Plot-level technical efficiency accounting for farm-level effects: Evidence from Chilean wine grape producers

Abstract

This paper extends previous work focusing on the analysis of technical efficiency (TE) for a sample of Chilean boutique wine grape producers. We apply alternative panel data models to a cross-sectional dataset that contains multi-plot observations for individual farms, where these plots have separate supervision according to specific requirements. Treating individual plots as being independent of each other reveals technical inefficiency differences across plots. However, when it is recognized that plots belonging to a particular farm are subject to an overall central (farm-level) management and farm-level unobserved heterogeneity is controlled for, no differences in intra-farm (i.e., plot level) inefficiencies are found. A Generalized True Random Effects model, which permits the separate identification of farm-level and plot-level inefficiencies while controlling for unobserved farm-level heterogeneity, shows that inefficiency differences exist at the farm level but not among plots within the same farm. This points to the importance of accounting for unobserved farm-level heterogeneity and farm-level inefficiency when data for individual plots within a farm unit is available. Geographical location was also found to matter for grape production. Moreover, agro-climatic conditions were found to influence production levels, with grape farms located on cooler zones producing significantly less than their counterparts in warmer zones, as expected.

JEL Classifications: C23; C52; D22; L23; M11; Q12

Key Words: Stochastic Production Frontiers, Cross Sectional Data, Panel Data Models, Technical Efficiency, Generalized True Random Effects, Wine Grapes, Chile.

Plot-level technical efficiency accounting for farm-level effects:
Evidence from Chilean wine grape producers

1. Introduction

Global production of wine has changed considerably in recent decades, where a notable development has been a significant rise in the so-called New World countries (Anderson, 2005; Anderson and Nelgen, 2011). Chile is one of these relative newcomers to the international wine market and has experienced rapid growth in land area planted with grapes (ODEPA, 2016). Such expansions have also occurred in several other countries including Australia, South Africa, New Zealand, China and the USA (Strohm, Dirksmeyer and Garming, 2014), and the overall effect has been increasing competition and all-out efforts to retain or expand market share. In 2017, most of the increase in exported wine went to China and Japan (Banfi Piazza, 2017).

The evolution of the Chilean wine industry over the past two decades has been driven chiefly by the rapid growth in international markets, and more than 60% of Chilean wine is exported (Banfi Piazza, 2017). The Chilean model from 2000 up to 2006 was based on penetrating and developing markets based on good quality wines at a low price. By comparison, over the 2007-2017 period, export earnings have increased faster than the total volume of bottled wine and thus the price per liter has risen (Buzzetti Horta, 2018).

Despite the rapid growth enjoyed by Chilean wines in foreign markets, the reliance on exports makes this sector vulnerable to FOB prices and exchange rate volatility, which are factors totally outside the control of producers. Therefore, productivity growth along the wine value chain is fundamental for the commercial success of this industry. A particularly vulnerable link in this value chain is growers that sell their highly perishable grapes to much larger wine makers. These growers have to focus carefully on minimizing costs while ensuring that their output satisfies the quality standards imposed by wine makers who in turn need to be responsive to the demands coming from international markets. Moreover, estimates indicate that grapes represent the most significant component in the cost structure of wine making in the Chilean industry (Arévalo and Martínez, 2006). Consequently, any efficiency gains at the vineyard level can have a significant bearing on the commercial success of the whole operation.

Chilean wine production comprises two primary groups of firms, namely large companies and family-owned estates, and both are oriented primarily towards the export market. However, whereas the large companies place emphasis on a product for mass consumption, a substantial number of family-owned estates, commonly referred to as *boutique* vineyards, are focused on a reduced-scale production of higher quality wines. Both types of firms have wine cellars vertically integrated to wine grape plantations, the large companies possessing larger areas than the smaller family-owned estates. In 1994, 12 of these *boutique* vineyards formed the Tecnovid consortium, a number that grew over the following years to 42 associates, which in 2006 accounted for 9% of all land area planted in the industry and 12% of total wine exports. Moreover, the farmers are located in all the main valleys where grapes are produced in Chile. Jano (2017) makes an interesting distinction between "quasi-subsistence" wine-grape farmers and "entrepreneurial" farmers in Chile. The former are low-income farmers that complement their subsistence income with earnings from wine-grape sales and/or rustic wine production, whereas the "entrepreneurial" farmers produce classic varieties that have the potential to produce high-quality wines. The producers associated with Tecnovid belong to the latter type of farmer.

In this paper, we use a unique cross-sectional data set to analyze the technical efficiency (TE) component of productivity for a sample of Chilean wine grape producers that belong to Tecnovid. Our work makes two main contributions. First, we add new evidence to the scant literature on the productivity of grape farming by presenting an analysis of TE for the boutique segment of Chilean wine-grape producers, a New World country that has acquired international standing in wine markets. Second, we make a methodological contribution by applying and comparing the performance of various panel data stochastic frontier models to analyze plot-level (i.e., sub-units within a given farm) and farm level TE while accounting for unobserved farm heterogeneity and several observed attributes. A plot is a well-defined land area with distinct technology, management, and varying levels of inputs chosen to achieve an expected quality of grapes. In other words, our data is a cross section of farms where each farm has several plots (data points) and we exploit this structure in our application of panel data models to provide a novel analysis of TE. Each plot produces a unique quantity and type of grape, with different levels of inputs and possesses specific observable characteristics. Hence, we identify the TE of each plot while also separately identifying, in our most refined model, farm level TE and unobserved heterogeneity. A way to think about this is that each farm has a central administration that provides management to the overall farm unit while individual plots are subject to separate management. Therefore, if the analysis supports the presence of unobserved farm-level heterogeneity but this is not identified or controlled for, then the individual plot TE levels would be biased. Our results do indicate that unobserved heterogeneity matters, so specifications that ignore such heterogeneity lead to biased results.

The paper proceeds as follows: Section 2 presents a brief review of the few published papers we have found that focus on grape farm TE. Section 3 presents the methodological framework, while Section 4 describes the data used. The results are presented in Section 5, and Section 6 offers some concluding remarks.

2. Related literature

Many studies have been published examining the TE component of productivity in farming (Bravo-Ureta, Jara-Rojas, Lachaud, Moreira, Scheierling and Treguer, 2016; Bravo-Ureta, Solís, Moreira, Maripani, Thiam and Rivas, 2007; Ogundari, 2014), but very few focus on wine grape production. Table 1 provides key features of 10 wine grape production studies. We trace this literature back to Townsend, Kirsten and Vink (1998), who analyzed the relationship between farm size, productivity and returns to scale for wine grape producers located in four regions of South Africa for the years 1992 to 1995, using a non-parametric approach. Another South African study, by Conradie, Cookson and Thirtle (2006), examined the relationship between TE and farm size for samples of producers situated in Western Cape Province. These authors estimated stochastic production frontier (SPF) models using panel data for wine grape farms located in the Robertson and Worcester regions for the years 2003 and 2004, and cross-sectional data for table grape farms located in De Doorns region for 2004.

Henriques, Carvalho and Fragoso (2009) used non-parametric techniques to measure TE for a sample of 22 wine grape farms from the Alentejo region of Portugal for the years 2001 and 2004. Guesmi, Serra, Kallas and Gil (2012) focused on the TE of organic and conventional grape farms in Catalonia, Spain, using an SPF along with cross sectional data for a sample of 141 farms.

Moreira, Troncoso and Bravo-Ureta (2011) examined the TE of wine grape production for a sample of Chilean firms for 2005/2006 using a standard cross sectional model. A Cobb-Douglas SPF was estimated using data for 38 farms for which input-output information is available at the plot level. Ma, Wu, Feng and Jiao (2012) use 1,020 farm level observations collected across 24 grape producing provinces in China to estimate a Cobb-Douglas SPF model. Coelli and Sanders (2013) used an unbalanced panel data set (2006/2007 to 2009/2010) for a sample of 135 farmers specializing in wine grape production located in the Murray and Murrumbidgee river basins in Australia. The authors used the translog functional form to fit SPF models based on the Battese and Coelli (1992) approach.

More recently, Manevska-Tasevska (2013) used a three-year (2006-2008) panel data set for a sample of 300 commercial grape producers from Macedonia, employing a Cobb-Douglas SPF model and a second-stage regression to analyze TE. Latruffe and Nauges (2014) employed panel data for the period 1999-2007 for French crop farms, including grape producers, to examine whether TE using conventional practices has an effect on the likelihood of converting to organic farming, using both a Cobb-Douglas SPF as well as nonparametric frontier methods. Finally, Piesse, Conradie, Thirtle and Vink (2017) estimated a translog production frontier to examine TE for a panel of 77 wine grape farms in South Africa observed between 2005 and 2015 and compared the efficiency levels of old established wine regions with newer regions.

From our review of studies that have examined the TE of wine grape producers, we can conclude that the literature so far has covered different continents and has used both non-parametric (3 cases) and parametric techniques (8 cases) and availed of both cross-section and panel data. The 10 wine grape production studies summarized in Table 1 reported an overall average TE of 67.5%. The limited available evidence in the literature points towards ample scope for improvements in TE amongst producers. Of particular relevance to the work presented in this paper is the fact that all the articles that we found focus on farm-level TE without considering possible intra-farm or plot-level variability and ignore unobserved heterogeneity even when panel data is employed.

