https://doi.org/10.17221/341/2017-AGRICECON

Effects of model specification, short-run, and long-run inefficiency: an empirical analysis of stochastic frontier models

Habtamu ALEM*

Norwegian Institute of Bioeconomy Research, Ås, Norway

*Corresponding author: habtamu.alem@nibio.no

Alem H. (2018): **Effects of model specification, short-run, and long-run inefficiency: an empirical analysis of stochastic frontier models**. Agric. Econ. – Czech, 64: 508–516.

Abstract: This paper examines the recent advances in stochastic frontier (SF) models and its implications for the performance of Norwegian crop-producing farms. In contrast to the previous studies, we used a cost function in multiple input-output frameworks to estimate both long-run (persistent) and short-run (transient) inefficiency. The empirical analysis is based on unbalanced farm-level panel data for 1991–2013 with 3 885 observations from 455 Norwegian farms specialising in crop production. We estimated seven SF panel data models grouped into four categories regarding the assumptions used to the nature of inefficiency. The estimated cost efficiency scores varied from 53–95%, showing that the results are sensitive to how the inefficiency is modeled and interpreted.

Keywords: agriculture, cost function, panel data, short and long-run inefficiency

In the economic literature of the last few decades, performance analysis in agriculture, involving both cross-sectional and panel data, has attracted considerable attention (Battese and Coelli 1988; Kumbhakar et al. 2014; Filippini and Greene 2016). Measuring the performance of the farm has been pivotal for the development of the agricultural sector and application of the econometric models of frontier functions (Battese and Coelli 1992).

The standard neoclassical frontier function applied in empirical efficiency models entails an assumption that all farms are fully efficient. Aigner et al. (1977) and Meeusen and van den Broeck (1977) proposed a stochastic frontier (SF) model for cross-sectional data, which diverges from the standard neoclassical production function model by including two distinct error components. One of the error components captures random noise that is beyond the control of the producer and can affect the output such as weather, disease, and pest infestation, which should not be considered as farm inefficiency. The second component captures inefficiency as reflected in the difference between the actual output and the maximal potential output, which is individual specific (i.e. the farm-effect) and is interpreted as one-sided inefficiency.

Since the introduction of one-sided inefficiency within the context of SF panel-data models, there has been considerable research to extend and apply the model to generate consistent and unbiased estimates. Several SF models have been developed based on assumptions about the temporal behaviour of the inefficiency, the specifications of the model, distributional assumptions and estimation techniques. For detail review of SF models, see, e.g. Kumbhakar and Lovell (2000), and Kumbhakar et al. (2015). Consequently, selecting SF models for performance measurement is not straightforward since none has an absolute advantage over the others. The choice is made more complicated by the fact that the models are not nested within one another, which implies that there are no statical criteria to discriminate among them (Karagiannis and Tzouvelekas 2009).

A few studies have been conducted that compare the performance of SF panel-data models using the same dataset. For instance, Karagiannis and Tzouvelekas (2009) estimated ten short-run/time-varying SF mod-

Supported by the Norwegian Research Council, Grant No. 225330/E40.

https://doi.org/10.17221/341/2017-AGRICECON

els for the Greek olive oil sector. Abdulai and Tietje (2007) estimated technical efficiency using seven SF panel data models for the northern German dairy farms. Kumbhakar et al. (2014) estimated six SF panel data models for the Norwegian grain farming. These studies found that efficiency results were sensitive to how the inefficiency is modeled and interpreted.

What is missing in all these studies is an estimation of the competitive SF panel models, allowing the multi-input and multi-output technologies. That is, the previous study estimations were based on a single-output technology specification which is not appropriate, as the construction of a single index of outputs can lead to aggregation problems (Kumbhakar et al. 2008). This study uses a cost function approach and estimates seven competitive SF panel data models within a multi-input output framework. To the best of our knowledge, no previous comparison has been undertaken for agricultural production data[1] within a multi-output technologies framework. Moreover, we estimate the long-run (persistent or time-invariant) and short-run (transient or time-varying) inefficiencies of crop-producing farms in Norway using a cost function framework.

APPROACHES TO MEASURING EFFICIENCY

There are two main approaches to measuring farm performance using panel data: a parametric (econometric) methods, such as that involving SF models, and non-parametric methods, such as data envelopment analysis (DEA). In both cases, the methods are based on the radial contraction/expansion connection of observed inefficiency points with reference points on the frontier (unobserved). Each approach owns pros and cons in measuring of the performance of a farm (Kumbhakar et al. 2015). The treatment of measurement error is the curtail distinction between the two approaches. SF models can accommodate stochastic noise, such as measurement errors due to weather, disease, and pest infestation that are likely to be significant in farming. The non-parametric (DEA) approach is sensitive to outliers since the model ignores the measurement error (Coelli et al. 2005).

