Productive efficiency of specialty and conventional coffee farmers in Costa Rica: accounting for the use of different technologies and self-selection.

Meike Wollni
Department of Applied Economics and Management, Cornell University
mw323@cornell.edu

Selected paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Portland, OR, July 29-August 1, 2007

As a result of considerable oversupply of green coffee in international markets, world coffee prices dropped to their lowest levels in 30 years giving rise to the most severe crisis experienced by the coffee sector (Ponte 2002). In many countries, coffee prices fell below average production costs causing widespread financial and social hardships among producers (Varangis, Siegel et al. 2003, Flores, Bratescu et al. 2002). Economic losses and the lack of viable income alternatives forced many farmers to abandon their coffee plantations and migrate to urban areas in search of employment. Overall, the effects of the crisis pose serious threats to the prospects for sustainable rural development (Chaveriat 2001, Damiani 2005, International Coffee Organization 2004).

In the face of this situation, policymakers and development agencies have shown their willingness to assist farmers in improving their production performance and thus their ability to cope with the crisis. To avoid wasting scarce resources, policy actions must be tailored to the needs of farmers. On this account, the paper seeks to identify the factors that determine farmers' technical efficiency in coffee production. As inefficiency in production results in a failure to maximize profits at the farm level, increases in productive efficiency enhance farmers' competitiveness and could help them to confront the adverse economic conditions caused by the coffee crisis. An empirical evaluation of the factors determining efficiency is critical to identify the constraints faced by farmers and to derive adequate policy measures.

Coffee has traditionally been marketed through a commodity system, in which the lowest cost production of a standardized product is typically rewarded (Lewin, Giovannucci et al. 2004). Farmers have no incentive to increase the quality of their produce as long as they do not receive price premiums for the added value of the product. During past years, an increasing number of specialty marketing channels has emerged satisfying increasingly diversified consumer demand patterns (Ponte 2002). Coffee marketed through specialty channels is subjected to various grades and standards aiming to ensure different aspects of sustainability and/or product quality. Farmers have to comply with these standards, if they wish to access specialty segments often requiring the adoption of sustainable and/or quality-enhancing production technologies (Muradian and Pelupessy 2005). Compared to other coffee producing countries, Costa Rica has favorable natural conditions for the production of high-quality coffee as well as a strong organizational structure throughout the production and marketing stages of the coffee sector. In the face of the crisis, the country has put emphasis on exploring this competitive advantage, motivating farmers to adjust their production to the requirements of specialty markets. Taking this important development into account, the household sample selected for the empirical analysis includes both farmers producing in the specialty segment as well as in the conventional segment. Farmers' efficiency levels and their determinants are then assessed respective to the technology applied on the farm.

The next two sections present the methods employed for the empirical analysis. Section four and five describe the data, empirical model specification and the explanatory variables included in the model. Results are presented and discussed in section six. Section seven summarizes findings and policy implications.

Measuring productive efficiency

Following Meeusen and van den Broeck (1977) and Aigner, Lovell and Schmidt (1977), the present study employs stochastic frontier analysis to estimate a production function[1] and to obtain farm-level technical efficiency estimates. By means of a composed error structure, the stochastic frontier model distinguishes technical inefficiency from the effects of random shocks. The basic stochastic frontier model for panel data can be expressed as

$$Y_{it} = f(X_{it};\beta)\exp(\varepsilon_{it}) \tag{1}$$

where Y_{it} is scalar output of farmer i at time t, X_{it} is a vector of input quantities, β is a vector of unknown parameters that define the production technology, and ε_{it} is a random error term composed of two independent components such that $\varepsilon_{it} \equiv V_{it} - U_{it}$. The V's are assumed to be identical and independently distributed as $N(0, \sigma_V^2)$ and reflect measurement error, omitted variables and statistical noise. The U's are a one-sided random variable independent of the V's and truncated at zero such that $U_{it} \geq 0$. U_{it} is assumed to represent technical inefficiency.

The farmer-specific technical inefficiency is the ratio of the observed output and the farmer-specific stochastic frontier output (Battese 1992). Accordingly, technical efficiency of farmer i at time t can be expressed as (Battese 1992):

$$\begin{aligned} TE_{it} &= Y_{it} / Y_{it}^{*} \\ &= f(X_{it};\beta)\exp(V_{it} - U_{it}) / f(X_{it};\beta)\exp(V_{it}) \\ &= \exp(-U_{it}) \end{aligned} \tag{2}$$

In order to identify the factors that explain differences in efficiency levels among farmers, the U's obtained from the stochastic frontier have to be related to farm-specific variables. Early approaches to the analysis of technical inefficiency effects have applied a two-step procedure. In the first step, a production frontier is estimated to obtain inefficiency estimates, and in the second step, these estimates are regressed on a range of exogenous farm-specific variables (e.g. Larson, Palaskas et al. 1999). However, this widely applied two-step procedure suffers from a serious bias. In the first step the U's are assumed to be identical and independently distributed, in the second step they are expected to depend on a number of farm-specific variables (Kumbhakar and Lovell 2000: 264). Wang and Schmidt (2002) prove that the two-step approach leads to inconsistent results. Kumbhakar, Gosh and McGuckin (1991) and Reifschneider and Stevenson (1991) were the first to incorporate the estimation of technical inefficiency effects into the estimation of a stochastic production frontier. Huang and Liu (1994) derived a non-neutral frontier model in which the technical inefficiency effects are allowed to interact with the inputs. Battese and Coelli (1993, 1995) expanded these models to panel data. For the present analysis the model proposed by Battese and Coelli (1993, 1995) is used, which allows U_{it} to be a function of several exogenous variables. The basic stochastic frontier model is the same as in (1). The U's are defined as a non-negative truncated normal distribution with mean μ_{it} and variance σ_u^2. Basically, the model allows μ_{it} to vary among farms by specifying that

$$\mu_{it} = \delta Z_{it} + W_{it} \qquad (3)$$

where Z_{it} is a vector of farm-specific variables that are expected to influence efficiency, δ is a vector of parameters to be estimated, and W_{it} is an i.i.d. random error term. Maximizing the log likelihood function for the model in (1) and (3) yields estimates of the slope parameters and the variance parameter γ. The variance parameters are defined in terms of $\gamma \equiv \sigma_u^2 / \sigma^2$ and $\sigma^2 \equiv \sigma_u^2 + \sigma_v^2$.

