

The role of risk mitigation in production efficiency: A case study of potato cultivation in the Bolivian Andes

*CATHERINE LAROCHELLE

Department of Agricultural and Applied Economics, Virginia Tech, Blacksburg, VA, USA

Email: claroche@vt.edu

**Corresponding author*

JEFFREY ALWANG

Department of Agricultural and Applied Economics, Virginia Tech, Blacksburg, VA, USA

Email: alwangj@vt.edu

The role of risk mitigation in production efficiency: A case study of potato cultivation in the Bolivian Andes

Catherine Laroche¹ and Jeffrey Alwang

Abstract: Using a stochastic production frontier to model potato production in Bolivia, we quantify the costs of environmental and activity diversification in the form of efficiency losses and yield forgone. We find that efficiency decreases with the number of fields in a geographic cluster, distance between the dwelling and a particular field, discontinuity between fields, and off-farm income. However, environmental diversification is more detrimental than activity diversification. Using spatial analysis of field and household efficiency measures we assess production vulnerability to climatic shocks and the potential of environmental diversification in mitigating shocks. We find important spatial clusters of low and high efficiency at the field-level suggesting that climatic shocks influence efficiency measures. Household-level efficiency measures exhibit random spatial patterns suggesting that on average households can mitigate the adverse effects of shocks through environmental diversification.

Key words: Stochastic Production Function, Risk Mitigation, Technical Efficiency, Potato Production, Bolivia

JEL Classifications: D13, Q12, O12

1. INTRODUCTION

In a country like Bolivia where formal insurance mechanisms are rare, small-scale farmers rely on a variety of strategies to manage risk. Many environmental risks such as frost, hail, or drought can be mitigated through self-insurance techniques. The literature distinguishes between two types of self-insurance: risk coping and risk management (Alderman and Paxson, 1992). Risk coping refers to strategies that smooth consumption either intertemporally or across households through risk sharing. Intertemporal consumption smoothing can be achieved through saving and borrowing or through asset accumulation and sales. Risk sharing is used to mitigate income shocks at a given time

¹ Department of Agricultural and Applied Economics, Virginia Tech, Blacksburg, VA 24601, USA. claroche@vt.edu and alwangj@vt.edu.

The authors thank two anonymous reviewers and Editor David Harvey for helpful suggestions. We are also thankful to George Davis for his comments and to researchers at the foundation PROINPA (Cochabamba, Bolivia) for advices, discussions, and help identifying potato plot locations. Funding for this research was provided by the SANREM CRSP, and was made possible by the United States Agency for International Development and the generous support of the American people through USAID Cooperative Agreement No. EPP-A-00-04-00013-00.

across households within a village. Risk management involves actions to reduce income variability, such as crop, field, and income source diversification.

We focus on the costs of risk management and on how activity and environmental diversification translate into efficiency losses in potato production. For small-scale farmers in the Bolivian Andes, potato is the main crop. Compared to other major potato-producing countries, potato yields in the region are low--about 10.6 tons per hectare--compared to 16.3 and 16.8 tons per hectare on average in Latin America and worldwide (Potato World, 2008). Shocks to production in the region include frost, hail, drought, pest infestation, and disease. In order to attenuate environmental risk exposure, producers diversify potato production by cultivating beans, cereals and vegetables, and raise livestock. Risk exposure is also reduced by cultivating potatoes across different microclimatic conditions within walking distance of the dwelling (Dorsey, 1975). Typically, producers cultivate potatoes in valleys where fields are relatively flat and at higher elevations where fields are sloped. Flat fields are easier to manage, but are more vulnerable to hail and frost shocks than sloped fields. Most households in our study area cultivate fields in different micro-regions.

The effectiveness of self-insurance depends on the nature of risk. In the West Africa context, Carter (1997) examines how activity and environmental diversification can reduce household risk by limiting the impact of microclimatic shocks on the production portfolio variance (see also, Alderman and Paxson, 1992). Here, activity diversification (crop diversification) refers to cultivating crops that respond differently to climatic shocks in the same environment (fields or plots), for example intercropping in a single plot a crop that does well in dry conditions with a crop that performs best in humid conditions. Carter (*Op. cit.*) finds mixed results; activity diversification was found to be effective in reducing risk exposure in only one of two regions studied. Environmental diversification (field or plot diversification) involves cultivating the same crop in different microenvironments where risks are not perfectly correlated. Environmental diversification was found to reduce household risk in both regions but by more where shocks are more severe.

Self-insurance techniques have the potential to reduce household vulnerability to environmental shocks, but these mechanisms, like formal insurance, are not costless. An important cost of informal insurance is forgone expected yield, including cultivation of

safer traditional varieties as opposed to riskier, high-yielding varieties. Alternatively, farmers may use purchased inputs less intensively in order to reduce financial risks (Morduch, 1995), using what Fafchamps (1993) describes as “flexible farming”. Fafchamps (*Ibid*) finds that small-scale farmers in Burkina Faso increase their labour effort in response to positive environmental shocks and reduce their labour effort in response to negative shocks. In situations of extreme negative shocks, leading to very low marginal productivity of labour, they may reallocate their labour into alternative activities. While there is no direct reference to the costs of dealing with environmental risk, there is strong evidence of flexibility in farming practices in environments characterized by high vulnerability to climatic shocks.

Other costs are associated with activity and environmental diversification. Gains from specialization are reduced through activity diversification. Costs of farm fragmentation can include time lost walking between fields and increased transportation costs (Carter, *Op. cit.*). While risk management is frequently discussed in the development literature, the costs associated with it are not commonly measured. Using a state-contingent approach, Chambers and Quiggin (2000) develop a general theoretical model of state-contingent choice to identify how producers reduce production uncertainty with costs reflected by foregone output. Carter (*Op. cit.*) estimates the insurance premium households are willing to pay to reduce production variability, using expected utility, risk aversion, and certainty equivalent concepts. However, he could not econometrically assess the cost of risk management in the form of yield forgone as he finds an insignificant relationship between the Simpson land index² and yield. Monchuk *et al.* (2010) would explain this lack of significance by noting that the Simpson land index captures two distinct components of field scattering: the number of plots and variation in plot sizes and that the Simpson’s index is increasing in the number of plots and decreasing in the variability of plot sizes. In order to better capture the impact of field scattering on yield, Monchuk *et al.* propose to replace the Simpson land index with two explanatory variables to distinguish between the effects of plot number and plot size variability. They find that

² The Simpson land index is a measure of land dispersion defined as $S_{it} = 1 - \frac{\sum (a_i^2/A^2)}$ where a_i is the size of a given plot and A is the summation over all plots ($A = \sum (a_i)$). The Simpson land index ranges from 0 to 1, where a value of 0 indicates cultivation in one single field and a value near 1, cultivation in various fields

plot number is a more powerful measure than plot size variability in capturing the impact of land fragmentation on productivity. They also come up with an alternative measure named “effective distance”, which measures the discontinuity between fields and between a particular field and the homestead. Our approach is intuitive and allows us to quantify costs of risk management in terms of efficiency losses, which are easily converted into yield losses.