3. Methodological framework

As discussed in Section 2, in this paper we use a unique dataset that contains plot level data for a sample of grape farms, which enables the identification of TE for each planted plot separately from unobserved heterogeneity and TE at the farm level. In a general context, Greene (2005a, 2005b) argues that it is important to disentangle time-variant TE from time-invariant heterogeneity because the true "…underlying production function might contain unmeasured firm-specific characteristics that reflect the technology in use, not inefficiency [in which case] the model estimated … is actually incomplete or mis-specified…" (Greene, 2005a, p. 270). More recently, a distinction has been made, as discussed in more detail below, between time-invariant unobserved heterogeneity and time-invariant TE (Colombi, 2010; Filippini and Greene, 2016; Tsionas and Kumbhakar, 2014). Considering the cross-sectional structure of our data set, the notion of "time-invariant" in the preceding discussion needs to be replaced by 'plot-invariant'.

As stated above, the plots in our data have unique output, inputs and other observable attributes to be discussed later. A way to think about the plot-invariant heterogeneity is that farm level management provides overall coordination across all plots, while there is also plot level supervision according to specific requirements. Another source of unobserved heterogeneity might stem from microclimatic and related agro-ecological variability across farms, and these are captured by the valley where the farms are located and by agro-ecological characteristics based on degree days.

In order to identify the plot-level TE separately from the farm-level unobserved heterogeneity we make use of the True Fixed Effects and True Random Effects (TFE and TRE) models introduced by Greene (2005a, 2005b). Despite the added flexibility offered by the TFE and TRE models, there appears to be limited applications in the agricultural economics literature, with Abdulai and Tietje (2007), Carroll, Newman and Thorne (2011) and Qi, Bravo-Ureta and Cabrera (2015) representing exceptions. To further identify farm-level TE from unobserved heterogeneity, we apply the Generalized True Random Effects (GTRE) model (Colombi, 2010; Filippini and Greene, 2016; Tsionas and Kumbhakar, 2014). Applications of the GTRE model in agriculture are few in number and have relied on conventional panel data sets. Examples are Yang (2014), Njuki and Bravo-Ureta (2015), and Lachaud, Bravo-Ureta and Ludeña (2015).

We now present each alternative frontier model used in the analysis below. These mirror a panel data structure, where plots, instead of time periods, can be assigned to farms. Thus, the alternative frontier models we consider below are variants of the following equation:

$$y_{jp} = \beta_j + f\left(x_{jp}; \delta\right) + v_{jp} - u_{jp} \tag{1}$$

where:

y_{jp} is the output quantity of the j^{th} farm from the p^{th} plot; $x_{jp} = (x_{1jp}, x_{2jp}, \ldots, x_{Njp})$ is the vector of all input quantities, where N is the number of inputs; δ is a vector of parameters; β_j represents the individual farm effects[1]; v_{jp} is the idiosyncratic error term with an expectation of zero; and $u_{jp} \geq 0$ is a one-sided error term capturing technical inefficiency.

We begin by estimating a **Pooled Stochastic Production Frontier (Pooled SF)**. This is the traditional composed error stochastic production frontier, as originally proposed by Aigner, Lovell and Schmidt (1977), where each observation (a plot in this study) is basically treated as independent from the others. This model can be written as:

$$y_{jp} = \beta_0 + f(x_{jp}; \delta) + v_{jp} - u_{jp} \tag{2}$$

where v_{jp} is the symmetric random error term which we assume normally distributed with mean zero, and $u_{jp} \geq 0$ is a one-sided error term capturing technical inefficiency assumed to follow a half-normal distribution, i.e., $u_{jp} \sim N^{+}(0, \sigma_u^2)$. In this traditional stochastic frontier (SF) model, and subsequent SF models, the TE component (TE_{jp}) is calculated applying the formula developed by Jondrow, Lovell, Materov and Schmidt (1982) and can be represented as $TE_{jp} = exp(-\hat{u}_{jp})$.

Turning now to panel data frontiers, the next two models are fixed effects specifications. The first is the **Fixed Effects Frontier (FE)**. This is a reinterpretation of the standard linear panel data fixed effects model proposed by Schmidt and Sickles (1984):

$$y_{jp} = \beta_0 + f(x_{jp}; \delta) + v_{jp} - u_j \tag{3}$$

[1] The individual farm effects, β_j, may be a constant intercept or vary across individual farms depending on the model under consideration. For example, in a non-frontier OLS model, β_j would be a common constant, while in a fixed effects models it would be a set of dummy variables for each farm, and so on.

where $u_j = max(\beta_j) - \beta_j \geq 0$. Thus, a farm's inefficiency, which is invariant across plots, is calculated by the difference between the largest estimated fixed effect for the whole sample and the fixed effect of the farm (Coelli, Rao, O'Donnell and Battese, 2005; Cornwell, Schmidt and Sickles, 1990; Mundlak, 1961; Schmidt and Sickles, 1984). As such, this is a 'deterministic' rather than 'stochastic' frontier as inefficiency is calculated from a linear fixed effects model. Note that all unobserved plot-invariant heterogeneity in this formulation is attributed to inefficiency.

The second fixed effects specification is the **True Fixed Effects (TFE) model**. The TFE stochastic frontier model was introduced by Greene (2005a), and in our context allows for the measurement of plot-variant TE through the term u_{jp} while also accounting for unobserved farm-level heterogeneity. This model can be expressed as:

$$y_{jp} = \beta_j + f(x_{jp}; \delta) + v_{jp} - u_{jp} \tag{4}$$

where $v_{jp} \sim N(0, \sigma_v^2)$ and $u_{jp} \sim N^+(0, \sigma_u^2)$. Cross-farm heterogeneity is captured by the vector β_j, which represents farm dummy variables, and inefficiency is allowed to vary across plots.

The FE and TFE specifications have the advantage of allowing for correlation between individual farm-specific effects and the explanatory variables. However, a potential drawback of the FE and TFE models is that they do not permit the inclusion of time/space-invariant variables (or, in our context, plot-invariant variables) which may be of particular interest. In our case, for example, key variables representing possible spatial heterogeneity which we want to account for are the agro-climatic zone and the valleys where the vineyards are located. These plot-invariant variables can be included in random effects frontier specifications, although these assume no correlation between the individual effects and the other regressors.

Our first random effects frontier is the **True Random Effects (TRE)** model. The TRE model, also introduced by Greene (2005a), is basically a random-effects counterpart to the fixed effects frontier, and is obtained by combining a conventional random-effects model with a one-sided stochastic term representing inefficiency. We can write this model as:

$$y_{jp} = \beta_j^* + f\left(x_{jp}; \delta\right) + v_{jp} - u_{jp} \tag{5}$$

where $v_{jp} \sim N(0, \sigma_v^2)$, $u_{jp} \sim N^+(0, \sigma_u^2)$, and the farm-specific individual effect is distributed as $\beta_j^* \sim N\left(0, \sigma_{\beta*}^2\right)$ which is a farm-specific random term that captures cross-farm heterogeneity, and should be uncorrelated with all other terms in the model. Unlike the TFE model, the TRE model has the advantage of allowing for the incorporation of farm-invariant regressors, a matter of interest in this paper. However, since unobserved factors may be correlated with some of the explanatory variables, the estimates of the production frontier coefficients may be biased (Abdulai and Tietje, 2007).

According to Mundlak (1978), the random effects specification is a mis-specified version of the FE (within) model since it ignores the possible correlation between individual effects and regressors (Debarsy, 2012). Mundlak (1978) proposed an approach to address this problem, which was to include the group-means of the explanatory variables as additional regressors. Farsi, Filippini and Kuenzle (2005) adopted Mundlak's proposal in the context of the true random effects model giving rise to the **Mundlak True Random Effects (MTRE) model**. The MTRE can be written as:

$$y_{jp} = \beta_j^* + \sum_{i=1}^{4} \varphi_i \overline{x_{ij}} + f\left(x_{jp}; \delta\right) + v_{jp} - u_{jp} \tag{6}$$

where $\overline{x_{ij}}$ are the group means of the explanatory variables (plot-level inputs for a given farm in our setting) and everything else is the same as in the TRE model represented in equation (5). An attractive feature of the MTRE model is that it controls for unobserved heterogeneity while permitting relevant plot-invariant variables to be included. As such, the MTRE captures the positive features of the TFE and of the TRE.

