UNIVERSITA' DEGLI STUDI DI BERGAMO
DIPARTIMENTO DI INGEGNERIA GESTIONALE
QUADERNI DEL DIPARTIMENTO[†]

Department of Economics and Technology Management

Working Paper

n. 01 – 2011

A stochastic frontier model with short-run and long-run inefficiency random effects

by

Roberto Colombi, Gianmaria Martini, Giorgio Vittadini

[†] Il Dipartimento ottempera agli obblighi previsti dall'art. 1 del D.L.L. 31.8.1945, n. 660 e successive modificazioni.

NOTE

I *Working Papers* della Collana dei Quaderni del Dipartimento di Ingegneria Gestionale costituiscono un servizio atto a fornire la tempestiva divulgazione dei risultati dell'attività di ricerca, siano essi in forma provvisoria o definitiva.

I *Working Papers* della Collana sono pubblicati nell'archivio istituzionale dell'Università degli studi di Bergamo (Aisberg): http://dspace-unibg.cilea.it/handle/10446/370.

A Stochastic Frontier Model with short-run and long-run inefficiency random effects

Roberto Colombi, Gianmaria Martini, Giorgio Vittadini

Abstract This paper presents a new stochastic frontier model for panel data. The model takes into account firm unobservable heterogeneity and short-run and long-run sources of inefficiency. Each of these features is modeled by a specific random effect. In this way, firms' latent heterogeneity is not wrongly modeled as inefficiency, and it is possible to disentangle a time-persistent component from the total inefficiency. Under reasonable assumptions, we show that the closed-skew normal distribution allows us to derive both the log-likelihood function of the model and the posterior expected values of the random effects. The new model is compared with nested models by analyzing the efficiency of firms belonging to different sectors.

Key words: Closed-Skew Normal Distribution, Longitudinal Data Analysis, Mixed Models, Stochastic Frontiers.

1 Introduction

The estimation of a stochastic frontier to evaluate firms' technical or cost efficiency has been the object of several contributions (see Coelli *et al.* 2005

Roberto Colombi
Università degli Studi di Bergamo, Department of Information Technology and Mathematical Methods, viale Marconi 5 24044 Dalmine, e-mail: roberto.colombi@unibg.it

Gianmaria Martini
Università degli Studi di Bergamo, Department of Economics and Technology Management, viale Marconi 5 24044 Dalmine, e-mail: gianmaria.martini@unibg.it

Giorgio Vittadini
Università degli Studi di Milano-Bicocca, Departments for of Quantitative Methods for Economics and Business, via Bicocca degli Arcimboldi 8, 20126 Milano, e-mail: giorgio.vittadini@unimib.it

and Kumbhakar and Lovell 2000, for recent overviews of the existing literature). According to this approach, the frontier estimation has to take into account that if the firm's observed output (cost) is not equal to the maximum (minimum) feasible output (cost) (i.e., the firm is not efficient), "not all the shortfall has to be attributed to technical inefficiency" (Kumbhakar and Lovell 2000, p. 65); exogenous random shocks may also have an impact. Hence, a typical stochastic frontier (SF) model separates the error terms into two components: (1) the random shocks error component, given by a random variable with normal distribution and zero mean; and (2) the firm's technical inefficiency, which is a non-negative random variable with different possible probability distributions (e.g., half-normal, exponential, truncated normal, and gamma).

In this paper, we introduce a new SF model in which the error term is instead split into four components to take into account different factors influencing the firms' possible shortfalls. The first component captures firms' latent heterogeneity, which has to be disentangled from the inefficiency effects; the second component identifies the impact on firms' performance of short-run sources of inefficiency (Greene 2005). The third component investigates the importance of long-run sources of inefficiency, while the last component captures random shocks.

The new SF model presented in this paper fills a gap in the literature on SF applications to panel data. Indeed the existing contributions have been developed in two directions (see Kumbhakar and Lovell 2000 and Greene 2005 for comprehensive surveys). One class of SF models investigates firms' performance over time by assuming that inefficiency is time variant. The other class assumes instead that firms' inefficiency is time invariant—i.e., if a firm is inefficient, this gap remains fixed over time. Both approaches present some drawbacks.

Time-varying inefficiency models treat inefficiency as a period-specific random variable without considering the possible presence of some elements leading to long-lasting (i.e., time-invariant) effects on firms' inefficiency. We define them as long-run components of inefficiency. Moreover, time-varying inefficiency SF models assume that the firm's random inefficiency at period t is independent from its previous inefficiency levels. The latter is somehow in contrast with the idea that a firm takes time to recover efficiency, given the presence of some rigidities both in its assets and internal organization of production. If inefficiency has a long-run effect, then it may have a certain degree of time dependence. For example, a hospital has more capacity (beds) than the optimal required level, but downsizing may be a long-run process due to social pressure. This implies that the hospital has a long-run inefficiency since this gap cannot be completely recovered in the short-run. Another example is an airport with a radar system that does not allow the optimal utilization rate of the other avionics assets (the runways' length, the available aircraft parking positions, etc.), so that its maximum flight capacity is rather low. However, the substitution costs of this asset could be greater than the market value of a new, updated radar system because the obsolete equipment has not completed its expected period of service yet. Hence, an updated radar system is not installed, and so the related inefficient utilization of the other assets has a multi-period impact. Consequently, in these circumstances it is hard to believe that a firm's inefficiency in period t is independent from that arising at period $t-1$.

