

Anselin, L. (1988). *Spatial Econometrics: Methods and Models* (4th ed.). Boston:

Dordrecht ; Boston : Kluwer Academic Publishers.

Anselin, L., & Le Gallo, J. (2006). Interpolation of Air Quality Measures in Hedonic House Price Models: Spatial Aspects. *Spatial Economic Analysis*. https://doi.org/10.1080/17421770600661337

Anselin, L., & Rey, S. J. (2014). *Modern spatial econometrics in practice: A guide to GeoDa, GeoDaSpace and PySAL*. *GeoDa Press LLC*.

Asare-marfo, D., Gonzalez, C., Perez, S., Schwarz, J., & Zeller, M. (2013). Prioritizing Countries for Biofortification Interventions Using Country-Level Data Dorene. *HarvestPlus*, (October).

Asare-Marfo, D., Herrington, C., Birachi, C., Birol, E., Tedla-Diressie, M., Funes, J., & Zeller, M. (2016). Assessing the Adoption of High Iron Bean Varieties and Their Impact on Iron Intakes and Other Livelihood Outcomes in Rwanda. *Report*.

Asselin, L. M. 2009. Analysis of Multidimensional Poverty: Theory and Case-Studies. *Economic Studies in Inequality, Social Exclusion and Well-being*, Volume No. 7, Springer

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*. https://doi.org/10.1080/00273171.2011.568786

Austin, P. C. (2014). A comparison of 12 algorithms for matching on the propensity score. *Statistics in Medicine*. https://doi.org/10.1002/sim.6004

Beaman, L., BenYishay, A., Magruder, J., & Mobarak, A. M. (2018). *Can Network Theory-based Targeting Increase Technology Adoption?* (No. 1062).

Cambridge, MA. https://doi.org/10.3386/w24912

Berti, P., Moursi, M., Kung, J., Tugirimana, P., Siekmans, K., Lubowa, A., & Boy, E. (2012). Inadequate diet, but low iron deficiency anemia, in rwandan women and children.

Besley, T., & Case, A. (1993). Modeling technology adoption in developing countries. *American Economic Review*.

Bitocchi, E., Nanni, L., Bellucci, E., Rossi, M., Giardini, A., Zeuli, P. S., … Papa, R. (2012). Mesoamerican origin of the common bean (Phaseolus vulgaris L.) is revealed by sequence data. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1108973109

Bouis, H. (2017). An Overview of the landscape and approach for Biofortification in Africa. *African Journal of Food, Agriculture, Nutrition and Development*. https://doi.org/10.18697/ajfand.78.harvestplus01

Bravo-Ureta, B. E., Greene, W., & 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*(1), 55–72. https://doi.org/10.1007/s00181-011-0491-y

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*. https://doi.org/10.1111/j.1467-6419.2007.00527.x

Case, A. (1992). Neighborhood influence and technological change. *Regional Science and Urban Economics*, *22*(3), 491–508. https://doi.org/10.1016/0166-0462(92)90041-X

Chagas, A. L. S., Toneto, R., & Azzoni, C. R. (2012). A Spatial Propensity Score Matching Evaluation of the Social Impacts of Sugarcane Growing on Municipalities in Brazil. *International Regional Science Review*, *35*(1), 48–69. https://doi.org/10.1177/0160017611400069

Chavas, J.-P., Petrie, R., & Roth, M. (2005). Farm Household Production Efficiency: Evidence from The Gambia. *American Journal of Agricultural Economics*, *87*(1), 160–179. https://doi.org/10.1111/j.0002-9092.2005.00709.x

Chirwa, E. W. (2005). Adoption of fertiliser and hybrid seeds by smallholder maize farmers in Southern Malawi. *Development Southern Africa*, *22*(1), 1–12. https://doi.org/10.1080/03768350500044065

Conley, T. G., & Udry, C. R. (2010). Learning about a New Technology: Pineapple in Ghana. *American Economic Review*, *100*(1), 35–69. https://doi.org/10.1257/aer.100.1.35

Croppenstedt, A., Demeke, M., & Meschi, M. M. (2003). Technology Adoption in the Presence of Constraints: the Case of Fertilizer Demand in Ethiopia. *Review of Development Economics*, *7*(1), 58–70. https://doi.org/10.1111/1467-9361.00175

Del Gatto, M., Di Liberto, A., & Petraglia, C. (2011). MEASURING PRODUCTIVITY. *Journal of Economic Surveys*, *25*(5), 952–1008. https://doi.org/10.1111/j.1467-6419.2009.00620.x

Dinar, A., Karagiannis, G., & Tzouvelekas, V. (2007). Evaluating the impact of agricultural extension on farms' performance in Crete: a nonneutral stochastic frontier approach. *Agricultural Economics*, *36*(2), 135–146. https://doi.org/10.1111/j.1574-0862.2007.00193.x

