A Case for Space: The Efficiency Spillover Effect of Iron Biofortified Beans in Rwanda

Jose Funes[1], Laixiang Sun[1], Todd Benson[2], Fernando Sedano[1], and Giovanni Baiocchi[1]

[1]Department of Geographical Sciences, University of Maryland, College park
[2]International Food Policy Research Institute, Washington, DC

Selected Paper prepared for presentation at the 2019 Agricultural and Applied Economics Association Annual Meeting, Atlanta, GA

Abstract

This paper provides an analysis on the technical efficiency of iron biofortified bean production in Rwanda, as well as recommendations for implementation of targeted biofortification programs. We employ a three-pronged approach to examine efficiency across matched pairs of treated (i.e., grew iron biofortified beans) and control (i.e., grew other improved or traditional beans) bean growing households. First, bean growing households (treatment and control) were matched by using the spatial propensity score matching (PSM) approach. Second, efficiency was estimated using spatial autoregressive models and non-spatial stochastic frontier models on the three data groups: treatment, control, and pool. Third, an econometric analysis was conducted to explore factors affecting farmers' efficiencies. All these steps were implemented separately for farmers who grew bush beans and for those who grew climbing beans. Results showed that for both types of bean growers, farmers who grew iron biofortified varieties were relatively more efficient and obtained greater bean production (by 3 percent for bush and 13 percent for climbing) than the control group. A spatial univariate clustering spatial analysis on farmers efficiency helped identify geographic areas in which farmers with high technical efficiency (i.e., hotspots) and low technical efficiency (i.e., cold spots) were located, and can be reached through targeted, tailor-made interventions.

Keywords: iron biofortified beans, efficiency, spillover, spatial econometrics

1 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 Rwandan population is at risk of iron deficiency, resulting in loss of economic development and growth. Beans have the highest share of crop-harvested area 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. There is therefore significant yield gains to be made from the introduction and scaling up of seeds of improved varieties of beans. Improved, iron biofortified beans (IBB) can not only help improve yields (Funes et al. 2019) and hence incomes, but also nutrition (Haas et al. 2016) and health outcomes e.g., cognitive (Murray-Kolb et al. 2017) and physical functions of consuming populations (Luna et al. 2015).

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, and Robinson 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 application of combination of analytical methods (e.g., for earlier examples see (Dinar, Karagiannis, and Tzouvelekas 2007; Bravo-Ureta, Greene, and Solís 2012)). Recent literature in productivity analysis and impact evaluation highlights the importance of measuring spillovers to non-beneficiaries e.g. (Gamerman and Moreira 2004; Schmidt et al. 2009; Affuso 2011).

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 ignores 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, Petrie, and Roth 2005; Simar and Wilson 2007). Greene (2005) 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, and Sickles 2016). In this paper we implement that latter method through a novel, three-pronged approach

The first prong assesses the importance of social networks in the adoption and diffusion of the technology in question (i.e., iron biofortified beans) (Funes 2018). The second prong uses propensity score matching (PSM) method to measure the impact of growing iron biofortified bean on farmers' bean yield and bean income (Funes et al. 2018). 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 paper contributes to the this growing literature by presenting a case study on an intervention that delivered a new technology, namely iron biofortified beans, in Rwanda. We developed an innovative multi-pronged approach and applied it to a cross-sectional, nationally representative data of bean farmers in Rwanda to estimate farmers' unbiased efficiency scores. We combine spatial econometrics and quasi-experimental methods to estimate a national technological frontier for all bean farmers, a frontier for iron biofortified bean growers, and a frontier for farmers that grow other improved or traditional bean varieties. We compare standard stochastic frontier models to spatial stochastic frontier models, which help us estimate efficiency spillovers among iron biofortified bean 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.

We organize this paper as follows. The next section describes the theoretical framework; defines efficiency and productivity, and introduces the tools used for analysis. Section three provides the conceptual framework, including a brief introduction to spatial and non-spatial stochastic frontier models to estimate efficiency and efficiency spillovers. This section also provides a brief description of DEA technique to estimate scale efficiency. Section four presents the results and highlights areas of further discussion. Section five concludes the paper with some programmatic implications.

2 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 $X \epsilon R_M^+$ into a set of outputs $Y \epsilon R_M^+$ (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, we evaluate efficiency for bush and climbing bean growers, separately.

To evaluate efficiency levels among bean farmers in 2015 Season B, we 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, we measure—through scale efficiency analysis—how close bean farming households are to operate at optimal scale. The larger the scale efficiency, the closer the farming household is to optimal scale.

