Jorge Andrés PERDOMO CALVO*,, Josu ARTECHE*† and Alberto ANSUATEGI ***†**

Returns to Scale and Technical Efficiency in Colombian Coffee Production: Implications for Colombia's Agricultural and Land Policies

This paper applies a parametric approach to estimate technical and scale (in)efficiencies using input and output data at the level of 850 individual Colombian coffee-farms. Different Stochastic Production Frontier functions are estimated using a two-step procedure that corrects the endogeneity that has been ignored in previous works, leading to more reliable (i.e. unbiased and consistent) results. We conclude that small and medium coffee farmers are technically inefficient and exhibit increasing returns to scale, whereas large coffee farmers are close to being quasi-technically efficient and exhibit decreasing returns to scale. The corrected-for-endogeneity estimation also indicates that small and medium-sized units must prioritise primarily the land factor, whereas large farms should concentrate their efforts on increasing the labour factor. Based on these results, several agricultural and land policy recommendations are made.

Keywords: coffee production, stochastic production frontier, endogeneity, technical efficiency, returns to scale

JEL classification: Q12

* PhD candidate at Doctoral Programme in Economics: Tools of Economic Analysis (University of the Basque Country UPV/EHU), Bilbao 48015, Spain.

** CEO/ General Director at Big-Analytics Consultants, Madrid, Spain. Corresponding author: gerencia@big-analytics.es

*** Associate Researcher, Basque Centre for Climate Change (BC3), Scientific Park, UPV/EHU, 48940 Leioa, Spain.

Received: 22 September 2022; Revised: 10 October 2022; Accepted: 12 October 2022.

Introduction

Colombia is the World's third largest producer of coffee after Brazil and Vietnam and the highest in terms of the Arabica bean (Giovanucci *et al.*, 2002; ICO, 2021). From the time the commercial production of coffee first began in 1870, coffee has traditionally played an important role in the economic growth of Colombia. Today, it plays a smaller economic role, but it is still a primary source of income for nearly half a million rural families.

A great deal of Colombian coffee is produced on small and medium-sized family farms. This may be the consequence of the existence at the beginning of the twentieth century of large quantities of unclaimed land on the slopes of (particularly) the central cordillera, the relative scarcity of large accumulations of capital, and the country's inability to attract foreign immigrants. Whatever the precise reason for the growth of small and medium-sized coffee farms in Colombia, currently units of small size comprise the great bulk of coffee farms of the nation. Thus, the National Federation of Coffee Growers of Colombia estimates that there are 560,000 coffee growing families, where small farmers with less than 5 hectares of land are responsible for approximately 69 percent of coffee production in Colombia. This feature can in the future be exacerbated by virtue of the peace deal signed by the Government of Colombia with the Revolutionary Armed Forces of Colombia (FARC) at the end of 2016, that pledges to address unequal land ownership and foster development in neglected rural areas hit hard by violence.

Some reports indicate that agricultural productivity in general, and the productivity of coffee plantations in particular, are relatively low in Colombia (OECD, 2015). Hence, it is essential to assess what possibilities exist for improving the efficiency of coffee production. It is particularly interesting to analyse if providing land to a wider share of the rural population has a positive effect in terms of improving the productivity of coffee plantations. For that analysis, it is important to focus on the relationship between land size and productivity in Colombian coffee production.

This study aims to shed some light in this direction by examining the technical efficiency of small, medium- and large-sized coffee farms as well as testing for economies of scale in each of these groups. For that purpose, we apply a parametric approach to estimate technical and scale (in) efficiencies using input and output data at the level of 850 individual farms (556 small, 200 medium and 94 large-sized) in the Departments of Risaralda, Caldas, and Quindío in Colombia in year 2003. As far as we know, this database is the most recent to have been applied to coffee farms and, although a more current database may be desirable, no updated database exists with the same level of detail.

This study draws on the extensive literature on technical efficiency and returns to scale in agricultural production in developing countries following the seminal finding by Sen (1962) that yields per acre and farm size were inversely related for small Indian farms. This inverse relationship has been confirmed by studies in Africa (Barrett, 1996; Kimhi, 2006), Asia (Carter, 1984; Heltberg, 1998; Akram-Lodhi, 2005; Besley and Burgess, 2000), Europe (Alvarez and Arias, 2004) and Latin America (Berry and Cline, 1979) and contested by others, such as Bhalla and Roy (1988), who have shown that when differences in land quality are taken into consideration this phenomenon disappears. Lamb (2003) has additionally attributed these findings to labour market imperfections and measurement errors. More recent studies have imposed greater theoretical structure on the empirical work and have found that large farms are more efficient and more productive than small farms (Adamopoulos and Restuccia, 2014).

A subset of the literature on technical efficiency and returns to scale has focused on coffee production. Thus, Data Envelopment Analysis (DEA) techniques have been used to compute farm-level technical efficiency measures in Costa

Rica by Mosheim (2002), in Côte d'Ivore by Binam *et al.* (2003), in Colombia by Perdomo and Mendieta (2007), and in Vietnam by Rios and Shively (2006) and Garcia and Shively (2011). Vedenov *et al.* (2007), Nchare (2007) and Perdomo and Hueth (2011), instead of using non-parametric mathematical programming, have made use of Stochastic Frontier Analysis (SFA) to estimate an input distance function and evaluate production efficiency in Mexico, Cameroon, and Colombia, respectively.