On the contrary, time-invariant SF models assume that the firm's inefficiency is constant over time, a questionable assumption in a sufficiently long panel, as pointed out by Greene (2005). Even if a firm may suffer from the presence of some rigidities in its production process, leading to long-run inefficiency, according to these models the firm is unable to eliminate in the short-run any sources of inefficiency. Again, this is something difficult to accept. It is more sensible to assume that the firm may recover part of its inefficiency by removing some sources having short-run effects, while some other sources of inefficiency stay with the firm over time. For example, a hospital with excess capacity may increase in the short-run its labor productivity because the top management rearranges the personnel among the different activities, so that part of the physicians' and nurses' daily working hours are employed in day hospital activities rather than being partially under-utilized in a full-time job allocation to acute discharges. This is a short-run improvement eliminating only one source of inefficiency (the job allocation within the hospital) that may also be completely independent from previous-period short-run inefficiency levels. Hence, the hospital continues to suffer from long-run inefficiency due to excess capacity, but it may have recovered part of its short-run inefficiency.

Pitt and Lee (1981) pointed out these drawbacks, and they argued that between the two extreme models—where all inefficiency stays with the firm over time, and where none of it stays with the firm over time—there is an intermediate case in which some inefficiency stays with the firm over time and some that instead does not. However, they fail to propose a proper SF model because of the difficulty they found in specifying a flexible multivariate distribution for the inefficiency error components.

Another important limitation of many SF models is that they do not take into account the impact of the firm's unobserved heterogeneity on its production process. By doing so, as observed by Kumbhakar and Heshmati (1995), SF models confound time-invariant aspects of inefficiency with producer-specific heterogeneity. For example, the physicians of a hospital located in a given region may adopt a highly efficient surgery technique that cannot be easily transferred to other physicians. Furthermore, latent heterogeneity among hospitals belonging to different countries may have an even greater effect. As a result, these heterogeneity effects have an impact on the production frontier and may explain hospitals' different performances, but they are not due to managerial inefficiency. Greene's (2005) "true random effect" model is a first attempt to deal with this problem. He splits the error term of the production frontier into three components: a producer-specific, time-varying inefficiency term; a producer-specific random-effect capturing latent heterogeneity; and the random shocks. However, Greene's true random-effect SF model considers any producer-specific, time-invariant component as unobserved heterogeneity. Hence, heterogeneity effects are now included in the SF model but at the cost of regarding the firm's inefficiency only as a per-period effect, completely independent from past levels and with no long-lasting impacts. In other words, long-run inefficiency is again confounded with latent heterogeneity, while only short-run inefficiency is allowed.

In order to fill these gaps, in Section 2 of the paper we present an SF model in which all possible sources of inefficiency are considered. Therefore, we split the error term into a producer-specific effect capturing latent heterogeneity; a producer-specific, time-invariant component, which is distinct from latent heterogeneity, capturing long-run sources of inefficiency; a producer-specific, time-variant component, capturing short-run inefficiency; and a random shock. The probabilistic assumptions on these random components are also introduced in Section 2. To the best of our knowledge, such an SF model has not been provided yet in the literature.

In Section 3, we show how well-known results about the closed-skew normal distribution (Arellano-Valle and Azzalini 2006, Gonzáles-Farías *et al.* 2004) are useful to derive the log-likelihood of the new SF model and to compute the posterior expected values of its random components. Computational aspects concerning the maximum likelihood estimator are also briefly examined in this section. In Section 4 the new SF model is applied to analyze firms' inefficiency using three longitudinal data sets regarding firms operating in different sectors (i.e., hospitals, rice producers, and airports). The performances of the new model and of nested models are compared, and we show that the proposed SF model improves the analysis of firms' inefficiency. More in detail, these applications display that the four-component SF model is particularly appropriate when firms are heterogeneous and when the panel is sufficiently long. In these cases, it is important to distinguish between latent heterogeneity and long-run inefficiency, and to investigate whether firms are able to partially improve their performances over time by removing some short-run inefficiency sources.