Dorosh, P., & Wang, H. (2010). Crop production and road connectivity in Sub-Saharan Africa: a spatial analysis. *World Bank Policy {…}*. https://doi.org/10.1596/1813-9450-5385

Dorosh, P., Wang, H. G., You, L., & Schmidt, E. (2012). Road connectivity, population, and crop production in Sub-Saharan Africa. *Agricultural Economics*. https://doi.org/10.1111/j.1574-0862.2011.00567.x

Duflo, E., Kremer, M., & Robinson, J. (2008). How High Are Rates of Return to Fertilizer? Evidence from Field Experiments in Kenya. *American Economic Review*, *98*(2), 482–488. https://doi.org/10.1257/aer.98.2.482

FAO. (2020). Rwanda at a glance.

Feder, Gershon and T. O'Mara, G. (1981). Farm Size and the Diffusion of Green Revolution Technology. *Economic Development and Cultural Change*, *30*(1), 59–76.

Feder, G., Just, R. E., & Zilberman, D. (1985). Adoption of Agricultural Innovations in Developing Countries: A Survey. *Economic Development and Cultural Change*, *33*(2), 255–298. https://doi.org/10.1086/451461

Foster, A. D., & Rosenzweig, M. R. (1995). Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture. *Journal of Political Economy*, *103*(6), 1176–1209. https://doi.org/10.1086/601447

Foster, A. D., & Rosenzweig, M. R. (2010). Microeconomics of Technology Adoption. *Annual Review of Economics*, *2*(1), 395–424. https://doi.org/10.1146/annurev.economics.102308.124433

Foster, A., & Rosenzweig, M. (2017). *Are There Too Many Farms in the World?*

Labor-Market Transaction Costs, Machine Capacities and Optimal Farm Size. Cambridge, MA. https://doi.org/10.3386/w23909

Gamerman, D., & Moreira, A. R. B. (2004). Multivariate spatial regression models. *Journal of Multivariate Analysis*, *91*(2), 262–281. https://doi.org/10.1016/j.jmva.2004.02.016

Gelman, A., Hwang, J., & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Statistics and Computing*. https://doi.org/10.1007/s11222-013-9416-2

Giné, X., & Yang, D. (2009). Insurance, credit, and technology adoption: Field experimental evidencefrom Malawi. *Journal of Development Economics*, *89*(1), 1–11. https://doi.org/10.1016/j.jdeveco.2008.09.007

Gonzales, R. A., Aranda, P., & Mendizabal, J. (2018). Is Microfinance Truly Useless for Poverty Reduction and Women Empowerment? A Bayesian Spatial-Propensity Score Matching Evaluation in Bolivia. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3167321

Greene, W. H. (2008). The Econometric Approach to Efficiency Analysis. In *The Measurement of Productive Efficiency and Productivity Change* (pp. 92–250). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195183528.003.0002

Greene, W. H. (2012). *Econometric analysis*. Boston: Pearson.

Haas, J. D., Luna, S. V, Lung'aho, M. G., Wenger, M. J., Murray-Kolb, L. E., Beebe, S., … Egli, I. M. (2016). Consuming Iron Biofortified Beans Increases Iron Status in Rwandan Women after 128 Days in a Randomized Controlled Feeding

Trial. *The Journal of Nutrition*. https://doi.org/10.3945/jn.115.224741

Han, J., Ryu, D., & Sickles, R. (2016a). How to measure spillover effects of public capital stock: A spatial autoregressive stochastic frontier model. *Advances in Econometrics*, *37*, 259–294. https://doi.org/10.1108/S0731-905320160000037017

Han, J., Ryu, D., & Sickles, R. (2016b). How to Measure Spillover Effects of Public Capital Stock: A Spatial Autoregressive Stochastic Frontier Model. In T. B. F. R. C. H. I. J. J. C. E. and E. Hillebrand (Ed.), *Advances in Econometrics* (pp. 259–294). Emerald Group Publishing. https://doi.org/10.1108/S0731-905320160000037017

Hazell, P., Poulton, C., Wiggins, S., & Dorward, A. (2010). The Future of Small Farms: Trajectories and Policy Priorities. *World Development*, *38*(10), 1349–1361. https://doi.org/10.1016/j.worlddev.2009.06.012

Holloway, G. (2002). Bayesian spatial probit estimation: a primer and an application to HYV rice adoption. *Agricultural Economics*, *27*(3), 383–402. https://doi.org/10.1016/S0169-5150(02)00070-1

Jamison, D. T., & Moock, P. R. (1984). Farmer education and farm efficiency in Nepal: The role of schooling, extension services, and cognitive skills. *World Development*, *12*(1), 67–86. https://doi.org/10.1016/0305-750X(84)90036-6