Perdomo and Hueth (2011) and Perdomo and Mendieta (2007) constitute two preliminary attempts to study the production function, returns to scale and technical efficiency of Colombian coffee farms using SFA and DEA. They found that small- and medium-sized coffee farms presented technical inefficiency and increasing returns to scale, whereas the larger coffee farms presented technical efficiency and decreasing returns to scale. Nevertheless, some authors have raised concerns about endogeneity in production function estimation (Kutlu, 2010; Tran and Tsionas, 2013). Stochastic production frontier models usually assume that input choices are independent of the efficiency and productivity terms. However, if a producer observes some factors – unobservable by the econometrician – that affect a farm's efficiency and/or its productivity, the input choices may also be influenced by this knowledge, resulting in an endogeneity problem in the stochastic production frontier estimation (Shee and Stefanou, 2015). This situation may therefore lead to a biased inference on input elasticities, economies of scale and technical efficiency. In this paper we follow Kutlu (2010) (see also Amsler *et al.*, 2016) to deal with endogeneity when estimating the SFA to assess the technical and scale (in)efficiencies of Colombian coffee farms.

The rest of the paper is organised as follows: the empirical model for the estimation of technical and scale efficiency is presented in the next section. The data set is described in the third section and the empirical results are discussed in the fourth section. Some recommendations for agricultural and land policies and concluding remarks follow in the fifth and sixth sections, respectively.

Empirical Model

Consider the following general form of the stochastic production frontier (SPF) function:

$$q_i = f(x_{i1}, \cdots, x_{im}, \beta)e^{\mu_i - \omega_i} \qquad i = 1, \cdots, n \tag{1}$$

where q_i is the observed output produced by the i-th farm, x_{ij} is the quantity of the j-th input used by the i-th farm ($j=1,\cdots,m$), β is a vector of parameters to be estimated, and $\mu_i - \omega_i$ is a composite error term. The μ_i term corresponds to the statistical noise (assumed to be independently and identically distributed) and ω_i is a non-negative random variable associated with technical inefficiency. Regarding $f(\cdot)$, the Transcendental and Cobb-Douglas functions are the two most commonly used functional forms in empirical studies of production, which include frontier analyses (Battese and Broca, 1997). The Cobb-Douglas stochastic frontier model takes the form:

$$q_i = A \prod_{j=1}^{m} x_{ij}^{\beta_j} e^{\mu_i - \omega_i} \tag{2}$$

which can be estimated as a linear relationship using the following expression:

$$ln\, q_i = \beta_0 + \sum_{j=1}^{m} \beta_j \, ln\, x_{ij} + \mu_i - \omega_i \tag{3}$$

Similarly, the logarithmic transformation of the Transcendental SPF model takes the following form:

$$\begin{aligned} ln\, q_i = \beta_0 &+ \sum_{j=1}^{m} \beta_j \, ln\, x_{ij} + \frac{1}{2}\sum_{j=1}^{m} \beta_{jj} (ln\, x_{ij})^2 + \\ &+ \sum_{j=1}^{m} \sum_{k>j}^{m} \beta_{jk} \, ln\, x_{ij} \, ln\, x_{ik} + \mu_i - \omega_i \end{aligned} \tag{4}$$

Note that the usual procedures for estimating SPF models depend on the assumption that the inputs are exogenous. However, in many situations this assumption is difficult to maintain because some inputs can be influenced by unobserved factors such as expected rainfall in the farm's location, managerial ability of the farmer etc. that obviously have an impact also on the produced output. To overcome this endogeneity problem, we follow Kutlu (2010) and Amsler *et al.* (2016) and estimate the SPF in a two-step procedure. In the first step, we estimate the reduced form of the inputs demand function system, where the endogenous variables ($x_{i1},\cdots,x_{im}$) are log- linear functions of their prices ($p_{x_{i1}}, \cdots, p_{x_{im}}$) and a set of unobserved factors, which have the characteristics of providing good instruments for the log inputs. Note that the error terms of such regressions, denoted as $\varepsilon_{i1}, \cdots, \varepsilon_{im}$, are possibly contemporaneously correlated, and consequently the system requires an estimation by means of seemingly unrelated regression (SUR) using iterative generalised least squares to obtain unbiased, consistent, and efficient estimators (Rosales *et al.*, 2013). In the second step the residuals from the SUR estimation, denoted as $\hat{\varepsilon}_{i1}, \cdots, \hat{\varepsilon}_{im}$, are used as controls in an operational version of equation (1):

$$q_i = f(x_{i1}, \cdots, x_{im}, \hat{\varepsilon}_{i1}, \cdots, \hat{\varepsilon}_{im}; \beta^*)e^{\mu_i - \omega_i} \tag{5}$$

Following Battese and Coelli (1992), the specification of the technical efficiency of production for the i-th farm (TE_i) is defined by:

$$TE_i = \frac{f(x_{i1}, \cdots, x_{im}, \hat{\varepsilon}_{i1}, \cdots, \hat{\varepsilon}_{im}; \beta^*)e^{\mu_i - \omega_i}}{f(x_{i1}, \cdots, x_{im}, \hat{\varepsilon}_{i1}, \cdots, \hat{\varepsilon}_{im}; \beta^*)e^{\mu_i}} = e^{-\omega_i} \tag{6}$$

$TE_i \in [0,1]$ provides a measure of the shortfall of observed output from maximum feasible output in an environment that allows for variation across farms.