The TFE, TRE and MTRE models all account for unobserved farm-specific heterogeneity (β_j), plot-varying inefficiency (u_{jp}), and plot level idiosyncratic error (v_{jp}). Our final model, the **Generalized True Random Effects (GTRE)**, provides further flexibility by also permitting the estimation of farm-specific inefficiency (u_j). The GTRE, which is an extension of the 'true-random' effects model, incorporates an error structure with four parts. In a standard panel data context, the GTRE allows for the separate identification of time-invariant unobserved firm-specific heterogeneity, persistent and transient inefficiency, and statistical errors (Colombi, Kumbhakar, Martini and Vittadini, 2014; Filippini and Greene, 2016; Kumbhakar, Lien and Hardaker, 2014; Tsionas and Kumbhakar, 2014). In our context, the GTRE model can be written as:

$$y_{jp} = \beta_j + f(x_{jp}; \delta) + v_{jp} - u_{jp} - \omega_j \tag{7}$$

where $\beta_j \sim N(0, \sigma_\beta^2)$, $v_{jp} \sim N(0, \sigma_v^2)$, $u_{jp} \sim N^+(0, \sigma_u^2)$, and ω_j is a non-negative farm-specific plot-invariant inefficiency term described by $\omega_j \sim N^+(0, \sigma_\omega^2)$. Thus, this model includes both farm-specific heterogeneity (β_j) and plot level idiosyncratic error (v_{jp}), as well as farm-specific inefficiency (ω_j) and plot-varying inefficiency (u_{jp}). In this model we can think of an individual farm effect, α_j, comprising farm-specific heterogeneity and plot invariant inefficiency $\left(\alpha_j = \beta_j - \omega_j\right)$. We will estimate this model using the maximum simulated likelihood method proposed by Fillipini and Greene (2016).

4. Data

The data used in the study was obtained from 38 Chilean wine grape producers that accepted to fully cooperate and provided all the data required for the study. Considering that the total number of Tecnovid producers when the survey was implemented was 42, the response rate is just over 90% and this is considerably higher than what can be gleaned from the limited related published evidence (Johansson, Effland and Coble, 2017; Meyer, Mok and Sullivan, 2015; Pennings, Irwin and Good, 2002; Weber and Clay, 2013). Unfortunately, four of the 42 farms had to be excluded because they lacked some key variables; however, the available information at that time indicated nothing unusual about these farms compared to the 38 that are in the data set used in the analysis below. So, the 38 farms are considered to be a representative sample of all Tecnovid farms.

The farms in the Tecnovid dataset exported more than 90% of their wine production and were classified as *boutique* wine producers. The data was collected by researchers from the Universidad de Talca in Central Chile and corresponds to the agricultural year 2005-2006. The total number of observations is 263, which is the total number of plots in the 38 farms. According to the descriptive statistics, presented in Table 2, the number of plots per farm goes from a low of two to a high of 17 with an average of seven. The size of the plots ranges from 0.2 ha. to 108.7 ha. The grapes are classified according to quality as Premium or Varietal and the number of plots is equally distributed between each category. Most of the grapes are produced in a simple cordon training system (73%), followed by a double cordon training system (13%).

We also have information on the agro-climatic zone and the valleys where the vineyards are located. All farms are located in Central Chile and distributed, from North to South, among the following 10 valleys: Limarí, Aconcagua, Casablanca, San Antonio, Maipo, Cachapoal, Rapel, Colchagua, Curicó and Maule. A standard agro-climatic criterion, commonly known as "degree-days", is used to classify farms according to the total number of days with temperatures above 10ºC over the period September-February, when the vine is active. We then classify zones as “Cool” when degree-days are less than or equal to 1,200 and “Warm” when greater than 1,200. According to this criterion, five farms come from a “Cool” agro-climatic zone, and the remaining 33 from a “Warm” zone.[2] Note that farms from the same valley can be classified into different agro-climatic zones. The regional location and corresponding agro-climatic zone of each farm are potentially important factors to be considered when analyzing grape production (Fraga et al., 2014; Ponti, Gutierrez, Boggia and Neteler, 2018).

In our econometric models, the dependent variable is *Grape Production* in kilograms. The inputs used are plot size measured in hectares (*Land*), annual expenditure on labor (*Labor*), annual expenditure on agrochemicals and fertilizers (*Chemical*s), and annual operating machinery costs (*Capital*), all measured at the plot level using market prices. Machinery costs include outlays for operating equipment associated with various field activities required in vineyards throughout the

[2] Based on the range of degree-days, viticulturists classify vine growing climatic zones in four categories: (1) < 800 degree-days; (2) 801-1000 degree-days; (3) 1,001-1,200 degree-days; and (4) 1,201-1,400 degree days. 33 of the 38 farms fell into category 4, with categories (1), (2) and (3) having one, one and three farms respectively. To avoid having groups with a single farm, we grouped the farms from categories (1)-(3) together.

growing season (e.g. tractor plus disking, spraying, fumigation, or mowing, etc.). These outlays include both variable and fixed costs. The variable costs are for fuel, lubricants, operator and incidentals. The fixed costs are for depreciation, maintenance, repairs, insurance and the opportunity cost of the capital invested in machinery.

Other explanatory variables include a series of dummy variables taking the value 1 (and 0 otherwise) for vines older than 5 years (*Age Vines*), grape color (*Red*), grape quality (*Premium*), training system (*Single Cordon, Double Cordon, Pergola*), agro-climatic zone (*Cool*), and the valley where the vineyard is located (*Aconcagua/Cachapoal*, *Colchagua/Rapel*, *Casablanca*, *Maipo*, *Curicó, Maule*). Monetary variables are expressed in US$ and the exchange rate used is the average for 2005-2006 at US$1 = Ch$ 542 (Central Bank of Chile, www.bcentral.cl).

5. Econometric specification and results

To implement the models presented in the previous section, we first need to specify an appropriate functional form for the production frontier. Selecting a functional form is an important consideration and the two most popular choices have been the translog and the Cobb-Douglas (Bravo-Ureta, Jara-Rojas, Lachaud, Moreira, Scheierling and Treguer, 2016; Bravo-Ureta, Solís, Moreira, Maripani, Thiam and Rivas, 2007). Another important consideration when relying on production function/frontier models has to do with identification or, to put it differently, with the possible endogeneity of inputs. The long-standing justification for the identification of such production models is "that…entrepreneurs maximize the mathematical expectation of profit" (Zellner, Kmenta and Drèze, 1966, p. 787). This justification was further elaborated by Hodges (1969) and Blair and Lusky (1975), and has been invoked regularly over the years in the agricultural economics literature either explicitly (e.g., Karagiannis and Kellermann, 2019; Picazo-Tadeo and Wall, 2011; Dawson and Lingard, 1982) or implicitly (e.g., Abdul-Rahaman and Abdulai, 2018; Piesse, Conradie, Thirtle and Vink, 2017).

Due to its flexibility, we choose the translog functional form and will test this specification against the Cobb-Douglas. Given the variables available in the dataset, the generic frontier specification presented in equation (1) can be expressed in translog form as:

$$\ln y_{jp} = \beta_j + \sum_{i=1}^{4} \delta_i \ln x_{ijp} + \frac{1}{2}\sum_{i=1}^{4}\sum_{k=1}^{4} \delta_{ij} \ln x_{ijp} \ln x_{kjp} + \sum_{l=1}^{6} \delta_l D_l + \sum_{z=1}^{2} \delta_z D_z + v_{jp} - u_{jp} \tag{8}$$

where, as stated earlier, y_{jp} is the output quantity of the j^{th} farm of the p^{th} plot and x_{ijp} is the quantity of the i^{th} input of the j^{th} farm of the p^{th} plot, v_{jp} is the idiosyncratic error term with an expectation of zero; and β_j and the δ's are parameters to be estimated. Symmetry restrictions are imposed so that $\delta_{ij} = \delta_{ji} \ \forall \ i,j$. Depending on the model, the individual farm effects, β_j, may comprise fixed effects or random effects capturing farm-specific heterogeneity. D_l is a dummy variable for the region in which the farm is located ($l = 1, \dots, 6$)[3], and D_z is a dummy variable to capture the farm's agro-climatic zone and is equal to 1 when the zone is $Cool$ and zero otherwise. In equation (8), the inefficiency term u is depicted as plot-varying. However, depending on the model used, it may be only farm-varying (u_j) or may encompass both plot-varying and farm-specific inefficiency ($u_{jp} - \omega_j$).

[3] The original 10 valleys were grouped into 7 because of the small number of observations for some of the valleys. Limarí is the omitted valley in the estimations.

As is common practice, when estimating the models, the logarithms of the inputs are transformed by subtracting their sample geometric mean (Coelli, Estache, Perelman and Trujillo, 2003). Consequently, the estimated first-order parameters can be interpreted as partial output elasticities for a representative farm characterized by an input endowment equal to the sample geometric mean.

Before presenting the estimates from the stochastic production frontiers, we present findings from non-frontier models, namely the Ordinary Least Squares (OLS)[4], Random Effects (RE) and Fixed Effects (FE) models.[5] For comparison purposes, we exclude the regional and agro-climatic variables as these are plot-invariant and therefore cannot be estimated in a fixed effects model. The results from these translog non-frontier models are presented in Table 3 and at the bottom we include a set of various specification tests. The first test is of the FE model against OLS and the latter is rejected. This shows the importance of unobserved farm-specific heterogeneity in our sample. The next test is a Hausman test of FE versus RE, which reveals that the RE is unsuitable. Thus, the FE model appears as the preferred model of the three. Given this, we test the validity of the translog specification versus the more restrictive Cobb-Douglas specification, and the Cobb-Douglas is strongly rejected. Looking more closely at the estimates in Table 3, it is notable that the first-order coefficients for *Chemicals* and *Land* have negative signs in the OLS and FE models respectively, implying negative partial output elasticities at the sample mean in contradiction to what is expected from economic theory. In the RE model, on the other hand, all

[4] In the context of the frontier literature, the OLS model is often referred to as an *average production function* (King, 1980; Timmer, 1971).