The objectives of this study are (i) to quantify the costs of environmental and activity diversification in the form of yield forgone, (ii) to spatially analyse production vulnerability to environmental shocks, and (iii) to assess the potential of environmental diversification as a self-insurance strategy. We estimate a stochastic production frontier and model the mean of inefficiency as a function of environmental and activity diversification. We find that efficiency decreases with the number of fields in a geographic cluster, distance between the dwelling and a particular field, discontinuity between fields, and off-farm income.

To show where shocks occur and assess the potential of environmental diversification in mitigating risk, spatial analyses of field and household efficiency measures are performed. We find important spatial clusters of low and high efficiency at the field-level, confirming the presence of climatic shocks and how those are microenvironment-specific. Household average efficiency measures exhibit random spatial patterns, supporting the hypothesis that households can mitigate adverse effects of shocks through environmental diversification.

2. THEORETICAL FRAMEWORK

Our theoretical framework describes how farmers manage environmental risk through activity and environmental diversification and respond to climatic shocks. We assume that households have a production portfolio Y defined as:

$$Y = \sum_{j=1}^N y_j \tag{1}$$

Y defines total output, y_j is the output associated with plot j , and N is the number of plots a given household cultivates. For each plot, households determine which crops to plant and the amount of inputs to allocate. These decisions are based on expectations about plot

productivity, yield variability, and desire to manage risk. Yield, measured in output per hectare, is given by:

$$\bar{y}_j = \frac{y_j}{h_j} \quad (2a)$$

where $h_j = \alpha_j \cdot H$, and α_j represents the land share for plot j and H is the total land.

$$\mu = E(Y) = \sum_{j=1}^N E(y_j) = E(\bar{y}_j \cdot \alpha_j \cdot H) = \alpha_j \cdot H \cdot E(\bar{y}_j) \quad (2b)$$

$$s^2 = \sum_{j=1}^N \alpha_j^2 s_j^2 + \sum_{j=1}^N \sum_{k \neq j}^N \alpha_j \alpha_k s_j s_k r_{j,k} \quad (2c)$$

Equation 2b indicates that the household production portfolio mean (μ) return is the sum of each plot's expected production (y_j), which depends on expected yield, land share, and total land. In 2c, the portfolio variance (σ^2) varies with the proportion of land area (α_j), plot yield variance (σ_j^2), and the correlation coefficient $\rho_{j,k}$, which gives the correlation in yield between two plots. By choosing a combination of activities that have low or negatively correlated returns (i.e. $-1 \leq \rho < 1$), the portfolio variance will be lower than the sum of individual field variances, implying that diversification can reduce risk exposure and production uncertainty. Households are assumed to be risk averse and prefer portfolios with lower variance for a given mean return.

The household objective is to maximize the expected utility of profit of the production portfolio (3a) subject to a cash (3b) and variance (3c) constraints.

$$EU\left(\sum_{j=1}^N y_j P_{y_j} - \sum_{j=1}^N x_j P_{x_j}\right) \quad (3a)$$

$$\sum_{j=1}^N y_j P_{y_j} \geq \sum_{j=1}^N x_j P_{x_j} \quad (3b)$$

$$\sum_{j=1}^N \alpha_j^2 s_j^2 + \sum_{j=1}^N \sum_{k \neq j}^N \alpha_j \alpha_k s_j s_k r_{j,k} \leq s \quad (3c)$$

P_{y_j} represents the price received for plot j 's production (y_j), and x_j and P_{x_j} are the input quantities and costs allocated to plot j . The cash constraint ensures that revenues ($\sum y_j \cdot P_{y_j}$) from the production portfolio are equal to the sum of all input costs ($\sum x_j \cdot P_{x_j}$). The variance constraint specifies that the production portfolio variance is less than or equal to s , where s

is the variance level that ensures that the production portfolio Y will yield (with a certain probability) sufficient returns to meet subsistence needs (Stanley, 2007). This constraint is similar to the safety-first principle introduced by Roy in 1952. In order to meet the variance constraint, households can resort to two risk management strategies: activity diversification and environmental diversification. For simplicity, we assume that households are concerned with managing production risk only³.

We hypothesize that resorting to diversification strategies as opposed to maximizing profits will result in greater inefficiencies in production (deviations below the optimal output level defined by the production frontier). This risk-efficiency hypothesis also requires that the dynamic structure of agricultural production is taken into account (Antle, 1983a). Modelling the dynamics of the production process reflects the potential that farmers resort to flexible farming practices (Fafchamps, 1993) and may reallocate their effort to maximize expected utility across states of nature (Chambers and Quiggin, 2000).

We model a field-level potato production function as a two-stage dynamic process (Antle, 1983a). We focus on the technical efficiency of potato production as plots devoted to potato cultivation represent a large share of the households' production portfolio. This specification allows us to assess the impact of activity diversification (such as bean and cereal cultivation, livestock production, and off-farm activities) on potato technical efficiency. During the first stage, land preparation and planting decisions are assumed to be made, and in the second stage, crops are managed and harvested. This dynamic progress is presented in equations 4a, 4b, and 4c.

$$y_{j1} = f(x_{j1}) \tag{4a}$$

$$y_{j2} = y_{j1} + f(x_{j2}/\theta_j) \tag{4b}$$

$$y_j = f(x_{j1}) + f(x_{j2}/\theta_j) \tag{4c}$$

In stage 1, the household allocates inputs x_{j1} to plot j to maximize expected profit given prices of x_{j1} and y_j , the variance constraint, and expectations about y_j . Producer expectations in stage 1, denoted as $E_1(y_j)$, have a probability distribution shaped by previous shocks and plot-specific agro-ecological conditions, such as elevation and

³ Farmers provide most of the inputs, such as seeds and labour, themselves and do not rely heavily on agricultural markets, being engaged mainly in semi-subsistence farming.

fertility. Once planting decisions have been made but before the start of period 2, field-specific shocks (θ_j) occur. After the shocks, producers update their expectations about y_j , and adjust farming practices accordingly. More precisely, in stage 2 households select the optimal combination of inputs to maximize expected profit based on input and output prices, the variance constraint, and its new expectations about y_j , $E_2(y_j)$, where $E_2(y_j)$ has a probability distribution conditioned by realization of the shocks. Output in the second decision stage (y_{j2}) depends on the realized output in the first stage (y_{j1}), which implicitly depends on previous input allocations (x_{j1}), and inputs in the second stage, which are shock specific, (x_{j2}/θ_j).