The elasticity of output[1] of the i-th farm with respect to the j-th input ($e_{q_i x_{ij}}$) is defined by:

$$e_{q_i x_{ij}} = \frac{\partial\, ln\, q_i}{\partial\, ln\, x_{ij}} \tag{7}$$

[1] Whereas the elasticity is constant for the Coob-Douglas specification, the form of the translog in equation (4) implies that the elasticity depends on the level of the inputs. Following general conventions (see Greene, 2012) the elasticity is here calculated at the average inputs as $\frac{\partial\, ln\, \hat{q}_i}{\partial\, ln\, x_{ij}} = \hat{\beta}_j + \sum_{k=1}^{m} \hat{\beta}_{jk} \, ln\bar{x}_{ik}$ where, $ln\, \bar{x}_{ij}$ and $ln\bar{x}_{ik}$ are the averages log-inputs.

As a result, the returns to scale (RTS) are expressed by:

$$RTS_i = \sum_{j=1}^{m} e_{q_i x_{ij}} \tag{8}$$

It measures the proportional change in output resulting from a unit proportional increase in all inputs. Then $RTS>1$ shows the presence of increasing returns to scale, $RTS<1$ indicates the existence of decreasing returns to scale and $RTS=1$ implies constant returns to scale.

Data Description

The data used in the present study are from a survey undertaken by the Department of Agricultural and Resource Economics (AREC) of the University of Maryland[2] (United States) during the year 2004 in the Departments of Risaralda, Caldas, and Quindío in Colombia. It contains information obtained from 850 coffee farms of which 556 are small-sized (below 2 hectares), 200 are medium-sized (between 2 and 7 hectares) and 94 are large-sized (above 7 hectares). The information collected corresponds to the 2003 crop year[3].

For the purposes of the present study, output is measured in annual arrobas[4] produced. Four inputs are included in the production frontier function, namely land measured in hectares, labour (including family, hired workers and coffee pickers) measured in full time equivalents, intermediate inputs (fertiliser and pesticides) measured in kilograms, and capital stock (machinery) measured through a synthetic index of capital intensity. We use this index because the information in the survey only includes the number of machines used by each farm, without discriminating between different types of machines. This index, called Index of Machinery Intensity (IMI), is constructed by means of Principal Component Analysis (PCA) and feature scaling or minmax scaler process as follows (see details in Johnson, 1998, Ch. 5 and Perdomo *et al*. 2016, p. 42-44). The relative weights across different factors of machinery used in coffee growing (total number of coffee pulper machines, water pump machines, coffee demucilager machines, motors, coffee silo machines, fumigation machines, scythes machines and chainsaws) were estimated with PCA, because their units of measurement are heterogeneous, so their direct aggregation or sum is unsuitable for determining machinery intensity (MI). Once MI is calculated, values are normalised (between zero and one) using feature scaling or min-max scaler (see details in Perdomo *et al*., 2016, p. 42) as

$$IMI_i = \frac{MI_i - MI_{min}}{MI_{max} - MI_{min}} \tag{9}$$

where MI_i are obtained from PCA, MI_{min} and MI_{max} are their minimum and maximum values and $IMI_i \rightarrow 1$ indicates more intensity of machinery.

Several additional variables have been included in the regression in the first step to obtain the residuals used as controls in the second step. First, the number of people per household is used as a proxy of rural population density. Second, three dummy variables have been used to indicate (i) if the farm obtains income from activities other than coffee production, (ii) if the main source of income comes from coffee activity, and (iii) if the farm has road access to the municipal centre. The sample mean of these, and the rest of variables are given in Table 1.

Empirical Results

Table A1 in the Appendix shows the SUR estimates (first stage) of the input demand functions. The residuals in this regression are incorporated in the SPF function in the second step. Tables 2, 3 and 4 show the maximum likelihood

Table 1: Sample mean values of model variables.

Variable	Small-sized farms	Medium-sized farms	Large-sized farms
Output (arrobas year)	160.31	481.97	2726.11
Land (hectares)	1.44	3.53	14.33
Labour (workers, full time equiv.)	9.09	21.87	99.02
Chemicals (Kgs)	3.48	23.59	102.33
Machinery (capital intensity index-IMI-)	0.13	0.23	0.18
Price of Land (US$ per hectare)	22,184.28	41,378.20	52,852.84
Price of Labour (US$ weekly per worker)	100.44	179.76	188.55
Price of Chemicals (US$ per Kg)	7.22	7.66	8.82
Price of Machinery (index)	0.87	0.93	0.83
Family size (persons)	4.00	3.92	3.24
Diversification (dummy variable)	0.28	0.42	0.50
Specialisation (dummy variable)	0.87	0.80	0.74
Road Access (dummy variable)	0.66	0.79	0.98
Sample size	556	200	94

Source: Own composition

[2] The survey strategy was conducted by Prof. Darrell Hueth.

[3] Unfortunately, similar surveys have not been conducted since then.

[4] Arroba is a Portuguese and Spanish unit of weight, mass, or volume, representing a weight of around 25 pounds or 12.5 kilograms.

estimates of the different specifications for the SPF function for small-, medium- and large-sized farms, respectively. The standard errors from the two-stage method employed here are inconsistent because the estimates are conditional on estimated standardised error terms from the first stage. Hence, we only present bootstrap standard deviations as proposed by Kutlu (2010). The tables also include values of the Hausman test indicating that endogeneity exists in equation (3) in the three groups of farms. The general significance of the control functions reinforces the hypothesis of endogeneity of the inputs. The results of the Sargan test evidence as well the validity of the instruments used in the first step to control for the endogeneity of the input variables. For the sake of comparison, the estimation of the SPF function with and without correction of endogeneity are included. Even though not all the inputs are individually significant, we keep them in all the functional specifications for comparative purposes.