[5] All estimations presented in this paper have been carried out using the software package *LIMDEP* (version 11). The commands used and output are available in an on-line appendix.

first-order input coefficients are positive. To check the effects of input quantities on output, for each input we carried out a Wald test of the joint significance of the parameters of all explanatory variables that include this input quantity. For the OLS, FE and RE models presented in Table 3, the null hypothesis that they were jointly zero was rejected at conventional levels for all inputs with the exceptions of capital in the OLS and RE models. The parameters for the age of the vines (*Age Vines*), grape color (*Red*) and grape quality (*Premium*) are significant at the 5% level in all models. However, the parameters for the training systems (*Single Cordon*, *Double Cordon*, *Pergola*) are significant for the first two systems in the OLS model but none of them are significant in the FE and RE models.

We also check the extent to which monotonicity conditions are fulfilled, i.e., positive marginal products. For the OLS model, for example, monotonicity is satisfied for 70% of all observations. Summarizing, the average compliance rate is 70.1% for all models presented in this paper, ranging from 66% for the FE model to 76.5% for the Pooled SPF with regional and agro-climatic variables. These compliances rates are quite similar to those presented by Perez-Mendez, Roibas and Wall (2019), who found that monotonicity conditions held for 74.1% of all observations when estimating a translog production function for Spanish dairy farms.

We now turn to production frontier estimates and Table 4 shows results for the Pooled SPF, the TFE and the TRE models. The first-order coefficients for *Chemicals* in the Pooled SPF model and those for *Land* in the TFE and TRE models are negative. As in the OLS models, a Wald test of the joint significance of the parameters of the variables containing *Capital* could not reject that they were jointly zero. For the Pooled SPF, the null hypothesis of zero inefficiency was strongly rejected on the basis of a generalized LR test.[6] The parameter $\lambda(=\sigma_u/\sigma_v)$ for the Pooled SPF indicates a relatively large contribution of inefficiency to overall variance. This parameter is much smaller in the TFE model and is equal to zero in the TRE model. Thus, plot-varying inefficiencies can be estimated in the Pooled SPF and TFE models, while no plot-varying inefficiency was found in the TRE model.

The TE scores corresponding to these models and to the FE frontier model (FEF, where the inefficiencies are calculated based on the estimates of the fixed effects) are presented at the top of Table 8, and we can observe some interesting differences. In particular, the FEF assigns the farm-specific effects completely to inefficiency and the efficiency scores have a very low average value (0.283) and are widely dispersed. The Pooled SPF and TFE models both allow for plot-varying inefficiency, and the latter also incorporates farm-specific heterogeneity. As we would expect, the efficiency scores for the TFE model are higher than those for the Pooled SPF and are less dispersed, consistent with the fact that the Pooled SPF assigns any unobserved farm-specific heterogeneity to inefficiency whereas the TFE distinguishes between unobserved farm-specific heterogeneity and inefficiency. In other words, part of the estimated differences in TE across farms in the Pooled SPF model can be attributed to unobserved heterogeneity; hence, when this unobserved heterogeneity is accounted for in the TFE model, the differences in TE across farms fall substantially.

Now we move to results from models that incorporate the plot-invariant variables capturing possible spatial heterogeneity, which are the regional location of each farm and the

[6] This was a likelihood-ratio test with a mixed χ^2 distribution of the test statistic. For the null hypothesis that $\sigma_u = 0$, the test statistic was 12.81, yielding $p < 0.001$.

corresponding agro-climatic zone. We begin by incorporating these variables into the simplest models, namely the OLS average production function and the Pooled SPF. Table 5 contains results from three models. The first column of results corresponds to the baseline OLS model with neither regional nor agro-climatic dummies that has already appeared in Table 3, the middle column shows the OLS model incorporating the regional and agro-climatic dummies, and the final column is the Pooled SPF also including the regional and agro-climatic variables. Comparing the OLS model with agro-climatic and regional variables with the baseline OLS model, we see that the region (valley) where the grape farm is located matters for productivity, as four out of the six coefficients of the regional dummy variables are highly significant. Incorporating the regional and agro-climatic effects also leads to improved *p*-values of the coefficients for the *Age of Vine* and *Red* variables. In addition, the coefficient for *Chemicals*, which is negative in the baseline model, becomes positive.[7] However, not only the region but also the agro-climatic zone matters for grape productivity, as the coefficient for the agro-climatic zone dummy variable *Cool* is highly statistically significant with a negative value. The estimated coefficient for *Cool* implies that farms in the relatively cooler agro-climatic zones of Chile produce 25% less, *ceteris paribus*, than their counterparts in warmer zones.

The OLS results indicate that regional and agro-climatic effects are indeed important in grape production. In the Pooled SPF model incorporating these effects, the parameters for the regional variables are significant at the 5% level for four regions and at the 10% level for another, while the agro-climatic variable is highly significant. As we would expect, incorporating the regional

[7] A further improvement of the pooled SPF over the baseline OLS model is that for the former, the Wald test of the null hypotheses that the parameters of all explanatory variables that include the *Capital* input quantity are jointly zero was rejected at the 5% level ($p = 0.041$).

and agro-climatic effects into the Pooled SPF also affects the TE estimates.[8] Comparing the TE estimates from the Pooled SPF with and without the regional and agro-climatic variables (Table 8) we can see that the average and minimum TE scores are higher in the Pooled SPF model that includes regional and agro-climatic variables.

We estimate the TRE model with agro-climatic and regional variables but, in line with the TRE model without these variables, the results show no evidence of plot-varying inefficiency. However, it may be the case that farm-specific inefficiency exists. To explore this, we now turn to the final set of estimates obtained from the Mundlak True Random Effects (MTRE) and Generalized True Random Effects (GTRE) models. When comparing the FE and RE models in Table 3, we observed that the RE is unsuitable on the basis of a Hausman test. Whereas the RE model assumes no correlation between inputs and unobserved heterogeneity, the MTRE model uses the Mundlak adjustment to relax the assumption of no correlation by incorporating the group means of the inputs into the model, as explained in Section 3. When we estimate the models with the group means of the four inputs, the models do not converge. After some experimentation, we obtain convergence of the MTRE when incorporating the group mean only for *Chemicals*. As the correlation issue can also arise with the GTRE model (Filippini and Greene, 2016), we estimate the latter incorporating the group mean of the *Chemicals* variable and label this model MGTRE.[9] The results from these estimations, with and without the regional and agro-climatic variables are presented in Tables 6 and 7.

[8] A likelihood ratio test of the null hypothesis that $\sigma_u = 0$, yielded a test statistic equal to 1.85. Given a mixed χ^2 distribution of the test statistic, the resulting *p*-value is 0.087.

[9] We also estimated MGTRE models including the group means of *Labor* and *Capital* alongside that of *Chemicals*, and the conclusions did not change. The group means were not significant, and the estimated TE scores were virtually identical to those estimated by the model including only the group mean for *Chemicals*.

In the MTRE and MGTRE results without regional and agro-climatic variables (Table 6), the first-order parameter for *Land* is negative in both models. In addition, both models reveal no evidence of plot-varying inefficiency. However, the MGTRE results reveal the presence of farm-specific inefficiency. Incorporating the regional and agro-climatic variables (Table 7), we see that the first-order coefficient for *Land* is negative in the MTRE model but positive in the MGTRE model.[10] As before, the MGTRE model shows evidence of farm-specific inefficiency.[11]

Descriptive statistics for the TE scores obtained from the different models are summarized in Table 8. In the models without the regional and agro-climatic variables, the average TE score from the MGTRE model is lower than when these variables are included.[12] Finally, Figure 1 illustrates the effect of controlling for firm-specific heterogeneity by comparing the average farm-level TE scores from the Pooled SPF model with those from the MGTRE model. This comparison includes the regional and agro-climatic variables in both cases. As can be seen, the TE scores from the MGTRE model, with very few exceptions, are consistently higher than the corresponding scores from the Pooled SPF. Focusing on the points below the 45º line, the vertical gap between these points and the 45º line can be interpreted as representing the

[10] It should be noted that the Wald tests for the contributions of the input quantities to output rejected the null hypothesis that the joint significance of the input parameters is zero in all models presented in Tables 6 and 7.