Since crop-producing farms in our study are sensitive to the external random shocks, we have selected the SF approach to estimate the cost efficiency score. From the cost point of view, the SF approach assumes that those farms identified as cost-efficient are the best practice farms and all the other farms operate inefficiently above the cost frontier (the minimum cost estimated). The basic panel-data cost function of the SF model is expressed logarithmically as:

$$\ln c_{it} = \beta_0 + h(y_{it}, w_{it}; \boldsymbol{\beta}) + v_{it} + u_i \qquad (1)$$

where c_{it} is the total cost incurred by the farm i in year t; w_{it} denotes the vector of input price of farm i in period t; y_{it} denotes the vector of output; $\boldsymbol{\beta}$ is the vector of the parameters to be estimated; and $h(y_{it}, w_{it}; \boldsymbol{\beta})$ is the cost function that represents the (dual) farm technology. Finally, v_{it} is the noise term, and $u_i \geq$ is the long-run/persistent inefficiency. Different models have been developed based on the assumption of the inefficiency terms. In this study, we estimated seven SF panel models grouped into four categories according to the assumptions connected to the inefficiency behaviour (Table 1). The following chapters review four categories of panel-data models commonly used in the literature.

We begin with the first category, which is the most restrictive in terms of the assumed behaviour of inefficiency (Models 1.1 and 1.2). Pitt and Lee (1981) developed the means of capturing the persistent (time-invariant) part of inefficiency and then interpreted the random effects of the panel data as inefficiency. In the persistent models, it is assumed that the inefficiency in Equation 1 is constant through time, that is, $u_i = u_{i1} = u_{i2} = \ldots = u_{iT}$. In this category, we estimated the random-effect model (Schmidt and Sickles 1984) for Model 1.1 and the truncated normal distribution model with time-invariant inefficiency (Battese and Coelli 1988) for Model 1.2.

The assumption of the long-run inefficiency might be reasonable in short panels; this might also be the case in some situations, such as where inefficiency is associated with managerial ability, and there is no change in the farm management for any of the farms during the study period (Kumbhakar et al. 2014). The main disadvantage of the first category models is that identification of inefficiency and individual heterogeneity are not considered separately. The inefficacy score over time is constant, and it is restrictive in a competitive economic environment especially when the number of the period is large.

[1]Filippini and Greene (2016) estimated the efficiency of a sample of Swiss railway companies using four SF models and a cost function framework.

https://doi.org/10.17221/341/2017-AGRICECON

We expect that farm management is dynamic and that farmers can learn from their own experience, and agricultural extension, to improve their management over time.