Table 2: Stochastic Production Frontier estimates (Second Stage) for small-sized farms.

Dependent Variable: (Coffee Production)	**Translog without endogeneity corrected**	**Translog with endogeneity corrected**	**Cobb Douglas without endogeneity corrected**	**Cobb Douglas with endogeneity corrected**
Explanatory Variables	**Coefficients (β)**	**Coefficients (β*)**	**Coefficients (β)**	**Coefficients (β*)**
Intercept	6.0258***	3.0936	3.5793***	2.2199**
Land	1.3551	1.5594	0.6171***	1.7572***
Labour	0.3007	0.8253	0.6322***	0.5204**
Chemicals	-0.2040	0.5583	0.1770***	0.7259***
Machinery	1.6073**	0.5774	-0.0076	-0.3369
Land^2	-0.4230	-0.3396	-	-
Labour ^2	-0.3609***	-0.4252***	-	-
Chemicals^2	-0.0750	-0.1472**	-	-
Machinery^2	0.1415	0.0448	-	-
Land*Labour	-0.0884	0.0654	-	-
Land*Chemicals	-0.2521**	-0.1670*	-	-
Land*Machinery	0.0431	-0.0833	-	-
Labour*Chemicals	0.1450**	0.1366***	-	-
Labour*Machinery	-0.4579*	-0.2568	-	-
Chemicals*Machinery	-0.1355	-0.1059	-	-
Residual first stage land	-	-1.0262*	-	-1.2548**
Residual first stage labour	-	-0.0145	-	0.1328
Residual first stage chemicals	-	-0.7090***	-	-0.6492***
Residual first stage machinery	-	0.2310	-	0.3001
Natural logarithm of v_i	-1.9760***	-2.1502***	-1.868013***	-2.0444***
Natural logarithm of u_i	-1.6106***	-1.7686***	-1.56359***	-1.7063***
AIC	950.76	860.62	974.04	886.66
Wald test (chi-square)	792.46***	1,690.14***	569.20***	1,562.62***
LR test of sigma_u=0 (chi-square)	55.25***	78.62***	47.97***	66.59***
Hausman test for endogeneity (chi-square)	-	62.97***	-	56.42***
Sargan test (F statistic)	-	0.01	-	0.06
RTS	2.15	3.17	1.42	2.67
TE (50th percentile)	0.75	0.75	0.74	0.75
Observations	555	550	555	550

Note: *, ** and *** Significant at 0.10, 0.05 and 0.01 levels, respectively
Source: Own composition

Table 3: Stochastic Production Frontier estimates (Second Stage) for medium-sized farms.

Dependent Variable: (Coffee Production)	**Translog without endogeneity corrected**	**Translog with endogeneity corrected**	**Cobb Douglas without endogeneity corrected**	**Cobb Douglas with endogeneity corrected**
Explanatory Variables	**Coefficients (β)**	**Coefficients (β*)**	**Coefficients (β)**	**Coefficients (β*)**
Intercept	6.1810***	7.0956***	4.1913***	5.5284***
Land	0.7045	2.1543**	0.5397***	1.7836***
Labour	-0.0810	-0.3845	0.5087***	0.1133
Chemicals	-0.6445***	-1.1043***	0.0655*	-0.2826**
Machinery	0.4215**	0.7649*	0.0551	0.5864***
Land^2	0.3019	0.2633	-	-
Labour ^2	-0.0544	-0.0969	-	-
Chemicals^2	0.0173	0.0493	-	-
Machinery^2	0.0692***	0.0336	-	-
Land*Labour	-0.0292	0.0360	-	-

Dependent Variable: (Coffee Production)	Translog without endogeneity corrected	Translog with endogeneity corrected	Cobb Douglas without endogeneity corrected	Cobb Douglas with endogeneity corrected
Explanatory Variables	**Coefficients (β)**	**Coefficients (β*)**	**Coefficients (β)**	**Coefficients (β*)**
Land*Chemicals	-0.1599	-0.1309	-	-
Land*Machinery	0.0165	0.0352	-	-
Labour*Chemicals	0.2848*	0.2533**	-	-
Labour*Machinery	-0.0410	-0.0501	-	-
Chemicals*Machinery	-0.0237	-0.0151	-	-
Residual first stage land	-	-1.6827***	-	-1.3040***
Residual first stage labour	-	0.4826**	-	0.4709***
Residual first stage chemicals	-	0.4527***	-	0.3273***
Residual first stage machinery	-	-0.5292***	-	-0.5967***
Natural logarithm of v_i	-2.0833***	-2.2288	-1.820015***	-1.8495***
Natural logarithm of u_i	-1.1813***	-1.2933	-1.294065**	-1.6848
AIC	307.72	289.64	310.17	288.78
Wald test (chi-square)	295.25***	570.08***	243.33***	215.98***
LR test of sigma_u=0 (chi-square)	5.66***	5.59***	3.05***	1.52
Hausman test for endogeneity (chi-square)	-	15.09***	-	26.82***
Sargan test (F statistic)	-	1.45	-	1.14
RTS	1.26	2.44	1.17	2.20
TE (50th percentile)	0.70	0.71	0.71	0.75
Observations	199	199	199	199

Note: *, ** and ***Significant at 0.10, 0.05 and 0.01 levels, respectively
Source: Own composition

Table 4: Stochastic Production Frontier estimates (Second Stage) for large-sized farms.