Controlling for self-selection

When estimating a production frontier the underlying assumption is that all farmers in the sample use the same production technology. In the present study, a sub-sample of farmers produces for specialty markets. This requires the adoption of quality-enhancing production techniques to increase the quality of the output and to meet specific product standards required by specialty marketing channels. To account for differences in the underlying technology separate production frontiers are estimated for each sub-sample of farmers. These sub-samples, however, are unlikely to represent unbiased representations of the population. If farmers choose to participate in one group or the other based on their expected performance under the chosen technology, the two sub-samples will systematically differ with respect to certain farm and household characteristics. Consequently, if self-selection is ignored in the estimation of separate production frontiers, coefficient estimates will be biased (Greene 1997: 975, Heckman 1979). Heckman (1979) shows that self-selectivity bias can be controlled for by including an error correction term. Heckman proposes a two-step procedure to obtain the inverse Mill's ratio, which is then inserted as a regressor in the second-stage model. Similarly, Lee (1978) controls for selection bias in the framework of an endogenous switching regression model.

Following Heckman (1979) and Lee (1978), the probability that a household chooses to produce in the specialty segment is estimated by means of a probit model. The inverse Mill's ratio (IMR) is obtained from the linear prediction of the probit model. According to Heckman (1979), it is defined as:

(4) $$IMR_{Sit} = -\frac{\phi(\beta' x_i)}{\Phi(\beta' x_i)}$$ if the quality-enhancing technology is adopted, and

(5) $$IMR_{Cit} = \frac{\phi(\beta' x_i)}{1 - \Phi(\beta' x_i)}$$ otherwise,

where ϕ and Φ denote the normal density and the cumulative normal distributions, respectively. In the second step, the inverse Mill's ratio is included among the exogenous variables of the production frontiers to correct for possible selection bias.

In the context of efficiency studies, selection bias has often been neglected when estimating separate production frontiers for farmers using different technology sets. Exemptions can be found in Sipiläinen and Lansink (2005) and Curtiss, Brümmer et al. (2006). A shortcoming of these papers is that they do not report adjusting standard errors, which is required in the context of two-step models (Greene 1997, Heckman 1979, Lee 1978).

Data and model specification

The empirical analysis is based on a sample of 216 coffee farming households that were randomly chosen from within two of the main coffee regions in Costa Rica, namely the Western Central Valley and the Brunca region in the South. A standardized questionnaire was used to collect data on coffee production as well as on the socio-economic characteristics of household members. The information collected covers the production periods 2003/04 and 2002/03 partly including recall data. The percentage of farmers in the sample participating in the specialty segment increased from 31% in 2002/03 to 49% in 2003/04.

In the first step of the analysis, data from both production periods is pooled to estimate the probability of participation in the specialty segment and to derive the inverse Mill's ratio. A farmer will choose whether to adopt the quality-enhancing technology subject to the specific attributes of the available production technologies and household-specific factors. Following a random utility framework, it is assumed that the unobservable indirect utility (U) that farmer i at time t derives from the technology choice is a sum of observable (V) and unobservable (u) portions:

$U_{it} = V_{it}(\beta' X_{it}) + u_{it}$.

In this framework, $E(U_{it}) = V_{it}(\beta' X_{it})$, which can be estimated as a function of exogenous farmer and technology-specific variables X and a vector of unknown parameters ß. The unobservable portion of the utility is represented by a random error term u, which is assumed to be independently and identically distributed with mean zero. The household will choose to adopt the quality-enhancing technology if the utility gained from participation in the specialty segment (U_{it}^S) is greater than the utility of producing in the conventional segment (U_{it}^C). Formally, the probability that farmer i at time t chooses to adopt the quality-enhancing technology can be expressed as:

$$\mathrm{Prob}(S_{it} = 1 \mid X) = \mathrm{prob}(U_{it}^S > U_{it}^C) = \mathrm{Prob}(\beta^{S\prime} X_{it} + u_{it}^S - \beta^{C\prime} X_{it} - u_{it}^C > 0 \mid X)$$

$$= \mathrm{Prob}((\beta^S - \beta^C)' X_{it} + u_{it}^S - u_{it}^C > 0 \mid X) = \mathrm{Prob}(\beta' X_i + u_i > 0 \mid X), \text{ for } i = 1,\ldots N$$

A standard normal distribution is assumed for the random error term u giving rise to the probit model, which will be used to obtain parameter estimates. To account for serial correlation in the scores across t, a robust variance estimator is used in the pooled probit model[2] (Wooldridge 2002: 482).

In the second step, two different functional forms are considered for the production frontier given in equation (1). The Cobb-Douglas form for the i-th farmer (i=1,…,n; n=216) and the t-th year (t=1, 2) can be specified as:

$$\ln Y_{it} = \alpha_0 + \sum_{j=1}^{4} \alpha_j \ln X_{jit} + \alpha_t t + \sum_{m=1}^{6} \beta_m D_{mit} + V_{it} - U_{it} \quad (8)$$

where Y is the amount of coffee cherries harvested in fanegas, X is a j*n matrix of input quantities, t is a time dummy controlling for unobserved factors that differ between the two years, such as technical change or weather conditions[3], D is a m*n matrix of dummy variables characterizing the production process, the α's and β's are unknown parameters to be estimated, V is a N(0, σ_V^2) distributed random error term, and U is a non-negative random variable representing technical inefficiency.

The Cobb-Douglas functional form imposes constant production elasticities and a constant rate of substitution on the data. This often proves too restrictive in empirical applications (e.g. Villano and Fleming 2004). The translog functional form includes second order terms and interactions between the input variables and is therefore more flexible. The translog model is specified as:

$$\ln Y_{it} = \alpha_0 + \sum_{j=1}^{4} \alpha_j \ln X_{jit} + 0.5 \sum_{j\leq}^{3} \sum_{k}^{3} \alpha_{jk} \ln X_{jit} \ln X_{kit} + \alpha_t t + \sum_{m=1}^{6} \beta_m D_{mit} + V_{it} - U_{it} \tag{9}$$

where the variables are as previously defined.

