CHAPTER 8

Determinants of Banana Productivity and Technical Efficiency in Uganda

Fredrick Bagamba, Ruerd Ruben, and Mariana Rufino

This chapter analyzes the technical efficiency of banana production among Ugandan smallholders by estimating a stochastic production frontier model with inefficiency effects. The empirical models are formulated within the overarching framework of the agricultural household specified in Chapter 2. The data design is also summarized in Chapter 2, although additional information about soil quality was collected for this analysis from three sites selected for this purpose. Hypotheses are tested about the regional effects on productivity relationships and returns to scale. Specific hypotheses are tested with respect to the impact of market access and household- and farm-specific factors on technical efficiency. This chapter examines banana productivity as it relates to two important constraints in banana production: soil fertility and labor; it therefore complements the analysis in Chapter 6, which treats biotic constraints and the potential adoption of resistant cultivars, and Chapter 7, which examines the adoption of soil fertility management practices by farmers.

Stochastic Frontier Production Function

A large body of theoretical and empirical literature has investigated the measurement of efficiency of farm enterprises, using various methods. Ali and Byerlee (1991) have emphasized that the focus in analyzing economic efficiency should be the performance of the whole production system, including farmers and institutional support systems. These results can be used to pinpoint the factors that impede the capacity of farmers to reach their productivity potential.

Technical efficiency (TE) can be estimated using one- or two-step approaches. In the two-step procedure, the production frontier is estimated first, and the technical efficiency of each firm is derived subsequently. In the second step, the derived technical efficiency variable is regressed against a set of variables that are hypothesized to influence the firm's efficiency (Kalirajan 1981; Pitt and Lee 1981). However, the two-stage procedure lacks consistency in assumptions about the distribution of the inefficiencies. In step one, it is assumed that inefficiencies are independently and identically distributed in order to estimate their values. In step two, estimated inefficiencies are assumed to be a function of a number of firm-specific factors, violating the assumption in step one (Coelli, Rao, and Battese 1998). To overcome this inconsistency, Kumbhakar, Gosh, and McGuckin (1991) suggest estimating all the parameters in one step. In a one-step procedure, which we adopt for this study, the inefficiency effects are defined as a function of the farm-specific factors and incorporated directly into the maximum likelihood (ML) estimate.

TE is measured as a ratio of actual to potential output (Aigner, Lovell, and Schmidt 1977; Meeusen and van den Broeck 1977). Approaches for measuring TE generally vary from programming (nonparametric) approaches to statistical estimation (parametric) approaches, depending on functional forms and techniques for estimating the potential output (Forsund, Lovell, and Schmidt 1980; Bauer 1990; Fried, Lovell, and Schmidt 1993; Coelli 1995; Kalirajan and Shand 1997). In analyzing farm-level data where measurement errors are substantial and weather is likely to have a significant effect, the stochastic frontier method is usually recommended (Coelli 1995).

Early frontier production functions that followed Farrell (1957) were deterministic in that they assumed a strict one-sided error term (Schmidt 1986; Coelli 1995). One of the major criticisms against deterministic frontier estimates is that no account is taken of the possible influence of the measurement errors and other data noises on the shape and the positioning of the estimated frontiers. All the observed deviations from the estimated frontier are assumed to be a result of technical inefficiency (TI; Coelli, 1995). Aigner, Lovell, and Schmidt (1977) and Meeusen and van den Broeck (1977) proposed a stochastic frontier production function, in which sources of data noise are accounted for by adding a symmetric error term to the non-negative error. The parameters of this model are estimated by ML, given suitable distributional assumptions for the error terms. The stochastic frontier is not, however, without problems. The major limitation is that one has to make arbitrary assumptions regarding the functional form of the frontier and the distributional form of the error. Moreover, as the model is estimated by ML, the solution obtained might not be optimal, because the likelihood function is not globally concave and allows for multiple local maxima (Maddala 1971).

Using the statistical estimation approach, we define a farm specific stochastic production frontier involving outputs and inputs as:

$$y_i^* = f(x_i)\exp(v_i) \qquad (1)$$

where y_i^* is the maximum possible stochastic potential output from the ith farm, x_i is a vector of m inputs, and v_i are statistical random errors assumed to be distributed as $N(0,\sigma_v^2)$. The production realized on the ith farm can be modeled as:

$$y_i = y_i^* \exp(-u_i) \qquad (2)$$

where y_i^* is the maximum possible stochastic potential output from the ith farm, x_i is a vector of m inputs, and v_i are statistical random errors assumed to be distributed as $N(0, \sigma_v^2)$. The production realized on the ith farm can be modeled as:

$$TE = \exp(-u_i) = \frac{y_i}{y_i^*}. \qquad (3)$$

Substituting equation (1) into equation (2) and taking logs on both sides, we get:

$$\ln y_i = \ln f(x_i) + v_i - u_i\,, \qquad (4)$$

where y_i denotes the production of the ith farm ($i = 1,2, \dots , n$); x_i is a $(1 \times k)$ vector of functions of input quantities used by the ith farm; each is assumed to be an independently and identically distributed random error independent of every u_i, and u_i is a one-sided error term representing the technical inefficiency of farm i.

Subtracting v_i from both sides of equation (4), the production of the ith farm can be estimated as:

$$1\text{n}\; y'_i = 1\text{n}\; f(x_i) - u_i\,. \qquad (5)$$

The efficient level of production can be defined as:

$$\ln \hat{y}_y = \ln f(x_i) \,. \tag{6}$$

From equations (5) and (6), we can compute TE given by:

$$\ln TE_i = \ln y' - \ln \hat{y} = -u_i \tag{7}$$

$TE_i = e^{-u_i}$ and is constrained to be between 0 and 1. When $u_i = 0$, then $TE = 1$ and production is said to be technically efficient.

The distribution of u_i could be half normal with zero mean, truncated normal (at mean μ), or based on conditional expectation of the exponential ($-u_i$). There are no a priori reasons for choosing a specific distributional form, because each has advantages and disadvantages (Coelli, Rao, and Battese 1998). The half normal and exponential distributions have a mode of zero, implying that most firms being analyzed are efficient. The truncated normal allows for a wide range of distributional shapes, including nonzero modes, but is computationally more complex (Coelli, Rao, and Battese 1998).

We adapt the model proposed by Battese and Coelli (1995), in which the TI effects are defined by:

$$u_i = \mathbf{z}_i \boldsymbol{\delta} + w_i \,, \tag{8}$$

where $\mathbf{z}_i$ is a $(1 \times m)$ vector of explanatory variables associated with the TI effects; $\boldsymbol{\delta}$ is a $(m \times 1)$ vector of unknown parameters to be estimated; and w_i is an unobservable random variable. The parameters indicate the impacts of variables in $\mathbf{z}$ on TE. A negative value suggests a positive influence on TE and vice versa. The frontier model may include intercept parameters in both the frontier and the model for the inefficiency effects, provided the inefficiency effects are stochastic and not merely a deterministic function of relevant explanatory variables (Battese and Coelli 1995).