Dependent Variable: (Coffee Production)	Translog without endogeneity corrected	Translog with endogeneity corrected	Cobb Douglas without endogeneity corrected	Cobb Douglas with endogeneity corrected
Explanatory Variables	**Coefficients (β)**	**Coefficients (β*)**	**Coefficients (β)**	**Coefficients (β*)**
Intercept	7.3622***	8.5448***	4.2502***	5.7954***
Land	0.2053	0.0035	0.2628***	0.7685***
Labour	-0.4440	-0.5065	0.6300***	0.2026**
Chemicals	-0.1954	-0.2082	0.0799**	-0.0210
Machinery	0.4661	0.3550	-0.0068	0.2878**
Land^2	-0.3158	-0.2698	-	-
Labour ^2	0.0611	0.0280	-	-
Chemicals^2	-0.0594	-0.1003	-	-
Machinery^2	-0.0118	-0.0611	-	-
Land*Labour	0.2351*	0.2633	-	-
Land*Chemicals	0.0560	0.0325	-	-
Land*Machinery	0.1685	-0.0305	-	-
Labour*Chemicals	0.0122	0.0431	-	-
Labour*Machinery	-0.1023	0.0773	-	-
Chemicals*Machinery	-0.1184**	-0.1313**	-	-
Residual first stage land	-	-0.4821	-	-0.6583***
Residual first stage labour	-	0.6750***	-	0.6954***
Residual first stage chemicals	-	0.0624	-	0.0797
Residual first stage machinery	-	-0.3359**	-	-0.3273***
Natural logarithm of v_i	-2.1385***	-3.0192	-2.412315***	-3.6032
Natural logarithm of u_i	-9.1123	-3.2178	-1.858953***	-1.7887
AIC	99.77	49.26	99.12	54.97
Wald test (chi-square)	519.94***	953.60***	409.40***	537.59***
LR test of sigma_u=0 (chi-square)	0	0.12	1.24	3.34***
Hausman test for endogeneity (chi-square)	-	35.37***	-	71.76***
Sargan test (F statistic)	-	0.27	-	1.79
RTS	0.90	0.99	0.97	1.24
TE (50th percentile)	0.99	0.86	0.77	0.77
Observations	94	94	94	94

Note: *, ** and ***Significant at 0.10, 0.05 and 0.01 levels, respectively.
Source: Own composition

Table 5 contains the estimated elasticities and RTS defined in equations (7) and (8), for small-, medium- and large-sized farms. The RTS obtained from the SPF without endogeneity correction is underestimated in every situation. Focusing on the estimates with endogeneity correction, Table 5 shows that small- and medium-sized farms are subject to increasing RTS. It is also noteworthy that land is by far the most important input, especially in small- and medium-sized farms, whereas labour is especially important in large farms.

Implications for Colombia's Agricultural and Land Policies

In general, agricultural policies in post-conflict situations prioritise improvements in productivity and competitiveness with the aim of increasing the incomes of households whose livelihoods come from agriculture and guaranteeing food production (Adam-Bradford *et al.*, 2020; Jimenez *et al.*, 2021). This is precisely why it is pertinent to analyse in detail what effect the land distribution measures proposed in the 2016 peace accord in Colombia could have on the strategic sector of coffee production in terms of productivity. The results shown in the previous section suggest that small and medium coffee farmers in Colombia are technically inefficient in their production process and moreover, these production units exhibit increasing returns to scale. The challenge for agricultural and land policies is therefore to increase the scale of these farms in a way that does not conflict with another major objective of the proposed reform, which is to establish a more equitable distribution of land in rural areas (Faguet *et al.*, 2017).

It is beyond the scope of this article to carry out a detailed study of the direct and indirect effects of the different ways in which land reform can be implemented in Colombia. However, it does seem pertinent to comment that some concrete proposals in the literature and in the peace agreement itself, such as the formalisation of communal property regimes in rural settings, can make it possible to reconcile the objective of expanding access to land ownership with ensuring that the scale of farms is not sub-optimal.

Another important challenge is to enable the largest farms to improve their productivity through better access to labour. In fact, some reports attribute a reduction in factor endowments to the decline in coffee output that the country has suffered in the past decades (Saenz *et al.*, 2021). With a large mass of potential workers fleeing conflict zones, the wages of the remaining rural workers rose, leading to higher costs for coffee producers (World Bank, 2002). In addition, rural labour shortages have complicated the control of crop pests and the harvesting of the crop at the optimal time (Ocampo-Lopez and Alvarez-Herrera, 2017).

The resolution of the armed conflict may alleviate to some degree the depopulation of these rural areas and reduce some labour supply tensions. However, there are many more issues that need to be resolved in order to improve labour productivity indicators, which is the way in which the economic performance of every farm, but primarily the large plantations, can be improved. There are several studies promoted by companies and associations in the coffee

Table 5: Production elasticities and RTS.