We 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. In this paper, we provide a brief literature review on these spillovers.

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 and Case 1993; Foster and Rosenzweig 2010), such as the exchange of information on the nutritional and agronomic benefits of iron biofortified beans. Conley and 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 taking into account the amounts allocated by the other farmers in the system.

Technology spillovers refers to the benefits that smallholder bean farmers receive from research efforts without incurring on shared costs. At a broader spatial scale, international spillovers from public agricultural research and development (RD) represent a high percentage of agricultural productivity growth (Alston et al. 2000). Specifically, agricultural RD 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, International Centre for Tropical Agriculture (CIAT) and Rwanda Agriculture Board (RAB). Parental lines of improved bean varieties are bred in CIAT 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 onto the spatial patterns of growing iron biofortified bean varieties and the determinants of farmers' technical efficiency, we conducted a spatial clustering analysis to better understand the geographic concentration of advanced farmers versus less advanced farmers.

The literature presents social or peer-effects as follows: (1) pure peer effect defined by their role in and between social organizations, (2) the impact of individual characteristics on peer effects, and (3) the structure of interaction across economic agents (Charness and Kuhn 2011; Herbst and Mas 2015). Seminal work by Manski 1993 shows standard econometric methods to model social interactions.

3 Conceptual framework

We used stochastic frontier analysis (SFA) and stochastic spatial frontier analysis (SSFA) to estimate bean farmers technical efficiency. In addition, we use 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 methods (Del Gatto, Di Liberto, and 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.

$$y_i = f(x_i, \beta) + \epsilon_i \tag{1}$$

$$\epsilon_i = v_i - u_i \tag{2}$$

where (5): $v_i \sim N(0, \sigma_v^2) \quad and \quad u_i \sim N(0, \sigma_u^2)$

Combining equation 1 and 2,

$$y_i = f(x_i, \beta) + v_i - u_i \tag{3}$$

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. ϵ_i 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 (2) 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, we are able to estimate the best practice production function rather than the average practice production function.

However, note that model (3) 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 (3) (Han, Ryu, and Sickles 2016). The first is endogenous effects which explain the dependence between the dependent variable, y_i and y_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_j. Third, interaction effects among the error terms equation (5). A full model with all types of spatial interaction effects are specified in equation (4).

$$y_i = \rho \sum_{j=1}^{N} W_{ij} y_i + \beta_0 + X_i \beta_1 + \gamma \sum_{j=1}^{N} W_{ij} x_j + \epsilon_i \tag{4}$$

$$\epsilon_i = \lambda \sum_{j=1}^{N} W_{ij} \epsilon_j + u_i \tag{5}$$

where, y is a $Nx1$ vector of observations on the dependent variable, W is an exogenous NxN 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 NxK matrix of observations on explanatory variables with associated $Kx1$ coefficient vector β_1. X_j the spatial lags of the covariates (independent variables) with coefficients γ. ϵ_i is an $Nx1$ vector of error terms as $i.i.d.$ disturbance for household h_i with zero mean and variance σ_ϵ^2. Restriction $\gamma = 0$ and $\lambda = 0$ give rise to the spatial autoregressive model (SAR) and restriction to $\rho = 0$ and $\gamma = 0$ give rise to the error spatial autoregressive model (Anselin 1988; Anselin and Rey 2014).

We implemented the first and third methods following the spatial specifications search suggested by Anselin and Rey 2014. We started with our 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, we report the OLS model.

For the efficiency analysis, we 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 we are unable to identify the term $-u$, we 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, Ryu, and Sickles 2016).

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), 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. We shift more importance to the second stage-censored analysis (Chavas, Petrie, and Roth 2005). In all models, we used the multiplicative form and conduct the estimation in log-linear form. The first order coefficients can be interpreted as partial elasticities.

3.1 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. We 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.

Truncated regression has shown robust results. A truncated regression is a distribution that occurs when some values above or below are omitted. Simar and 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 truncated regression was less unstable.

3.2 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, we use Local Moran's I statistics (Anselin 1995). 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.

4 Results

4.1 Bush bean farmers

Table A.1 presents the results of the fitted models for the pool, treated (iron biofortified bush bean growers), and control (other bush bean growers) data groups. The starting point is the production function estimated with OLS. We test whether the technical efficiency of farming households is stochastic. We 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 A.1 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, we estimated separate technology frontiers for each data groups. The bottom of table 1 shows 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. It is worth highlighting that the parameter rho (0.115) or global spatial multiplier is significant for the spatial stochastic frontier model - pool (SSFA-P). This parameter reveals the link or spillover effect in the system between treated and control groups. Bean farmers who did not grow iron biofortified beans got an indirect benefit of 12 % in terms of total bean production because they interact with neighbors who grew iron biofortified beans. In addition, the IBB adoption parameter, in the SFA-pool and SSFA-pool models is positive and significant ($p<0.10$).