In our model, input demands depend on farmers' expectations about output. Since $E_t(y_j)$ is nonstochastic, we can assume that input and output are independent and can estimate this sequential decision-making process with a single-equation as long as the error terms between the input demand functions and production function are independent (Antle, 1983b). This assumption is plausible since the input demand function error terms are likely to reflect human acts, such as human mistakes, while the production function random error term is more likely to reflect natural variation (Zellner *et al.*, 1966).

To best represent the first component of the risk-efficiency hypothesis (risk affects technical efficiency), we employ a stochastic production frontier to model the dynamic nature of potato production. Stochastic frontier analyses, since first introduced by Aigner *et al.* (1977) and Meeusen and van den Broeck (1977), have evolved substantially and various specifications are now available. We employ the stochastic production frontier proposed by Kumbhakar *et al.* (1991), Huang and Liu (1994), and Battese and Coelli (1995), referred to as the KGMHLBC model. We assume that the production technology takes the form of a Cobb-Douglas stochastic production frontier⁴ (5a). Additional assumptions for the KGMHLBC model are: (i) the random error term v_j has a normal distribution with mean zero and variance σ^2_v (5b); (ii) the inefficiency term u_j has a truncated-normal distribution with a mean expressed as linear combination of the covariates z_j , and a variance equal to σ^2_u (5c).

⁴ A translog production was estimated but the Cobb-Douglas production is preferred due to collinearity and loss of degrees of freedom caused by the multiple interaction terms included in the translog function. In addition, returns to scale are likely to be rare in subsistence farming, making the homothetic assumption appropriate.

$$\ln y_j = \ln f(x_{j1}, x_{j2} | \theta_j; \beta_{j1}, \beta_{j2}) + v_j - u_j \quad (5a)$$

$$v_j \sim N[0, \sigma_v^2] \quad (5b)$$

$$u_j \sim N^+[\delta z_j, \sigma_u^2] \quad (5c)$$

Using a stochastic production frontier specification allows us to estimate the costs of environmental and activity diversification strategies. 5c stipulates that the mean of u_j can be modelled as a function of exogenous variables z_j , such that $u_j = \delta z_j$, referred to as the inefficiency model. The variables z_j influence the efficiency with which inputs are converted into outputs. For example, if efficiency across farms is believed to vary according to manager abilities, manager education and experience will enter the inefficiency model. We hypothesize that household ability to manage a given plot depends, inter alia, on the degree of activity and environmental diversification. Using this assumption, we model the mean of u_j as a function of activity and environmental diversification measures.

On and off-farm activity diversification reduces specialization in potato production, adversely affecting a household's ability to manage its potato fields. On average, environmental diversification is expected to reduce production efficiency, but in particular states of nature it might increase it⁵.

We hypothesize that resorting to diversification strategies result in additional deviations below the frontier, also referred as inefficiency compared to pure profit maximization. Our first *hypothesis* is that managing the production portfolio through activity and environmental diversification influences the efficiency of the production process and can explain variations in the mean of u_j , which we test through the joint significance of the z_j variables (5c) associated with activity and environmental diversification, using a likelihood-ratio (LR) test. Since the coefficients in the inefficiency model provide information only on the direction and not on the magnitude of the

⁵ For example, field scattering, the result of environmental diversification, increases transaction costs associated with field management reducing productivity. A pest outbreak could go unnoticed at its early stage in more distant fields if they are monitored less frequently than fields located near the house. Moreover, after walking long distances to reach more distant fields, labour may not be as productive as when working on nearby fields. Farming activities might be performed less frequently (but not necessary less intensively) in remote fields causing farm management practices to be less effective. Six hours of weeding accomplished over a three-week period at a rate of two hours a week will not have the same impact on yield as six hours of weeding accomplished in a single day.

influence, we estimate marginal effects and elasticities of technical efficiency with respect to the z_j variables (Frame and Coelli, 2001; Rahman and Rahman, 2008).

With distributional assumptions on v_j and u_j (5b and 5c), we obtain measures of production efficiency, eff_j , based on the relationship that $eff_j = E\{exp(-u_j)/e_j\}$, where eff_j is the efficiency measure of plot j and $e_j = v_j - u_j$. Efficiency measures can take values between zero and one and correspond to the ratio of observed production to the maximum feasible output (the production frontier) given a set of inputs. In line with Antle (1983a), we argue that the reallocation of inputs in stage 2 impacts the measure of production efficiency. Consider a farmer who manages two fields: *fields j and k*. Assume that in $t=1$ the production functions in the fields are identical such that farmer's expectations about outputs are the same for both fields, and thus equal marginal products across fields. Assume that shocks occurring between $t=1$ and $t=2$ cause *field k's* production function to shift downward. Consequently in $t=2$, *field k* expected output $E_2(y_k) = f(x_{k1}) + f(x_{k2}/\theta_k)$ is lower than *field j* expected output $E_2(y_j) = f(x_{j1}) + f(x_{j2}/\theta_j)$, resulting in a lower marginal product of inputs in *field k* across the whole range of inputs. The optimizing producer will reduce input application in this field, resulting in fewer inputs applied in *field k* in comparison to *field j* in stage 2. Nevertheless, the output produced will never lie on the production frontier as inputs applied in $t=1$ did not yield the anticipated output $E_1(y_k)$ and efficiency measures are influenced by inputs applied in both stages. As Chambers and Quiggin (2000) show, households can trade-off technical efficiency in one state of nature to achieve higher efficiency in another state of nature. Efficiency in risk management is thus disguised as productive inefficiency as producers can consciously select a level of technical efficiency as a means of managing risk.

We assume that a negative shock will be captured in the u_j term and appear as production inefficiency. We expect to observe relatively high measures of efficiency on plots where positive shocks occurred (or for plots not affected by the shock) where the household also owns plots experiencing negative shocks. Household reallocating inputs, such as labour, to a given field following the occurrence of negative shocks in its other plots would be consistent with higher measures of technical efficiency on the unaffected plot. This suggests that in an environmental diversification setting, particular states of

nature, such as simultaneous occurrence of good and bad states, can increase technical efficiency on some fields (Chambers and Quiggin, 2003).