Table 1. Detail specification of seven fitted panel data models

	Model specifications	Assumption on inefficiency*	Estimation method	Inefficiency
Model 1.1 Schmidt and Sickles (1984)	$\ln c_{it} = \beta_0 + h(y_{it}, w_{it}; \boldsymbol{\beta}) + v_{it} + u_i =$ $= (\beta_0 - \mu) + h(y_{it}, w_{it}; \boldsymbol{\beta}) + v_{it} + u_i^* =$ $= \alpha^* + h(y_{it}, w_{it}; \boldsymbol{\beta}) + v_{it} + u_i^*$	time-invariant u_i is random number distributional assumption of u_i $v_{it} \sim N[0, \sigma_v^2]$	FGLS	long-run/persistent inefficiency $\widehat{u_i} = \widehat{\alpha_i} - \min_i \{\widehat{\alpha_i}\}$
Model 1.2 Battese and Coelli (1988)	$\ln c_{it} = \beta_0 + h(y_{it}, w_{it}; \boldsymbol{\beta}) + v_{it} + u_i$ $\varepsilon_{it} = v_{it} + u_i$	time-invariant $u_i \sim N^+[\mu, \sigma_u^2]$ $v_{it} \sim N[0, \sigma_v^2]$	ML	long-run/persistent inefficiency $\mathrm{E}(u_{it} \mid v_{it} + u_{it})$
Model 2.1 Kumbhakar (1990)	$\ln c_{it} = \beta_0 + h(y_{it}, w_{it}; \boldsymbol{\beta}) + v_{it} + u_{it}$ $u_{it} = G(t)\, u_i$ G is a function for time	time-varying $u_i \sim N^+[\mu, \sigma_u^2]$ $v_{it} \sim N[0, \sigma_v^2]$	ML**	short-run/transient inefficiency $\mathrm{E}(u_{it} \mid v_{it} + u_{it})$
Model 2.2 Battese and Coelli (1992)	$\ln c_{it} = \alpha_{it} + h(y_{it}, w_{it}; \boldsymbol{\beta}) + v_{it} + u_{it}$ $\varepsilon_{it} = v_{it} + u_{it}$ $u_{it} = G(t)\, u_i$ G is a function for time	time-varying $u_i \sim N^+[\mu, \sigma_u^2]$ $v_{it} \sim N[0, \sigma_v^2]$	ML	short-run/transient inefficiency $\mathrm{E}(u_{it} \mid v_{it} + u_{it})$
Model 3.1 Greene (2005) TFE	$\ln c_{it} = \alpha_i + h(y_{it}, w_{it}; \boldsymbol{\beta}) + \varepsilon_{it}$ $\varepsilon_{it} = v_{it} + u_{it}$ and $\alpha_i = (\alpha + \mu_i)$ μ_i treated as fixed effect	time-varying $u_i \sim N^+[\mu, \sigma_u^2]$ $v_{it} \sim N[0, \sigma_v^2]$	ML***	short-run/transient inefficiency $\mathrm{E}(u_{it} \mid v_{it} + u_{it})$
Model 3.2 Greene (2005) TRE	$\ln c_{it} = \alpha + \mu_i + h(y_{it}, w_{it}; \boldsymbol{\beta}) + \varepsilon_{it}$ $\varepsilon_{it} = v_{it} + u_{it}$ μ_i treated as random effect	time-varying $u_i \sim \exp(\theta)$, $\sigma_u = \frac{1}{\theta}$ $v_{it} \sim N[0, \sigma_v^2]$, $\mu_i \sim N[0, \sigma_\mu^2]$	SML	short-run/transient inefficiency $\mathrm{E}(u_{it} \mid v_{it} + u_{it})$
Model 4 Kumbhakar et al. (2014)	$\ln c_{it} = \alpha_0 + h(y_{it}, w_{it}; \boldsymbol{\beta}) + \varepsilon_{it} + \alpha_i$ $\varepsilon_{it} \equiv v_{it} + u_{it}$ $\alpha_i \equiv \mu_i + \tau_i$ μ_i is the firm specific effect	time-varying and -invariant $u_{it} \sim N^+[0, \sigma_u^2]$ $v_{it} \sim N[0, \sigma_v^2]$ $\mu_i \sim N[0, \sigma_\mu^2]$ $\tau_i \sim N^+[0, \sigma_\tau^2]$	MME****	long-run/persistent inefficiency $\mathrm{E}(\tau_i \mid v_{it} + u_{it})$ short-run/transient inefficiency $\mathrm{E}(u_{it} \mid v_{it} + u_{it})$

FGLS – feasible generalized least square; ML – maximum likelihood; SML – simulated maximum likelihood; MME – method of moments estimator; *all models assume individual specific effect; **Cornwell et al. (1990) follow a quadratic pattern over time ($\alpha_i = \alpha_{it} = \alpha_{0i} + \alpha_{1i}t + \alpha_{2i}t^2$) and it is estimated using a modified-least square dummy variable method without specifying the distribution assumption of inefficiency; ***it can be estimated using adding dummy variable to the model to accommodate μ_i, we estimated by ML based on Chen et al. (2014); ****economic literature proposed different methods to estimate Model 4 – Colombi et al. (2014) used a full maximum likelihood method (FML) and Filippini and Greene (2016) showed that it is possible to estimate Model 4 using SML estimation recently; for further explanations of variables see chapter Empirical model

Source: author's own elaboration based on economics literature

https://doi.org/10.17221/341/2017-AGRICECON

$$\ln c_{it} = \beta_0 + \sum_{j=2}^{J} \beta_j \ln w_{jit} + \sum_{m=1}^{K} \beta_m \ln y_{mit} + \sum_{m=1}^{M} \sum_{j=1}^{J} \beta_{mj} \ln y_{mit} \ln w_{jit} + \\ + \frac{1}{2}\left[\sum_{m=1}^{M} \sum_{n=1}^{M} \beta_{mm} \ln y_{mit} \ln y_{mit} + \sum_{j=2}^{J} \sum_{k=2}^{J} \beta_{jj} \ln w_{jit} \ln w_{jit} + bt^2\right] + \\ + \sum_{m=1}^{M} \beta_{mt} \ln y_{mit}\, t + \sum_{j=2}^{J} \beta_{jt} \ln w_{jit}\, t + bt + \alpha_i + \varepsilon_{it} \quad (2)$$