With respect to the pool specifications of the SSFA, land area, hired labor, and economic active population positively contributed the most to bean production. For the control group (SSFA-C) land area and hired labor positively contributed the most whereas for the treatment group (SSFA-T) hired labor, seed, and land area had a positive significant and most influence on bean production. Interpretation of the coefficient requires the estimation of direct and indirect impact effects which we do not discuss in this paper. 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 the table, we include two summary measures of economic performance: relative efficiency and relative efficiency weighted by bean production (kg). In the spatial stochastic production frontier 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% higher output than bean farmers with similar level of inputs growing other bean varieties.

4.2 Climbing bean farmers

As table A.2 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 is significant, suggesting the presence of spatial spillovers. The rho parameters for the pool (SSFA-C) was significant. The rho (0.09) parameter reveals 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% in terms of total bean production. The sum of all partial production elasticities is larger than 1 which indicates increasing returns to scale (IRS) in all models. Given an IRS, climbing bean growers are not operating at an efficient scale.

The most influential parameters for both pool models, SFA and SSFA, were the management index, land area, and seed. For the control - SFA model land area and economic active household members were the most influential parameters. The most influential parameters for the control SSFA model were the management index and land area. The treatment SSFA model shows something new compare to the control model. After the management index, 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 total land 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 A.2, we 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 farmers' relative efficiency than the control group of bean farmers. In addition of having a higher relative efficiency, IBB climbing growers had 13% higher output than bean farmers with similar level of inputs who grew other bean varieties.

4.3 Truncated – second stage analysis

In contrast with the input variables used in the first stage to estimate bean farmers' relative efficiency, covariates listed in table A.3 are sources of inefficiency. One out of the four covariates is an indicator of regional endowment, for example proximity to extension services. We found a negative relationship between travel time to extension services and farmers' efficiency. Bean farmers closer to extension services had higher efficiency than those farmers farther away that may incur in higher transportation cost to access basic agricultural inputs and advise from extension services. Bean farmers with lower transportation cost are more efficient connecting to markets to meet the demand for staple food crops in 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 of 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 grow IBB varieties brings about a significant change in crop production efficiency, which in turn improves farm household's income (Funes et al. 2019).

We found crop specialization is associated with higher efficiency. The parameter of crop diversification (i.e., number of crops grown) was negative and significant. This evidence suggests that bean farmers' efficiency increases as they cultivate fewer crops. In Rwanda, bean farmers tend to grow up to six crops. The fact that most farms are dual or multicrop farms suggests that the benefits of diversification are significant in Rwandan agriculture, which we 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 2008). From a nutritional security and suitability (transaction cost) point of view, diversification can be efficient and sustainable in agriculture but might have an impact on yields of any one crop.

The drought index parameter is significant and has the most influence on farmers' efficiency. 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 to farmers' efforts to adapt to climate change by adjusting their farming management plans to grow their crops.

4.4 Clustering analysis

Figure 1 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 2 splits the global spatial association into five observation groups: (1) no significant spatial association; two set of clusters, (2) hot spots and (3) cold spots; and two set of spatial outliers: (4) low-high and (5) high-low.