The null hypothesis that the TI effects are not random is expressed by H_0: $\sigma_v = 0$. Accepting the null hypothesis that $\sigma_v = 0$ would indicate that σ_u^2 is zero and thus the term u_i should be removed from the model, leaving the specification that can be consistently estimated by ordinary least squares (Coelli 1994). Further, the null hypothesis that the impact of the variables included in the inefficiency effects model in equation (8) on the TI effects is zero is expressed by H_0: $\boldsymbol{\delta}' = \mathbf{0}$, where $\boldsymbol{\delta}'$ denotes the vector ($\boldsymbol{\delta}$) with the constant term (δ_0) omitted, given that it is included in the expression $\mathbf{z}_i\boldsymbol{\delta}$ (Battese and Broca 1997).

Factors Affecting Technical Efficiency

In crop production, TE is likely to be affected by a wide range of farm- and village-specific factors. Forsund, Lovell, and Schmidt (1980) argue that inefficiency is typically related to factors that are associated with farm management practices. Such factors include education, family size and composition, experience, proximity to markets, and access to credit. Education, which is directly related to management skills, has received adequate attention in the efficiency literature (Weir 1999; Tian and Wan 2000; Weir and Knight 2000; Binam et al. 2003). The results of the impact of education on TE are mixed, with some studies showing positive impact (Belbase and Grabowski 1985; Kalirajan and Shand 1986; Bravo-Ureta and Pinheiro 1997) and others showing a negative impact (Kalirajan 1984, 1991; Phillips and Marble 1986; Bravo-Ureta and Evenson 1994). Education increases the household's ability to utilize existing technologies and attain higher efficiency levels (Battese and Coelli 1995). In our study, we use education of household as a proxy for management skills and age of household head as a proxy for experience (learning by doing). TE is expected to increase with age as the farmer gains experience, but at a decreasing rate as the farmer becomes elderly. Access to resources (land, labor, and capital) is one of the reasons for this type of behavior. Young households are deficient in resources and might not be able to apply inputs or implement certain agronomic practices sufficiently quickly. Timely application of inputs and implementation of management is expected to enhance efficiency. The other factor that explains the quadratic relationship between age and efficiency reflects access to information. Elderly farmers are less likely to have contacts with extension and training programs, and are therefore less willing to adopt new practices and modern inputs (Hussain 1989).

Gender of the household head is expected to have significant effects on TE. Farms managed by men are expected to attain higher TE than those managed by women. Men have more access to resources and information and are more likely to obtain credit, which increases production efficiency.

The effect of household size on TE has not been widely reported in the literature. Household size is expected to influence TE through its effect on the labor endowments of households (including child labor). Large households are expected to be more technically efficient, because they can implement activities on time, attaining higher output with the same or less labor input. The effect of more adults per household on TE is expected to produce mixed results. On the one hand, an increase in the number of adults in the family could increase TE if it results in increased labor devoted to banana production. On the other hand, the effect could be negative if adults have higher chances of obtaining off-farm employment. The effect could be insignificant if labor withdrawn from the farm into off-farm employment is substituted with capital inputs.

Another factor for which the effect on TE has been infrequently reported in the literature is proximity to factor markets. Households located nearer factor markets are expected to have higher TE than those located in remote areas. Proximity to good roads increases access to training and extension programs, from which farmers can attain information and skills for better crop management. Proximity to markets also increases farmers' access to credit facilities and income-generating activities (such as off-farm employment) that enable them to buy and apply inputs on time. By contrast, access to markets may increase the access farmers have to alternative employment with higher returns than from farming, leading them to reallocate labor from farm to nonfarm activities. Households located in remote areas that have greater access to farm labor are expected to attain higher efficiencies than do households in close proximity to nonfarm labor markets.

Production Function

A number of functional forms have been used in the empirical estimation of frontier models. The simplest, the Cobb-Douglas, is specified in logarithmic form as:

$$\ln y = \ln A + b_1 \ln x_1 + b_2 \ln x_2 . \quad (9)$$

The transcendental production function, which is a generalized Cobb-Douglas function, is:

$$\ln y = \ln A + b_1 \ln x_1 + b_2 \ln x_2 + c\, {}^{x_i}\!/_{x_2} . \quad (10)$$

A more complex form, the transcendental logarithmic (translog) form is:

$$\ln y = \ln A + \sum_i b_i \ln x_i + \tfrac{1}{2} \sum_i \sum_j b_{ij} \ln x_i \ln x_j . \quad (11)$$

The most commonly used function forms are the translog and Cobb-Douglas function. Often preferred for its simplicity, the Cobb-Douglas imposes restrictions on returns to scale and elasticities; the translog imposes no restrictions on returns to scale but suffers from multicollinearity and degrees-of-freedom problems. In any case, the impact of functional form on estimated efficiency has been reported to be very limited (Kopp and Smith 1980). Battese and Broca (1997) recommend approaches in which more general model specifications and assumptions are made and simpler formulations are formally tested. In our estimations of the frontier production functions, we use each of the three function forms to estimate the production of cooking bananas. We then compare the results of the inefficiency effects across the three forms.

Accounting for Soil Nutrients and Organic Matter

Agricultural production in Uganda, as in many other developing agricultural economies, depends largely on land and labor input, with little or no external inputs used. The soils are poor in nutrients and rely on recycling of nutrients from soil organic matter (SOM) to maintain crop productivity. The soil's ability to retain and supply nutrients to a crop depends on the cation exchange capacity (CEC)—soils with high CEC are able to bind more cations, such as K^+, to the exchange sites of clay and SOM particle surfaces. Soils with high CEC also have a greater battering capacity and thus the ability to resist changes in pH. Thus soils with high amounts of clay and/or SOM typically have higher CEC and buffering capacities than do more silty or sandy soils. Soil pH also affects nutrient retention and availability to crops. Soils with high pH have low concentrations of H^+, which enables more base cations to be on the particle exchange sites, thus making the soil less susceptible to leaching. With the exception of P, which is most available within a pH range of 6 to 7, other macronutrients (N, K, Ca, Mg, and S) are more available within a pH range of 6.5 to 8. High rainfall can result to soil acidity (Tisdale et al., 1993). Rufino (2003) found that unfavorable soil pH limits maximum yield in 42 percent of the banana plots in Bamunanika, Kisekka, and Ntungamo, which is indicative of other soil fertility problems. In the same sites, soil K was a limiting factor for 19 percent of the banana plots, N was limiting in 12 percent, whereas P was not a limiting factor. Exchangeable K is determined by the neutral ammonium acetate method (Thomas 1982). Available P is determined by the Olsen method (Olsen et al. 1954).

There is a need to take into considerations the interrelations between N, K, SOM, soil texture, and chemical characteristics in modeling production behavior. First, SOM is affected by the soil texture and drainage (sand content), C:N ratios of organic materials, climate, and cropping practices. The SOM content can be estimated as:

$$\ln SOM = \alpha_0 + \alpha_1 sand + \alpha_2 D1 + \alpha_3 D2\,, \quad (12)$$

where ln *SOM* is the natural log of soil organic matter content (percent), $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ are parameters to be estimated, *sand* is the ratio of sand to (sand + clay + silt; percent), and $D1$ and $D2$ are village dummies for measuring the impact of differences in climate and cropping practices. Equation (12) can be estimated by ordinary least squares to obtain the estimates of $\alpha_0, \alpha_1, \alpha_2, \alpha_3$.