To account for the use of different production technologies, separate models will be estimated for specialty coffee farmers (denoted by subscript S) and conventional coffee farmers (denoted by subscript C). Potential selectivity bias will be controlled for by the inclusion of the inverse Mill's ratio that is obtained from the first-stage pooled probit model. Here, the example of the Cobb-Douglas functional form is given, however, the same applies to the translog form.

$$\ln Y_{Sit} = \alpha_{S0} + \sum_{j=1}^{4} \alpha_{Sj} \ln X_{Sjit} + \alpha_{St} t + \sum_{m=1}^{6} \beta_{Sm} D_{Smit} + \lambda_S \left(- \frac{\phi(\beta' x_i)}{\Phi(\beta' x_i)} \right) + \omega_{Sit} - U_{Sit} \tag{10}$$

if $SPEC_{it} = 1$, where $E(\omega_{Sit} | SPEC_{it} = 1) = 0$, and

$$\ln Y_{Cit} = \alpha_{C0} + \sum_{j=1}^{4} \alpha_{Cj} \ln X_{Cjit} + \alpha_{Ct} t + \sum_{m=1}^{6} \beta_{Cm} D_{Cmit} + \lambda_C \left(\frac{\phi(\beta' x_i)}{1 - \Phi(\beta' x_i)} \right) + \omega_{Cit} - U_{Cit} \tag{11}$$

if $SPEC_{it} = 0$, where $E(\omega_{Cit}|SPEC_{it} = 0) = 0$.[4]

Given the functional specifications presented above, the technical inefficiency effects model for farmer i, time t, and production technology p (p=S,C) is defined as:

$$\mu_{pit} = \delta_{p0} + \sum_{j=1}^{11} \delta_{pj} Z_{pjit} + \delta_{pt} t + W_{pit} \tag{12}$$

where Z is a vector of farm-specific variables that are expected to determine technical efficiency levels, t is a time dummy that accounts for changes in mean technical efficiency, the δ's are unknown parameters[5], and W is a normally distributed random error term with mean zero and variance σ_u^2 , where σ_u^2 is defined such that $U_{it} \geq 0$.

In the second step, the usual procedure to obtain standard errors is incorrect, if selection bias is present (Heckman 1979). Therefore, standard errors of the production frontier are adjusted using the Murphy-Topel estimate of variance (Murphy and Topel 1985)[6].

Explaining variables in the production frontier and inefficiency effects model

Given the technology choice of the farmer, output is a function of land, labor, and other input factors as well as farmers' management capabilities. The input vector X in the production frontier model includes the classical production factors land, labor, and capital. Land refers to the area planted with coffee trees and is measured in hectares. Labor is measured in hours and includes all maintenance activities realized on the coffee plantation[7]. Capital is measured as the value of materials including chemical and organic fertilizers, pesticides and herbicides in Costa Rican Colones. Due to different concentrations of nutrients and active components, quantity is not a meaningful indicator, but the value of material inputs is assumed to reflect concentration and quality of the input. Furthermore, the age of the coffee trees[8] is included to reflect the farmer's investment in the renovation of the plantation[9]. To yield a more accurate specification of the model, a range of dummy variables are included that characterize the production system. According to Battese (1997) zero values in input variables can lead to biased estimates. The author suggests the inclusion of a dummy variable that assumes one if the input variable equals zero. In the present data set, there are 40 cases that do not apply any material input and seven cases that do not use any maintenance labor. A dummy variable is included that assumes one if labor or capital equal zero. Including separate dummies for each of the input variables leads to multicollinearity as the non-use of labor is highly correlated with the non-use of material. The second dummy variable included in the model assumes one if the farmer uses motorized equipment to perform the maintenance tasks on her plantation. This variable acts as a technology shifter moving the frontier up if a higher level of mechanization is achieved on the farm. Therefore, the expected sign of this variable is positive. Furthermore, a dummy that assumes one if the coffee varieties *Caturra* or *Catuai* are planted on the farm is included. Due to higher performance levels and suitability for local agro-ecological conditions, the use of these superior plant varieties should result in higher output levels. The pruning of coffee trees, while necessary to maintain plant productivity in the long run, is expected to reduce output levels in the following years. Two dummy variables – one for pruning in the current and one for pruning in the previous year – are included in the model. Furthermore, a regional dummy that assumes one if the farm is located in the Western Valley is used to reflect regional differences in production systems and natural conditions.

Productive efficiency is commonly associated with the management skills of the farmer (Coelli, Rao et al. 1998). The efficient organization of the production process depends on the availability of relevant technical knowledge as well as on the access to productive resources. If access to resources is constrained, optimal production choices are limited (Sadoulet and de Janvry 1995). Therefore, the explanatory variables included in vector Z in the inefficiency effects equation reflect farmers' management capabilities, their access to knowledge as well as to productive resources. The first two explanatory variables refer to education and experience in coffee farming. EDUCSEC is a dummy variable assuming one if the household head completed secondary school; EXPER reflects the farmer's experience in coffee growing measured in years. Both education and experience are expected to have a positive effect on farmers' management skills and thus on efficiency. The variable BOOK indicates whether the farmer keeps an account of the expenditures and labor activities related to the coffee plantation. Allowing for closer monitoring of input use and timing, this should increase the efficiency level of farmers. Similarly, the number of extension visits received by the farmer (ASSIST) is expected to positively contribute to productive efficiency. The variables AGE, FEMALE and FAMILY reflect the structure of the household. The number of children[10] and adults available for working on the coffee plantation reflect the household's access to family labor, which, if labor markets are imperfect, is expected to have an efficiency-enhancing effect (Eswaran and Kotwal 1986). Female-headed households are expected to face more difficulties in accessing markets and as a result display lower levels of efficiency. Similarly, the age of the household head, is expected to negatively influence efficiency levels. Reflecting households' endowments, total farm size in hectares is included in the inefficiency effects model. The hypothesized effect of farm size on efficiency is ambiguous. If labor market imperfections are severe, farm size is likely to be negatively related to efficiency. On the other hand, if financial markets are constrained, farm size as a proxy for overall wealth and credit access (Binswanger and Sillers 1983) is expected to have a positive influence on efficiency. Similarly, the variable ACT indicating whether a household pursues other income-generating activities than coffee does not have an unambiguous effect on efficiency. The effect is likely to be negative, if the diversion of labor from coffee cultivation to other activities results in maintenance activities being delayed or ignored. On the other hand, farmers working off-farm often have better access to information. Furthermore, additional income can help farmers to overcome liquidity constraints and thus to buy inputs in a timely manner, even if income from coffee is low. Finally, a variable that assumes one if the farmer is a member of a coffee cooperative is included. It is hypothesized that cooperatives help farmers to reduce transaction costs, thereby increasing their access to resources and improving their productive efficiency (Shaffer 1987, Deininger 1995). The last variable included in vector Z is a regional dummy that accounts for regional heterogeneity that might influence the achievement of technical efficiency. These factors include differences in the agro-ecological environment, institutional setting and level of competitive pressure. The ultimate effect of the regional dummy depends on which factors predominate. Summary statistics for the dependent and independent variables included in the stochastic frontier and in the inefficiency effects models are given in table one.