Figure 1: Moran's I bean farmers' technical efficiency

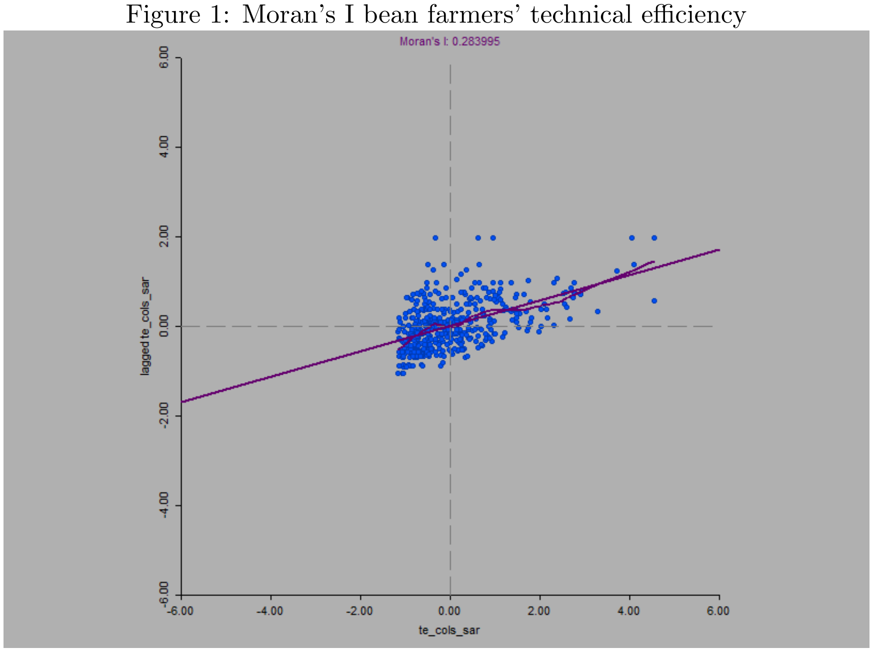
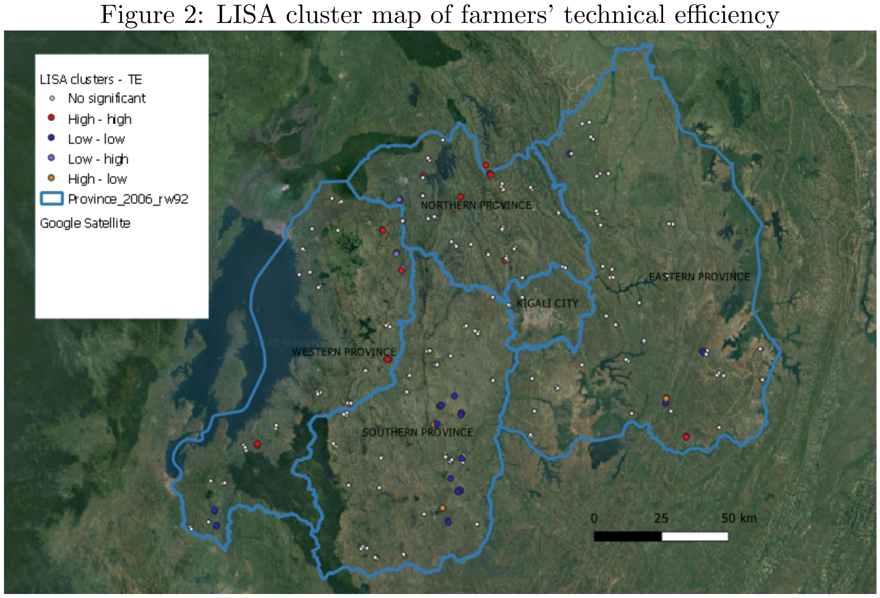


Figure 2: LISA cluster map of farmers' technical efficiency



The first cluster are hotspots of bean farmers with high efficiency scores surrounded by farmers with similar efficiency scores. These hotspots are on the neighboring districts of Burera, Gakenke, and Musanze in the Northern region and on 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. Cold spots 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 next section focuses on the spatial outliers. The first 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-farmers with high efficiency surrounded by farmers that are lower efficiency scores. 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 diffusion of new technologies, such as biofortified iron bean varieties.

5 Conclusion

In this paper, we 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 propensity score matching approach was used to create two comparable matched groups: treatment group (which included farmers who grew iron biofortified bean in 2015 Season B) and a control group (farmers 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, we fit a truncated model to validate and explore the sources of inefficiency. This paper contributes to the literature by controlling for self-selection bias, missing counter-factual and for 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 technical 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 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 (IRS) in all models. In the second stage analysis, we discovered that IBB growers that had market linkages were located in areas less vulnerable to weather shocks, and had better access to extension services 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 reached out through targeted, tailored-made interventions. These analyses provide valuable input into targeting strategies and resource allocation for scaling up of such interventions.

References

Affuso, E. 2011. *Spatial Autoregressive Stochastic Frontier Analysis: An Application to an Impact Evaluation Study.* Technical report, Technical report.

Alston, J. M., T. J. Wyatt, P. G. Pardey, M. C. Marra, and C. Chan-Kang. 2000. *A meta-analysis of rates of return to agricultural R D.*

Anselin, L. 1988. *Spatial Econometrics: Methods and Models.* 4th ed. 272. Boston: Dordrecht ; Boston : Kluwer Academic Publishers.

Anselin, L., and S. Rey. 2014. *Modern Spatial Econometrics in Practice.* 1st ed. 359. Chicago: GeoDa Press LLC.