The second category represent Models 2.1 and 2.2, for which we assume the inefficiency effect to be individual-specific and time-varying (short-run or transient). Short-run inefficiency models allow the likelihood that the inefficiency changes over time, $u_{it} = u_i f(t)$ in Equation 1. Various models were developed based on this general specification. We used the method of Kumbhakar (1990) for Model 2.1 and Battese and Coelli (1992) for Model 2.2. The main drawback of Models 2.1 and 2.2 is that the unobserved factors are assumed to be random over time. This implies that time-invariant factors such as soil type and quality are confounded into the inefficiency, and so the performance of the farm is underestimated (Greene 2005).

The third category consists of the Model 3.1 and Model 3.2, for which the error term is split into three components to separate latent heterogeneity (farm-effect) from the inefficiency effect that is the error term, the farm effect, and the inefficiency term. The first component captures a random noise, the second component time-invariant unmeasured and unobserved heterogeneity, and the third component a farm-specific inefficiency term. We have estimated Greene's (2005) 'true' fixed effect (TFE) in Model 3.1 and the 'true' random effect (TRE) in Model 3.2.

The SF models estimate either long-run or short-run part of the farm efficiency; however, being able to estimate both levels of inefficiency is important. Fortunately, a four-component error-term SF model (Model 4) developed by Colombi et al. (2014), Kumbhakar et al. (2014), and Kumbhakar and Tsionas (2014) is the latest model in the efficiency literature; it allows the estimation of the persistent and transient parts of inefficiency simultaneously from the same data. The first component captures the random shocks that are out of the control of the farm manager (weather, disease, and pest infestation). The second component captures latent heterogeneity, which is distinguished from the inefficiency (Greene 2005). The third component captures long run inefficiency, for instance, quality of the land or farm management rigidity within a farm organisation and production process. The last component captures short-run inefficiency. Even in the presence of farm management rigidity in the production processes, a farm could be able to improve performance in the short-run (Kumbhakar et al. 2014; Filippini and Greene 2016). Both the first and fourth components vary across the farms and over time (i.e., observation specific).

EMPIRICAL MODEL

As the models in categories one to three above are similar to Model 4, though less comprehensive, we firstly describe Model 4 and then show how the other models are related to Model 4. Model 4 is specified as translog cost function with a three-output and four-input as in Equation 2.

In Equation 2, c_{it} stands for the total cost incurred by farm i in year t; w_j represents the price of inputs j; and y_m is the quantity of output m. The error terms ε_{it} and α_i are split into four components, namely, $\varepsilon_{it} \equiv v_{it} + u_{it}$ and $\alpha_i = \mu_i + \tau_i$. As discussed in chapter Empirical model, the u_{it} component captures transient cost inefficiency[2]; v_{it} is the idiosyncratic error term capturing random shocks; μ_i captures latent heterogeneity; and the τ_i component captures persistent inefficiency. All Greek letters are parameter estimates. The trend variable t, introduced to capture the effect of technological change, starts at $t = 1$ for 1991 and increases by one annually.

By inserting μ_i and u_{it} in Equation 2, we can estimate the persistent inefficiency models (the first category models), such as those of Schmidt and Sickles (1984) and Pitt and Lee (1981). By inserting μ_i in Equation 2, we can estimate transient inefficiency models (the second category models), such as those of Cornwell et al. (1990), Kumbhakar (1990), and Lee and Schmidt (1993). Finally, it is possible to estimate the Greene

[2]There is a confusion in the economics literature in using terms cost effectiveness and cost efficiency (Farsi and Filippini 2009). Cost-effectiveness (pareto efficiency) is a technique for defying the least cost option for meeting specific objectives or outcome (Balana et al. 2011). However, consistent with the economics literature on production theory (Farrel 1957), in this article the term cost efficiency is a relative concept, which measures relative to the best performance defined by the production technology, for detail review refer to Kumbhakar and Lovell (2000) and Coelli et al. (2005).

https://doi.org/10.17221/341/2017-AGRICECON

(2005) TRE and TFE transient inefficiency models (the third category models) by inserting τ_i in Equation 2.