Our *second hypothesis* is that shocks and input reallocation following these shocks influence production efficiency measures eff_j , either positively or negatively. To assess our *second hypothesis*, we analyse the spatial patterns of efficiency measures based on the assumption that fields located in the same microenvironment are affected by similar shocks and have similar efficiency measures. By examining spatial patterns of efficiency, we observe where shocks occurred. We expect to observe spatial clusters of low efficiency where negative shocks occurred and high efficiency where positive shocks occurred. These patterns are likely to be reinforced if, following shocks, inputs (especially labour) are reallocated. We use Global Moran's I statistic (a global statistic for spatial autocorrelation based on variable locations and values) to test for the presence of spatial clustering within the study area. The null hypothesis is that the data do not exhibit any spatial pattern. Rejection of the null with a positive z-score indicates that observations with similar values are clustered spatially while rejecting the null with a negative z-score indicates dispersion of similar observations (Spatial Autocorrelation (Morans I) (Spatial Statistics)). Rejecting the null with a positive z-score would indicate that field efficiencies are spatially clustered, supporting the hypothesis that shocks affect efficiency and that households respond internally to these shocks. Since a global statistic does not answer the question of where the spatial clusters are located, the Local Getis-Ord G_i^* (hot spot analysis) is used to visualize clusters of high and low efficiency when a positive z-score for the General Moran's I statistic is obtained (Hot Spot Analysis (Getis-Ord G_i^*) (Spatial Statistics)).

Our *third hypothesis* is that environmental diversification can be an effective strategy in attenuating climatic shocks affecting potato production. We expect fields located in different microenvironments to be affected by different shocks, and as a result, yields between fields to be weakly or negatively correlated. To explore our third hypothesis, we exploit the differences between spatial patterns of field efficiency and

household efficiency. Since households cultivate generally more than one plot⁶, a measure of household efficiency can be calculated.

$$eff = \sum_{j=1}^n \alpha_j eff_j \quad (6)$$

Equation 6 indicates that the household-specific efficiency measure, *eff*, depends on share of land area devoted to potato production (α_j), and plot efficiency measures (*eff_j*), where *n* is the total number of potato plots cultivated by the household. While we expect fields located near each other to have similar efficiency measures, we do not expect households located nearby to have correlated measures of efficiency. Even if adjacent households have similar characteristics, they are unlikely to cultivate potatoes in the same microenvironments, especially for fields located at higher elevations. Therefore, we expect less pronounced spatial patterns of efficiency at the household-level compared to the field-level. As with our second hypothesis, we employ the Global Moran's I to test for spatial autocorrelation where the variables of interest are household location and its corresponding measure of efficiency.

3. DATA

In 2006/7, 284 Bolivian producers in Tiraque Province, Cochabamba Department, were randomly selected and interviewed. Steep mountainous terrain with slopes ranging from 10 to 40 per cent and elevation between 3000 and 4500 meters characterize the area. Households are organized into 14 communities that comprise approximately 3,000 inhabitants. The 14 communities are located on each side of a paved road between Cochabamba and Santa Cruz, two major cities. Ease of access to the communities and dwellings is variable and depends on their location relative to the paved road. Off the road, transportation is limited and dirt roads are of poor quality. Consequently, isolation increases with distance to the paved road.

Prior to the 1953 land reform, the Tiraque area was dominated by large haciendas typical of rural Bolivia. Under the hacienda system, colonos (workers) were given

⁶More precisely, 41 households cultivate one plot, 36 households cultivate two plots, 26 households cultivate three plots, and 20 households cultivate four plots or more.

usufruct rights to small plots of land in return for their labor. These plots were generally located close to worker households; the remainder of the land formed the hacienda. Beginning in the late 1930s, indigenous campesinos began to form sindicatos (syndicates), which agitated for worker rights. Following the 1953 reform, the colonos were guaranteed access to their original plots, while the sindicatos continued to press for additional land for their members. As a result of this process, access to land followed an erratic process. Some campesinos received additional plots with titles, others farmed additional plots with insecure formal rights, while others purchased parcels (Dorsey, 1975). By the mid-1960s, land holdings in the area were fragmented with most farming more than one parcel, most of which were spread over space (Pienado Sotomayor, 1971). Since the early 1970s, fragmentation has continued with sub-divisions among families and irregular sales of small parcels to younger farmers.

The survey included information on household demographics and composition, agricultural activities and equipment, household revenues and expenses, and gender division of labour. The longitude and latitude of each dwelling were recorded. In order to obtain the geographical coordinates of the potato fields, additional fieldwork was required, to locate farmers' fields on satellite-based maps⁷ of the area. Farmers' unwillingness to reveal these details reduced the final sample size to 292 geo-referenced potato fields⁸ belonging to 124 households.

4. EMPIRICAL SPECIFICATION

⁷ The satellite image, from the Instituto Militar de Ingenieria in Bolivia, is a raster dataset of IMAGINE Image with cells of one meter squared resolution. The area of each potato plot was digitalized on this satellite image (ArcGIS 9.3.1), and field longitude and latitude were extracted based on the plot center. Field coordinates were combined with two GIS data layers: i) a digital elevation map (DEM) downloaded from the Shuttle Radar Topography Mission (SRTM) website (<http://srtm.usgs.gov/>) and interpolated using the Spline methodology to obtain cells of 30 meters resolution (Her and Heatwole, 2008), and ii) a shape file of the soil characteristics of the area. By combining field coordinates with GIS data, we obtain the elevation and severity of soil erosion for each plot. We also used the satellite image to digitize the dirt roads and compute travel path-based distance measures, such as distance between fields and distance between the dwelling and a particular field.

⁸ In the area in Bolivia where the research was conducted, farmers use the word "parcela" which is translated as parcel or plot. Parcels are not divided into sub-parcels and we use the terms plot and field interchangeably.

The technology for potato production is represented by a Cobb-Douglas stochastic production frontier (7a); the inefficiency model is defined by equation 7b.

$$\ln \bar{y}_j = \beta_0 + \sum_{j1}^k \beta_{j1} \ln x_{j1} + \sum_{j2=k+1}^l \beta_{j2} \ln x_{j2} | \theta + \sum_{m=l+1}^n \beta_m \ln q_{m-1} + v_j - \mu_j \quad (7a)$$

$$u_j = d_0 + \overset{\circ}{\underset{n=1}{\hat{a}}}_n d_n ED_j + \overset{\circ}{\underset{k+1}{\hat{a}}}_k d_k AD + \overset{\circ}{\underset{l+1}{\hat{a}}}_l d_l HHH \quad (7b)$$

\bar{y}_j represents potato yield in kilograms per hectare (kg/ha) obtained in the j^{th} plot and is a function of agricultural inputs applied in both periods. Inputs considered in the first stage are seeds (kg/ha), fertilizer (kg/ha), and labour (hours/ha). Inputs in stage 2 are the number of pesticide applications, fertilizer, and labour. We control for the role of field-specific agro-ecological conditions, which affect both yield and risk exposure, by including in the production frontier the elevation and level of soil erosion of each plot. We also include a variable for seed size to quantify the role of seed quality on production. These three variables are denoted by the symbol q_j in Equation 7a. Studies from Bolivia show that higher elevation leads to higher potato yields in all departments (Terrazas *et al.*, 1998) because of lower late blight infestation at higher altitude, reflecting more recent cultivation and drier conditions. However, population pressures have pushed households to cultivate plots at ever-higher altitudes (which also becomes possible due to warmer climate), making the influence of elevation on yield unknown, as these recently cultivated plots are more subject to frost damage due to their very high elevations. To capture the synergy between elevation and reduced pest pressure, an interaction term between elevation and the number of pesticide applications is included in the model. Soil erosion in the study area varies from light to moderate, moderate, and moderate to heavy. A dummy variable representing the last category is included in the frontier to quantify its effect on yield. Since seed quality is a crucial determinant of potato yield, we include a dummy variable for seed tuber size; small tubers tend to produce higher yields than large tubers since the number of lateral buds (or ‘eyes’) increases only slightly as tuber size increases. Since large seed tubers are cut into pieces, large tuber pieces will have fewer buds than small tubers, potentially reducing yield. Moreover, cutting tubers might result in blind

seed pieces (Bohl *et al.*, 1995). Definitions of the variables included in the stochastic production frontier and inefficiency model are shown in Table 1.