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

Berti, P., M. Moursi, J. Kung'u, P. Tugirimana, K. Siekmans, A. Lubowa, and E. Boy. 2012. *Inadequate Diet, but Low Iron Deficiency Anemia, in Rwandan Women and Children.*

Besley, T., and A. Case. 1993. "Modeling technology adoption in developing countries." *The American Economic Review* 83 (2): 396–402.

Bravo-Ureta, B. E., W. Greene, and D. Solís. 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. doi:10.1007/s00181-011-0491-y.

Charness, G., and P. Kuhn. 2011. "Lab Labor: What Can Labor Economists Learn from the Lab?" In *Handbook of Labor Economics,* 229–330. doi:10.1016/S0169-7218(11)00409-6.

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

Conley, T. G., and C. R. Udry. 2010. "Learning about a New Technology: Pineapple in Ghana." *American Economic Review* 100 (1): 35–69.

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

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

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

Foster, A., and M. Rosenzweig. 2010. "Microeconomics of Technology Adoption." *Annual Review of Economics* 2 (1): 395–424.

Funes, J. E. 2018. "The Role of Social Interaction in the Adoption and Geographic Diffusion of an Agricultural Technology: The Case of High Iron Bean (Phaseolus Vulgaris) in Rwanda," 1–36. 2018 AAEA annual meeting program.

Funes, J. E., L. Sun, T. Benson, F. Sedano, G. Baiocchi, and E. Birol. 2019. "Impact of iron biofortified beans on yields and farmers' incomes: The case of Rwanda," 1–30. 2019 AAEA annual meeting program.

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

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

Haas, J. D, S. V. Luna, M. G. Lung'aho, M. J. Wenger, L. E. Murray-Kolb, S. Beebe, J-B Gahutu, and I. M. Egli. 2016. "Consuming Iron Biofortified Beans Increases Iron Status in Rwandan Women after 128 Days in a Randomized Controlled Feeding Trial." *The Journal of Nutrition* 146 (8): 1586–1592.

Han, J., D. Ryu, and R. Sickles. 2016. "How to Measure Spillover Effects of Public Capital Stock: A Spatial Autoregressive Stochastic Frontier Model." In *Advances in Econometrics,* edited by Jeliazkov; Juan Carlos Escanciano; Hillebrand Thomas B. Fomby; R. Carter Hill; Ivanpublisher = Profanity Institute and Eric, 259–294. Emerald Group Publishing. doi:10.1108/S0731-905320160000037017.

Herbst, D., and A. Mas. 2015. "Peer effects on worker output in the laboratory generalize to the field." *Science* 350 (6260): 545–549. doi:10.1126/science.aac9555.

Luna, S., M. Lung, J. B. Gahutu, and J. Haas. 2015. "Effects of an Iron-biofortification Feeding Trial on Physical Performance of Rwandan Women." *European Journal of Nutrition & Food Safety* 5 (June 2014): 5641.

Manski, C. F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *The Review of Economic Studies* 60 (3): 531.

Murray-Kolb, L. E, M. J. Wenger, S. P. Scott, S. E. Rhoten, M. G. Lung'aho, and J. D. Haas. 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* 147 (11): 2109–2117.

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

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

Appendices

A Frontier analysis tables

Table A.1: Estimation results of a OLS, a non-spatial frontier analysis (SFA) model and a spatial frontier model (SSFA): pooled, control and treatment – bush bean growers

	OLS - pool	*SFA-pool*	*SSFA-pool*	*OLS - C*	*SFA-C*	*SSFA-C*	*OLS-T*	*SFA-T*	*SSFA-T*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log(Bean area)	0.514***	0.6413***	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 (dummy)	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 $(\rho)/Lambda(\lambda)$			0.115*			0.185*			0.27**
LMerr	1.385			1.456			6.123***		
LMLag	4.351**			3.321*			1.104		
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
Note:	*p<0.1; **p<0.05; ***p<0.01								

Table A.2: Estimation results of a OLS, a non-spatial frontier analysis (SFA) model and a spatial frontier model (SSFA): pooled, control and treatment – climbing bean growers

	OLS - pool	*SFA-pool*	*SSFA-pool*	*OLS - C*	*SFA-C*	*SSFA-C*	*OLS-T*	*SFA-T*	*SSFA-T*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(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	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)	(0.101)	(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 (ρ)/*Lambda*(λ)			0.091			0.196**			0.236*
LMLag	3.094*			4.217			0.943		
LMerr	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 (df = 1)			2.983*			4.692**			4.240**
TE			0.262			0.239			0.299
Average TE			0.395			0.340			0.472

Note: *p<0.1; **p<0.05; ***p<0.01

Table A.3: 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

Note: *p<0.1; **p<0.05; ***p<0.01