Economic theory requires the imposition of price homogeneity and symmetry restrictions to the parameters. Symmetry restrictions require that $\beta_{nj} = \beta_{jn}$ and $\beta_{mn} = \beta_{nm}$. The cost function is homogeneous of degree 1 in input prices, so the following restrictions apply:

$$\sum_{j=1}^{J}\beta_j = 1 \text{ and } \sum_j \beta_{jn} = \sum_j \beta_{mj} = 0.$$

An easier way to impose price homogeneity is to divide the quantity of all input prices and cost by an arbitrarily selected input price. Thus, in Equation 2, the left-hand side is redefined as:

$\ln c = \ln\left(c/w_j\right)$, and all input prices are redefined as:

$\ln w_j = \ln\left(w_j/w_J\right)$,

that is, we divided all input prices and the total cost by wages before estimating the translog cost function. The wage coefficient can be obtained by subtracting the estimated input price coefficient from 1.

Equation 2 can be estimated using a single-stage simulated maximum likelihood (ML) estimator based on Filippini and Greene (2016) or using a multi-step method of moments estimators following Kumbhakar et al. (2014). We use the latter method and estimate the model in three steps. In the first step, the standard random-effect panel regression is used to estimate parameters and predict $\hat{\varepsilon}_{it}$ and $\hat{\alpha}_i$. In the second step, we use $\hat{\varepsilon}_{it}$ as a regressor in an SF model and estimate the transient inefficiency $\left(\exp\left(-\hat{u}_{it}\right)\right)$, which is the Jondrow et al. (1982) estimator of u_{it} (Model 4-T). In the final step we use $\hat{\alpha}_i$ as a regressor in an SF model and estimate the $\exp\left(-\widehat{\tau_i}\right)$ persistent component following a procedure similar to the second step (Model 4-P). The overall efficiency (reported as Model 4-OE) is a product of Model 4-P and Model 4-T. Equation 2 is specified in logs so that the inefficiency terms can be interpreted as the percentage deviations of the observed performance.

Data and the definition of variables

The data used in this study is a farm-level unbalanced panel dataset for the period 1991–2013 with a total of 3 885 observations from 455 farms specialising in the production of grain and forage crops. The dataset includes production and economic data collected annually by NIBIO. To accommodate panel features in the estimation, we included only those farms for which at least three consecutive years of data were available. Most farms that are engaged in crop production are located in the eastern and central regions of Norway. The 2012 statistics show that 286 000 hectares of land are under cultivation. Of this, 81% is located in the eastern region and 17% in the central areas (Statistics Norway 2013). Thus, to obtain a homogenous group of farms, only farms from the eastern and central regions of Norway specialising in crop production that had reported their accounting data for the 1991–2013 period were selected.

The outputs are grain production in kilograms, adjusted for quality to indicate feed units milk (FUm)[3] (y_1), forage production in FUm (y_2), and the value of other crop outputs in Norwegian Kronor (y_3). Grain output is an aggregate of four principal species: barley, wheat, oats, and oilseed species. The cost function in Equation 2 is specified by means of the four input prices described below (w_j). Land prices are based on the market price for land in terms of rent paid for land at the farm level. The price of labour is the wage for hired labour. We computed the implicit prices (opportunity costs) of owned land and family labour based on data for farm-level rent and wages provided by NIBIO. The prices of other variable inputs and the price of capital costs are constructed as Laspeyres indices, based on figures provided by NIBIO (BFJ 2016). All prices are deflated to 2013 levels using agricultural price index data provided by NIBIO.

EMPIRICAL RESULTS

Testing model specification

Table 2 presents the estimated coefficients of the translog cost function for seven SF panel data models. A series of hypotheses about the nature of the frontier model and the consistency of the cost function along with its properties were tested using likelihood ratios (LRs). The goodness of fit, which measures the log

[3]A feed unit milk (FUm) is a measure of physical output adjusted for differences in the quality of output. FUm is defined as 1 kg of grain with 15% water content. Thus, the output measure is a quality-adjusted yield of all crops in kilograms per year.

https://doi.org/10.17221/341/2017-AGRICECON

Table 2. Estimates of parameters in the translog cost function (TL) for seven models (number of observations = 3 885)