Soil N is highly correlated to SOM, organic amendment (mainly animal manure), and regional characteristics and can be estimated as:

$$\ln N = \theta_0 + \theta_1 \ln SOM + \theta_2 M + \theta_3 D1 + \theta_4 D2\,, \quad (13)$$

where ln *N* is the natural log of soil N content (percent), $\theta_0, \theta_1, \theta_2, \theta_3, \theta_4$ are parameters to be estimated, and *M* is animal manure input (kg/year). The remainder of the variables are as already defined. Equation (13) can be estimated using a two-stage least squares model in which ln *SOM* is instrumented by *sand*.

Availability of soil K is affected by soil pH, SOM content, and additions of crop residues and can be estimated as:

$$\ln K = \delta_0 + \delta_1 pH + \delta_2 \ln SOM + \delta_3 C + \delta_4 D1 + \delta_5 D2\,, \quad (14)$$

where ln *K* is the natural log of available soil K (meq/100 g soil), δ_{0-5} are parameters to be estimated, *pH* is soil pH, and *C* is crop residue input (kg/year). Equation (14) is estimated using two-stage least squares, again instrumenting ln *SOM* with the *sand* variable.

Crop output is determined by labor input, area allocated to the crop, and nutrient availability (mainly N and K for bananas). Organic amendment (animal manure, grass mulch, and crop residues) contribute to soil nutrients but also to the physical and chemical properties of soil, enabling a given land area to produce higher output. Crop output can be modeled as:

$$\begin{aligned} \ln Y = \beta_0 + \beta_1 \ln A + \beta_2 \ln L + \beta_3 M \cdot \\ \beta_4 C + \beta_5 \ln SOM + \beta_6 \ln K \cdot \\ \beta_7 D1 + \beta_8 D2 \qquad (25) \end{aligned}$$

where ln Y is the natural log of crop output (kg/year), β_{0-8} are parameters to be estimated, ln A is the natural log of the area allocated to crop (cooking bananas; in acres), and ln L is the natural log of labor input (hours/year). Equation (15) can be estimated using two-stage least squares, with *sand* as the instrument for *SOM* and *pH* as that for for *K*.

To obtain efficient estimates, equations (12), (14), and (15) are estimated simultaneously using a three-stage least squares method, which is the most appropriate technique to use to estimate a system of equations with endogenous variables included on the right hand side.

The three equations (12), (14), and (15) can be collapsed into a reduced-form equation:

$$\begin{aligned} \ln Y = \beta_0 + \beta_1 \ln A + \beta_2 \ln L + \beta_3 M \\ \beta_4 C + \beta_5 sand + \beta_6 pH \\ \beta_7 D1 + \beta_8 D2. \qquad (16) \end{aligned}$$

Endogeneity

Equation (16) is estimated using ordinary least squares. A problem could arise if labor input were endogenously determined. We test for endogeneity by first estimating the labor equation with wage rate, output price, household characteristics, and opportunities included on the right hand side. The residual obtained from the estimated labor equation is then included on the right hand side in the production function estimation. If the effect of the residual turns out to be significant (5 percent), then labor input is confirmed as endogenously determined. An instrumental variable approach or two-stage least squares would be the approach to use to obtain efficient and consistent estimates if valid instruments are available. If the soil quality variables are included in equation (9), ordinary least squares is valid for obtaining consistent and efficient estimates of manure and other organic amendments. When soil quality variables (*sand* and *pH*) are missing in equation (16), the manure and crop residue variables can be treated as endogenous, because farmers would tend to apply these inputs where soils are poor, and no application is carried out where the soil is fertile. We lack sufficient and valid instruments for manure and crop residues. Therefore the estimates for manure and crop residue should be interpreted with care. In the absence of endogenous variables on the right hand side, equation (16) can be consistently estimated using a stochastic frontier approach.

Data

The data design is described in Chapter 2 and Appendix D. Of the total sample of 660 farmers surveyed in Uganda, data for 508 were usable in the analysis. The production function is estimated for cooking bananas, whereas the whole sample was selected for farmers that grow bananas. Some farmers, especially in the lower elevation areas, had banana plots that were less than 2 years old and harvested no output. Others had abandoned plots and did not allocate labor to them. These farms were not included in the estimation. Some households had missing cases in some of the variables, and therefore were excluded from the sample.

The final sampling frame consisted of 27 subcounties, of which 3 were selected (Ntungamo, Bamunanika, and Kisekka) to complement soil analyses. Ntungamo subcounty represents the high-elevation region, where banana production levels are high and there are no evident signs of decline, whereas Kisekka and Bamunanika subcounties represent the low-elevation region, where there is serious decline in yield and production. The sample stratification enables us to capture elevation-related effects (differences in rainfall, temperature, pests, and disease pressure) on banana production (Chapter 2 and Appendix D).

Definitions of variables and summary statistics are shown in Table 8.1. The average area under cooking bananas is 0.58 acres. More area is allocated to cooking bananas in the high-elevation compared with that of the low-elevation region. The trend for allocation of labor and other inputs (manure, grass mulch, and crop residues) is the same, with farmers at high elevations applying greater quantities of inputs compared to those applied by farmers at low elevations.

Farm sizes are considerably larger in the low-elevation region compared to the highlands. Banana plantations are longer lasting at high elevation than at low elevation. Black Sigatoka and weevil scores are somewhat higher for low elevations compared to those for the high-elevation region. Heads of households are slightly more aged and more educated, and more households are located in remote areas in the low-elevation than in the high-elevation region. The average distance from the highway to farms is 14.6 km for low elevation and 10.3 km for high elevations. Farmers in the low-elevation region have higher access to credit compared to those in the highlands. The average amount received in remittances and rent is lower for the low-elevation region compared to the amount received in the high-elevation region.

Results

The hypothesis that labor is endogenously determined in the production of cooking bananas is rejected in the case of the high-elevation region but not for low elevations. The residual variable, included in the second stage of the two-stage least squares method, is found to have no significant effect in the case of high elevation, leading to the rejection of the hypotheses that labor is endogenous (Table 8A.1 in the Supplementary Tables section of this chapter). The endogeneity hypothesis assumes a two-way causal relationship in which farmers are thought to rely on the expected output in deciding about the amount of labor to allocate to production of cooking bananas. At the same time, the amount of labor allocated would determine the output obtained from the production process. Rejection of the endogeneity hypothesis implies that labor used in the production of cooking bananas is exogenously determined, independently of the expected output. However, for low elevations, labor input is most likely not predetermined, and estimates from the frontier production function need to be interpreted cautiously.