Jolliffe, D. (2002). Whose education matters in the determination of household income? Evidence from a developing country. *Economic Development and Cultural Change*. https://doi.org/10.1086/322880

Kelejian, H. H., & Prucha, I. R. (2001). On the asymptotic distribution of the Moran I test statistic with applications. *Journal of Econometrics*. https://doi.org/10.1016/S0304-4076(01)00064-1

King, G., & Nielsen, R. (2019). Why Propensity Scores Should Not Be Used for Matching. *Political Analysis*. https://doi.org/10.1017/pan.2019.11

Krishnan, P., & Patnam, M. (2014). Neighbors and Extension Agents in Ethiopia: Who Matters More for Technology Adoption? *American Journal of Agricultural Economics*, *96*(1), 308–327. https://doi.org/10.1093/ajae/aat017

Lacombe, D. J., & LeSage, J. P. (2018). Use and interpretation of spatial autoregressive probit models. *The Annals of Regional Science*, *60*(1), 1–24. https://doi.org/10.1007/s00168-015-0705-x

Larochelle, C., Alwang, J., Norton, G., Katungi, E., & Labarta, R. (2014). Crop improvement, adoption, and impact of improved varieties in food crops in Sub-Saharan Africa: in T. S. Walker & Alwang (Eds.) (pp. 314–337). New York: CABI.

Larson, D. F., Otsuka, K., Matsumoto, T., & Kilic, T. (2014). Should African rural development strategies depend on smallholder farms? An exploration of the inverse-productivity hypothesis. *Agricultural Economics (United Kingdom)*. https://doi.org/10.1111/agec.12070

LeSage, J. P. (2008, September 1). An introduction to spatial econometrics. *Revue d'Economie Industrielle*. Editions Techniques et Economiques. https://doi.org/10.4000/rei.3887

LeSage, J. P. (2009). *An Introduction to Spatial Econometrics*. (W. R. Balakrishnan, N; Schucany, Ed.), *Introduction to Spatial Econometrics*. Taylor & Francis

Group, LLC.

LeSage, J. P., Kelley Pace, R., Lam, N., Campanella, R., & Liu, X. (2011). New Orleans business recovery in the aftermath of Hurricane Katrina. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *174*(4), 1007–1027. https://doi.org/10.1111/j.1467-985X.2011.00712.x

Luna, S., Lung'ago, M., Gahutu, J., & Haas, J. (2015). Effects of an Iron-biofortification Feeding Trial on Physical Performance of Rwanda Women. *European Journal of Nutrition & Food Safety*, *5*(5).

Maertens, A., & Barrett, C. B. (2013). Measuring Social Networks' Effects on Agricultural Technology Adoption. *American Journal of Agricultural Economics*, *95*(2), 353–359. https://doi.org/10.1093/ajae/aas049

Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, *60*(3), 531. https://doi.org/10.2307/2298123

Meenakshi, J. V., Banerji, A., Manyong, V., Tomlins, K., Mittal, N., & Hamukwala, P. (2012). Using a discrete choice experiment to elicit the demand for a nutritious food: Willingness-to-pay for orange maize in rural Zambia. *Journal of Health Economics*, *31*(1), 62–71. https://doi.org/10.1016/j.jhealeco.2012.01.002

Murekezi, A, Birol, E., Asare-Marfo, D., & Ktasvairo, L. (2013). Farmer Feedback Study on High Iron Bean Seed Delivery in Rwanda. *AAEA CAES Joint Annual Meeting*.

Murekezi, Abdoul, Oparinde, A., & Birol, E. (2017). Consumer market segments for biofortified iron beans in Rwanda: Evidence from a hedonic testing study. *Food*

Policy. https://doi.org/10.1016/j.foodpol.2016.11.005

Murray-Kolb, L. E., Wenger, M. J., Scott, S. P., Rhoten, S. E., Lung'aho, M. G., & Haas, J. D. (2017). Consumption of Iron-Biofortified Beans Positively Affects Cognitive Performance in 18- to 27-Year-Old Rwandan Female College Students in an 18-Week Randomized Controlled Efficacy Trial. *The Journal of Nutrition*. https://doi.org/10.3945/jn.117.255356

NISR. (2015). Rwanda Demographic and Health Survey. *Demographic and Health Survey*, *1*, 1:24.