Our multiple-random-component SF model is related to the SF model introduced by Dominguez-Molina *et al.* (2004) and to the linear mixed models proposed by Lin and Lee (2005) and Arellano-Valle *et al.* (2005). Dominguez-Molina *et al.* (2004) were the first to recognize the relevance of the closed-skew normal distribution in SF analysis, but they did not examine multiple-random-component SF models. Lin and Lee (2005) and Arellano-Valle *et al.* (2005) used the closed-skew normal distribution to relax the normality assumption in the mixed-regression models. However, the framework of mixed-regression models is different from ours, and we think that skew normality arises more naturally in the context of SF models than in that of mixed-regression models.

Some generalizations of the new SF model are presented in Section 5, which concludes the paper.

2 A four-random-component SF model for Panel Data

We consider the stochastic frontier model:

$$y_{it} = \beta_0 + \boldsymbol{x}'_{it}\boldsymbol{\beta} + b_i - u_{it} - u_{i0} + e_{it}, \tag{1}$$

where the index i, $i = 1, 2, ...n$, denotes n units and t, $t = 1, 2, ..., T$, the T time points at which every unit is observed. The variable y_{it} is the logarithm of the i-th unit's output at time t, $\boldsymbol{x}'_{it}$ is a row vector of p regressors, and $\boldsymbol{\beta}$ is a column vector of unknown parameters. The random variable e_{it} is the idiosyncratic random component capturing random shocks, and b_i is the random unit effect for latent heterogeneity. Furthermore, u_{it} is the time-dependent stochastic inefficiency capturing short-run inefficiency effects, and u_{i0} is the time-invariant stochastic inefficiency for long-run sources of inefficiency.

Many interesting models can be obtained from Eq. (1) by omitting one or more random components. By doing so, every model is identified by a three-letter label (without taking into account the random shock term). The first letter pertains to the presence (T = True) or absence (F = False) in the SF model of the random-unit effect, the second letter (again, T or F) is related to the presence/absence of the time-dependent inefficiency term, and the third letter indicates the presence/absence of the time-independent inefficiency term. Hence, the new model presented in this contribution is labeled as TTT (see Eq. (1)); the Greene's (2005) true random-effect SF model is obtained by dropping u_{i0} from Eq. (1), and it is labeled as TTF; the Pitt and Lee (1981) model I is obtained by dropping the terms b_i and u_{it} and is labeled as FFT; and the pooled SF model (Pitt and Lee, 1981, model II) is labeled as FTF (i.e., b_i and u_{i0} are dropped). Furthermore, other two new SF models can be analyzed: TFT—i.e., an SF model with latent unit heterogeneity and long-run inefficiency—and FTT—i.e., a model with both short-run and long-run inefficiency terms but without latent heterogeneity. Finally, by dropping both of the inefficiency components, the random-intercept regression model TFF (with only latent heterogeneity) and the classical regression model FFF (with no latent heterogeneity and inefficiency terms) are obtained.

Our purpose is to test how the new TTT SF model improves the analysis of firms' inefficiency in panel data by comparing its statistical performances with those of the other less general SF models. Testing one of the previous models against the general one (i.e., TTT) is a non-standard problem because under the null hypothesis one or more parameters are on the boundary of the parametric space. In fact, under reasonable assumptions, in this case the asymptotic distribution of the log-likelihood ratio test statistic is a mixture of chi-square distributions known as chi-bar-square distribution (Silvapulle and Sen 2005). For example, in the case of models differing for the presence of only one random component (i.e., the TTT and FTT models), the log-likelihood ratio test statistic is asymptotically distributed as a 0.5 mixture of a chi-square distribution with zero degrees of freedom and a chi-square distribution with one degree of freedom, and the p-value is found by dividing by two the p-value corresponding to a chi-square distribution with one degree of freedom.

For the sake of simplicity, we consider only the case of balanced panel data with a fixed number of observations per unit, but the results of the following sections can be easily extended to unbalanced panels.

We assume that:

(A1a) for $i = 1, 2, ..., n$, the $2 \times (T + 1)$ random variables u_{i0}, b_i, u_{it}, e_{it} $(t = 1, 2, ..., T)$ are independent in probability;
(A1b) the random vectors $(b_i,\ u_{i0},\ u_{i1}, ...,\ u_{iT},\ e_{i1}, ...,\ e_{iT})$, $i = 1, 2, ..., n$ are independent in probability;
(A2) for every i, u_{i0} is a normal random variable with null expected value and variance σ^2_{1u} left-truncated at zero, and b_i is a normal random variable with null expected value and variance σ^2_b;
(A3) for every i and t, u_{it} is a normal random variable with null expected value and variance $\sigma^2_{2u,t}$ left-truncated at zero, and e_{it} is a normal random variable with null expected value and variance σ^2_e;
(A4) the $\boldsymbol{x}_{it}$ are vectors of known constants.

The following matrix representation of model (1) will be useful in the next sections.