Results of the frontier function are shown in Table 8.2. Results from the Cobb-Douglas function show that output responds positively to area and labor in both regions, consistent with expectations. The results show that labor contributes more to productivity compared with crop area. The labor/crop area (*L*/*A*) variable has significant effect in the transcendental function for the low-elevation region but not for high elevations. Manure has a positive and significant effect on productivity. The effects of grass mulch and crop residues are only significant in the high-elevation region, where the effects are positive and significant. Farm size has a positive influence on productivity but the effect is only significant for the high-elevation region. For a given size of banana plot, farmers with larger farms have higher banana yields. Large farmers are more likely to be committed to farming than are small farmers, who are more likely to diversify into off-farm wage employment. Extension visits have a positive effect and are significant (1 percent) in the high-elevation but not in the low-elevation region. Interaction with

Table 8.1 Variable definitions and summary statistics for cooking bananas, productivity and technical efficiency analysis

		Overall sample		Low elevation		High elevation		Case study	
		n **= 512**		***n*** **= 374**		***n*** **= 138**		***n*** **= 157**	
Variable	**Definition**	**Mean**	**Standard deviation**	**Mean**	**Standard deviation**	**Mean**	**Standard deviation**	**Mean**	**Standard deviation**
Y	Cooking bananas output (kg/year)	3,109.7	4,919	1,581	1,958	7,252	7,494	4,125.8	5,870.9
Variables in the production function									
A	Area (acres) under cooking bananas	0.58	0.76	0.504	0.771	0.801	0.688	0.622	0.569
L	Labor input (hours/year)	636.3	649.7	470	509	1,088	769	521.9	5,208
M	Manure input (kg/year)	495	2707	292	1245	1,045	4,765	528.2	2,610.3
G	Grass mulch input (kg/year)	194	1,461	95	672	461	2,577		
C	Crop residue input (kg/year)	331	1,660	280	1,114	469	2,620	205.2	1,008.5
N	Soil nitrogen (percent)							0.127	0.07
K	Available soil potassium (meq/100 g soil)							1.165	1.124
SOM	Soil organic matter (percent)							4.839	1.806
pH	Soil pH							5.975	0.619
sand	Ratio of sand to (sand + clay + silt) (percent)							59.35	9.99
farm size	household farm size (acres)	4.023	8.567	4.45	9.39	2.866	5.635		
Ext	Extension visits in 6 months	0.702	1.912	0.69	2.05	0.73	1.487		
plotage	Age of banana plot (years)	20	23	11.9	13.73	41.8	28.3		
plotage^2	Plotage squared	926	1,996	329.5	900.7	2,544	3,006		
Sigatoka	Sigatoka score[a]	0.163	0.272	0.22	0.3	0.02	0.07		
weevils	Weevil score[a]	0.394	0.333	0.42	0.34	0.33	0.32		
Technical efficiency variables									
Age	Age of household head (years)	45.2	16	45.72	16.46	43.7	14.6		
Age^2	Age squared	2,295	1,610	2,360	1,673	2,118.3	1,417.4		
edhh	Education household head (years)	5.39	4.09	5.55	4.12	4.93	3.99	5.981	4.1
D	Distance to tarmac road (km)	13.46	18.7	14.6	20.3	10.3	12.97		
hhsz	Household size	5.89	2.65	5.84	2.70	6.02	2.52	5.329	2.323
depr	Persons >64 or <14 years old/family size	0.497	0.239	0.498	0.252	0.494	0.201		
hplot	Plot managed by husband	0.764	0.425	0.73	0.445	0.854	0.354	0.839	0.369
kk	Amount credit obtained (thousand Ush)	14	92.3	17.83	107.3	3.66	17.15		
sk	Remittances + rent (thousand Ush)	90	368	80.65	306	115.3	500.1		
wp	Real wage rate (wage/price bananas)	2.67	1.08	2.72	1.21	2.53	0.58		

[a]Farmers were asked to score the presence of the disease or pest on a particular plot and the number of years the disease or pest had been observed on the plot. Presence was scored as 1 and absence as 0. The final score of the disease or pest was computed taking into consideration the number of years it had been observed on the plot and the size of the plot. For example, if the household has three plots with disease scores 0 for all the years, 1 for 3 years out of 5 years, and 1 for 7 out of 10 years and the corresponding areas of each plot are 0.5, 0.9, and 1.5 acres, the final score is $(0 \times 0.5 + 0.6 \times 0.9 + 0.7 \times 1.5)/(0.5 + 0.9 + 1.5) = 0.548$.

Table 8.2 Results of the frontier function

	Overall sample			Low elevation			High elevation		
Variable	**Eq1**	**Eq2**	**Eq3**	**Eq1**	**Eq2**	**Eq3**	**Eq1**	**Eq2**	**Eq3**
N	512	512	512	374	374	374	138	138	138
Constant	5.79***	5.158***	4.262***	5.694***	5.064***	4.279***	6.637***	6.259***	3.526
	(18.99)	(13.73)	(5.27)	(15.75)	(11.32)	(4.45)	(14.52)	(9.68)	(0.78)
$\ln(A)$	0.332***	0.207***	–0.211	0.332***	0.206***	–0.233	0.264***	0.199**	0.065
	(8.04)	(3.44)	(–0.77)	(6.45)	(2.81)	(–0.68)	(5.79)	(2.19)	(0.10)
$\ln(L)$	0.384***	0.489***	0.864***	0.395***	0.5***	0.812***	0.282***	0.343***	1.189
	(8.87)	(8.64)	(3.49)	(7.87)	(7.55)	(2.69)	(4.24)	(3.43)	(0.93)
$\ln(A)^2$			–0.065**			–0.064*			–0.096**
			(–2.47)			(–1.93)			(–2.06)
$\ln(L)^2$			–0.038*			–0.031			–0.065
			(–1.85)			(–1.19)			(–0.72)
$\ln(L) \times \ln(A)$			0.065***			0.069			0.013
			(1.71)			(1.45)			(0.13)
L/A		–0.0001***			-7×10^{-5}**			-3×10^{-5}	
		(–2.82)			(–2.39)			(–0.82)	
M	3×10^{-5}*	3×10^{-5}*	3×10^{-5}**	8×10^{-5}**	8×10^{-5}**	8×10^{-5}**	2×10^{-5}	2×10^{-5}***	2×10^{-5}***
	(1.84)	(1.86)	(1.99)	(2.00)	(2.08)	(2.06)	(2.77)	(2.82)	(3.18)
G	1×10^{-5}	1×10^{-5}	2×10^{-5}	-1×10^{-5}	-1×10^{-5}	-1×10^{-5}	2×10^{-6}*	2×10^{-5}*	3×10^{-5}**
	(0.47)	(0.48)	(0.75)	(–0.16)	(–0.2)	(–0.17)	(1.66)	(1.74)	(2.26)
C	2×10^{-5}	2×10^{-5}	3×10^{-5}	2×10^{-5}	1×10^{-5}	2×10^{-5}	3×10^{-5}**	3×10^{-5}***	3×10^{-5}***
	(1.15)	(1.14)	(1.38)	(0.40)	(0.3)	(0.42)	(2.58)	(2.6)	(3.00)
farm size	0.008*	0.008*	0.01*	0.002	0.002	0.003	0.014**	0.014**	0.016***
	(1.68)	(1.70)	(1.95)	(0.36)	(0.36)	(0.55)	(2.21)	(2.22)	(2.61)
Ext	0.026	0.024	0.025	–0.009	–0.01	–0.01	0.134***	0.135***	0.134***
	(1.2)	(1.14)	(1.18)	(–0.35)	(–0.36)	(–0.38)	(4.83)	(4.82)	(4.68)
plotage	0.017***	0.017***	0.018***	0.028***	0.027***	0.028***	0.0153***	0.0148***	0.016***
	(3.2)	(3.21)	(3.38)	(3.00)	(2.82)	(2.87)	(3.68)	(3.55)	(3.79)
plotage^2	–0.0001	-9×10^{-5}	-9×10^{-5}*	–0.0003**	–0.0002*	–0.002*	-7×10^{-5}**	-7×10^{-5}*	–0.0001*
	(–1.56)	(1.54)	(–1.66)	(–2.03)	(–1.71)	(1.69)	(–1.77)	(–1.72)	(–1.86)
Sigatoka	–0.21	–0.204	–0.214	–0.183	–0.167	–0.17	–0.737	–0.742	–0.744
	(–1.59)	(–1.56)	(–1.64)	(–1.18)	(–1.08)	(–1.09)	(–1.59)	(1.6)	(–1.61)
weevils	–0.035	–0.063	–0.088	0.049	0.012	–0.0009	–0.161	–0.153	–0.172*
	(–0.34)	(–0.60)	(–0.83)	(0.36)	(0.09)	(–0.01)	(–1.53)	(1.46)	(–1.68)
High elevation	0.54***	0.52***	0.523***						
	(5.12)	(4.98)	(5.23)						
Log likelihood	–624.9	–621.2	–620.6	–500.1	–497.4	–497.7	–51.4	–51.1	–48.2
TE	0.475	0.478	0.474	0.437	0.442	0.44	0.705	0.703	0.706