[Table 1]

Results of the efficiency model

The results of the probit model indicate that the probability of participation in specialty markets increases with farmers' experience in coffee cultivation, education, farm size, and membership in coffee cooperatives (see table 5 in the annex). Furthermore, farmers who have received extension in quality-enhancing cultivation practices are more likely to produce in the specialty segment. In contrast, if farmers dedicate their time to other income-generating activities, the probability of participation decreases. It is apparent that some significant differences exist in terms of farm and household characteristics between specialty and conventional farmers[11]. In the following sections, the results of the efficiency analysis are presented for each sub-group of farmers, while controlling for potential selection bias.

Model specification tests

In order to select the best model specifications, a number of null hypotheses were tested using the one-sided likelihood ratio test. Coelli (1995) showed that this test performs superior to a range of other tests when investigating the existence of inefficiency effects. Test results for the two separate models are presented in table two. The first null-hypothesis assumes that the Cobb-Douglas functional form is an adequate representation of the data. In the case of specialty coffee farmers, the null-hypothesis is accepted. In contrast, for the conventional coffee farmers the null-hypothesis is rejected at the 1% probability of error and the more flexible translog form is adopted. The next three tests refer to the inefficiency effects (Battese and Coelli 1995). The first test assesses the null-hypothesis that the inefficiency effects are absent from the model. If this was true, an average response function would fit the model. However, the null-hypothesis is rejected for both models at the 1% probability of error indicating that the stochastic frontier model is a more appropriate representation of the data than OLS. The second null-hypothesis assumes that the inefficiency effects are not stochastic, which would imply that they should be included in the frontier model and gamma would equal zero. This null-hypothesis is also rejected in both cases at the 1% probability of error. Finally, a test is run on whether the inefficiency estimates are indeed a linear combination of the exogenous variables included in the inefficiency effects model. The null-hypothesis that they are not related is rejected at the 1% probability of error in the case of the conventional coffee farmers and at the 5% probability of error in the case of the specialty coffee farmers.

[Table 2]

The translog functional form that is used in the case of conventional coffee farmers achieves greater flexibility by including second order terms and interactions between the inputs. Unlike the Cobb-Douglas functional form, it does not a priori impose a restrictive structure on the data. Hence, monotonicity is not necessarily fulfilled and has to be a posteriori tested for. In order for monotonicity to be fulfilled marginal products have to be positive with respect to all inputs (Sauer and Hockmann 2005). Production elasticities of land, labor and capital inputs are calculated for every farm household and t-tests are conducted to test whether these elasticities differ significantly from zero. Results indicate that partial production elasticities for land and labor are non-negative for all farmers in the sample. Of those farmers that use fertilizers or other agro-chemicals, one farmer displays negative production elasticities with respect to this input variable. Although it is inconsistent with theory that a farmer uses additional inputs if these reduce output, Chambers (1988) points out that this behavior may be observed in practice as a result of uncertainties faced in agricultural decision-making.

Results of the stochastic production frontiers

Table 3 presents parameter estimates from the two production frontier models. The inverse Mill’s ratio (IMR) is significant in both models indicating that selection bias is indeed present. The negative sign of the IMR in the specialty coffee model indicates that the average output of specialty farmers is larger than it would be if all farmers cultivated specialty coffee. In contrast, the negative sign of the IMR in the model for conventional farmers indicates that average output of conventional farmers is smaller than it would be if all farmers were using that technology. This can result from specialty coffee farmers having larger plantations, using inputs more intensively, or achieving higher levels of efficiency.

A range of dummy variables was included in the models to characterize the production process. In the case of specialty coffee farmers, the variable MOTOR is positive and significant at the 1% probability of error. As expected, output is higher for farmers who use motorized equipment. The variable SUPERIOR is significant at the 10% probability of error indicating that farmers who have superior plant varieties on their farm achieve higher output levels. According to the time dummy, which is significant at the 1% probability of error and has a negative sign, output is lower in 2003. As regards the model of conventional coffee farmers, the variables MOTOR, L_PRUNE, and the time dummy are significant at an error probability of 5%. As in the model of specialty coffee farmers, farmers who use motorized equipment achieve higher output levels. Also, output decreased in 2003 as compared to the previous year. Furthermore, if farmers pruned their coffee trees in the previous year, they achieve lower output levels. Pruning in the same year also has a negative effect on output, but is not significant.

In the case of the Cobb-Douglas form, the coefficients of land, labor, and capital reveal the partial production elasticities of these inputs. For specialty coffee farmers, partial production elasticities of land and capital are 0.593 and 0.262, respectively. Partial production elasticities of labor are 0.089, but according to the t-statistic not significantly different from zero. This may result from the fact that quality of output is not accounted for in the analysis. As specialty coffee farmers have to invest additional effort to increase the quality of their produce, they might be using additional units of labor even though this is not reflected in an increase in output quantity.