Let $\mathbf{1}_T$ be a vector of ones, $\mathbf{0}_T$ a vector of zeros, and $\boldsymbol{I}_T$ the identity matrix of dimension T. Moreover, $\boldsymbol{y}_i$ is a vector of the T observations on the *i-th* unit; $\boldsymbol{X}_i$ is the $T \times p$ matrix with rows $\boldsymbol{x}'_{it}$, $\boldsymbol{u}_i$ is the $(T + 1)$ vector with components $u_{i0}, u_{i1}, u_{i2}, ..., u_{iT}$; and $\boldsymbol{e}_i$ is the vector of the idiosyncratic random components of the *i-th* unit. From (1), it follows that: $\boldsymbol{y}_i = \mathbf{1}_T(\beta_0 + b_i) + \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{A}\boldsymbol{u}_i + \boldsymbol{e}_i$, where the matrix $\boldsymbol{A}$ is defined as: $\boldsymbol{A} = -[\,\mathbf{1}_T\ \boldsymbol{I}_T\,]$.

In the next section, we will derive the joint density function of the random components $\mathbf{1}_T b_i + \boldsymbol{A}\boldsymbol{u}_i + \boldsymbol{e}_i$.

3 Statistical properties of the TTT model

Some important consequences of assumptions (A1)-(A4) are examined here in order to derive the log-likelihood of the TTT model presented in Eq. (1) and to derive the posterior expected values of the random components.

With $\phi_q(\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{\Omega})$, we denote the density function of a q-dimensional normal random variable with expected value $\boldsymbol{\mu}$ and variance $\boldsymbol{\Omega}$, while $\bar{\Phi}_q(\boldsymbol{\mu}, \boldsymbol{\Omega})$

is the probability that a q-variate normal random variable of expected value $\boldsymbol{\mu}$ and variance matrix $\boldsymbol{\Omega}$ belongs to the positive orthant.

3.1 The closed-skew normal distribution

For an easy reference, we report the definition of the closed-skew normal distribution (Arellano-Valle and Azzalini 2006; Gonzáles-Farías *et al.* 2004).

Definition 1. A random vector $\boldsymbol{x}$ has a (p, q) closed-skew normal distribution with parameters $\boldsymbol{\mu}$, $\boldsymbol{\Gamma}$, $\boldsymbol{D}$, $\boldsymbol{\nu}$, $\boldsymbol{\Delta}$ if its probability density function is:

$$f(\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{\Gamma}, \boldsymbol{D}, \boldsymbol{\nu}, \boldsymbol{\Delta}, p, q) = \frac{\phi_p(\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{\Gamma})\bar{\Phi}_q(\boldsymbol{D}(\boldsymbol{y} - \boldsymbol{\mu}) - \boldsymbol{\nu}, \boldsymbol{\Delta})}{\bar{\Phi}_q(-\boldsymbol{\nu}, \boldsymbol{\Delta} + \boldsymbol{D}\boldsymbol{\Gamma}\boldsymbol{D}')}. \tag{2}$$

The moment-generating function of the previous random variable is:

$$E(\exp\{\boldsymbol{t}'\boldsymbol{x}\}) = \frac{\bar{\Phi}_q(\boldsymbol{D}\boldsymbol{\Gamma}\boldsymbol{t} - \boldsymbol{\nu}, \boldsymbol{\Delta} + \boldsymbol{D}\boldsymbol{\Gamma}\boldsymbol{D}')}{\bar{\Phi}_q(-\boldsymbol{\nu}, \boldsymbol{\Delta} + \boldsymbol{D}\boldsymbol{\Gamma}\boldsymbol{D}')} \exp\{\boldsymbol{t}'\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{t}'\boldsymbol{\Gamma}\boldsymbol{t}\}. \tag{3}$$

3.2 Closed-skew normality in the TTT model

For our purposes it is useful to introduce the following matrices:

$$\begin{aligned}
&\boldsymbol{V} = \begin{bmatrix} \sigma^2_{1u} & \boldsymbol{0}'_T \\ \boldsymbol{0}_T & \boldsymbol{\Psi} \end{bmatrix}, \quad \boldsymbol{\Sigma} = \sigma^2_e \boldsymbol{I}_T + \sigma^2_b \boldsymbol{1}_T \boldsymbol{1}'_T \\
&\boldsymbol{\Lambda} = \boldsymbol{V} - \boldsymbol{V}\boldsymbol{A}'(\boldsymbol{\Sigma} + \boldsymbol{A}\boldsymbol{V}\boldsymbol{A}')^{-1}\boldsymbol{A}\boldsymbol{V} = (\boldsymbol{V}^{-1} + \boldsymbol{A}'\boldsymbol{\Sigma}^{-1}\boldsymbol{A})^{-1}, \\
&\boldsymbol{R} = \boldsymbol{V}\boldsymbol{A}'(\boldsymbol{\Sigma} + \boldsymbol{A}\boldsymbol{V}\boldsymbol{A}')^{-1} = \boldsymbol{\Lambda}\boldsymbol{A}'\boldsymbol{\Sigma}^{-1},
\end{aligned}$$

where $\boldsymbol{\Psi}$ is the diagonal matrix with the variances $\sigma^2_{2u,t}$ $(t = 1, 2, ..., T)$ on the main diagonal.