[11] We estimated an alternative 3-stage version of the GTRE model (Kumbhakar, Lien and Hardaker, 2014), incorporating the Mundlak adjustment, which confirmed the results of the 1-stage MGTRE model. The procedure involves first estimating the standard RE model. The predicted values for the estimated error term are then used as dependent variables in a standard SF model to estimate plot-varying inefficiency. Similarly, the estimated individual effects are used as the dependent variable in a standard SF model to estimate farm-specific technical inefficiency. As with the 1-stage model, no evidence of plot-varying inefficiency was found, but evidence was found of farm-specific inefficiency. Mean efficiency was 0.834, very similar to the 1-stage model, though the 3-stage showed somewhat greater dispersion of technical inefficiencies. We do not report these results to save space, but they are available on request.

[12] Recall that no evidence of plot-varying inefficiency was found in the MGTRE models. Therefore, the average overall farm-level efficiency scores in Table 8 only reflect farm-level efficiency.

farm-level heterogeneity which would be incorrectly attributed to inefficiency if such heterogeneity is ignored.

6. Concluding remarks

We have analyzed wine grape production in Chilean boutique vineyards by estimating a series of production models, both average and frontier specifications. The dataset used has the advantage of containing information on inputs and output at the plot level (i.e., a sub-unit within a farm), which allows us to distinguish overall farm management influences from plot level management performance. The use of panel-data frontier models allows us to take account of unobserved farm-specific heterogeneity when estimating plot-level inefficiency. Our results show that unobserved heterogeneity is relevant, lending support to the use of panel-data specifications in cases such as ours where individual plots can be assigned to specific farms.

When carrying out efficiency analysis in a context where data are available at farm and sub-unit levels, e.g. plots, care must be taken to control for unobserved heterogeneity besides management effects that may affect efficiency at the various levels. In the case of wine grape production, plot-invariant spatial variables, such as the valley in which the farm is located and its agro-climatic zone, are potentially relevant. Estimates from pooled stochastic frontier models with and without the latter variables exhibit inefficiency differences between plots from the same farm. The pooled model that includes regional and agro-climatic variables reveals that technical efficiency (TE) is higher than when such variables are omitted, implying that part of technical inefficiency measured in the latter model is wrongly attributed to differences in regional and agro-climatic effects. The results from our different model specifications show that farms located in a cooler agro-climatic zone, *ceteris paribus*, have a significantly lower grape production than farms in warmer zones.

While our pooled models show evidence of intra-farm (i.e., inter-plot) differences in efficiency, these plot-level differences disappear when panel data frontier models are used to account for farm-specific heterogeneity. Thus, estimates from the true fixed effects (TFE) model yield an average TE of 0.941 compared to an average of 0.724 in the corresponding pooled model where there is very little variation in efficiency across plots. Estimates from the true random effects (TRE) model and the TRE model with a Mundlak adjustment (MTRE) show no TE differences between plots. Indeed, all plots were found to be equally efficient in these models. Thus, the conclusion of efficiency differences between plots obtained from the pooled frontier models is reversed, even when controlling for regional and agro-climatic effects.

While the TRE and MTRE produce no evidence of inefficiency in plots, and by extension no evidence of inefficiency in farms, these models do not control for possible farm-specific inefficiency. In particular, they assign all the differences between farms captured by the individual farm-specific effects to unobserved heterogeneity. However, part of these differences may be due to differences in farm-level inefficiency. To address the question of whether farm level inefficiencies are present, we estimated the generalized true random effects model with a Mundlak adjustment (MGTRE), which yielded an average TE of 0.791; hence, these results do show evidence of farm-level inefficiency and this holds even when controlling for regional and agro-climatic effects (average TE of 0.859).

Overall, our results reveal that efficiency on a given farm in our sample is determined by management at the farm level and not by plot-level management differences. However, disparities in TE between farms exist, even when controlling for location and agro-climatic conditions. This implies, in accordance with the existing literature, that wine grape producers have the potential to augment productivity significantly by improving TE. This in turns suggests that strategies designed to improve farm-level managerial performance are likely to pay high dividends in the increasingly competitive wine value chain. It should be kept in mind that our sample represents a particular segment of the Chilean wine grape market –family-owned 'boutique wine' farms - so our results are not necessarily generalizable to the market as a whole. Nevertheless, we have shown that panel data models can be adapted to cases where the available data is cross-sectional but composed of observations at different levels (e.g., plots or sub-units) for each individual unit (e.g., a farm). These models can be easily estimated and can provide rich insights regarding unit and subunit performance.

REFERENCES

Abdul-Rahaman, A., & Abdulai, A. (2018). Do farmer groups impact on farm yield and efficiency of smallholder farmers? Evidence from rice farmers in northern Ghana. *Food Policy, 81*, 95-105. https://doi.org/10.1016/j.foodpol.2018.10.007

Abdulai, A., & Tietje, H. (2007). Estimating technical efficiency under unobserved heterogeneity with stochastic frontier models: Application to Northern German dairy farms. *European Review of Agricultural Economics, 34*(3), 393-416. https://doi.org/10.1093/erae/jbm023

Aigner, D., Lovell, C. A. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics, 6*(1), 21-37. https://doi.org/10.1016/0304-4076(77)90052-5

Anderson, K. (2005). *The worlds wine markets: Globalization at work*: Edward Elgar Publishers.

Anderson, K., & Nelgen, S. (2011). *Global wine markets, 1961 to 2009: A statistical compendium*. www.adelaide.edu.au/press/titles/global-wine

Arévalo, M. A., & Martínez, J. M. (2006). *Estimación de costos mediod de producción de vino a granel y márgenes de comercialización, case de estdio: Cooperativa agrícola y vitivinícola Loncomilla.* (Pregrado), Universidad de Talca,

Banfi Piazza, S. (2017). Antecedentes de los mercados del vino y de la uva vinífera. http://www.enologo.cl/images/PDF/mercadoVino2017.pdf

Battese, G. E., & Coelli, T. J. (1992). Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India. *Journal of Productivity Analysis, 3*(1-2), 153-169. https://doi.org/10.1007/BF00158774

Blair, R. D., & Lusky, R. (1975). A note on the influence of uncertainty on estimation of production function models. *Journal of Econometrics, 3*(4), 391-394. https://doi.org/10.1016/0304-4076(75)90056-1

Bravo-Ureta, B. E., Jara-Rojas, R., Lachaud, M., Moreira, V. H., Scheierling, S. M., & Treguer, D. O. (2016). *Farm-level technical efficiency and water: A meta-analysis of the frontier function literature*. Paper presented at the International Association of Agricultural Economists Inter-Conference Symposium, Almaty Kazakhstan.

Bravo-Ureta, B. E., Solís, D., Moreira, V. H., Maripani, J. F., Thiam, A., & Rivas, T. E. (2007). Technical efficiency in farming: A meta-regression analysis. *Journal of Productivity Analysis, 27*(1), 57-72. https://doi.org/10.1007/s11123-006-0025-3

Buzzetti Horta, C. (2018). Boletín del vino: Producción, precios y comercio exterior. Avance a mayo de 2018. https://www.odepa.gob.cl/wp-content/uploads/2018/06/Boletin_vino_201806-1.pdf

Carroll, J., Newman, C., & Thorne, F. (2011). A comparison of stochastic frontier approaches for estimating technical inefficiency and total factor productivity. *Applied Economics, 43*(27), 4007-4019. https://doi.org/10.1080/00036841003761918

Coelli, T., Estache, A., Perelman, S. & Trujillo, L. (2003). *A primer on efficiency measurement for utilities and transport regulators*. Washington D.C. The World Bank.

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., & Battese, G. E. (2005). *An introduction to efficiency and productivity analysis* (2nd ed.). New York, USA: Springer. 0-387-25895-7

Coelli, T. J., & Sanders, O. (2013). *The technical efficiency of wine grape growers in the Murray–darling basin in Australia*: Palgrave Macmillan.

Colombi, R. (2010). *A skew normal stochastic frontier model for panel data*. Paper presented at the Proceedings of the 45-th scientific meeting of the Italian statistical society, Paua, Italy.