Notes: ***, **, * indicate statistical significance at the 1, 5, and 10 percent levels, respectively. Eq1 is the Cobb-Douglas technology, Eq2 is the transcendental production function, and Eq3 the transcendental logarithmic function (translog). See Table 8.1 for definitions of the variables.

extension agents could enable farmers to adopt new farming techniques and to raise their production frontier. However, it is possible that the extension agents visit the most productive farmers, without necessarily increasing farmers' use of new technologies and practices.

The effect of age of a banana plot is significant (positive for young plots and negative for old ones) for all the cases. Age of the banana plot was included in the estimation to account for the low yields observed in young plantations and old ones. Results for black Sigatoka show that it has a negative effect on banana production, but the effect is not significant. The effect of weevils is also not significant. The insignificant results obtained for black Sigatoka and weevils for the overall sample might be due to correlation between the disease/pest and the location dummy (elevation). Excluding the location dummy variable from the estimation makes the coefficients of the black Sigatoka score and weevil score significant at 5 percent and 10 percent, respectively (Table 8.3).

The regional dummy variable (location in the high-elevation region) has a positive and significant effect (1 percent) on output. This result shows that elevation-related effects are important in determining the productivity of bananas. The dummy was included in the equations for the overall sample to capture regional differences and to capture the differences in biotic and abiotic factors characterizing the two different regions.

The TE scores reveal the presence of inefficiency especially for low-elevation region. The TE score obtained for the high-elevation was higher than that for the low-elevation region, implying that inefficiency contributes to the low output realized in the region. The TE scores obtained by using different function forms were very close, implying that model specifications for the frontier function have no impact on the predicted technical efficiencies for the sample farmers. This finding is consistent with what is reported in the literature (Kopp and Smith 1980).

Table 8.4 shows elasticities of production, with respect to labor and land, and the returns to scale for cooking bananas production in the two regions. The elasticity of labor is higher for the low-elevation region, while the opposite is true for crop area. The sum of the elasticities of labor and land are all below 1 in all the cases, which indicates decreasing returns to scale. The implication of this result is that farmers would loose efficiency if they increase scale of production. The decrease in efficiency as a result of the increase in scale of production is most likely due to differences in soil quality between small and large plots. This result is consistent with that obtained for farm size. Given farm size, increasing plot size leads to a decrease in banana productivity.

The three functional forms (Cobb-Douglas, transcendental, and translog) yield different results in terms of elasticities. The elasticities of labor obtained from the transcendental function are much higher than those obtained from the Cobb-Douglas and translog forms. However, the returns to scale obtained from all functions are quite close. The Cobb-Douglas seems to be a consistent and appropriate function for assessing production technology in the two different regions.

The data support rejection of the null hypothesis that farmers growing cooking bananas are technically efficient in all cases (Table 8.5). The results for factors influencing TE are shown in Table 8.6.

The effect of age on TE is not significant, perhaps as a consequence of multicollinearity. However, the relationship between age and TE is not as expected. TE first decreases as age increases in the early years but later starts to increase, as shown by the negative effect of the quadratic term. The reason for this behavior could be associated with the reproduction process of the household. Young families may allocate more of their time to raising children, so that more time is available for farm production as the

Table 8.3 Cobb-Douglas production estimates for the overall sample (location dummies excluded)

Variable	Coefficient	*t*-value
Production function estimates		
Constant	5.773***	18.55
ln(*A*)	0.34***	8.17
ln(*L*)	0.422***	9.66
M	0.00003*	1.65
G	0.00002	0.88
C	0.00002	1.28
farm size	0.006	1.32
Ext	0.017	0.74
plotage	0.024***	4.62
plotage^2	–0.0001**	–2.10
Sigatoka	–0.304**	–2.36
weevils	–0.167*	–1.65
Log likelihood	–639.1	
Wald X^2	726.7	
TE	0.449	
Technical inefficiency estimates		
Constant	0.511	0.71
Age	0.024	0.82
Age^2	–0.0002	–0.76
Hplot	–0.533***	–3.13
Edhh	0.018	1.02
Hhsz	–0.053*	–1.83
depr	0.368	1.26
Kk	–0.001	–1.00
D	–0.005	–0.84
σ_V (standard error)	0.38	0.048

Notes: ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively. See Table 8.1 for definitions of the variables.

children grow up. The children also contribute labor, enabling the farmers to implement timely management decisions.

The education variable gives mixed results, as expected. In the low-elevation region, the impact of education on TE is negative. This finding is consistent with our hypothesis that educated households are less efficient if education increases the opportunity cost of labor, so that farmers reallocate resources from farm to nonfarm activities. The impact of education on TE in the high-elevation region is positive. In this region, education appears to increase farmers' management capabilities and their ability to utilize technologies.

The results for the relationship between the distance variable and TE show a negative relationship but is only significant for high elevations. Distance to paved roads is

Table 8.4 Production elasticities

Region	Elasticities of production: Labor	Elasticities of production: Land	Returns to scale
Overall sample			
Cobb-Douglas	0.384	0.332	0.716
Transcendental	0.489	0.207	0.696
Translog	0.343	0.314	0.658
Low elevation			
Cobb-Douglas	0.395	0.332	0.727
Transcendental	0.5	0.206	0.706
Translog	0.377	0.316	0.693
High elevation			
Cobb-Douglas	0.282	0.264	0.546
Transcendental	0.343	0.199	0.542
Translog	0.296	0.261	0.557

Table 8.5 Test for the null hypothesis that $\sigma_u = 0$

Region	χ^2	P	Outcome
Overall sample	45.08	0.000	Reject null
Low elevation	29.52	0.000	Reject null
High elevation	8.9	0.004	Reject null

positively correlated with TE. The distant farms are more technically efficient, most likely because of access to cheap labor, which enables them to implement timely management decisions.