SMALL-SIZED FARMS	Translog with endogeneity corrected	Translog without endogeneity corrected	Cobb Douglas with endogeneity corrected	Cobb Douglas without endogeneity corrected
Output elasticity of land	1.59***	0.758	1.76***	0.62***
Output elasticity of labour	0.67***	0.64***	0.52***	0.63**
Output elasticity of chemicals	0.85***	0.20	0.73***	0.18***
Output elasticity of machinery	-0.15	0.313	-0.34	-0.01
RTS	2.95***	1.90**	2.67***	1.42***
MEDIUM-SIZED FARMS	**Translog with endogeneity corrected**	**Translog without endogeneity corrected**	**Cobb-Douglas with endogeneity corrected**	**Cobb-Douglas without endogeneity corrected**
Output elasticity of land	2.20***-	0.58-	1.78***	0.54***
Output elasticity of labour	0.06-	0.46**	0.11	0.51***
Output elasticity of chemicals	-0.40-	0.05	-0.28**	0.07*
Output elasticity of machinery	0.57**-	0.15	0.59***	0.06
RTS	2.43***-	1.24***-	2.20***	1.17***
LARGE-SIZED FARMS	**Translog with endogeneity corrected**	**Translog without endogeneity corrected**	**Cobb-Douglas with endogeneity corrected**	**Cobb-Douglas without endogeneity corrected**
Output elasticity of land	0.62**	0.27	0.26***	0.77***
Output elasticity of labour	0.24	0.63***	0.63***	0.20**
Output elasticity of chemicals	-0.05	0.02	0.08**	-0.02
Output elasticity of machinery	0.24	0.04	-0.01	0.29**
RTS	1.05**	0.95**	0.97***	1.24***

Source: Own composition

sector that point to another series of factors as determinants to address the shortage and low productivity of labour in the Colombian coffee sector (Rocha, 2014).

In order to increase labour productivity on all types of plantations, but especially the larger ones, it is necessary to implement the following actions: (1) accompanying policies to hire more salaried workers with the formalisation of contracts that are more in line with labour regulations, (2) develop strategies to improve competitiveness in international markets that allow for wage improvements, and (3) offer training programmes to encourage specialisation among workers in the sector and prevent them from having to combine their activity with other complementary activities[5], without any signs of considering coffee growing as a long-term activity.

Conclusions

Two main contributions have been made in this work. On the one hand, the analysis of returns to scale, elasticities and technical efficiency previously carried out by other authors has been refined, correcting for endogeneity biases through a two-stage process to estimate the stochastic production frontier, in line with the proposal of Kutlu (2010) and Amsler *et al*. (2016). The correction for endogeneity is crucial, as it substantially conditions the conclusions of the analysis. We show that small and medium coffee farmers in Colombia are technically inefficient in their production process. In addition, these production units exhibit increasing returns to scale. Besides, large coffee farmers are close to being technically efficient and exhibit decreasing returns to scale. The corrected-for-endogeneity results also indicate that the input intensity that small and medium-sized units must prioritise in their agricultural activity is primarily the land factor, whereas large farms should concentrate their efforts on increasing the labour factor.

On the other hand, in this paper we try to translate these empirical results into agricultural and land policy recommendations in a context as special as the current one, where peace talks revolve around proposals to facilitate access to agricultural land for the poorest peasants in violence-affected areas.

We are aware that there are many aspects and challenges affecting the coffee sector in Colombia that are not addressed in this analysis and that could be analysed in future extensions of this paper. To the productivity analysis in this article should be added an analysis of competitiveness in international markets, as some of the aforementioned challenges relate to the need to attract investment from international suppliers, to accommodate the rapid expansion of coffee farms in low-income areas that have largely remained remote and isolated from international markets, as well as to cope with coffee's high dependence on foreign exchange rates.

[5] It should be borne in mind that many small farmers are in fact usually part farmers, part workers. The income of small farmers is based partly on the sale of crops and livestock, and partly on wage employment, whether on a farm or plantation or in some other rural occupation. Therefore, a sustainable development strategy for the coffee sector must also take into account, as a component, the wages of workers in coffee plantations.

Acknowledgements

Jorge Perdomo expresses special thanks of gratitude to Professor Darrell L. Hueth. Josu Arteche acknowledges financial support from the Spanish Ministry of Science and Innovation grant PID2019-105183GB-I00 and UPV/EHU Econometrics Research Group (Basque Government grant IT1359-19). Alberto Ansuategi acknowledges financial support from the Spanish Ministry of Economy and Competitiveness grant RTI2018-093352-B-100, Basque Government grant IT-353-19 and UPV/EHU grant GIU18/136.

References

Adamopoulos, T. and Restuccia, D. (2014): The Size Distribution of Farms and International Productivity Differences. American Economic Review, **104** (6), 1667–1697. https://doi.org/10.1257/aer.104.6.1667

Adam-Bradford, A., Elkahlout, G., Byrne, R., Wright, J. and Rahman, M. (2020): Stabilisation agriculture: reviewing an emerging concept with case studies from Afghanistan and Iraq. CABI Reviews, **15,** 42. https://doi.org/10.1079/PAVSNNR202015042

Akram-Lodhi, A.H. (2005): Vietnam's Agriculture: Processes of Rich Peasant Accumulation and Mechanisms of Social Differentiation. Journal of Agrarian Change, **5** (1), 73–116. https://doi.org/10.1111/j.1471-0366.2004.00095.x.

Alvarez, A. and Arias, C. (2004): Technical efficiency and farm size: a conditional analysis. Agricultural Economics, **30** (3), 241–250. https://doi.org/10.1016/j.agecon.2003.04.001

Amsler, C., Prokhorov, A. and Schmidt, P. (2016): Endogeneity in stochastic frontier models. Journal of Econometrics, **190** (2), 280–288. https://doi.org/10.1016/j.jeconom.2015.06.013

Barrett, C.B. (1996): On price risk and the inverse farm size–productivity relationship. Journal of Development Economics, **51** (2), 193–215. https://doi.org/10.1016/S0304-3878(96)00412-9

Battese, G.E. and Coelli, T.J. (1992): Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India. Journal of Productivity Analysis, **3,** 153–169. https://doi.org/10.1007/BF00158774

Battese, G.E. and Broca, S.S. (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,** 395–414. https://doi.org/10.1023/A:1007736025686

Berry, A. and Cline, W. (1979): Agrarian Structure and Productivity in Developing Countries. Baltimore: Johns Hopkins University.