Elasticities	Model 1.1	Model 1.2	Model 2.1	Model 2.2	Model 3.1	Model 3.2	Model 4
y_1	0.16*** (0.01)	0.14*** (0.01)	0.14*** (0.02)	0.15*** (0.01)	0.09*** (0.13)	0.14*** (0.01)	0.09*** (0.01)
y_2	0.13*** (0.02)	0.13*** (0.01)	0.13*** (0.02)	0.13*** (0.01)	0.10*** (0.02)	0.13*** (0.01)	0.10*** (0.02)
y_3	0.23*** (0.02)	0.22*** (0.01)	0.22*** (0.03)	0.23*** (0.01)	0.17*** (0.03)	0.22*** (0.01)	0.17*** (0.01)
w_2	0.04** (0.02)	0.03** (0.01)	0.03* (0.01)	0.03** (0.01)	0.03 (0.02)	0.03*** (0.01)	0.02** (0.01)
w_3	0.35*** (0.13)	0.37*** (0.13)	0.49*** (0.14)	0.38*** (0.13)	0.40*** (0.13)	0.36** (0.13)	0.43*** (0.13)
w_4	0.55*** (0.14)	0.54*** (0.13)	0.41*** (0.15)	0.53*** (0.13)	0.51*** (0.19)	0.54*** (0.13)	0.48*** (0.13)
Gamma (γ)	0.57	0.72	0.76	0.76	0.96	0.34	0.66
Eta (η)	–	–	–	–0.01*** (0.00)	–	–	–
Mu (μ)	–	0.67*** (0.08)		0.61*** (0.04)	–111.67*** (6.15)	–	–
Theta (θ)	–	–	–	–	–	0.26*** (0.01)	–
R^2	–	1 035	977	1 047	2 006	1 041	0.76

statistically significant at *, **, *** $p < 0.05$, 0.01, 0.001, respectively; standard errors in parentheses; y_1 = grain output, y_2 = forage output, and y_3 = other crop output; w_2 = rent, w_3 = price of variable inputs, and w_4 = price of capital inputs, all in log form ; R^2 – log likelihood/adjusted R-square; Model 1.1 and Model 1.2 – Schmidt and Sickles (1984) random-effect and Battese and Coelli (1988) models, respectively; Model 2.1 – Kumbhakar (1990); Model 2.2 – Battese and Coelli (1992); Model 3.1 and Model 3.2 – 'true' fixed effect (TFE) and 'true' random effect (TRE) models of Greene (2005), respectively; Model 4 – Kumbhakar et al. (2014); the second-order parameter in the translog cost function is not reported, but available from the author on request

Source: own calculation based on farm level data

of the likelihood function (for model category 2–3), is satisfactory.

Empirical model specification

The test of skewness returned a p-value of less than 0.001, showing that the null hypothesis of no skewness can be rejected with confidence. Hence, we have found support for a right-skewed error distribution, and obtained evidence for the SF specification of the model. This implies that the null hypothesis of an ordinary least squares (OLS) specification is rejected at the 0.01% significance level. Hence, we estimate the model with the parametric distributional assumptions of α_i and u_{it}. Table 3 shows that a simplification of the translog to Cobb-Douglas is rejected.

The second step concerns the distribution of the inefficiency effects using the null hypothesis H_0: $\mu = 0$. Table 2 shows that the coefficient of μ is statically different from zero, which implies that models assuming the truncated normal distribution are more appropriate than models assuming the half-normal distribution. The other component we consider is the value of eta (η). If η is both significant and non-zero, then cost efficiency can be said to vary over time. Since we estimate a cost function when $\eta < 0$, the degree of inefficiency increases over time; while when $\eta > 0$, the degree of inefficiency decreases over time. In our analysis, time-varying models are preferred. The value of η is a negative sign and is significant at the 1% level, thus cost inefficiency increases from 1991 to 2013. In Table 2, the estimate of η is approximately 0.01, which suggests that, on average, cost inefficiency increases at a rate of 1% per year.

The other parameter considered is the value of theta (θ), which shows the farm-specific heterogeneity (the unobserved heterogeneity). In our estimate, θ is statically significant and different from zero, which implies that Greene's (2005) models are preferred among time-varying models. The Hausman test is the common tool used to assist researchers to select which model (Model 3.1 – fixed effect; Model 3.2 – random effect) suits to our data better. The results in Table 4 show that the chi-square is considerable. Therefore, we can reject the null hypothesis, the random-effect (RE) models, in favour of the fixed-effect model. That is to say that the assumption of orthogonality in RE does not work. We also tested the characteristics of the technology, with the result that a Cobb-Douglas technology specification is rejected. The farm's inefficiency is

https://doi.org/10.17221/341/2017-AGRICECON

Table 3. Properties of grain and forage production technology

Restrictions	Parametric restrictions	Wald test statistics
CD technology	H_0: all interaction terms are zero	8.64
Breusch and Pagan Lagrangian (multiplier test for random effects)	test: variance inefficiency term = 0	5 034.68
Normality test/test return of skewness	Schmidt and Lin (1984)	11 850.97
Hausman test fixed and random effect	chi2 (33)	273.86
Log likelihood ratio (LR)	LR test for random effects	39 235.22

p-value – 0.00

Source: own calculation based on farm level data

estimated using the conditional mean of the cost-inefficiency term proposed by Jondrow et al. (1982) and is discussed in detail in the next sub-section.