The family size variable is positively related to TE but is only significant, at 10 percent, for the whole sample and the low-elevation region. Households with big families are more technically efficient, most likely because they strive to achieve higher output to meet their subsistence requirements. Moreover, large families have more labor endowment (including children) needed to implement management decisions. The impact of the dependence ratio on TE is negative but only significant for the high-elevation region.

The gender of the plot manager (husband = 1 and wife = 0) has a positive and significant impact on TE for the whole sample (1 percent) and for the low-elevation region (5 percent). In the high-elevation region, the relationship is negative but not significant. Higher efficiency in plots managed by husbands can be explained by differential access to production resources and hence timing of input application and management practices.

Results suggest that access to credit improves efficiency in production of cooking bananas in all cases, but the effect is not significant. This result implies that liquidity constraints affect farmers' ability to apply inputs and implement farm management decisions on time.

Table 8.6 Factors influencing technical efficiency

	Overall sample			Low elevation			High elevation		
Variable	**Eq1**	**Eq2**	**Eq3**	**Eq1**	**Eq2**	**Eq3**	**Eq1**	**Eq2**	**Eq3**
Constant	0.415	0.28	0.263	0.438	0.263	0.267	–2.375*	–2.312	–2.358
	(0.54)	(0.36)	(0.35)	(0.50)	(0.29)	(0.30)	(1.3)	(–1.28)	(–1.33)
Age	0.02	0.027	0.025	0.033	0.04	0.039	0.039	0.04	0.042
	(0.67)	(0.86)	(0.83)	(0.95)	(1.13)	(1.09)	(0.51)	(0.53)	(0.59)
Age^2	–0.002	–0.0003	–0.0003	–0.0004	–0.004	–0.0004	–0.0003	–0.0003	–0.0004
	(–0.70)	(–0.89)	(–0.83)	(–1.05)	(–1.22)	(–1.17)	(0.43)	(–0.46)	(–0.5)
edhh	0.016	0.015	0.019	0.013	0.013	0.015	–0.037	–0.043	–0.044
	(0.9)	(0.85)	(1.03)	(0.62)	(0.62)	(0.70)	(–0.86)	(–0.99)	(–1.00)
D	–0.001	–0.002	0.0008	–0.006	–0.006	–0.005	–0.064***	–0.065***	–0.061***
	(–0.22)	(–0.28)	(0.13)	(–0.95)	(–0.97)	(–0.70)	(–2.88)	(–2.98)	(–2.79)
hhsz	–0.054*	–0.056*	–0.055*	–0.059	–0.062*	–0.061*	–0.049	–0.047	–0.032
	(–1.78)	(–1.86)	(–1.86)	(–1.65)	(–1.73)	(–1.71)	(–0.57)	(0.55)	(–0.39)
depr	0.443	0.436	0.436	0.506	0.506	0.504	1.505*	1.513*	1.378*
	(1.42)	(1.40)	(1.42)	(1.39)	(1.38)	(1.39)	(1.73)	(1.78)	(1.66)
hplot	–0.567***	–0.579***	–0.577***	–0.522**	–0.535**	–0.53**	0.293	0.288	0.184
	(–3.21)	(–3.27)	(–3.29)	(–2.54)	(–2.59)	(–2.58)	(0.63)	(0.63)	(0.41)
kk	–0.001	–0.001	–0.001	–0.001	–0.001	–0.001	–0.015	–0.018	–0.022
	(–1.04)	(–1.04)	(–1.03)	(–1.28)	(–1.28)	(–1.27)	(–1.04)	(1.06)	(–0.74)
sk	0.0002	0.0002	0.0002	0.0004	0.0004	0.0004			
	(1.15)	(1.12)	(1.15)	(1.40)	(1.36)	(1.4)			
σ_V (standard error)	0.429	0.43	0.414	0.483	0.491	0.483	0.205	0.201	0.192
	(0.049)	(0.05)	(0.05)	(0.069)	(0.071)	(0.074)	(0.032)	(0.032)	(0.035)

Notes: ***, **, * indicate statistical significance at the 1, 5, and 10 percent levels, respectively. Eq1 is the Cobb-Douglas technology, Eq2 is the transcendental production function, and Eq3 is the transcendental logarithmic production function (translog). See Table 8.1 for definitions of variables. A positive sign on a coefficient implies a negative effect on efficiency and vice versa.

The results on the interaction between SOM and K, and physical (sand) and chemical (pH) characteristics and the effect on productivity are presented in Table 8A.3. The estimates from a three-stage least squares method show that the proportion of sand in the soil negatively affects SOM content. The results also show that the SOM content is higher in Masaka, implying that differences in regional characteristics affect SOM accumulation and decomposition. It should be noted that SOM is highly correlated with N content in the soil. Availability of K is positively influenced by the SOM content in the soil, pH, and additions of crop residues. In turn, K availability positively affects the yield of cooking bananas, as expected, but the effect is not significant. However, the effect of SOM on cooking-banana yield is negative, but only significant at 10 percent. This result can be explained by the conditions that favor accumulation of SOM, but are not favorable for the production of cooking bananas. SOM tends to accumulate faster in clay soils, which are not good for production of cooking bananas because of physical impediment of banana root growth. Another reason could be related to the C:N ratio of materials used in the formation of the SOM. SOM with high C:N ratios can affect availability of nutrients through immobilization of the nutrients during the SOM decomposition. Animal manure has a positive and significant (10 percent) effect on yields of cooking bananas. The effect of plot age is significant at 1 percent (positive for young plots and negative for older plots). The effect of black Sigatoka is negative and significant at 5 percent.

Finally, we estimate the reduced form using the frontier function approach (Table 8.7). The elasticities of labor and crop area are positive, as expected. The sum of the elasticities, from the Cobb-Douglas function, indicates constant returns to scale. This result contrasts with the result obtained for the main sample, which displays decreasing returns to scale. Most likely the case study sites are not representative of the main sample; hence the difference in the results obtained for returns to scale. However, the case study results sheds some light on the contribution of biophysical characteristics to the shift of banana production from the low- to the high-elevation region.

Animal manure has a positive effect on productivity, with a statistical significance of 1 percent. The effect of sand on the productivity of cooking bananas is positive and significant at 1 percent. The effect of pH is positive and significant at 5 percent for all the model specifications. The average TE obtained (44.9 percent to 45.6 percent, depending on function form) from the case study is close to those obtained for the main sample.

Conclusions

One of the objectives of this research report is to assess the impact of improved banana technology on smallholder farmers in the lake region of Uganda and Tanzania. In this chapter, we use the stochastic production functions to analyze productivity and efficiency of smallholder farmers in Uganda. In the chapter, we assess the impact of plot, farm, and regional characteristics on the productivity and efficiency of banana farmers. We also analyze the impact of soil organic amendments on banana productivity. Findings show that to improve the productivity of small farmers, much more will need to be done in terms of access to basic inputs, farm credit, information, and education.