The relevance of the closed-skew normal density function in the context of the TTT model stems from the following proposition (from now on, we will assume that $q = T + 1$).

Proposition 1. *Under the assumptions A1a, A2-A4, the random vector* $\boldsymbol{y}_i$ *has a (T,q) closed-skew normal distribution with the parameters:* $\boldsymbol{\nu} = \boldsymbol{0}$, $\boldsymbol{\mu} = \boldsymbol{1}_T\beta_0 + \boldsymbol{X}_i\boldsymbol{\beta}$, $\boldsymbol{\Gamma} = \boldsymbol{\Sigma} + \boldsymbol{A}\boldsymbol{V}\boldsymbol{A}'$, $\boldsymbol{D} = \boldsymbol{R}$, *and* $\boldsymbol{\Delta} = \boldsymbol{\Lambda}$.

Proof. See the Appendix.

From Definition 1 and Proposition 1, it follows that the density of $\boldsymbol{y}_i$ is

$$f(\boldsymbol{y}_i) = \phi_T(\boldsymbol{y}_i, \mathbf{1}_T\beta_0 + \boldsymbol{X}_i\boldsymbol{\beta}, \boldsymbol{\Sigma} + \boldsymbol{A}\boldsymbol{V}\boldsymbol{A}')\frac{\bar{\Phi}_q(\boldsymbol{R}(\boldsymbol{y}_i - \boldsymbol{X}_i\boldsymbol{\beta} - \mathbf{1}_T\beta_0), \boldsymbol{\Lambda})}{2^{-q}}$$

and that its moment-generating function is

$$E(\exp\{\boldsymbol{t}'\boldsymbol{y}_i\}) = \frac{\bar{\Phi}_q(\boldsymbol{V}\boldsymbol{A}'\boldsymbol{t}, \boldsymbol{V})}{2^{-q}}\exp\{\boldsymbol{t}'(\mathbf{1}_T\beta_0 + \boldsymbol{X}_i\boldsymbol{\beta}) + \frac{1}{2}\boldsymbol{t}'(\boldsymbol{\Sigma} + \boldsymbol{A}\boldsymbol{V}\boldsymbol{A}')\boldsymbol{t}\}.$$

When $\boldsymbol{y}_i$ is a vector of logarithms of outputs and $\boldsymbol{t}$ is the k-th column of the identity matrix of dimension T, the previous moment-generating function gives the expected value of the k-th component of the vector $\exp\{\boldsymbol{y}_i\}$ of the outputs.

It can be easily checked that a (T,T) closed-skew normal distribution is obtained when we only consider SF models without time-invariant inefficiency (e.g., Greene's 2005 TTF model; Pitt and Lee's 1981 FTF model). Finally, when the time-dependent inefficiency is omitted, a $(T,1)$ closed-skew normal density arises. When the unit random component is omitted, the joint distribution, is given by the previous results with $\sigma_b^2 = 0$.

The following Proposition is an immediate consequence of assumption A1b and of Proposition 1.

Proposition 2. *Under the assumptions A1-A4, the log-likelihood of nT observations from model (1) is:*

$$L = \textstyle\sum_{i=1}^{n}(\ln\phi_T(\boldsymbol{y}_i - \boldsymbol{X}_i\boldsymbol{\beta}, \mathbf{1}_T\beta_0, \boldsymbol{\Sigma} + \boldsymbol{A}\boldsymbol{V}\boldsymbol{A}') + \\ + \ln\bar{\Phi}_q(\boldsymbol{R}(\boldsymbol{y}_i - \boldsymbol{X}_i\boldsymbol{\beta} - \mathbf{1}_T\beta_0), \boldsymbol{\Lambda})) + nq\ln(2), \quad (4)$$

which is the log-likelihood of the n independent closed-skew normal random variables $\boldsymbol{y}_i - \boldsymbol{X}_i\boldsymbol{\beta}$.

Note that the computational complexity of the maximum-likelihood estimator, which maximizes the log-likelihood (4), comes from the multiple integrals $\bar{\Phi}_q(\boldsymbol{R}(\boldsymbol{y}_i - \boldsymbol{X}_i\boldsymbol{\beta} - \mathbf{1}_T\beta_0),\ \boldsymbol{\Lambda})$ $(i = 1, 2, ..., n)$. There are, however, remarkable exceptions. For the models TFT and FFT, the above integral is one dimensional, while for the pooled model FTF it is a product of T one-dimensional integrals. Thus, the computational problem is limited to the SF models where the time-dependent inefficiency is present together with the unit-specific component or with the time-invariant inefficiency. For a survey of efficient numerical and Monte Carlo methods to compute the previous multi-normal integrals, see Genz and Bretz (2009).