Colombi, R., Kumbhakar, S. C., Martini, G., & Vittadini, G. (2014). Closed-skew normality in stochastic frontiers with individual effects and long/short-run efficiency. *Journal of Productivity Analysis*, *42*, 123-136. https://doi.org/10.1007/s11123-014-0386-y

Conradie, B., Cookson, G., & Thirtle, C. (2006). Efficiency and farm size in Western Cape grape production: Pooling small datasets. *South African Journal of Economics, 74*(2), 334-343. https://doi.org/10.1111/j.1813-6982.2006.00061.x

Cornwell, C., Schmidt, P., & Sickles, R. C. (1990). Production frontier with cross-sectional and time-series variation in efficiency levels. *Journal of Econometrics, 46*(1-2), 185-200. https://doi.org/10.1016/0304-4076(90)90054-W

Dawson, P. J., & Lingard, J. (1982). Management bias and returns to scale in a Cobb-Douglas production function for agriculture. *European Review of Agricultural Economics, 9*(1), 7-24. https://doi.org/10.1093/erae/9.1.7

Debarsy, N. (2012). The Mundlak approach in the spatial durbin panel data model. *Spatial Economic Analysis*, *7*(1), 109-131. https://doi.org/10.1080/17421772.2011.647059

Farsi, M., Filippini, M., & Kuenzle, M. (2005). Unobserved heterogeneity in stochastic cost frontier models: An application to Swiss nursing homes. *Applied Economics, 37*(18), 2127 - 2141. https://doi.org/10.1080/00036840500293201

Filippini, M., & Greene, W. (2016). Persistent and transient productive inefficiency: A maximum simulated likelihood approach. *Journal of Productivity Analysis, 45*(2), 187-196. https://doi.org/10.1007/s11123-015-0446-y

Fraga, H., Malheiro, A. C., Moutinho-Pereira, J., Cardoso, R. M., Soares, P. M. M., Cancela, J. J., & Santos, J. A. (2014). Integrated analysis of climate, soil, topography and vegetative growth in Iberian viticultural regions. *PLoS One, 9*(9), e108078. https://doi.org/10.1371/journal.pone.0108078

Greene, W. H. (2005a). Fixed and random effects in stochastic frontier models. *Journal of Productivity Analysis, 23*(1), 7-32. https://doi.org/10.1007/s11123-004-8545-1

Greene, W. H. (2005b). Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics, 126*(2), 269-303. https://doi.org/10.1016/j.jeconom.2004.05.003

Guesmi, B., Serra, T., Kallas, Z., & Gil Roig, J. M. (2012). The productive efficiency of organic farming: The case of grape sector in Catalonia. *Spanish Journal of Agricultural Research, 3*(10), 552-566. https://doi.org/10.5424/sjar/2012103-462-11

Henriques, P. D. d. S., Carvalho, M. L. d. S., & Fragoso, R. M. d. S. (2009). Technical efficiency of Portuguese wine farms. *New Medit: Mediterranean Journal of Economics, Agriculture and Environment, 8*(1), 4-9.

Hodges, D. J. (1969). A note on estimation of Cobb-Douglas and ces production function models. *Econometrica, 37*(4), 721-725. https://doi.org/10.2307/1910448

Jano, P. A. (2017). Quality choice and market access: Evidence from Chilean wine grape production. *Agribusiness, 33*(3), 324-338. https://doi.org/10.1002/agr.21468

Johansson, R., Effland, A., & Coble, K. (2017). Falling response rates to usda crop surveys: Why it matters. Department of agricultural and consumer economics, university of illinois at urbana-champaign. *Farmdoc Daily, 7*(9), 6. http://farmdocdaily.illinois.edu/2017/01/falling-response-rates-to-usda-crop-surveys.html

Jondrow, J., Lovell, C. A. K., Materov, I. S., & Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics*, *19*(2-3), 233-238. https://doi.org/10.1016/0304-4076(82)90004-5

Karagiannis, G., & Kellermann, M. (2019). Stochastic frontier models with correlated effects. *Journal of Productivity Analysis, 51*(2), 175-187. https://doi.org/10.1007/s11123-019-00551-y

King, R. A. (1980). The frontier production function: A tool for improved decision making. *Journal of the Northeastern Agricultural Economic Council, IX*(2), 1-10. https://doi.org/10.22004/ag.econ.159577

Kumbhakar, S. C., Lien, G., & Hardaker, J. B. (2014). Technical efficiency in competing panel data models: A study of Norwegian grain farming. *Journal of Productivity Analysis, 41*, 321-337. https://doi.org/10.1007/s11123-012-0303-1

Lachaud, M. A., Bravo-Ureta, B. E., & Ludeña, C. E. (2015). Agricultural productivity growth in Latin America and the Caribbean and other world regions: An analysis of climatic effects, convergence and catch-up. *Inter-American Development Bank, IDB Working Paper Series No IDB-WP-607, Washington DC*.

Latruffe, L., & Nauges, C. (2014). Technical efficiency and conversion to organic farming: The case of France. *European Review of Agricultural Economics, 41*(2), 227-253. https://doi.org/10.1093/erae/jbt024

Ma, C., Mu, W., Feng, J., & Jiao, W. (2012). Assessing the technical efficiency of grape production in open field cultivation in China. *Journal of Food, Agriculture and Environment, 10*(1), 345-349.

Manevska-Tasevska, G. (2013). Farmers' knowledge attributes contribute to attaining higher farm technical efficiency: A transition economy case. *The Journal of Agricultural Education and Extension, 19*(1), 7-19. https://doi.org/10.1080/1389224X.2012.746001

Meyer, B. D., Mok, W. K. C., & Sullivan, J. X. (2015). Household surveys in crisis. *Journal of Economic Perspectives, 29*(4), 1-29. https://doi.org/10.1257/jep.29.4.1

Moreira L., V. H., Troncoso C., J. L., & Bravo-Ureta, B. E. (2011). Technical efficiency for a sample of Chilean wine grape producers: A stochastic production frontier analysis. *Ciencia e Investigación Agraria, 38*(3), 321-329. https://doi.org/10.4067/S0718-16202011000300001

Mundlak, Y. (1961). Empirical production function free of management bias. *Journal of Farm Economics, 43*(1), 44-56. https://doi.org/10.2307/1235460

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica, 46*(1), 69-85. https://doi.org/10.2307/1913646

Njuki, E., & Bravo-Ureta, B. E. (2015). The economic costs of environmental regulation in U.S. Dairy farming: A directional distance function approach. *American Journal of Agricultural Economics, 97*(4), 1087-1106. https://doi.org/10.1093/ajae/aav007

ODEPA. (2016). Vides: Superficie y producción. http://www.odepa.cl/vides-superficie-y-produccion-3/

Ogundari, K. (2014). The paradigm of agricultural efficiency and its implication on food security in Africa: What does meta-analysis reveal? *World Development, 64*, 690-702. https://doi.org/10.1016/j.worlddev.2014.07.005

Pennings, J. M. E., Irwin, S. H., & Good, D. L. (2002). Surveying farmers: A case study. *Review of Agricultural Economics, 24*(1), 266-277. https://doi.org/10.1111/1467-9353.00096

Perez-Mendez, J. A., Roibas, D., & Wall, A. (2019). The influence of weather conditions on dairy production. *Agricultural Economics, 50*(2), 165-175. https://doi.org/10.1111/agec.12474

Picazo-Tadeo, A. J., & Wall, A. (2011). Production risk, risk aversion and the determination of risk attitudes among Spanish rice producers. *Agricultural Economics, 42*(4), 451-464. https://doi.org/10.1111/j.1574-0862.2011.00537.x

Piesse, J., Conradie, B., Thirtle, C., & Vink, N. (2017). Efficiency in wine grape production: Comparing long-established and newly developed regions of South Africa. *Agricultural Economics*, *49*(2), 203-212. https://doi.org/10.1111/agec.12409

Ponti, L., Gutierrez, P. A., Boggia, A., & Neteler, M. (2018). Analysis of grape production in the face of climate change. *Climate*, *6*(2). https://doi.org/10.3390/cli6020020

Qi, L., Bravo-Ureta, B. E., & Cabrera, V. E. (2015). From cold to hot: Climatic effects and productivity in Wisconsin dairy farms. *Journal of Dairy Science*, *98*, 8664-8677. https://doi.org/10.3168/jds.2015-9536

Schmidt, P., & Sickles, R. C. (1984). Production frontiers and panel data. *Journal of Business and Economic Statistics*, *2*(4), 364-374. https://doi.org/10.2307/1391278

Strohm, K., Dirksmeyer, W., & Garming, H. (2014). International analysis of the profitability of wine grape production. In *8th international conference academy of wine business research, hochschule Geisenheim, june 28th - 30th, 2014, Geisenheim, Germany* (pp. 1-12).

Timmer, C. P. (1971). Using a probabilistic frontier production function to measure technical efficiency. *Journal of Political Economy*, *79*(4), 776-794. https://doi.org/10.1086/259787

Townsend, R., Kirsten, J., & Vink, N. (1998). Farm size, productivity and returns to scale in agriculture revisited: A case of wine producers in South Africa. *Agricultural Economics*, *19*, 175-180. https://doi.org/10.1016/S0169-5150(98)00033-4

Tsionas, E. G., & Kumbhakar, S. C. (2014). Firm heterogeneity, persistent and transient technical inefficiency: A generalized true random-effects model. *Journal of Applied Econometrics*, *29*, 110-132. https://doi.org/10.1002/jae.2300

Weber, J. G., & Clay, D. M. (2013). Who does dot respond to the agricultural resource management survey and does it matter? *American Journal of Agricultural Economics, 95*(3), 755-771. https://doi.org/10.1093/ajae/aas171

Yang, H. (2014, 2014//). *Land rental market and agricultural production efficiency: A Bayesian perspective.* Paper presented at the The Contribution of Young Researchers to Bayesian Statistics, Cham.