[Table 3]

In the case of the translog functional form, production elasticities cannot be directly obtained from the model. They are derived from the first order and second order terms of the inputs. The following formula can be used to calculate the partial production elasticities at the sample mean (Greene 2000: 286):

$$\delta \ln Y / \delta \ln X_k = \beta_k + \beta_{kk} \overline{\ln X_k} + \sum_{j \neq k}^{3} \beta_{kj} \overline{\ln X_j} \tag{13}$$

The respective standard errors can be obtained from the variance-covariance matrix. For example in the case of land, standard errors are calculated as the square root of:

$$Est.Var[\delta \ln Y / \delta \ln L] = w'(Est.Var[b_L])w \tag{14}$$

where w = (1, $\overline{\text{lnL}}$, $\overline{\text{lnA}}$, $\overline{\text{lnC}}$) is a vector of mean values and $b_L = (\beta_L, \beta_{LL}, \beta_{LA}, \beta_{LC})$ is the relevant partition of the maximum likelihood coefficient vector (Greene 2000: 286).

The production elasticities computed at the sample mean and their approximate standard errors can be found at the bottom of table 3. Partial production elasticities of land, labor and materials are positive and significant.

Results of the inefficiency effects models

Results of the two inefficiency effects models are presented in table 4. Mean technical efficiency of specialty farmers is estimated to be 81% and of conventional coffee farmers 61%. These percentages represent relative measures of technical efficiency referring to the most efficient farmers in the respective sub-sample as a benchmark to which all other farmers are compared. Accordingly, specialty farmers achieve higher levels of efficiency relative to the best-practice farmers using the technology. In contrast, there are more conventional coffee farmers that operate with low efficiency levels compared to their technology-specific standard. Intuitively, one would expect to find lower efficiency levels in the new market segment, where farmers are in a process of learning about the new technology. Yet, the probit analysis revealed that farmers with higher levels of education and more experience in coffee cultivation are more likely to participate in the specialty segment. This correlation between education and experience and the adoption of the new technology would likely explain at least to some extent the higher efficiency levels observed in the sub-sample of specialty coffee farmers.

Several factors were identified to have an influence on farm-specific technical efficiency levels in the case of specialty and conventional coffee farmers, respectively. It is important to note that a negative sign on a coefficient means that the predicted effect on inefficiency is negative, i.e. the variable has a positive effect on technical efficiency.

In the case of specialty coffee, the experience and age of the household head are significant. As expected, efficiency increases with experience in coffee cultivation and decreases with age. Unlike expected, the number of household members available for work on the coffee plantation has a significantly negative effect. This may indicate that family labor is used on the plantation beyond optimal levels, but it can also mean that family labor is used to increase the quality of the coffee, which is not reflected in the analysis. Furthermore, book keeping has an efficiency-enhancing effect, as hypothesized. With respect to households' endowments, total farm size is significant at the 5% probability of error having a negative effect on efficiency. Hence, larger farms are less efficient, which may be due to labor supervision problems. Finally, households pursuing other income-generating activities besides coffee display higher efficiency levels. This indicates that the positive liquidity effect outweighs the negative effect of labor diversion on efficiency. Apparently, in times of low coffee prices income from other activities is used to subsidize coffee cultivation and contributes to guarantee the timely and adequate application of inputs. Additionally, farmers working off-farm may have better access to relevant information (Mathijs and Vranken 2001).

With respect to conventional coffee farmers, the number of family members available for work on the coffee plantation as well as the availability of other income-generating activities are also significant at the 5% probability of error and have the same sign as in the specialty coffee model. Unlike in the case of specialty coffee farmers, conventional coffee farmers display higher levels of efficiency if they are member of a coffee cooperative and if they are located in the Brunca region. Experience, book keeping, and total farm size are not found to be significant in the model of conventional coffee cultivation.

[Table 4]

Conclusions and policy implications

This paper analyzes the determinants of farm-level technical efficiency in coffee production for a sample of 216 conventional and specialty coffee farmers in Costa Rica. This is done by simultaneously estimating a stochastic frontier model and the effects of a range of farm-specific variables on technical efficiency levels. Given that farmers in the sample use different sets of technologies, two separate production frontiers are estimated for farmers in each sub-sample. Unlike previous efficiency studies, the present approach controls for potential selectivity bias when estimating separate production frontiers. The results indicate that self-selection is present emphasizing the importance of taking selectivity bias into account when estimating different production functions for sample subsets.

The paper presents an empirical investigation of the factors that determine productive efficiency in coffee cultivation. The results allow for the derivation of adequate policy measures that can help farmers to improve their competitiveness in coffee production and to confront the adverse economic conditions caused by the coffee crisis.

In the case of specialty coffee farmers, it has been shown that efficiency decreases with farm size. This might be interpreted as an advantage of small-scale farms in the cultivation of specialty coffee, although it has to be kept in mind that, overall, small-scale farmers are less likely to participate in the specialty segment. Average farm size of specialty farmers is 17.9 ha compared to an average farm size of 8.6 ha in the case of conventional farmers. Within the sub-sample of specialty farmers, however, small-scale farmers are more efficient suggesting that larger farms experience labor supervision problems. Labor supervision is especially important in the case of quality-enhancing production techniques as the careful execution of labor tasks is critical for coffee quality, but less easily observed. Therefore, product differentiation may be well suited for small-scale producers, once the entry barriers they face can be overcome (Lewin, Giovannucci et al. 2004). Furthermore, efficiency increases in the specialty sub-sample, if farmers keep book of their activities and expenditures. This indicates that especially in the context of more sophisticated production techniques accounting methods are promising tools to increase farmers' efficiency.

In the case of conventional farmers, model results reveal that membership in cooperatives significantly contributes to the achievement of technical efficiency at the farm level. This is not the case for specialty farmers, which may be a result of low variability of that indicator, as most specialty farmers are members of cooperatives (92% as opposed to 79% in the sub-sample of conventional farmers). The analysis has revealed that multiple objectives can be accomplished by fostering coffee cooperatives. They play an important role in connecting farmers with specialty markets and in helping farmers (at least in the conventional segment) to organize their production process more efficiently.

In both models, the effect of other income-generating activities on efficiency is positive, which is likely to be a result of better access to liquidity and information of those farmers who have additional income sources. This underscores the need for alternative income opportunities in rural areas that can provide farmers with additional income in periods of low coffee prices and also give them the possibility to diversify out of coffee. The creation of feasible income opportunities should be fostered by conducting market research and facilitating the development of small and medium enterprises in rural areas.