Besley, T. and Burgess, R. (2000): Land Reform, Poverty Reduction, and Growth: Evidence from India. The Quarterly Journal of Economics, **115** (2), 389–430. https://doi.org/10.1162/003355300554809

Bhalla, S.S. and Roy, P. (1988): Mis-specification in Farm Productivity Analysis: The Role of Land Quality. Oxford Economic Papers, **40** (1), 55–73. https://doi.org/10.1093/oxfordjournals.oep.a041846

Binam, J.N., Sylla, K., Diarra, I. and Nyambi, G. (2003): Factors Affecting Technical Efficiency among Coffee Farmers in Côte d'Ivoire: Evidence from the Centre West Region. African Development Review, **15** (1), 66–76. https://doi.org/10.1111/1467-8268.00063

Carter, M.R. (1984): Identification of the Inverse Relationship between Farm Size and Productivity: An Analysis of Peasant Agricultural Production. Oxford Economic Papers, **36** (1), 131–145. https://doi.org/10.1093/oxfordjournals.oep.a041621

Faguet, J.-P., Sánchez Torres, F. and Villaveces, M.J. (2017): The paradox of land reform, inequality and development in Colombia. Working Paper Series, No. 17-181, London School of Economics and Political Science (LSE), Department of International Development, London.

Garcia, A.F. and Shively, G.E. (2011): How Might Shadow Price Restrictions Reduce Technical Efficiency? Evidence from a Restricted DEA Analysis of Coffee Farms in Vietnam. Journal of Agricultural Economics, **62** (1), 47–58. https://doi.org/10.1111/j.1477-9552.2010.00269.x

Giovannucci, D., Leibovich, J., Pizano, D., Paredes, G., Montenegro, S., Arévalo, H. and Varangis, P. (2002): Colombia: coffee sector study. Working Paper, Universidad de los Andes, Facultad de Economía, Documento CEDE 2002-15.

Greene, W. (2012): Econometric Analysis. Seventh Edition, Essex: Pearson

Heltberg, R. (1998): Rural market imperfections and the farm size - productivity relationship: Evidence from Pakistan. World Development, **26** (10), 1807–1826. https://doi.org/10.1016/S0305-750X(98)00084-9

ICO (2021): Total Production by all exporting countries. Retrieved from: https://ico.org/historical/1990%20onwards/PDF/1a-total-production.pdf (Accessed on 12 May 2022).

Jimenez, D., Saldarriaga-Isaza, A. and Cicowiez, M. (2021): Distributional and economy-wide effects of post-conflict agricultural policy in Colombia. European Review of Agricultural Economics, **49** (3), 644–667. https://doi.org/10.1093/erae/jbab020

Johnson, D. (1998): Applied Multivariate Methods for Data Analysis. Duxbury Press, University of Michigan

Kimhi, A. (2006): Plot size and maize productivity in Zambia: is there an inverse relationship? Agricultural Economics, **35** (1), 1–9. https://doi.org/10.1111/j.1574-0862.2006.00133.x

Kutlu, L. (2010): Battese-coelli estimator with endogenous regressors. Economics Letters, **109** (2), 79–81. https://doi.org/10.1016/j.econlet.2010.08.008

Lamb, R.L. (2003): Fertilizer Use, Risk and Off-Farm Labor Markets in the Semi-Arid Tropics of India. American Journal of Agricultural Economics, **85** (2), 359–371. https://doi.org/10.1111/1467-8276.00125

Mosheim, R. (2002): Organizational Type and Efficiency in the Costa Rican Coffee Processing Sector. Journal of Comparative Economics, **30** (2), 296–316. https://doi.org/10.1006/jcec.2002.1773

Nchare, A. (2007): Analysis of Factors Affecting Technical Efficiency of Arabica Coffee Producers in Cameroon. African Economic Research Consortium, Research Paper 163. Nairobi, Kenya.

Ocampo Lopez, O.L. and Alvarez-Herrera, L.M. (2017): Tendencia de la producción y el consumo del café en Colombia. Apuntes del Cenes, **36** (64), 139–165. https://doi.org/10.19053/01203053.v36.n64.2017.5419

OECD (2015): OECD Review of Agricultural Policies: Colombia 2015. OECD Review of Agricultural Policies, OECD Publishing, Paris, France.

Perdomo, J., Gamboa, L.F., Perea, D., Rodríguez, E.F. and Guzmán, S.G. (2016): Avances sociales en Bogotá y municipios aledaños entre 2011 y 2014: calidad de vida, segregación, capacidad de pago y focalización. Ediciones Secretaría Distrital de Planeación, Bogotá, Colombia.

Perdomo, J. and Hueth, D. (2011): Funciones de Producción, Análisis de Economías a Escala y Eficiencia Técnica en el Eje Cafetero Colombiano: Una Aproximación con Frontera Estocástica. Revista Colombiana de Estadística, **34** (2), 377–402.

Perdomo, J.A. and Mendieta, J.C. (2007): Factores que Afectan la Eficiencia Técnica y Asignativa en el Sector Cafetero Colombiano: una Aplicación con Análisis Envolvente de Datos. Revista Desarrollo y Sociedad, **60,** 1–46. https://doi.org/10.13043/dys.60.1

Rios, A.R. and Shively, G.E. (2006): Farm Size, Irrigation Infrastructure, and the Efficiency of Coffee Production in Vietnam. Forests, Trees, and Livelihoods, **16** (4), 397–412. https://doi.org/10.1080/14728028.2006.9752575

Rocha, R. (2014): Informalidad laboral cafetera: rasgos, determinantes y propuestas de política. Archivos de Economía, Documento 418. Departamento Nacional de Planeación (DNP), Colombia.