Zellner, A., Kmenta, J., & Drèze, J. (1966). Specification and estimation of Cobb-Douglas production function models. *Econometrica: Journal of the Econometric Society, 34*(4), 784-795. https://doi.org/10.2307/1910099

Table 1. Summary of Key Features of 10 Winegrape Production Studies

	Country	Region	Period	Methodology	Mean TE	Focus of study
Townsend et al. (1998)	South Africa	Western Cape Province	1992 to 1995	Non-parametric	N.A.	Relationship between farm size, productivity and returns to scale
Conradie et al. (2006)	South Africa	Robertson and Worcester regions De Doorns	2003 and 2004 2004	SPF	1st quartile 65% 2nd quartile 71% 3rd quartile 75% 4th quartile 76%	Relationship between TE and farm size
Henriques et al. (2009)	Portugal	Alentejo region	2001 and 2004	Non-parametric	60.7%	TE component of productivity
Moreira et al. (2011)	Chile	Several regions	2005/2006	SPF	77.2%	TE component of productivity
Guesmi et al. (2012)	Spain	Catalonia	2008	SPF	Organic 80% Conventional 64%	Organic and conventional grape farms
Ma et al. (2012)	China	24 provinces	2009	SPF	South 76% Bohai Bay Rim area 71%	TE levels and effect of inputs elements on profit
Coelli and Sanders (2013)	Australia	Murray and Murrumbidgee river basins	2006/2007 to 2009/2010	SPF	79%	TE in wine grape production of Australia
Manevska-Tasevska (2013)	Macedonia	Tikvesh vineyard district	2006 to 2008	SPF	69%	Effect of farmers' knowledge on TE
Latruffe and Nauges (2014)	France	Data from FADN database	1999 to 2007	SPF and non-parametric	Conventional 38% 67% and 72% Organic 35% 57% and 72%	Effect of conventional practices on the likelihood of converting to organic farming
Piesse et al. (2017)	South Africa	Nine regions	2005 to 2015	SPF	Old regions 72.9% New regions 72.9%	Compare the efficiency levels of old established wine regions with newer regions
Average Technical Efficiency (TE)					**67.5%**	

Table 2. Descriptive Statistics

Variable	Unit	Average	Min.	Max.
Number of farms	Farms	38		
Size of farms	Ha	86.8	4.0	414.0
Number of plots per farm	Plots	7	2	17
Size of plots	Ha	12.5	0.2	108.7
Grape production	1,000 Kg	134.4	1.2	1,155
Labor	US$	13,388	212	187,664
Capital	US$	4,362	43	40,474
Chemical inputs	US$	3,844	45	34,865
Age vines	Years	16	3	118
Type of Wine Produced				
- Red	%	71		
- White	%	29		
Grape Quality				
- Premium	%	50		
- Varietal	%	50		
Training System				
- Simple cordon	%	73		
- Double cordon	%	13		
- Pergola	%	7		
- Other	%	7		
Location-Valley (from North to South)				
- Limarí	%	4		
- Aconcagua	%	6		
- Casablanca	%	8		
- San Antonio	%	2		
- Maipo	%	15		
- Cachapoal	%	11		
- Rapel	%	2		
- Colchagua	%	30		
- Curicó	%	6		
- Maule	%	16		
Agro-Climatic Zone				
- Cool	Farms	5		
- Warm	Farms	33		

Table 3. Comparison of Non-Frontier Models: OLS, Fixed Effects (FE) and Random Effects (RE) Models

	OLS			FE			RE		
Variable	*Coeff.*	*S.E.*	*p-value**	*Coeff.*	*S.E.*	*p-value*	*Coeff.*	*S.E.*	*p-value*
Labor	0.502	0.087	0.000	0.721	0.145	0.000	0.612	0.105	0.000
Capital	0.067	0.084	0.422	0.794	0.243	0.001	0.329	0.132	0.013
Land	0.477	0.131	0.000	-0.516	0.219	0.019	0.078	0.159	0.626
Chemicals	-0.020	0.033	0.550	0.029	0.048	0.540	0.006	0.038	0.873
$0.5 \times Labor^2$	1.037	0.203	0.011	0.552	0.263	0.294	0.526	0.222	0.236
$0.5 \times Capital^2$	0.272	0.079	0.086	0.196	0.316	0.757	0.559	0.121	0.021
$0.5 \times Land^2$	1.376	0.415	0.099	3.528	0.527	0.001	2.490	0.420	0.003
$0.5 \times Chemicals^2$	-0.073	0.014	0.009	0.030	0.021	0.476	-0.024	0.017	0.473
Labor×Capital	-0.279	0.383	0.466	0.488	0.404	0.228	0.158	0.323	0.625
Labor×Land	-0.887	0.516	0.087	-1.649	0.610	0.007	-1.102	0.500	0.027
Labor×Chemicals	-0.105	0.097	0.279	0.483	0.123	0.000	0.270	0.106	0.011
Capital×Land	-0.210	0.379	0.579	-0.965	0.689	0.163	-0.860	0.430	0.045
Capital×Chemicals	0.127	0.101	0.209	0.211	0.149	0.158	0.091	0.118	0.442
Land×Chemicals	0.028	0.157	0.861	-0.731	0.209	0.001	-0.346	0.174	0.046
Age Vines	0.291	0.128	0.023	0.335	0.110	0.003	0.327	0.103	0.002
Red	-0.116	0.052	0.027	-0.111	0.038	0.004	-0.126	0.037	0.001
Premium	-0.255	0.048	0.000	-0.167	0.037	0.000	-0.185	0.036	0.000
Single Cordon	-0.263	0.095	0.006	0.093	0.115	0.421	-0.006	0.101	0.956
Double Cordon	-0.545	0.114	0.000	0.130	0.127	0.309	-0.094	0.114	0.411
Pergola	-0.008	0.123	0.950	0.056	0.138	0.685	0.054	0.123	0.662
Constant	11.289	0.159	0.000				10.940	0.149	0.000

H_0: OLS vs FE $\chi^2_{(37)} = 272.36$ $p < 0.001$ H_0: RE vs FE $\chi^2_{(20)} = 46.15$ $p < 0.001$

H_0: Cobb-Douglas vs Translog (FE model) $\chi^2_{(10)} = 39.00$ $p < 0.001$

* *p*-values of "0.000" in this and the remaining tables signify that p < 0.001.

Table 4. Comparison of Frontier Models: Pooled SPF, TFE and TRE

	Pooled SPF			TFE			TRE		
Variable	*Coeff.*	*S.E.*	*p-value*	*Coeff.*	*S.E.*	*p-value*	*Coeff.*	*S.E.*	*p-value*
Labor	0.439	0.081	0.000	0.719	0.143	0.000	0.707	0.062	0.000
Capital	0.091	0.078	0.244	0.632	0.325	0.052	0.416	0.060	0.000
Land	0.498	0.123	0.000	-0.352	0.279	0.207	-0.119	0.108	0.269
Chemicals	-0.010	0.031	0.756	0.048	0.063	0.447	0.021	0.023	0.355
0.5×Labor^2	1.001	0.185	0.007	0.641	0.256	0.211	0.554	0.121	0.022
0.5×Capital^2	0.177	0.078	0.253	0.082	0.530	0.938	0.685	0.049	0.000
0.5×Land^2	1.660	0.381	0.029	3.057	0.537	0.004	2.883	0.278	0.000
0.5×Chemicals^2	-0.061	0.013	0.020	0.007	0.033	0.920	0.000	0.009	-0.994
Labor×Capital	-0.080	0.351	0.820	0.492	0.553	0.374	0.157	0.229	0.494
Labor×Land	-1.097	0.480	0.022	-1.661	0.702	0.018	-1.220	0.362	0.001
Labor×Chemicals	-0.040	0.090	0.657	0.329	0.117	0.005	0.375	0.067	0.000
Capital×Land	-0.267	0.331	0.420	-0.704	0.789	0.372	-1.002	0.243	0.000
Capital×Chemicals	0.100	0.088	0.257	0.068	0.129	0.601	0.108	0.069	0.119
Land×Chemicals	-0.029	0.148	0.846	-0.432	0.189	0.022	-0.492	0.111	0.000
Age Vines	0.279	0.108	0.010	0.500	0.064	0.000	0.329	0.082	0.000
Red	-0.136	0.048	0.005	-0.233	0.054	0.000	-0.119	0.034	0.001
Premium	-0.233	0.044	0.000	-0.031	0.064	0.630	-0.181	0.029	0.000
Single Cordon	-0.257	0.084	0.002	0.108	0.190	0.570	0.035	0.064	0.592
Double Cordon	-0.507	0.101	0.000	0.148	0.188	0.432	-0.042	0.080	0.600
Pergola	0.019	0.110	0.864	0.141	0.210	0.502	0.058	0.091	0.526
Constant	11.640	0.136	0.000				10.859	3.240	0.001
Log Likelihood	-77.27			-40.04			-26.20		
σ_u^2	0.189			0.006			0.000		
σ_v^2	0.041			0.039			0.218		
$\lambda\ (=\sigma_u/\sigma_v)$	2.138			0.387			0.000		

Table 5. Comparison of OLS Models with and without Regional and Agro-climatic Variables with Pooled SPF with Regional and Agro-climatic Variables