[Table 1]

Inefficiency in production (7b) is modelled as a function of environmental diversification (ED), activity diversification (AD), and characteristics of the household head (HHH). The first measure of environmental diversification is the number of clusters cultivated by a given household. We define clusters as circles of 600 meters in diameter⁹, equivalent to 282,744 m². Having two plots at 800 meters away from the dwelling, one to the east and another to the west, is expected to have a different impact on productivity than having two plots at 800 meters away from the dwelling but adjacent to each others. Households normally have clusters at different distances to the main residence. We commonly observe one cluster of fields nearby the dwelling and a second at higher elevation. In a region characterized by steep mountains, a variation of 600 meters can be associated with important fluctuations in agro-climatic conditions such as temperature, soil fertility, and rainfall. The greater the number of clusters a household cultivates, the greater the environmental diversification. The second measure of environmental diversification is the number of fields per cluster, which captures the impact of plot fragmentation *within a* cluster. Monchuk *et al.* (2010) show that land fragmentation can have a detrimental effect on output. However if fragmentation occurs within a given area, its impact might be negligible as it is easier to allocate inputs among geographically clustered fragmented fields. By including variables for the number of clusters and cluster fragmentation, we provide additional insights about how different types of land fragmentation affect efficiency. In our sample, the average cluster contains two plots although some “clusters” are comprised of one plot only, reflecting the important distance certain households have to walk between plots.

We modify the concept of effective distance introduced by Monchuk *et al.* (2010) since their measure captures discontinuity between fields *as well as* between the dwelling and a particular field, while we are interested in the role both measures have on efficiency. For this reason, we include in the inefficiency model the distance between the dwelling

⁹ While determining clusters in ArcGIS, we ensure that for a given household clusters do not overlap.

and a particular field as one variable and the effective distance as a second variable. We define effective distance as a measure of discontinuity between fields only, which is calculated as follows.

$$Eff_Dist_j = \frac{\sum_{k=1, k \neq j}^n dist_{j,k}}{n-1} \quad (8)$$

$dist_{j,k}$ represents the distance (in kilometres) between plot j and plot k and n , the number of plots the household devotes to potato production. A small effective distance indicates that a particular plot is located near or connected to other household potato plots, where a large effective distance implies that a particular field is disconnected from other potato plots. As effective distance increases, transaction costs related to field monitoring and input transportation increase, which can adversely influence farming practices and efficiency. The effective distance measure complements the two previous variables (the number of clusters and the number of fields per cluster), as a household might have two clusters at one kilometre apart when its neighbour has two clusters at three kilometres apart¹⁰. Therefore, effective distance reflects the effort needed to reach all fields and the inefficiency that might arise from fields being scattered geographically. A small effective distance indicates that fields are located within the same clusters or clusters are not far from each other. The squared terms of the four measures of environmental diversification are included to control for potential nonlinearities between these variables and inefficiency.

Three variables capturing the influence of activity diversification are included in the inefficiency model: i) the number of non-potato plots, indicating the level of specialization in potato production relative to other crops (and its square term), ii) revenue from livestock production (Bs), and iii) revenue from off-farm activities (Bs)¹¹.

To conform with previous studies on inefficiency and control for managerial abilities, we include characteristics of the household head (age, education, and gender) in the inefficiency model. The age of the household head would decrease inefficiency if older

¹⁰ These variables measure different dimensions of environmental diversification. Correlation between them is relatively low. The highest correlation is between effective distance and the number of clusters, with a correlation coefficient of 0.47

¹¹ The two revenue variables are normalized by potato revenue to indicate the importance of these activities relatively to potato production since income varies greatly amongst households.

farmers are more experienced and knowledgeable about agricultural production than younger farmers (Battese *et al.*, 1996; Ahmed *et al.*, 2002). Alternatively, age could increase inefficiencies if older farmers are more reluctant to adopt new technologies while younger farmers welcome these innovations (Villano and Fleming, 2004; Boshraadi *et al.*, 2006). In addition, older farmers might have more difficulties coping with the rigors of potato cultivation in the Andes. We expect a positive relation between efficiency and household head education, where education is proxied by a literacy dummy variable (Ahmed *et al.*, 2002).

To examine our first hypothesis, a LR test is performed to determine whether the influences of environmental and activity diversification on inefficiency are jointly significant ($H_0: \delta_n = \dots = \delta_k = \delta_{k+1} = \dots = \delta_l = 0$). Rejecting the null hypothesis indicates that activity and environmental diversification significantly influences production efficiency.

5. RESULTS

The unknown parameters β and δ in equations 7a and 7b are obtained by estimating simultaneously¹² the stochastic production frontier and inefficiency model through maximum likelihood (table 2). The coefficients of the Cobb-Douglas¹³ production frontier represent the output elasticity with the exception of pesticide application and elevation because of inclusion of the interaction term between the two. Output elasticities with respect to these two variables are reported in Table 3 with the marginal effects¹⁴ and elasticities of production and inefficiency.

[Table 2]

Potato yield is highly responsive to the quantity of seed, as average yield would increase by 6.3 kg/ha for an increase in seed of one kg/ha. The coefficient for labour devoted to land preparation and planting is insignificant, implying that labour applied

¹² Wang and Schmidt (2002) have shown the biases that can result from two-step estimation.

¹³ When estimating the Cobb-Douglas production function, the explanatory variables with zero values are handled as suggested in Battese (1997).