Rosales, R., Perdomo, J., Morales, C. and Urrego, J. (2013): Fundamentos de econometría intermedia: teoría y aplicaciones. Bogotá: Ediciones Uniandes.

Saenz, M., Alvarez, D. and Brock, G. (2021): Lessons from long-run (1975–2017) structural change in Colombia's coffee production. Agricultural and Resource Economics Review, **50** (2), 201–225. https://doi.org/10.1017/age.2021.1

Sen, A. (1962): An Aspect of Indian Agriculture. The Economic Weekly, **14,** 243–246.

Shee, A. and Stefanou, S.E. (2015): Endogeneity Corrected Stochastic Production Frontier and Technical Efficiency. American Journal of Agricultural Economics, **97** (3), 939–952. https://doi.org/10.1093/ajae/aau083

Tran, K.C. and Tsionas, E.G. (2013): GMM Estimation of Stochastic Frontier Models with Endogenous Regressors. Economic Letters, **118** (1), 233–236. https://doi.org/10.1016/j.econlet.2012.10.028

Vedenov, D., Houston, J. and Cardenas, G. (2007): Production Efficiency and Diversification in Mexican Coffee-Producing Districts. Journal of Agricultural and Applied Economics, **39** (3), 749–763. https://doi.org/10.1017/S1074070800023397

World Bank (2002): Colombia: Coffee Sector Study 2002. Washington, D.C, USA.

Appendix

Appendix 1: SUR model first stage results.

Dependent variable: LN(chemicals)	Farm size			Dependent variable: LN(machinery)	Farm size		
	Small	Medium	Large		Small	Medium	Large
Explanatory Variables	**Coefficients**	**Coefficients**	**Coefficients**	**Explanatory Variables**	**Coefficients**	**Coefficients**	**Coefficients**
Intercept	-2.6856***	2.4852	11.4980**	Intercept	-1.8679***	-5.0786**	-7.2970*
LN (price land)	0.0816***	0.2327***	0.3093**	LN (price land)	0.0059	0.1334*	0.2799**
LN(price labour)	0.1029	0.0512	-0.4464*	LN(price labour)	0.0054	0.1107	-0.0729
LN(price chemicals)	0.0734	-0.5258***	-0.5917**	LN(price chemicals)	-0.0335	-0.1401	-0.0310
LN(price machinery)	-0.1247	-0.0618	-0.0409	LN(price machinery)	-0.3825***	-0.1098	-0.3022**
Specialisation (dummy variable, yes=1 and no=0)	0.1606**	0.0270	-0.1620	Specialisation (dummy variable, yes=1 and no=0)	-0.0687**	0.3221*	0.3203
Road access (dummy variable, yes=1 and no=0)	0.1759***	0.0132	-1.4296*	Road access (dummy variable, yes=1 and no=0)	0.0130	0.4808***	0.6713
Diversification (dummy variable, yes=1 and no=0)	0.0315	0.2114	0.1909	Diversification (dummy variable, yes=1 and no=0)	-0.0689**	0.3780**	0.1635
LN(family size)	-	-	-0.3102	LN(family size)	-	-	0.2463
Global fit (chi-square)	68.09***	36.91***	20.15***	Global fit (chi-square)	126.29***	35.94***	14.90*
Observations	551	200	94	Observations	551	200	94

Dependent variable: LN(land)	Farm size			Dependent variable : LN(labour)	Farm size		
	Small	Medium	Large		Small	Medium	Large
Explanatory Variables	**Coefficients**	**Coefficients**	**Coefficients**	**Explanatory Variables**	**Coefficients**	**Coefficients**	**Coefficients**
Intercept	-0.4910	0.7622	2.0774	Intercept	2.0310**	5.0348	12.6031
LN (price land)	0.0380*	0.0401	0.1838**	LN (price land)	0.1366***	0.0748	0.2808***
LN(price labour)	0.0579***	-0.0188	-0.1025	LN(price labour)	-0.0167	-0.3373***	-0.9355***
LN(price chemicals)	-0.0602	-0.0022	-0.1047	LN(price chemicals)	-0.2160***	0.0561	-0.1737
LN(price machinery)	-0.0558	0.0221	-0.1557*	LN(price machinery)	-0.1147	-0.0185	-0.1501
Specialisation (dummy variable, yes=1 and no=0)	-0.0160	-0.0633	0.2205	Specialisation (dummy variable, yes=1 and no=0)	-0.1649*	-0.1929	0.6042***
Road access (dummy variable, yes=1 and no=0)	0.0282	-0.0539	-0.7360*	Road access (dummy variable, yes=1 and no=0)	0.0664	0.3056**	-0.1615
Diversification (dummy variable, yes=1 and no=0)	-0.0574	0.1249*	0.2022	Diversification (dummy variable, yes=1 and no=0)	-0.1237*	0.2988**	0.4497**
LN(family size)	-	-	-0.2391**	LN(family size)	-	-	-0.2370
Global fit (chi-square)	26.41***	9.27	21.81***	Global fit (chi-square)	32.78***	47.64***	66.36***
Observations	551	200	94	Observations	551	200	94

Note: *, ** and ***Significant at 0.10, 0.05 and 0.01 levels, respectively.
Source: Own composition