Figures and Tables

Table 1: Summary statistics for the variables included in the efficiency analysis

Variable	Description	Specialty coffee			Conventional coffee		
		N	Mean	Std. Dev.	N	Mean	Std. Dev.
Dependent variable							
Coffee	Total amount of coffee cherries harvested (in fan)	173	203.6	319.1	258	89.3	107.6
Input variables							
Land	Total area cultivated with coffee (in ha)	173	7.7	9.2	258	4.0	4.1
Labor	Total labor hours used for the maintenance of coffee plantations	173	926.8	1823.9	258	478.9	543.9
Capital	Total value of fertilizers and agro-chemicals (in Costa Rican Colones)	173	763220	1232442	258	287667	475736
Agetree	Average age of the coffee trees	171	11.7	7.9	258	13.2	7.4
Dummy variables							
Inp_d	Dummy that assumes 1 if capital = 0 or labor = 0	173	0.04	0.2	258	0.07	0.3
Motor	Dummy that assumes 1 if hh uses motorized equipment	173	0.8	0.4	258	0.6	0.5
Prune	Dummy that assumes 1 if hh pruned in current year	173	0.2	0.4	258	0.3	0.5
L_prune	Dummy that assumes 1 if hh pruned in previous year	173	0.3	0.4	258	0.2	0.4
Superior	Dummy that assumes 1 if hh has superior coffee varieties	173	0.9	0.3	258	0.9	0.3
Region	Dummy that assumes 1 if hh is located in Western Valley (0 = Brunca)	173	0.98	0.1	258	0.45	0.5
Time	Time dummy (1 = 2003)	173	0.6	0.5	258	0.4	0.5
IMR	Inverse Mill's Ratio	173	0.6	0.4	258	0.4	0.4
Inefficiency effects							
Educsec	Hh head completed secondary school (1 = yes)	173	0.2	0.4	258	0.1	0.2
Exper	Experience in coffee cultivation (in years)	173	39.6	14.1	258	34.3	15.5
Age	Age of the hh head	173	56.1	12.5	258	54.9	14.0
Female	Hh is female-headed (1 = yes)	173	0.1	0.2	258	0.1	0.3
Family	No of family members available to work in coffee (children weighted by 0.5)	173	1.5	0.9	258	1.7	1.1
Book	Hh keeps book about the coffee activity (1 = yes)	170	0.4	0.5	257	0.2	0.4
Comem	Hh is member of coffee cooperative (1 = yes)	173	0.9	0.3	258	0.8	0.4
Act	Hh has income from other activities (1 = yes)	173	0.8	0.4	258	0.8	0.4
Size	Total farm size (in ha)	173	17.9	34.5	258	8.6	16.8
Assist	No. of extension visits received during the last year	173	1.3	0.9	258	0.9	1.2

Notes: hh = household, no. = number, fan = fanegas

Table 2: Hypotheses tests for the efficiency model specification

		Specialty coffee			Conventional coffee		
Null hypothesis	Restrictions	Critical χ^2 value	Test value λ		Critical χ^2 value	Test value λ	
H_0: $\beta_{ij} = 0$, $i \le j = 1...3$	6	10.64	4.20		16.81	17.52	***
H_0: $\gamma = \delta_0 = \ldots = \delta_s = 0$	14	28.49[a]	52.36	***	28.49	52.22	***
H_0: $\gamma = 0$	4	12.48[a]	22.96	***	12.48	17.28	***
H_0: $\delta_1 = \ldots = \delta_s = 0$	12	21.03	23.36	**	26.22	27.24	***

**(***) The null-hypothesis is rejected at a level of significance of p=0.95 (0.99).
a) Critical values are obtained from the mixed χ^2 distribution (see Kodde and Palm 1986)

Table 3: Parameter estimates from the production frontier

		Specialty coffee			Conventional coffee		
Variable	Parameter	ML estimate			ML estimate		
Constant	β_0	0.382	(.5680)		7.957	(2.0878)	***
Linp_d	β_1	2.938	(.4990)	***	-5.079	(2.1295)	**
Agetree	β_2	0.007	(.0654)		-0.020	(.0633)	
Motor	β_3	0.315	(.0949)	***	0.171	(.0713)	**
Prune	β_4	-0.077	(.0860)		-0.093	(.0676)	
Lag_prune	β_5	-0.007	(.0784)		-0.152	(.0776)	**
Superior	β_6	0.249	(.1316)	*	-0.004	(.1076)	
Time	β_7	-0.254	(.0913)	***	-0.171	(.0859)	**
Region	β_8	-0.493	(.3717)		-0.033	(.0957)	
IVM	β_9	-0.246	(.0982)	**	-0.219	(.0929)	**
Land	β_L	0.593	(.0722)	***	1.217	(.3066)	***
Labor	β_A	0.089	(.0585)		0.015	(.1651)	
Capital	β_C	0.262	(.0384)	***	-1.081	(.3747)	***
0.5*Land²	β_{LL}				0.281	(.1228)	**
0.5*Labor²	β_{AA}				0.089	(.0670)	
0.5*Capital²	β_{CC}				0.128	(.0343)	***
Land*Labor	β_{LA}				-0.167	(.0753)	**
Land*Capital	β_{LC}				-0.007	(.0318)	
Labor*Capital	β_{AC}				-0.021	(.0180)	
Production elasticities							
Land					0.480	(.0774)	***
Labor					0.126	(.0489)	***
Capital					0.197	(.0352)	***

*(**)[***] The null-hypothesis is rejected at a level of significance of p=0.90 (0.95) [0.99].
Note: Standard errors are adjusted using the Murphy-Topel variance estimate (Murphy and Topel 1985).