¹⁴ While overall the marginal effects are small, they are of similar magnitude as those reported in Frame and Coelli (2001) and Rahman and Rahmann (2008).

during the first stage of production has no measurable effect on yield. However, an additional hour of labour (per ha) in the second stage of production increases potato yield by 4 kg/ha. Additional application of pesticide at the sample mean of 3.7 applications increases average potato yield by 217 kg/ha. Cultivating potato at 100 meters higher elevation than at the 3,652 meters sample mean results in a gain of 151 kg/ha. This result, combined with the significant interaction term between elevation and pesticide application, supports previous findings (Terrazas, V. Suarez et al., 1998) regarding reduced pest pressure at higher altitude. Switching from large to small tubers could increase potato yield by 24 per cent, corresponding to an increase of 2,592 kg/ha at the sample mean.

[Table 3]

Before discussing the efficiency costs of risk management, an overview of field efficiency measures is provided. The average field-level efficiency measure is 56%, which implies that potato yield could be increased by 80% $[(1 - 0.5568)/0.5568]$ if inefficiencies were to be eliminated. The minimum (maximum) efficiency measure is 5% (95%). The low efficiency level may reflect the welfare cost associated with coping with environmental risk.

The LR test statistic has a chi-square value of 45.34 with 12 degrees of freedom, which corresponds to a p-value near zero, strongly rejecting the null hypothesis that the diversification variables are jointly zero, implying that self-insurance in the form of environmental and activity diversification strategies significantly influence production efficiency.

Of the 12 variables representing risk management, six have significant coefficients. The number of fields per cluster and its squared term suggest that inefficiency increases at a decreasing rate with the number of fields per cluster. While this provides evidence of the detrimental effect of land fragmentation on production efficiency, the effect is small. Efficiency would decrease by 0.1% if one plot were added to a given cluster, an average yield loss of about 22 kg/ha- suggesting that land fragmentation when occurring within a small area impedes input allocation and efficiency only minimally. Inefficiency increases linearly with the distance between the dwelling and a particular

field, suggesting that transaction costs associated with moving labour and other agricultural inputs from the dwelling to the field results in time lost and output forgone. An additional kilometre between the dwelling and a particular field decreases efficiency on that field by 0.4%, representing a potato loss of 74 kg/ha. Effective distance and its squared term are both significant, suggesting that discontinuity between fields causes inefficiency. An increase in one kilometre in effective distance would decrease average efficiency by 0.5%, a loss of 87 kg/ha.

For activity diversification, only the coefficient for off-farm income is statistically significant, and this at the 10 per cent level. While the effect is of very small magnitude, it supports the belief that off-farm income negatively affects production efficiency. Non significance of on-farm diversification (such as mixing beans and cereals cultivation and livestock production with potato-growing activities) could be because labour allocated to these activities does not compete with labour devoted to potato production, especially compared with labour devoted to off-farm activities.

Of the three variables capturing the effects of household head characteristics on inefficiency, two are significant and the null hypothesis that effects of these three variables are jointly zero is strongly rejected.

We assess our second and third hypotheses using spatial analyses (Global Moran's I^{15}) of field- and household-level efficiency measures. The null hypothesis that field-level efficiencies are randomly distributed is strongly rejected, showing that fields located near each other have correlated measures of efficiency. This result supports our second hypothesis that environmental shocks affect production efficiency. Since spatial autocorrelation of field efficiency measures is confirmed, a hot spot analysis is conducted to visualize clusters of high and low efficiency. Negative z-score values, represented by the square points in Figure 1, indicate clusters of low efficiency (cold spots). High z-score values, symbolized by the triangles, indicate clusters of high efficiency (hot spots). There are three clusters of high efficiency, one large and two small, all located south of the

¹⁵ The zone of indifference was selected as the type of spatial relationship, which it is a mixed method between fixed distance band and inverse distance. The spatial weights were standardized based on row standardization, which is recommended when the distribution of the features is potentially biased due to sampling design or because of imposed aggregation scheme. Euclidean distance was the selected distance method.

paved road. Clusters of low efficiency are found mainly in the eastern part of the study area.

[Figure 1]

High efficiency clusters are expected to be located where micro-climatic conditions were favourable and should come with increased labour effort in stage 2. Low-efficiency clusters are expected to be found in micro-regions affected by negative shocks and should be associated with a reduction in labour effort in the second stage of production (Fafchamps, 1993). Our data concerning labour effort in the second stage of production are consistent with Fafchamps' findings. On average, households devoted 606 hours of labour per hectare in stage 2. However, for fields located in low efficiency clusters (fields with significant negative z-score), this average drops to 532 hours while for fields located in high efficiency clusters (fields with significant positive z-score), labour effort increases to 666 hours per hectare and this difference is statistically significant at a p-value of 0.003. In addition, a local agronomist reported favourable growing conditions during the 2006/7 agricultural season in the region where the largest hot spot is found. This local expert also mentioned that potato production in this rather flat region is highly vulnerable to frost and hail and frequently damaged. These qualitative observations further suggest that the interactions between agro-ecological conditions and microclimatic shocks are important determinants of production efficiency and highlight the importance of spatial diversification in vulnerable production environments.

Averaging field efficiency measures at the household-level, as stated in equation 6, leads to an average household efficiency measure of 53%. This measure is significantly lower than the average efficiency at the field level (p-value of 0.06). This finding supports the hypothesis that households trade-off efficiency for reductions in risk. The standard deviation of average household efficiency (18) is significantly lower (p-value of 0.01) than the standard deviation of field efficiency (22) resulting in an efficiency distribution that is less widely spread when computed at the household-level as compared to the field-level. This finding provides evidence that environmental diversification is effective in reducing the variability of the household production portfolio.

A similar conclusion is reached when analysing spatial patterns of household efficiency measures. According to a Global Moran's I statistic based on household

location and household efficiency measures, the null hypothesis that household efficiency measures are randomly distributed is accepted. Finding no spatial cluster of efficiency at the household-level supports our third hypothesis that spatial diversification can be a useful strategy to attenuate the adverse effects of microclimatic shocks. This result also provides evidence that households as a group can be well endowed to manage risk. By cultivating plots in different micro-environments, households within a given community are not affected by the same idiosyncratic shocks (as suggested by the lack of spatial correlation among household efficiency measures), indicating that risk sharing can be an effective strategy in reducing seasonal production fluctuations.

6. CONCLUSION

In an environment where formal insurance is rare and vulnerability to climatic risk is high, households resort to self-insurance mechanisms. Adoption of these mechanisms reduces the variance of household production but at a cost of increased apparent inefficiency in production. The average measure of efficiency in the study area is low (56%), which is consistent with an environment characterized by high vulnerability to climatic shocks.

Combining fields and households' geographical coordinates with GIS data allowed us to depict the production environment and to control for agro-ecological conditions that affect both risk exposure and efficiency. GIS technology enabled the creation of powerful variables to capture the effects of environmental diversification on production inefficiency. The spatial analyses showed that field-level efficiencies are clustered over space, indicating the influence of shocks and suggesting the relevance of environmental diversification in the studied area.