Cost efficiency score

The estimates of cost efficiency scores of the farms for each model specification are presented in Table 4. The results indicate that there is a significant variation in estimating mean cost efficiency scores ranging from 53–95%. The analysis shows that cost efficiency results are sensitive to how the inefficiency is modelled and interpreted. Karagiannis and Tzouvelekas (2009) and Kumbhakar et al. (2014) support this finding in their efficiency score studies.

Comparing the estimated models, the average cost-efficiency score was about 53 % of the long-run fixed-effect model (Model 1.2). The average cost efficiency scores of short-run models only (0.66 for Model 2.1 and 0.57 for Model 2.2) were relatively greater compared to the estimated the long-run models only (Model 1.1 and Model 1.2). The average estimated cost efficiency scores for Greene 'true' fixed and random effect model (Model 3.1 and Model 3.1) were 93% and 94%, respectively. This coincides with our expectation, with regard to which, as discussed above, the model separates some of the time-invariant heterogeneities (farm-effect) from the inefficiency term. However, the model overestimates the performance of the farm because the model could not separate time-invariant heterogeneity from time-invariant (persistent) inef-

Table 4. Cost efficiency scores of farms for each model specification for the seven models (number of observations = 3 885)

Model[a,b]	Mean	Standard deviation	Minimum	Maximum
Model 1.1	0.55**	0.13	0.30	1.00
Model 1.2	0.53**	0.13	0.25	0.95
Model 2.1	0.66***	0.15	0.31	0.99
Model 2.2	0.57**	0.14	0.28	0.97
Model 3.1	0.93***	0.04	0.51	0.98
Model 3.2	0.94***	0.05	0.51	0.99
Model 4-T	0.94***	0.01	0.83	0.99
Model 4-P	0.95***	0.00	0.93	0.96
Model 4-OE	0.89***	0.01	0.79	0.94

statistically significant at *,**,*** $p < 0.05$, 0.01, 0.001, respectively; [a]Model 1.1 and Model 1.2 – Schmidt and Sickles (1984) random-effect and Battese and Coelli (1988) models, respectively; Model 2.1 – Kumbhakar (1990); Model 2.2 – Battese and Coelli (1992); Model 3.1 and Model 3.2 – 'true' fixed effect (TFE) and 'true' random effect (TRE) models of Greene (2005), respectively; Model 4 – Kumbhakar et al. (2014); [b]Model 4-T, Model 4-P, and Model 4-OE show mean short–run (time-varying/ transient), long-run (time-invariant/ persistent), and overall efficiency for Model 4 in Kumbhakar et al. (2014) multi-stage estimation, respectively

Source: own calculation based on farm level data

https://doi.org/10.17221/341/2017-AGRICECON

ficiency. Similar results were reported in Kumbhakar et al. (2014).

Model 4 shows that the mean persistent (Model 4-P) and transit efficiency (Model 4-T) scores were 0.95 and 0.94, respectively (Table 4). In terms of the persistent cost inefficiency, an average crop producing farm incurred costs that are about 5% ((1/0.95) – 1) above the minimum cost defined by the frontier. The implication is that the average actual cost could be reduced by 5%, without reducing the output, if persistent (time-invariant) inefficiency could be removed. On the other hand, the actual costs could be reduced by 6% if transient (time-varying) inefficiency could be removed. The overall efficiency estimate for Model 4 (Model 4-OE), which is the interaction of persistent and transit efficiency was 0.89. The implication is that the average actual cost could be reduced by 11%, without reducing the output, if both inefficiencies could be removed in the Norwegian crop production.

CONCLUSION

We estimated seven alternative SFA panel data models grouped into four categories based on the assumptions applying to the inefficiency component. The empirical results show that the mean cost-efficiency score varied from 53% to 95%. The range found shows that the cost-efficiency score is sensitive to how the inefficiency is modelled and interpreted. We also distinguished the level of both persistent and transient inefficiency of crop farms using the four-component model. The overall efficiency score of Norwegian crop farms based on the latest model was 89%. The estimated persistent cost (long-run) efficiency was 95%, while transient (short-run) efficiency was 94%.