Table 4: Results of the inefficiency effects model

Variable	Parameter	Specialty coffee ML estimate			Conventional coffee ML estimate		
Constant	δ_0	0.452	(0.999)		-0.860	(1.080)	
Educsec	δ_1	1.304	(0.868)		0.796	(0.574)	
Exper	δ_2	-0.085	(0.039)	**	-0.004	(0.012)	
Age	δ_3	0.058	(0.030)	*	0.024	(0.015)	
Female	δ_4	0.121	(0.906)		0.237	(0.476)	
Family	δ_5	0.825	(0.325)	**	0.282	(0.132)	**
Book	δ_6	-2.133	(1.063)	**	-0.492	(0.420)	
Comem	δ_7	-0.906	(0.930)		-0.819	(0.385)	**
Act	δ_{10}	-2.136	(0.882)	**	-0.937	(0.426)	**
Size	δ_{11}	0.012	(0.005)	**	-0.013	(0.010)	
Assist	δ_{12}	-0.082	(0.242)		-0.133	(0.154)	
Region	δ_9	-1.549	(1.217)		0.797	(0.468)	*
Time	δ_8	-0.234	(0.430)		-0.382	(0.349)	
Variance parameter							
SIGMA²	σ^2	0.743	(0.171)	***	0.936	(0.361)	***
GAMMA	γ	0.825	(0.066)	***	0.942	(0.030)	***
Log likelihood function		-99.018			-184.174		
Mean efficiency		0.812			0.610		

*(**)[***] The null-hypothesis is rejected at a level of significance of p=0.90 (0.95) [0.99].

Annex

Table 5: Results of the pooled probit model on participation in specialty markets

Variables	Description	Coefficient		Robust standard errors
EXPER	Experience in coffee cultivation (in years)	0.021	***	0.006
EDUC	Level of education of the household head (1=no formal education, 6= university degree)	0.319	***	0.077
LAND	Total area of land cultivated with coffee (in ha.)	0.045	***	0.019
COMEM	Household is member of coffee cooperative (0/1)	0.413	**	0.206
ALT	Altitude of the coffee plantation (in meters)	0.006	***	0.001
QUAL	Whether household received training in quality enhancing practices (0/1)	0.861	***	0.263
NONAG	Number of non-agricultural income-generating activities household members are engaged in	-0.266	***	0.085
MEN	Number of male adults in the household (>= 14 years)	-0.095		0.089
WOMEN	Number of female adults in the household (>= 14 years)	0.034		0.083
CHILD	Number of children in the household (age below 14)	-0.081		0.065
TIME	time dummy (0 = 2002, 1 = 2003)	0.670	***	0.149
CONST	Constant	-8.585	***	0.863
N		431.000		
Log pseudo-likelihood		-193.851		
Wald chi² (11)		122.590	***	
Pseudo R^2		0.332		

**(***) significant at p=0.05 (0.01)

References

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

Battese, G.E. 1992. Frontier production functions and technical efficiency: a survey of empirical applications in agricultural economics. *Agricultural Economics* 7: 185-208.

Battese, G.E. 1997. A note on the estimation of Cobb-Douglas production functions when some explanatory variables have zero values. *Journal of Agricultural Economics* 48 (2): 250-252.

Battese, G.E., and T.J. Coelli. 1993. A stochastic frontier production function incorporating a model for technical inefficiency effects. Working Papers in Econometrics and Applied Statistics 69/1993. Armidale, Australia: Department of Econometrics, Univ. of New England.

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: 325-332.

Binswanger, H.P., and D.A. Sillers. 1983. Risk aversion and credit constraints in farmers' decision-making: a reinterpretation. *Journal of Development Studies* 20: 5-21.

Chambers, R.G. 1988. *Applied production analysis: a dual approach*. Cambridge, UK: Cambridge Univ. Press.

Charveriat, C. 2001. Bitter coffee: how the poor are paying for the slump in coffee prices. Oxford, UK: Oxfam.

Coelli, T.J. 1995. Estimators and hypothesis tests for a stochastic frontier function: a Monte Carlo analysis. *The Journal of Productivity Analysis* 6: 247-268.

Coelli, T.J., Rao, D.S.P., and G.E. Battese. 1998. *An introduction to efficiency and* *productivity analysis*. Boston, Dordrecht, London: Kluwer Acad. Publ.

Curtiss, J., Brümmer, B., Medonos, T., and R. Weaver. 2006. Can prosocial behavior affect efficiency and structural change? Evidence from Czech Agriculture. 96th Seminar of the European Association of Agricultural Economists, Jan. 10-11, 2006, Tänikon, Switzerland.

Damiani, O. 2005. Adversidad y cambio: estratégias exitosas de pequeños productores de café en Centroamérica. San José, Costa Rica: RUTA.

Deininger, K. 1995. Collective agricultural production: a solution for transition economies? *World Development* 23 (8): 1317-1334.

Eswaran, M., and A. Kotwal. 1986. Access to capital and agrarian production organisation. *The Economic Journal* 96: 482-498.

Flores, M., Bratescu, A., Martinez, J.O., Oviedo, J.A., and A. Acosta. 2002. Centroamérica: El impacto de la caída de los precios del café. Estudios y Perspectivas 9. Mexico, D.F.: UN/ECLAC.

Greene, W.H. 2000. *Econometric analysis*. London et al.: Prentice Hall Internat.

Greene, W.H. 1997. *Econometric analysis*. London et al.: Prentice Hall Internat.

Heckman, J. 1979. Sample selection bias as a specification error. *Econometrica* 47: 153-162.

Huang, C.J. and J.-T. Liu. 1994. Estimation of a non-neutral stochastic frontier production function. *The Journal of Productivity Analysis* 5: 171-180.

International Coffee Organization (ICO). 2004. Lessons from the world coffee crisis: a serious problem for sustainable development. UNCTAD XI Conference, June 13-18, 2003, Sao Paulo, Brazil.

Kodde, D.A. and F.C. Palm. 1986. Wald Criteria for jointly testing equality and inequality restrictions. *Econometrica* 54 (5): 1243-1248.

Kumbhakar, S. 2004. Productivity and efficiency measurement using parametric econometric methods. XIII International Tor Vergata Conference on Banking and Finance, Dec. 1-3, 2004, Rome, Italy.

Kumbhakar, S., Gosh, S., 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.

Kumbhakar, S.C., and C.A.K. Lovell. 2000. *Stochastic frontier analysis*. Cambridge, UK: Cambridge Univ. Press.