The cost of risk management is reflected by increased technical inefficiency. However, these differences are relatively small. A one-unit increase in the number of fields per cluster decreases yield by 22 kg/ha. Yield decreases by 74 kg/ha as the distance between the dwelling and a particular field increases by one kilometer. A one-kilometer increase in the measure of field effective distance results in a yield loss of 87 kg/ha. These results suggest that potato farmers make state-contingent farming decisions and reallocate

farming resources as states of nature are revealed. These strategies, however, imply real costs.

One possible avenue to attenuate costs linked to environmental diversification with a minimal amount of additional risk would be through reciprocity. Labour exchange can reduce inefficiency by reducing labour time losses and lowering input transportation costs (Carter, 1997). Better transportation infrastructure would reduce inefficiency related to travel distances between the dwelling and a particular field and between fields, augmenting potato yield. Moreover, the costs of environmental diversification could be reduced if households could achieve greater flexibility in their farming practices. This could occur if agricultural tasks such as planting, weeding, and harvesting do not have to be performed during the same time window for fields located in different microenvironments. New production technologies such as irrigation schemes, and drought and pest resistant varieties could better allow households to manage their resources over time and reduce vulnerability to environmental shocks.

REFERENCES:

- Ahmed, M.M., Gebremedhin, B., Benin, S. and Ehui, S. 'Measurement and sources of technical efficiency of land tenure contracts in Ethiopia', *Environment and Development Economics*, Vol. 7, (2002) pp. 507-527.
- Aigner, D.J., Lovell, C.A.K. and Schmidt, P. 'Formulation and estimation of stochastic frontier production function models', *Journal of Econometrics*, Vol. 6(1), (1977) pp. 21-37.
- Alderman, H. and Paxson, C.H. 'Do the poor insure? A synthesis of literature on risk and consumption in developing countries', *Working paper, The World Bank*, Vol. WPS 1008, (1992) pp. 44.
- Antle, J.M. 'Incorporating Risk in Production Analysis', *American Journal of Agricultural Economics*, Vol. 65(5), (1983a) pp. 1099-1106.
- Antle, J.M. 'Sequential Decision Making in Production Models', *American Journal of Agricultural Economics*, Vol. 65(2), (1983b) pp. 282-290.
- Battese, G.E. 'A note on the estimation of Cobb-Douglas production functions when some explanatory variables have zero values', *Journal of Agricultural Economics*, Vol. 48(2), (1997) pp. 250-252.
- Battese, G.E. and Coelli, T.J. 'A model for the technical inefficiency effects in a stochastic frontier production function with panel data', *Empirical Economics*, Vol. 20, (1995) pp. 325-332.
- Battese, G.E., Malik, S.J. and Gill, M.A. 'An investigation of technical inefficiencies of production of wheat farmers in four districts of Pakistan', *Journal of Agricultural Economics*, Vol. 47(1), (1996) pp. 37-49.
- Bohl, W.H., Nolte, P., Kleinkopf, G.E. and Thornton, M.K. (1995). Potato Seed Management: Seed Size and Age. University of Idaho, College of Agriculture, Cooperative Extension System Retrieved June 16, 2011, 2011, from <http://www.cals.uidaho.edu/edComm/pdf/CIS/CIS1031.pdf>.
- Boshraadi, H.M., Villano, R. and Fleming, E. 'Production Relation and Technical Inefficiency in Pistachio Production Systems in Kerman Province of Iran', *Working Paper Series in Agricultural and Resource Economics*, (2006-8), (2006) pp. 1-28.
- Carter, M.R. 'Environment, Technology, and the social articulation of Risk in West African agriculture', *Economic Development and Cultural Change*, Vol. 45(3), (1997) pp. 557-590.
- Chambers, R.G. and Quiggin, J. *Uncertainty, production, choice, and agency : the state-contingent approach* Cambridge; New York, Cambridge University Press, 2000.
- Chambers, R.G. and Quiggin, J. 'Price Stabilization and the Risk-Averse Firm', *American Journal of Agricultural Economics*, Vol. 85(2), (2003) pp. 336-347.
- Dorsey, J.F. *A case study of ex-hacienda Toralapa in the Tiraque region of the upper Cochabamba valley*, Land tenure center, University of Wisconsin, Madison 1975.
- Fafchamps, M. 'Sequential labor decisions under uncertainty: an estimable household model of west-african farmers', *Econometrica*, Vol. 61(5), (1993) pp. 1173-1197.
- Frame, W.S. and Coelli, T.J. 'U.S. Financial services consolidation: The case of corporate credit unions', *Review of Industrial Organization*, Vol. 18, (2001) pp. 229-242.

- Her, Y. and Heatwole, C. 'Assessment of Interpolation methods for refining SRTM and DEM', *ASABE Annual International Meeting*, Providence, Rhode Island, USA, 2008, pp. 1-9.
- Hot Spot Analysis (Getis-Ord Gi*) (Spatial Statistics). ArcGIS Resource Center Retrieved June 16, 2011, from [http://resources.esri.com/help/9.3/arcgisdesktop/com/gp_toolref/spatial_statistics_tools/hot_spot_analysis_getis_ord_gi_star_spatial_statistics .htm](http://resources.esri.com/help/9.3/arcgisdesktop/com/gp_toolref/spatial_statistics_tools/hot_spot_analysis_getis_ord_gi_star_spatial_statistics.htm).
- Huang, C.J. and Liu, J.-T. 'Estimation of a non-neutral stochastic frontier production function', *Journal of Productivity Analysis*, Vol. 5(2), (1994) pp. 171-180.
- Kumbhakar, S.C., Ghosh, S. and McGuckin, J.T. 'A Generalized Production Frontier Approach for Estimating Determinants of Inefficiency in U.S. Dairy Farms', *American Statistical Association*, Vol. 9(3), (1991) pp. 279-286.
- Liu, Y. and Myers, R. 'Model selection in stochastic frontier analysis with an application to maize production in Kenya', *Journal of Productivity Analysis*, Vol. 31, (2009) pp. 33-46.
- Meeusen, W. and Broeck, J.v.d. 'Efficiency estimation from Cobb-Douglas production functions with composed error', *International Economic Review*, Vol. 18(2), (1977) pp. 435-444.
- Monchuk, D., Deininger, K. and Nagarajan, H. 'Does land fragmentation reduce efficiency: Micro evidence from India', *Agricultural and Applied Economics Association Annual Meeting*, Denver Colorado, 2010.
- Morduch, J. 'Income smoothing and consumption smoothing', *The journal of economic perspectives*, Vol. 9(3), (1995) pp. 103-114.
- Pienado Sotomayor, M. *Land reform in three communities of Cochabamba, Bolivia*, Land tenure center, University of Wisconsin, Madison, 1971.
- Potato World. (2008). International Year of the Potato Retrieved June 15 2011, from <http://www.potato2008.org/en/world/index.html>.
- Rahman, S. and Rahman, M. 'Impact of land fragmentation and resource ownership on productivity and efficiency: the case of rice producers in Bangladesn', *Land Use Policy*, Vol. 26, (2008) pp. 25-103.
- Spatial Autocorrelation (Morans I) (Spatial Statistics). ArcGIS Resource Center Retrieved June 16, 2011, from [http://resources.esri.com/help/9.3/arcgisengine/java/gp_toolref/spatial_statistics_tools/spatial_autocorrelation_morans_i_spatial_statistics .htm](http://resources.esri.com/help/9.3/arcgisengine/java/gp_toolref/spatial_statistics_tools/spatial_autocorrelation_morans_i_spatial_statistics.htm).
- Stanley, D.L. 'Risk Management in Gathering Economies', *Journal of Development Studies*, Vol. 43(6), (2007) pp. 1009-1036.
- Terrazas, F., V.Suarez, G.Gardner, Thiele, G., Devaux, A. and Walker, T. 'Diagnosing Potato Productivity in Farmer's Fields in Bolivia', *International Potato Center Social Science Department Working paper No. 1998-5*, (1998) pp. 1-71.
- Villano, R. and Fleming, E. 'Analysis of technical Efficiency in Rainfed Lowland Rice Environmetn in Central Luzon Philippines Using a Stochastic Frontier Production Function with a Heteroskedastic Error Structure', *Working Paper Series in Agricultural and Resource Economics*, (2004-15), (2004) pp. 1-28.
- Wang, H.-J. and Schmidt, P. 'One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels', *Journal of Productivity Analysis*, Vol. 18, (2002) pp. 129-144.