The empirical analysis shows that the magnitude of persistent cost inefficiency (5%) was lower than the transient inefficiency level (6%). It is possible to reduce crop-production costs by on average 5% if we reduce shortfalls in farmers' managerial capabilities such as lack of experience. Moreover, it is possible to reduce crop production costs by 6% if we improve transient inefficiency such as the debt/asset ratio.

We estimated the levels of persistent cost and transient inefficiency, but we have not investigated the determinants of persistent and transient inefficiency in Norwegian crop farms. Thus, the limitations of this study suggest important topics that could benefit from further research.

REFERENCES

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: 393–416.

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

Balana B.B., Vinten A., Slee B. (2011): A review on cost-effectiveness analysis of agri-environmental measures related to the EU WFD: Key issues, methods, and applications. Ecological Economics, 70: 1021–1031.

Battese G.E., Coelli T.J. (1988): Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. Journal of Econometrics, 38: 387–399.

Battese G.E., Coelli T.J. (1992): Frontier production functions, technical efficiency, and panel data: with application to paddy farmers in India. In: Thomas R., Gulledge Jr. T.R., Lovell C.A.K. (eds): International applications of productivity and efficiency. Springer Netherlands.

Battese G.E., Corra G.S. (1977): Estimation of a production frontier model: with application to the pastoral zone of Eastern Australia. Australian Journal of Agricultural and Resource Economics, 21: 169–179.

Chen Y.-Y., Schmidt P., Wang H.-J. (2014): Consistent estimation of the fixed effects stochastic frontier model. Journal of Econometrics, 181: 65–76.

Coelli T.J., Rao D.S.P., O'Donnell C.J., Battese G.E. (2005): An Introduction to Efficiency and Productivity Analysis. Springer Science & Business Media.

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.

Cornwell C., Schmidt P., Sickles R.C. (1990): Production frontiers with cross-sectional and time-series variation in efficiency levels. Journal of Econometrics, 46: 185–200.

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

Farsi M., Filippini M. (2009): An analysis of cost efficiency in Swiss multi-utilities. Energy Economics, 31: 306–315.

Filippini M., Greene W. (2016): Persistent and transient productive inefficiency: a maximum simulated likelihood approach. Journal of Productivity Analysis, 45: 187–196.

Greene W. (2005): Fixed and random effects in stochastic frontier models. Journal of Productivity Analysis, 23: 7–32.

https://doi.org/10.17221/341/2017-AGRICECON

Jondrow J., Lovell C.A.K., Materov I.S., Schmidt P. (1982): On the estimation of technical inefficiency in stochastic frontier production function model. Journal of Econometrics, 19: 233–238.

Karagiannis G., Tzouvelekas V. (2007): A flexible time-varying specification of the technical inefficiency effects model. Empirical Economics, 33: 531–540

Karagiannis G., Tzouvelekas V. (2009): Parametric measurement of time-varying technical inefficiency: Results from competing models. Agricultural Economics Review, 10: 50–79.

Kumbhakar S.C. (1990): Production frontiers, panel data, and time-varying technical inefficiency. Journal of Econometrics, 46: 201–211.

Kumbhakar S.C., Lovell C.K. (2000): Stochastic Frontier Analysis. Cambridge University Press.

Kumbhakar S.C., Lien G., Flaten O., Tveterås R. (2008): Impacts of Norwegian milk quotas on output growth: a modified distance function approach. Journal of agricultural Economics, 59: 350-369.

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.

Kumbhakar S.C., Tsionas E.G. (2011): Some recent developments in efficiency measurement in stochastic frontier models. Journal of Probability and Statistics, 2011: 25.

Kumbhakar S.C., Wang H.-J., Horncastle A. (2015): A Practitioner‘s Guide to Stochastic Frontier Analysis Using Stata. Cambridge University Press, Oxford.

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

NIBIO report (BFJ) (2016): Totalkalkylen for jordbruket. Budget Committee for Agriculture, Aggregate Accounts for Agriculture. Norwegian Institute of Bioeconomy Research, Norway.

Statistics Norway (2013): Agriculture statistics. Available at https://www.ssb.no/jord-skog-jakt-og-fiskeri?de=Landbrukstellinger (accessed May, 2014).

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

Schmidt P., Lin T.-F. (1984): Simple tests of alternative specifications in stochastic frontier models. Journal of Econometrics, 24: 349–361.

Schmidt P., Sickles R.C. (1984): Production frontiers and panel data. Journal of Business & Economic Statistics, 2: 367–374.

Received November 22, 2017
Accepted February 2, 2018
Published online November 21, 2018