	OLS Baseline			OLS + Regional and Agro-climatic Dummies			Pooled SPF+ Regional and Agro-climatic Dummies		
Variable	*Coeff.*	*S.E.*	*p-value*	*Coeff.*	*S.E.*	*p-value*	*Coeff.*	*S.E.*	*p-value*
Labor	0.502	0.087	0.000	0.554	0.095	0.000	0.538	0.090	0.000
Capital	0.067	0.084	0.422	0.128	0.077	0.098	0.118	0.073	0.110
Land	0.477	0.131	0.000	0.286	0.125	0.023	0.320	0.119	0.007
Chemicals	-0.020	0.033	0.550	0.051	0.034	0.140	0.046	0.032	0.153
0.5×Labor^2	1.037	0.203	0.011	1.272	0.213	0.003	1.224	0.195	0.002
0.5×Capital^2	0.272	0.079	0.086	-0.116	0.074	0.435	-0.101	0.071	0.477
0.5×Land^2	1.376	0.415	0.099	1.731	0.409	0.035	1.797	0.383	0.019
0.5×Chemicals^2	-0.073	0.014	0.009	-0.047	0.015	0.110	-0.044	0.014	0.108
Labor×Capital	-0.279	0.383	0.466	0.042	0.343	0.902	0.041	0.319	0.897
Labor×Land	-0.887	0.516	0.087	-1.629	0.542	0.003	-1.576	0.504	0.002
Labor×Chemicals	-0.105	0.097	0.279	0.081	0.097	0.403	0.091	0.091	0.320
Capital×Land	-0.210	0.379	0.579	0.139	0.337	0.680	0.070	0.311	0.821
Capital×Chemicals	0.127	0.101	0.209	-0.169	0.104	0.107	-0.115	0.094	0.221
Land×Chemicals	0.028	0.157	0.861	0.102	0.158	0.519	0.036	0.148	0.808
Age Vines	0.291	0.128	0.023	0.255	0.114	0.026	0.244	0.104	0.019
Red	-0.116	0.052	0.027	-0.141	0.046	0.002	-0.147	0.043	0.001
Premium	-0.255	0.048	0.000	-0.224	0.044	0.000	-0.211	0.041	0.000
Single Cordon	-0.263	0.095	0.006	-0.169	0.087	0.052	-0.159	0.082	0.053
Double Cordon	-0.545	0.114	0.000	-0.239	0.108	0.028	-0.223	0.102	0.029
Pergola	-0.008	0.123	0.950	-0.026	0.111	0.813	0.003	0.103	0.979
Aconca./Cachap.				0.279	0.122	0.023	0.201	0.115	0.079
Colchagua/Rapel				0.394	0.120	0.001	0.325	0.110	0.003
Casablanca				-0.178	0.150	0.236	-0.248	0.138	0.073
Curicó				0.425	0.144	0.003	0.354	0.132	0.007
Maipo				-0.013	0.114	0.909	-0.083	0.107	0.436
Maule				0.281	0.122	0.022	0.224	0.113	0.048
Cool				-0.253	0.077	0.001	-0.206	0.074	0.005
Constant	11.29	0.159	0.000	11.00	0.176	0.000	11.29	0.163	0.000
Log Likelihood							-41.66		
σ_u^2							0.086		
σ_v^2							0.050		
$\lambda\ (= \sigma_u/\sigma_v)$							1.312		

Table 6. Comparison of MTRE and MGTRE without Regional and Agro-Climatic Variables

	MTRE			MGTRE		
Variable	*Coeff.*	*S.E.*	*p-value*	*Coeff.*	*S.E.*	*p-value*
Labor	0.707	0.064	0.000	0.698	0.065	0.000
Capital	0.417	0.060	0.000	0.436	0.064	0.000
Land	-0.120	0.115	0.297	-0.138	0.119	0.246
Chemicals	0.021	0.033	0.531	0.029	0.034	0.404
0.5×Labor^2	0.554	0.123	0.024	0.454	0.239	0.057
0.5×Capital^2	0.686	0.049	0.000	0.770	0.107	0.000
0.5×Land^2	2.886	0.280	0.000	2.954	0.564	0.000
0.5×Chemicals^2	0.000	0.010	0.988	0.008	0.020	0.700
Labor×Capital	0.157	0.230	0.495	0.189	0.231	0.414
Labor×Land	-1.221	0.362	0.001	-1.147	0.359	0.001
Labor×Chemicals	0.375	0.067	0.000	0.372	0.068	0.000
Capital×Land	-1.004	0.247	0.000	-1.132	0.255	0.000
Capital×Chemicals	0.109	0.073	0.139	0.121	0.076	0.110
Land×Chemicals	-0.492	0.118	0.000	-0.508	0.122	0.000
Age Vines	0.328	0.083	0.000	0.338	0.082	0.000
Red	-0.119	0.034	0.001	-0.114	0.035	0.001
Premium	-0.181	0.030	0.000	-0.183	0.030	0.000
Single Cordon	0.035	0.065	0.592	0.016	0.065	0.804
Double Cordon	-0.042	0.080	0.601	-0.045	0.082	0.582
Pergola	0.058	0.092	0.529	0.050	0.095	0.600
Constant	0.707	0.064	0.000	10.883	4.529	0.016
Mean-Chemicals	0.002	0.031	0.954	-0.024	0.031	0.441
Log Likelihood	-26.20			-25.672		
σ_u^2	0.000			0.000		
σ_v^2	0.219			0.047		
$\lambda\ (=\sigma_u/\sigma_v)$	0.000			0.000		
σ_β *(heterogeneity)*				0.312	0.020	0.000
σ_ω *(inefficiency)*				1.502	0.143	0.000

Table 7. Comparison of MTRE and MGTRE with Regional and Agro-Climatic Variables

	MTRE			MGTRE		
Variable	*Coeff.*	*S.E.*	*p-value*	*Coeff.*	*S.E.*	*p-value*
Labor	0.685	0.078	0.000	0.704	0.078	0.000
Capital	0.316	0.069	0.000	0.282	0.076	0.000
Land	-0.005	0.124	0.968	0.005	0.131	0.970
Chemicals	0.029	0.036	0.434	0.033	0.037	0.369
$0.5\times Labor^2$	0.732	0.150	0.014	0.681	0.297	0.022
$0.5\times Capital^2$	0.288	0.057	0.012	0.264	0.123	0.032
$0.5\times Land^2$	2.332	0.353	0.001	2.268	0.727	0.002
$0.5\times Chemicals^2$	-0.015	0.014	0.583	-0.012	0.028	0.678
Labor×Capital	0.117	0.275	0.672	0.100	0.284	0.724
Labor×Land	-1.332	0.476	0.005	-1.268	0.483	0.009
Labor×Chemicals	0.337	0.085	0.000	0.339	0.085	0.000
Capital×Land	-0.470	0.292	0.107	-0.446	0.312	0.153
Capital×Chemicals	-0.002	0.086	0.985	0.016	0.088	0.852
Land×Chemicals	-0.331	0.147	0.025	-0.355	0.149	0.018
Age Vines	0.328	0.086	0.000	0.332	0.086	0.000
Red	-0.133	0.036	0.000	-0.133	0.036	0.000
Premium	-0.182	0.035	0.000	-0.182	0.035	0.000
Single Cordon	-0.022	0.078	0.775	-0.030	0.080	0.706
Double Cordon	-0.015	0.090	0.869	-0.018	0.093	0.849
Pergola	0.012	0.102	0.908	0.003	0.102	0.977
Aconca./Cachapoal	0.371	0.095	0.000	0.386	0.096	0.000
Colchagua/Rapel	0.533	0.090	0.000	0.530	0.091	0.000
Casablanca	-0.180	0.115	0.119	-0.193	0.116	0.097
Curicó	0.456	0.115	0.000	0.447	0.117	0.000
Maipo	0.072	0.093	0.438	0.090	0.093	0.332
Maule	0.461	0.095	0.000	0.535	0.097	0.000
Cool	-0.214	0.063	0.001	-0.232	0.064	0.000
Constant	10.666	4.144	0.010	10.640	5.869	0.070
Mean-Chemicals	-0.003	0.038	0.940	0.005	0.039	0.900
Log Likelihood	-12.83			-12.51		
σ_u^2	0.000			0.000		
σ_v^2	0.048			0.049		
$\lambda\ (=\sigma_u/\sigma_v)$	0.000			0.000		
σ_β *(heterogeneity)*				0.206	0.017	0.000
σ_ω *(inefficiency)*				0.444	0.113	0.000

Table 8. Summary Statistics of Technical Efficiency Scores

Model	Mean	Min.	Max.	Std. Dev.
Without Regional and Agro-climatic Variables				
Pooled SPF	0.724	0.331	0.929	0.133
FEF	0.283	0.079	1.000	0.189
TFE	0.941	0.900	0.970	0.010
TRE	1.000	1.000	1.000	0.000
MTRE	1.000	1.000	1.000	0.000
MGTRE	0.791	0.768	0.799	0.006
With Regional and Agro-climatic Variables				
Pooled SPF	0.797	0.413	0.937	0.084
MTRE	1.000	1.000	1.000	0.000
MGTRE	0.859	0.834	0.872	0.008

Figure 1. Comparison of farm-level TE scores: MGTRE model vs Pooled SPF model with regional and agro-climatic variables