The productivity of cooking bananas depends on the climate and soil characteristics of regions, and, as hypothesized, it is higher in the high-elevation region. Labor and crop area respond positively to output, but the scale elasticity is below 1 in all cases, implying that farmers cannot increase scale of production without losing efficiency. Labor productivity is higher in the low-elevation region, where most agronomic practices (such as crop sanitation) are carried out minimally. Animal manure, grass mulch, and crop residues have a positive and significant effect on productivity, especially in the high-elevation region. Soil pH has a positive and significant effect on productivity, whereas the effect of soil organic matter and soil texture (sand) is not significant. For a given banana plot size, banana yields increase with farm size. Alternatively, keeping farm size constant, an increase in plot size decreases the productivity of bananas. Extension visits enhance productivity in the high- but not in the low-elevation region.

Findings illustrate substantial inefficiencies in the production of cooking bananas, especially in the low-elevation region. Education improves technical efficiency in the highlands but not in the low-elevation region. Market access (using distance to paved road as a proxy) reduces efficiency, especially for farmers in the high-elevation region. Household size is positively related to efficiency. Banana production appears to be more efficient when managed by men than by women, probably because of underlying differentials in access to resources. Access to credit increases efficiency, whereas rent and remittances reduce efficiency.

Policies to improve production efficiency include investments in education and extension services, and improving access to production credit. Clearly, policies will need to be tailored to the local conditions, as demonstrated by the different production and efficiency profiles of the two regions analyzed here.

Table 8.7 Incorporating soil factors in the production frontier function (*n* = 154)

	Cobb-Douglas		Transcendental		Translog	
Variable	Coefficient	*z*-value	Coefficient	*z*-value	Coefficient	*z*-value
Stochastic frontier function						
Constant	2.612***	2.75	2.533**	2.35	4.509	1.56
ln(*A*)	0.446***	4.52	0.426**	2.56	0.624	0.84
ln(*L*)	0.549***	6.79	0.568***	3.88	–0.18	–0.17
$\ln(A)^2$					–0.048	–0.59
$\ln(L)^2$					0.062	0.71
$\ln(L) \times \ln(A)$					–0.049	–0.4
L/A			–0.00002	–0.15		
M	0.0001***	3.8	0.0001***	3.69	0.0001***	3.27
C	0.00004	0.44	0.00004	0.45	0.00001	0.13
sand	0.02***	3.42	0.245***	3.39	0.264***	3.38
pH	0.245**	2.43	0.02**	2.44	0.021**	2.46
plotage	0.033***	3.89	0.033***	3.78	0.034***	3.86
$plotage^2$	–0.0003***	–3.27	–0.0003***	–3.21	–0.0003***	–3.28
Sigatoka	–1.5***	–3.83	–1.511***	–3.8	–1.517***	–3.79
Log likelihood	–193.8		–193.7		–193.2	
TE (standard deviation)	0.451		0.449		0.456	
Factors influencing technical inefficiency						
Constant	0.453	0.34	0.456	0.34	0.304	0.22
Age	–0.001	–0.02	–0.001	–0.02	0.007	0.12
Age^2	0.0001	0.23	0.0001	0.22	0.00005	0.09
hplot	–0.355	–0.94	–0.35	–0.92	–0.399	–1.03
edhh	–0.017	–0.45	–0.017	–0.45	–0.018	–0.48
hhsz	0.044	0.89	0.045	0.9	0.035	0.68
depr	–0.096	–0.17	–0.094	–0.17	–0.051	–0.09
σ_V (standard error)	0.383	0.1	0.378	0.105	0.395	0.22

Notes: ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively. TE is technical efficiency. A blank entry indicates that the variable was not included in the regression. See Table 8.1 for definitions of the variables.

Supplementary Tables

Table 8A.1 Labor demand estimates (first stage of the production function estimation)

	Overall sample		Low elevation		High elevation	
Variable	**Coefficient**	***t*-value**	**Coefficient**	***t*-value**	**Coefficient**	***t*-value**
Constant	6.425***	16.19	6.247***	12.59	6.808***	15.62
ln(*A*)	0.513***	13.62	0.498***	10.73	0.36***	6.48
ln(*w/p*)	–0.397***	–3.93	–0.46***	–3.82	–0.296*	–1.7
D	–0.009***	–3.94	–0.01***	–3.90	–0.06	–1.24
hhsz	0.027	1.57	0.03	1.37	0.052**	2.52
depr	–0.453***	–2.61	–0.58***	–2.64	0.018	0.09
Age	–0.005	–0.34	–0.008	–0.40	–0.005	–0.27
Age^2	0.0001	0.72	0.0002	0.92	0.0001	0.43
hplot	0.264***	2.8	0.314***	2.68	0.164	1.32
edhh	0.025**	2.51	0.034***	2.79	0.009	0.83
plotage	0.025***	4.52	0.018*	1.78	0.009	1.46
$plotage^2$	–0.0002***	–3.2	–0.0002	–1.38	–0.0001	–1.53
Sigatoka	–4.22***	–2.68	–0.331*	–1.88	–0.578	–0.96
weevils	–0.005	–0.04	0.266*	1.74	–0.261*	–1.85
Adjusted R^2	0.482		0.404		0.408	

Notes: ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels. See Table 8.1 for definitions of the variables.

Table 8A.2 Two-stage least squares estimates of the production function for cooking bananas (endogeneity test)

	Overall		Low elevation		High elevation	
Variable	**Coefficient**	***t*-value**	**Coefficient**	***t*-value**	**Coefficient**	***t*-value**
Constant	6.44***	7.48	6.229***	7.39	5.426***	3.37
ln(*A*)	0.532***	6.53	0.516***	6.10	0.256**	2.48
ln(*L*)	0.104	0.76	0.112	0.81	0.427*	1.8
M	0.00003*	1.77	0.0001**	2.28	0.00002**	2.32
G	0.00001	0.41	0.00006	0.76	0.000001	0.48
C	0.00003	1.17	0.00003	0.59	0.00003*	1.83
farm size	0.01**	2.03	0.003	0.52	0.018***	2.6
Ext	0.016	0.76	–0.0001	–0.01	0.124***	4.78
plotage	0.032***	4.36	0.039***	3.59	0.005	0.92
plotage^2	–0.0002***	–2.78	–0.0003**	–2.00	0.00002	0.31
Sigatoka	–0.096	–0.58	0.015	0.08	–0.542	–0.98
weevils	–0.108	–0.85	0.01	0.06	–0.114	–0.86
High elevation	0.715***	5.99				
Residual[a]	0.384**	2.60	0.383**	2.57	–0.145	–0.58
Adjusted R^2	0.662		0.491		0.687	

Notes: ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively. See Table 8.1 for definitions of the variables. A blank entry indicates that the variable was not included in the regression. See Table 8.1 for definitions of the variables.

[a]Residual is from the first-stage estimation (the labor equation); variables included in the labor equation are area under cooking bananas, wage/price ratio, distance to paved roads, household characteristics (size, composition, age education, and gender), and village characteristics representing opportunities (farm wage income, nonfarm wage income, salary income, and self-employment earnings).