Oparinde, A., Birol, E., Murekezi, A., Katsvairo, L., Diressie, M. T., Nkundimana, J. d. amou., & Butare, L. (2016). Radio Messaging Frequency, Information Framing, and Consumer Willingness to Pay for Biofortified Iron Beans: Evidence from Revealed Preference Elicitation in Rural Rwanda. In *Canadian Journal of Agricultural Economics* (Vol. 64, pp. 613–652). Blackwell Publishing Ltd. https://doi.org/10.1111/cjag.12105

Petry, N., Olofin, I., Hurrell, R. F., Boy, E., Wirth, J. P., Moursi, M., … Rohner, F. (2016). The proportion of anemia associated with iron deficiency in low, medium, and high human development index countries: A systematic analysis of national surveys. *Nutrients*. https://doi.org/10.3390/nu8110693

Rogers, E. M., Singhal, A., & Quinlan, M. M. (2019). Diffusion of innovations. In *An Integrated Approach to Communication Theory and Research, Third Edition* (pp. 415–433). Taylor and Francis. https://doi.org/10.4324/9780203710753-35

Rosenbaum, P. R. (2005). Observational Study. In *Encyclopedia of Statistics in Behavioral Science*. https://doi.org/10.1002/0470013192.bsa454

Rosenbaum, P. R. (2010). *Design of Observational Studies*. New York: Springer.

Rubin, D. B. (2005). Causal Inference Using Potential Outcomes. *Journal of the American Statistical Association*. https://doi.org/10.1198/016214504000001880

Rue, H., Martino, S., & Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*. https://doi.org/10.1111/j.1467-9868.2008.00700.x

Saltzman, A., Birol, E., Bouis, H. E., Boy, E., De Moura, F. F., Islam, Y., & Pfeiffer, W. H. (2013, March). Biofortification: Progress toward a more nourishing future. *Global Food Security*. https://doi.org/10.1016/j.gfs.2012.12.003

Saltzman, A., Birol, E., Oparinde, A., Andersson, M. S., Asare-Marfo, D., Diressie, M. T., … Zeller, M. (2017). Availability, production, and consumption of crops biofortified by plant breeding: current evidence and future potential. *Annals of the New York Academy of Sciences*, *1390*(1), 104–114. https://doi.org/10.1111/nyas.13314

Schmidt, A. M., Moreira, A. R. B., Helfand, S. M., & Fonseca, T. C. O. (2009). Spatial stochastic frontier models: accounting for unobserved local determinants of inefficiency. *Journal of Productivity Analysis*, *31*(2), 101–112. https://doi.org/10.1007/s11123-008-0122-6

Simar, L., & Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, *136*(1), 31–64. https://doi.org/10.1016/j.jeconom.2005.07.009

Simtowe, F. P., Zeller, M., & Diagne, A. (2009). The impact of credit constraints on the adoption of hybrid maize in Malawi. *Review of Agricultural and Environmental Studies - Revue d'Etudes En Agriculture et Environnement*, *90*(1), 5–22. Retrieved from http://econpapers.repec.org/RePEc:rae:jourae:v:90:y:2009:i:1:p:5-22

Stein, A. J. (2010). Global impacts of human mineral malnutrition. *Plant and Soil*. https://doi.org/10.1007/s11104-009-0228-2

Sunding, D., & Zilberman, D. (2001). Chapter 4 The agricultural innovation process: Research and technology adoption in a changing agricultural sector. In *Handbook of Agricultural Economics* (pp. 207–261). https://doi.org/10.1016/S1574-0072(01)10007-1

Tessema, Y., Asafu-Adjaye, J., Kassie, M., & 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*(3), 211.

Vaiknoras, K., & Larochelle, C. (2018). The Impact of Biofortified Iron Bean Adoption on Productivity, and Bean Consumption, Purchases and Sales. In *2018 AAEA annual meeting program*.

Verhofstadt, E., & Maertens, M. (2015). Can agricultural cooperatives reduce poverty? Heterogeneous impact of cooperative membership on farmers' welfare in Rwanda. *Applied Economic Perspectives and Policy*. https://doi.org/10.1093/aepp/ppu021

Walker, T. ., & Alwang, J. (2015). *Crop Improvement, Adoption and Impact of Improved Varieties in Food Crops in Sub-Saharan Africa*. New York, NY:

CABI.

Ward, P. S., & 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*(2), 225–241. https://doi.org/10.1111/1467-8489.12058

Watanabe, S. (2013). A widely applicable bayesian information criterion. *Journal of Machine Learning Research*.

Weir, S., & Knight, J. (2000). *Education Externalities in Rural Ethiopia: Evidence from Average and Stochastic Frontier Production Functions*.

Weir, S., & Knight, J. (2007). Production Externalities of Education: Evidence from Rural Ethiopia1. *Journal of African Economies*, *16*(1), 134–165. https://doi.org/10.1093/jae/ejl026

Wilhelm, S., & de Matos, M. G. (2013). Estimating spatial probit models in R. *R Journal*. https://doi.org/10.32614/rj-2013-013

World Bank. (2015). World development indicators. data retrieved from World Development Indicators.

Zhou, X., & Xie, Y. (2016). Propensity Score–based Methods Versus MTE-based Methods in Causal Inference: Identification, Estimation, and Application. *Sociological Methods and Research*. https://doi.org/10.1177/0049124114555199