Since the maximization of the log-likelihood reported in Eq. (4) is a complex and time-consuming task, in Section 4 the following two-step procedure will be applied to obtain reasonable starting values. Let ϵ_i be the unit-specific random component $\epsilon_i = b_i - u_{i0}$ and ϵ_{it} be the random component $\epsilon_{it} = -u_{it} + e_{it}$. In the first step, the procedure computes the estimate $\tilde{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ by a standard estimation method for panel data models or mixed models applied to the following random intercept model:

$$y_{it} = \beta_0 + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \epsilon_i + \epsilon_{it}. \tag{5}$$

From A.1-A.4, it follows that $\tilde{\boldsymbol{\beta}}$ is a consistent estimator of $\boldsymbol{\beta}$ (β_0 is not consistently estimated because the random components of the previous model have non-null expected values). In the second step, since $\boldsymbol{y}_i - \boldsymbol{X}_i\boldsymbol{\beta}$ ($i = 1, 2, ...n$) are independent closed-skew normal random variables, the following pseudo log-likelihood function is maximized with respect to β_0, σ_b^2, σ_{1u}^2, σ_e^2 and $\boldsymbol{\Psi}$:

$$\tilde{L} = \sum_{i=1}^{n} (\ln \phi_T(\boldsymbol{y}_i - \boldsymbol{X}_i\tilde{\boldsymbol{\beta}}, \boldsymbol{1}_T\beta_0, \boldsymbol{\Sigma} + \boldsymbol{A}\boldsymbol{V}\boldsymbol{A}) + \ln \bar{\Phi}_q(\boldsymbol{R}(\boldsymbol{y}_i - \boldsymbol{X}_i\tilde{\boldsymbol{\beta}} - \boldsymbol{1}_T\beta_0))). \tag{6}$$

The properties of these estimators follows from the general theory of two-step M-estimators (Parke 1986).

3.3 Prediction of the random components

When $\boldsymbol{y}_i$ is a vector of logarithms of outputs, an important topic in applied research is forecasting the unit random components by the expected values $E(\exp\{b_i\}|\boldsymbol{y}_i)$ and the unit random inefficiencies by $E(\exp\{-\boldsymbol{u}_i\}|\boldsymbol{y}_i)$. In doing so, it is convenient to introduce the following definitions:

$$\boldsymbol{r}_i = \boldsymbol{y}_i - \boldsymbol{X}_i\boldsymbol{\beta} - \boldsymbol{1}_T\beta_0, \quad \tilde{\sigma}_b^2 = \sigma_b^2 - \sigma_b^4\boldsymbol{1}'_T\boldsymbol{\Delta}\boldsymbol{1}_T \tag{7}$$

$$\boldsymbol{\Delta} = (\boldsymbol{\Sigma} + \boldsymbol{A}\boldsymbol{V}\boldsymbol{A}')^{-1}, \quad \tilde{\boldsymbol{\Lambda}} = \boldsymbol{\Lambda} - \boldsymbol{R}\boldsymbol{1}_T\boldsymbol{1}'_T\boldsymbol{R}'\frac{\sigma_b^4}{\tilde{\sigma}_b^2}. \tag{8}$$

The following proposition shows how the relevant expected values can be computed.

Proposition 3. *a) Conditionally on $\boldsymbol{y}_i$, the unit-specific random component b_i has a (1,q) closed-skew normal distribution with density*

$$f(b_i|\boldsymbol{y}_i) = \phi(b_i, \sigma_b^2\boldsymbol{1}'\boldsymbol{\Delta}\boldsymbol{r}_i, \tilde{\sigma}_b^2)\frac{\bar{\Phi}_q(\boldsymbol{R}\boldsymbol{r}_i - \boldsymbol{R}\boldsymbol{1}_T\sigma_b^2\tilde{\sigma}_b^{-2}(b_i - \sigma_b^2\boldsymbol{1}'_T\boldsymbol{\Delta}\boldsymbol{r}_i), \tilde{\boldsymbol{\Lambda}})}{\bar{\Phi}_q(\boldsymbol{R}\boldsymbol{r}_i, \boldsymbol{\Lambda})};$$

b) Conditionally on $\boldsymbol{y}_i$, the random inefficiency vector $\boldsymbol{u}_i$ is a left-truncated normal random variable with density:

$$f(\boldsymbol{u}_i|\boldsymbol{y}_i) = \frac{\phi_q(\boldsymbol{u}_i, \boldsymbol{R}\boldsymbol{r}_i, \boldsymbol{\Lambda})}{\bar{\Phi}_q(\boldsymbol{R}\boldsymbol{r}_i, \boldsymbol{\Lambda})}, \quad \boldsymbol{u}_i > \boldsymbol{0};$$

c)