Table 8A.3 Production function estimates, three-stage least squares estimate

	ln(*SOM*)		ln(*K*)		ln(*Y*)	
Variable	**Coefficient**	***t*-value**	**Coefficient**	***t*-value**	**Coefficient**	***t*-value**
Constant	2.37***	18.09	–4.498***	–13.73	5.601***	3.58
ln(*A*)					0.32***	2.77
ln(*L*)					0.679***	6.83
M					9.0×10^{-5}***	2.76
C			7.0×10^{-5}**	2.29		
ln(*SOM*)			1.135***	4.06	–1.479	–1.54
ln(*K*)					0.415	1.05
sand	–0.014***	–6.23				
pH			0.442***	4.99		
plotage					0.035***	3.48
plotage^2					–0.0003***	–2.82
Sigatoka					–0.971**	–2.18
High elevation	–0.058	–1.23				
Adjusted R^2	0.23		0.59		0.59	

Notes: *** and ** indicate statistical significance at the 1, 5, and 10 percent levels, respectively. The *sand* variable instruments *SOM,* and *pH* instruments *K*. See Table 8.1 for definitions of the variables.

References

Aigner, D. J., C. A. K. Lovell, and P. Schmidt. 1977. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics* 6 (1): 21–37.

Ali, M., and D. Byerlee. 1991. Economic efficiency of small farmers in a changing world: A survey of recent evidence. *Journal of International Development* 3 (1): 1–27.

Battese, G. E., and S. S. Broca. 1997. Functional forms of stochastic frontier production functions and models for technical inefficiency effects: A comparative study for wheat farmers in Pakistan. *Journal of Productivity Analysis* 8 (4): 395–414.

Battese, G. E., and T. J. Coelli. 1995. A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics* 20 (2): 325–332.

Bauer, P. W. 1990. Recent development in econometric estimation of frontiers. *Journal of Econometrics* 46 (1–2): 39–56.

Belbase, K., and R. Grabowski. Technical efficiency in Nepalese agriculture. *Journal of Development Areas* 19 (4): 515–525.

Binam, J. N., K. Sylla, I. Diarra, and G. Nyambi. 2003. Factors affecting technical efficiency among coffee farmers in Cote d'Ivoire: Evidence from the centre west region. *R&D Management* 15 (1): 66–76.

Bravo-Ureta, B. E., and R. E. Evenson. 1994. Efficiency in agricultural production: The case of peasant farmers in Eastern Paraguay. *Agricultural Economics* 10 (1): 27–37.

Bravo-Ureta, B. E., and E. Pinheiro. 1997. Technical, economic and allocative efficiency in peasant farming: Evidence from the Dominican Republic. *Developing Economies* 35 (1): 48–67.

Coelli, T. J. 1994. A guide to Frontier 4.1: A computer program for stochastic frontier production and cost function estimation. Department of Econometrics, University of England, Armidale, Australia. Mimeo.

———. 1995. Recent developments in the frontier modeling and efficiency measurement. *Australian Journal of Agricultural Economics* 39 (3): 219–245.

Coelli, T. J., D. S. P. Rao, and G. E. Battese. 1998. *An introduction to efficiency and productivity analysis.* Amsterdam: Kluwer Academic Publishers.

Farrell, M. J. 1957. The measurement of productive efficiency. *Journal of the Royal Statistical Society,* Series A 120 (3): 253–290.

Forsund, F. R., C. A. K. Lovell, and P. Schmidt. 1980. A survey of frontier production functions and their relationships to efficiency measurement. *Journal of Econometrics* 13 (1): 5–25.

Fried, H. O., C. A. K. Lovell, and S. S. Schmidt. 1993. *The measurement of productive efficiency: Techniques and applications.* Oxford: Oxford University Press.

Hussain, S. S. 1989. Analysis of economic efficiency in Northern Pakistan: Estimation, causes and policy implications. Ph.D. dissertation, University of Illinois, Urbana–Champaign.

Kalirajan, K. P. 1981. An econometric analysis of yield variability in paddy production. *Canadian Journal of Agricultural Economics* 29 (2): 283–294.

———. 1984. Farm-specific technical efficiencies and development policies. *Journal of Economic Studies* 11 (3): 3–13.

———. 1991. The importance of efficiency use in adoption of technology: A micro panel data analysis. *Journal of Productivity Analysis* 2 (2): 113–126.

Kalirajan, K. P., and R. T. Shand. 1986. Estimating location specific and farm specific technical efficiency: An analysis of Malaysian agriculture. *Journal of Economic Development* 11 (2): 147–160.

———. 1997. Modeling and measuring technical efficiency: An alternative approach. In *Issues in agricultural competitiveness: Markets and policies,* ed. R. Rose, C. Tanner, and M. A. Bellamy. Proceedings of the 22nd International Conference of Agricultural Economists, August 22–29, 1994, Harare, Zimbabwe. Hants, U.K.: Dartmouth Publishing.

Kopp, R. J., and V. K. Smith. 1980. Frontier production function estimates for steam electric generation: A comparative analysis. *Southern Economics Journal* 47 (1): 1049–1059.

Kumbhakar, S. C., S. Gosh, and J. T. McGuckin. 1991. A generalized production frontier approach for estimating determinants of inefficiency in U.S. dairy farms. *Journal of Business and Economic Statistics* 9: 279–286.

Maddala, G. S. 1971. The use of variance models in pooling cross section and time series data. *Econometrica* 39 (2): 341–358.

Meeusen, W., and van den Broeck, J. 1977. Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review* 18 (4): 435–444.

Olsen, S. R., C. V. Cole, F. S. Watanabe, and L. A. Dean. 1954. Estimation of available phosphorus in soils by extraction with sodium bicarbonate. U.S. Department of Agriculture Circular 939. Washington, D.C.: U.S. Government Printing Office.

Phillips, M. J., and R. P. Marble. 1986. Farmer education and efficiency: A frontier production function approach. *Economy of Education Review* 5: 257–264.

Pitt, M. M., and L. F. Lee. 1981. The measurement and sources of technical inefficiency in the Indonesian weaving industry. *Journal of Development Economics* 9 (1): 43–64.

Rufino, M. 2003. On-farm analysis of nematode infestation and soil fertility as constraints to the productivity of banana-based production systems in Uganda. M.S. thesis, Plant Sciences Department, Production Systems Group, Wageningen University, The Netherlands.

Schmidt, P. 1986. Frontier production functions. *Econometric Review* 4 (2): 329–334.

Thomas, G. W. 1982. Exchangeable cations. In *Methods of soil analysis*. Part 2: *Chemical and biological properties*, ed. A. L. Page, R. H. Miller, and D. R. Keeney. Madison, Wis., U.S.A.: American Society of Agronomy and Soil Science Society of America.

Tian, W., and G. H. Wan. 2000. Technical efficiency and its determinants in China's grain production. *Journal of Productivity Analysis* 13 (2): 159–174.

Tisdale, S. L., W. L. Nelson, J. D. Beaton, and J. L. Havlin. 1993. *Soil fertility and fertilizers*. Fifth edition: Upper Saddle River, N.J.: Prentice-Hall.

Weir, S. 1999. *The effects of education and farmer productivity in rural Ethiopia*. Working Paper CSAE WPS99-7. Oxford: Centre for the Study of African Economics, University of Oxford.

Weir, S., and J. Knight. 2000. *Education externalities in rural Ethiopia: Evidence from average and stochastic production functions*. Working Paper CSAE WWPS2000-4. Oxford: Centre for the Study of African Economics, University of Oxford.