Larsen, K., and K. Foster. 2005. Technical efficiency among organic and conventional farms in Sweden 2000-2002: A counterfactual and self-selection analysis. American Agricultural Economics Association Annual Meeting, Providence, RI, USA.

Larson, J.B.M., Palaskas, T., and G.J. Tyler. 1999. Land titling and technical efficiency among small coffee producers in Honduras. *Canadian Journal of Development Studies* 20 (2): 361-82.

Lee, L. 1978. Unionism and wage rates: a simultaneous equations model with qualitative and limited dependent variables. *International Economic Review* 19 (2): 415-433.

Lewin, B., Giovannucci, D., and P. Varangis. 2004. Coffee Markets - New paradigms in global supply and demand. Agriculture and Rural Development Discussion Paper 3. Washington, DC, USA: The World Bank.

Mathijs, E. and L. Vranken. 2001. Farm restructuring and production efficiency in transition agriculture. Working Paper. Leuven, Belgium: Catholic University of Leuven.

Meeusen, W., and J. van den Broeck. 1977. Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review* 18 (2): 435-444.

Muradian, R., and W. Pelupessy. 2005. Governing the Coffee Chain: The Role of Voluntary Regulatory Systems. *World Development* 33 (12): 2029-2044.

Murphy, K.M., and R.H. Topel. 1985. Estimation and inference in two-step econometric models. *Journal of Business and Economic Statistics* 3 (4): 370-379.

Ponte, S. 2002. The "latte revolution"? Regulation, markets and consumption in the global coffee chain. *World Development* 30 (7): 1099-1122.

Reifschneider, D.L., and R.E. Stevenson. 1991. Systematic departures from the frontier: a framework for the analysis of firm inefficiency. *International Economic Review* 32 (3): 715-723.

Sadoulet, E. and A. de Janvry. 1995. *Quantitative Development Policy Analysis*. Baltimore, MD, USA: John Hopkins Univ. Press.

Sauer, J., and H. Hockmann. 2005. The need for theoretically consistent efficiency frontiers. 11th Congress of the European Association of Agricultural Economists, Aug. 24-27, 2005, Copenhagen, Denmark.

Shaffer, J.D. 1987. Thinking about farmers' cooperatives, contracts, and economic coordination. In *Cooperative Theory: New Approaches*. ACS Service Report 18: 61-86. Washington, DC, USA: Agricultural Cooperative Service, US Department of Agriculture.

Sipiläinen, T., and A.O. Lansink. 2005. Learning in organic farming - an application on Finnish dairy farms. XI Congress of the European Association of Agricultural Economists, Aug. 24-27, 2005, Copenhagen, Denmark.

StataCorp. 2001. *Stata Reference Manual, Release 7.0, Vol. 4 Su-Z.* College Station, TX, USA: Stata Press.

Varangis, P., Siegel, P., Giovannucci, D., and B. Lewin. 2003. Dealing with the coffee crisis in Central America: impacts and strategies. Policy Research Working Paper 2993. Washington, DC, USA: The World Bank.

Villano, R., and E. Fleming. 2004. Analysis of technical efficiency in a rainfed lowland rice environment in Central Luzon Philippines using a stochastic frontier production function with a heteroskedastic error structure. Working Paper Series in Agricultural and Resource Economics No. 2004-15. Armidale, Australia: Univ. of New England.

Wang, H.-J., and P. Schmidt. 2002. One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels. *Journal of Productivity Analysis* 18: 129-144.

Wollni, M., and M. Zeller. 2007. Do farmers benefit from participating in specialty markets and cooperatives? The case of coffee marketing in Costa Rica. *Agricultural Economics* 37 (2/3), forthcoming.

Wooldridge, J.M. 2002. *Econometric analysis of cross section and panel data*. Cambridge, MA, USA: Massachusetts Institute of Technology Press.

Zellner, A., Kmenta, J., and J. Dreze. 1966. Specification and estimation of Cobb-Douglas production function models. *Econometrica* 34 (4): 784-795.

---

[1] The direct estimation of a production frontier is criticized for its susceptibility to simultaneous equation bias that results if farmers select the levels of input and output that maximize profits for given prices (Coelli, Rao et al. 1998: 54). In this matter, it is referred to Zellner, Kmenta and Dreze (1966), who show that the estimation of a production function does not suffer from simultaneous equation bias, if expected rather than actual profit is maximized.

[2] The random-effects probit model was disregarded because the Gauss-Hermite quadrature was not stable and there is no alternative method for calculating the random-effects model in Stata (StataCorp. 2001: 421)

[3] As the data for 2002 was obtained by recall, the dummy variable also reflects the measurement error that is likely to be higher for 2002 as compared to 2003.

[4] Technically, the selection bias results from the fact that $E(V_{Sit}|S_{it} = 1) \neq 0$ and $E(V_{Cit}|S_{it} = 0) \neq 0$. The terms $\lambda_S [\phi(\beta' x_i) / \Phi(\beta' x_i)]$ and $\lambda_C [\phi(\beta' x_i) / (1-\Phi(\beta' x_i))]$ are in fact the means $E(V_{Sit}|S_{it} = 1)$ and $E(V_{Cit}|S_{it} = 0)$, respectively (Lee 1978).

[5] According to Battese and Coelli (1995), restricting the model to not include an intercept parameter may lead to biased parameter estimates associated with the z-variables.

[6] Asymptotically, the Murphy-Topel estimate gives the same results as the Heckman correction (see Greene 1997: 981).

[7] Labor input used for the application of fertilizers and agro-chemicals has been excluded as it is correlated with the amount of these materials applied. Harvesting is also excluded as workers are hired under a piece-rate payment scheme, so that expenditures on harvest labor are highly correlated with total output.

[8] The squared term was excluded due to insignificance.

[9] Lagged labor input was also included as a proxy for investment in the plantation. Maintenance activities performed in one year are likely to have a positive effect on output in the following year as well. However, the indicator was not significant and therefore excluded from the model.

[10] Children below the age of 14 are weighted by 0.5 as they are usually assigned only the easier tasks on the plantation.

[11] For a more detailed discussion of the factors influencing participation in specialty markets see Wollni and Zeller 2007.