Zellner, A., Kmenta, J. and Dreze, J. 'Specification and Estimation of Cobb-Douglas Production Function Models', *Econometrica*, Vol. 34(4), (1966) pp. 784-795.

Table 1: Summary statistics of the variables included in the stochastic production frontier and inefficiency model

Variables	Description	Mean	Std. Dev	N
YIELD	Potato yield (kg/ha)	10,647.5	5,377.1	287
<i>Stochastic production frontier</i>				
SEED	Quantity of seed (kg/ha)	1,383.3	300.6	287
FERT_T1	Quantity of fertilizer (N-K-P kg/ha) in stage 1	212.1	170.9	287
FERT_T2	Quantity of fertilizer (N-K-P kg/ha) in stage 2	136.8	127.8	287
LABOUR1	Quantity of labour in stage 1 (hours/ha)	496.7	314.1	287
LABOUR2	Quantity of labour in stage 2 (hours/ha)	606.0	345.5	287
PESTAPL	Number of pesticide applications	3.7	1.6	287
ELEVATION	Elevation (meters)	3,652.2	151.4	287
DEROSION	Dummy whether erosion is heavy (1/0)	0.2	0.4	287
DSEEDS	Dummy whether seeds tuber are small (1/0)	0.6	0.5	287
<i>Inefficiency model</i>				
NBCLUSTER_600m	Number of clusters (600 meters diameter)	1.7	0.8	123
NBFIELD_600m	Number of fields per cluster	1.9	1.1	287
DIST_F_HH	Distance between field and residence (km)	1.6	2.0	287
EFF_DIST	Effective distance (km)	1.3	1.7	287
NONP_PLOT	Number of non-potato plots	1.8	1.4	123
OFF_FARM_I	Off-farm income (Bs. Normalized by potato income)	128.0	741.4	123
LIVES_I	Income from livestock (Bs. Normalized by potato income)	54.5	557.8	123
AGEH	Household head age	45.3	14.1	123
LITERACYH	Household head literacy (1.Literate/0.Illiterate)	0.9	0.4	123
GENDERH	Household head gender (1.Female/0.Male)	0.2	0.4	123

*The following filters were to eliminate potential outliers: i) yield exceeding 30 000 kg/ha; ii) seeding rates exceeding 2 600 kg/ha; iii) fertilizer applications exceeding 2 000 kg/ha; iv) labour applications equals to zero or exceeding 5 000 hours/ha. Our final sample includes 123 households and 287 fields.

Table 2: Results of the stochastic production frontier and inefficiency model

Variables	PRODUCTION FRONTIER		INEFFICIENCY	
	Coefficients	P-values	Coefficients	P-values
LN(SEED)	0.81	0.00		
LN(FERT_T1)	0.04	0.50		
LN(FERT_T2)	-0.05	0.52		
LN(LABOUR1)	0.03	0.69		
LN(LABOUR2)	0.22	0.00		
LN(PESTAPL)	21.16	0.04		
LN(ELEVATION)	3.74	0.01		
DEROSION	-0.08	0.18		
DSEEDS	0.24	0.00		
LN(PESTAPL) X LN(ELEVATION)	-2.57	0.04		
NBCLUSTER_600m			0.37	0.41
NBCLUSTER_600m SQ			-0.19	0.10
NBFIELDS_600m			0.40	0.06
NBFIELDS_600m SQ			-0.09	0.03
DIST_F_HH			0.14	0.05
DIST_F_HH SQ			-0.01	0.19
EFF_DIST			0.27	0.04
EFF_DIST SQ			-0.04	0.06
NONP_PLOT			0.04	0.32
NONP_PLOT_SQ			-0.01	0.37
OFF_FARM_I			0.01*	0.07
LIVES_I			0.01*	0.22
AGEH			0.01	0.02
LITERACYH			-0.26	0.13
GENDERH			-0.39	0.04
CONSTANT	-28.50	0.01	-0.29	0.65
Gamma	0.95			
LN(sigma2)	-1.08			
Log likelihood	-153.38			
Number of observations	287			

Note: * means that coefficients were multiple by 100 (coefficients x 100)

Table 3: Elasticity and marginal effect for the production variables and elasticity, marginal effect, and yield effect for the efficiency variables

Variables	PRODUCTION		EFFICIENCY		
	Elasticity (%)	Marginal effect (kg/ha)	Elasticity (%)	Marginal effect (%)	Yield effect (kg/ha)
SEED	0.82	6.26			
LABOUR2	0.22	3.90			
PESTAPL	0.08	215.90			
ELEVATION	0.52	1.51			
DSEEDS	0.24	2591.78			
NBFIELD_600 m			-0.04	-0.1*	-22.27
DIST_F_HH			-0.01	-0.4*	-73.84
EFF_DIST			-0.02	-0.5*	-87.45
OFF_FARM_I			-0.06*	0.001*	-0.06
AGEH			-0.02	-0.03*	-5.30
GENDERH			0.3*	0.01	217.40

Note: * means that coefficients were multiple by 100 (coefficients x 100)

Figure 1: Hot spot analysis for field-level efficiency