$$E(\exp\{b_i\}|\boldsymbol{y}_i) = \frac{\bar{\Phi}_q(\boldsymbol{R}\boldsymbol{r}_i - \boldsymbol{R}\mathbf{1}_T\sigma_b^2, \boldsymbol{\Lambda})}{\bar{\Phi}_q(\boldsymbol{R}\boldsymbol{r}_i\boldsymbol{\Lambda})} \exp\{\sigma_b^2\mathbf{1}_T'\boldsymbol{\Delta}\boldsymbol{r}_i + \frac{1}{2}\tilde{\sigma}_b^2\};$$

d)

$$E(\exp\{\boldsymbol{t}'\boldsymbol{u}_i\}|\boldsymbol{y}_i) = \frac{\bar{\Phi}_q(\boldsymbol{R}\boldsymbol{r}_i + \boldsymbol{\Lambda}\boldsymbol{t}, \boldsymbol{\Lambda})}{\bar{\Phi}_q(\boldsymbol{R}\boldsymbol{r}_i, \boldsymbol{\Lambda})} \exp\{\boldsymbol{t}'\boldsymbol{R}\boldsymbol{r}_i + \frac{1}{2}\boldsymbol{t}'\boldsymbol{\Lambda}\boldsymbol{t}\}.$$

Proof. See the Appendix.

If $-\boldsymbol{t}'$ is the k-th row of the identity matrix of dimension q, result d) of Proposition 3 gives the conditional expected value of the k-th component of the inefficiency vector $\exp\{-\boldsymbol{u}_i\}$. In particular, for $k = 1$ we get the conditional expected value of the time-invariant inefficiency. We highlight also that, conditionally on the observation $\boldsymbol{y}_i$, the random effect b_i has not a normal distribution as happens in standard random-effect models.

4 Applications and comparison of different SF models

In this section, we investigate the efficiency of firms belonging to three different sectors and compare the results provided by the TTT model with those of the nested models presented in Section 2.

To test whether the variances of one or more random components are zero, we will use the log-likelihood ratio test statistic. The log-likelihood ratio test statistic G^2 has, in this case, a non-standard asymptotic chi-bar-square distribution that is a mixture of chi-square distributions (Silvapulle and Sen, 2005). Moreover, in order to explain the effect of the elimination of a random component from the model TTT, we will investigate whether the omission of a random error (e.g., the latent heterogeneity component) has the impact of inflating the variances of the other random terms (i.e., the idiosyncratic error and the short-run and long-run inefficiency terms). The latter outcome would imply that the eliminated random component was capturing a relevant inefficiency (or heterogeneity) effect, that is confounded, in a less general model, with the remaining random terms.

Table 1 displays the main features of our data sets. We apply the above models to three different sectors (i.e., hospitals, rice producers, and airports), with sample data differing also for their geographic extension (the hospitals, and the rice producers data sets regard firms belonging to a single region, while the airport data set presents information at the country level). Furthermore, the sample size differs, in that the hospital data set covers 134 units, the rice producer data set 43 units, and the airport data set only 34 units. The time horizon is rather large for the hospital data set ($T = 10$) and for the rice producer one ($T = 8$), while it is relatively short for the airport data set ($T = 4$). In the production function estimation, we consider

5 inputs both for the hospital and airport cases, and only 3 inputs for the rice producers. Finally, the impact of exogenous factors (e.g., ownership, size) that may affect the production levels is considered only in the hospital and airports cases, where we believe they may play a more important role than in the rice producers' sector; in the latter case, firms have the same type of ownership and size and are also closely located.

Table 1 Features of the three data sets

Sector	Geographic extension	n (firms)	T (years)	p (inputs)	Exogenous factors
Hospitals	Region (Lombardy, Italy)	134	10	5	Ownership Specialization
Rice producers	Region (Tarlac, Philip.)	43	8	3	—
Airports	Country (Italy)	34	4	5	Size

The different time horizons of the three data sets may have an impact on the relative importance of the short-run and long-run inefficiency terms. For instance, if the time horizon is short it is less difficult to accept the hypothesis that the inefficiency is time invariant (Greene 2005), since firms may at most slightly improve (or deteriorate) their performances in such a short time period. Hence, in the presence of short panel data, not considering the long-run inefficiency component may involve omitting a factor capturing most of the firms' inefficiency level. Moreover, the omission of firms' latent heterogeneity may have a mild impact in the explanation of firms' shortfalls from the estimated frontier if the sample covers geographically concentrated firms producing a rather standardized good (e.g., rice producers).

In the first application, we examine a data set composed of $n = 134$ hospitals located in the Lombardy region of Northern Italy.[1] The data were collected every year from 1998 to 2007. The output variable is given by a case-mix adjusted number of outpatient *discharges*, and the inputs considered are the number of *beds, wards, physicians, nurses*, and *administrative workers*.[2] A log-linear frontier, that is to say a linearized Cobb-Douglas functional form, has been used as in all of the following applications. Hence, all variables are transformed in logarithm. In order to control for the impact on the hospitals' activity of some exogenous factors, we include in the linearized production function a dummy variable *Ownership* (Public or Private) and a dummy variable *Specialistic*—i.e., whether the hospital concentrates its activity on a single major diagnostic category—in this case, the dummy

[1] A full description of the data set is provided in Berta *et al.* 2009.

[2] The data set does not include day-hospital discharges. The case-mix is an index specifying the complexity of a discharge, based on the DRG classification.

variable is equal to 1—or not. In the TTT model, the random component b_i models unobserved differences among hospital optimal outputs, the random component u_i models hospitals' long-run inefficiency, and the random component u_{it} captures hospitals' short-run inefficiency. Table 2 presents the results obtained by fitting all the models presented in Section 2.[3]

Table 2 Skew normal stochastic frontiers for the hospital data set

Model	Max. log-lik.	σ_b^2	$\sigma_{2u,1}^2$	σ_{1u}^2	σ_e^2
TTT	744.773	0.032	0.0148	0.006	0.009
TFT	364.362	0.028	-	0.000	0.027
FTT	704.207	-	0.087	0.121	0.001
TTF	729.029	0.035	0.177	-	0.006
FFT	364.033	-		0.089	0.027
FTF	175.908	-	0.011	-	0.004
TFF	364.316	0.029	-	-	0.027
FFF	49.341	-	-	-	0.054

Table 3 shows the value of the log-likelihood ratio test statistic G^2 for the FTT, TFT, and TTF models and the corresponding p-values; furthermore, the last column reports whether the null hypothesis H_0 can be accepted.

Table 3 Log-likelihood ratio tests for the hospital data set

H_1	H_0	G^2	p-value	decision
TTT	TFT	760.822	0	Rejected
TTT	FTT	81.132	0	Rejected
TTT	TTF	31.132	0	Rejected

All the random components are clearly relevant and cannot be dropped. Furthermore, when we compare the two models TTT and FTT it is evident that the omission of the unit random effect capturing latent heterogeneity inflates the estimated variance of the time-independent inefficiency but not that of the time-dependent inefficiency. Hence, omitting the latent heterogeneity has the effect of overstating the firms' long-run inefficiency but not the short-run inefficiency. If instead the nested model TTF is adopted, not considering the long-run inefficiency component has the effect of inflating the variance of the short-run inefficiency terms. Finally, if the model without the short-run inefficiency terms is adopted, the variance of the random-shock

[3] In all estimates, and also in the following applications the R (R Development Core Team 2009) functions SNF_twostage and SNF_maxlik, developed by the authors, were used to fit the models. For the short-run inefficiency component, only the variance for the first year is reported.

component is inflated, so that short-run sources of inefficiency are confounded with random shocks.

These results point out that in the hospital sector the new model TTT improves the analysis of firms' inefficiency, since it identifies the relevance of short-run inefficiencies and highlights long-run sources of inefficiency that should not be confounded with latent heterogeneity. For the model TTT, the estimates of the time-varying variances $\sigma^2_{2u,t}$, $t = 1, 2, ...T$, are reported in Table 4. Notice that these variances decrease over time; this means that hospitals have almost eliminated short-run sources of inefficiency during the observed period.

The predicted random inefficiencies (or inefficiency scores) for four hospitals are reported in Fig. 1. The expected values $E(\exp\{-u_{i0}\}|\boldsymbol{y}_i)$, $E(\exp\{-u_{it}\}|\boldsymbol{y}_i)$, $E(\exp\{-u_{i0} - u_i t\}|\boldsymbol{y}_i)$ (predicted total inefficiency) were computed as shown in Section 3. From the graphs, it clearly emerges that at the beginning of the period almost all inefficiency was due to the time-varying component, while at the end of the period almost all inefficiency was due to the time-invariant component—i.e., the hospitals succeeded in removing almost all short-run sources of inefficiency.

Table 4 Estimated $\sigma^2_{2u,t}$ in model TTT for the HOSPITAL data set

year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
$\sigma^2_{2u,t}$	1.4787	0.0887	0.0402	0.0195	0.0052	0.0009	0	0	0	0

The second application is in the rice producers' sector[4]. In this data set, the output is given by the *annual rice production* (measured in tonnes) of $n = 43$ rice producers located in the Tarlac region of the Philippines, from 1990 to 1997.[5] We include as inputs the *area planted* (measured in hectares), the *labor* employed (measured in days per worker), and the amount of *fertiliser* used in production (measured in kilograms). In this application, the random component b_i models unobserved differences among producers. Table 5 displays the relevant informations on the fitted models.

The log-likelihood ratio tests to compare the model TTT with nested models are detailed in Table 6. The G^2 statistic to test the model FTT against the model TTT is equal to 0.206, and the p-value is equal to 0.350. It follows that model FTT, when compared with the more general TTT model, cannot be rejected. The G^2 statistic to test model TTF against model TTT is 5.286, and the p-value is equal to 0.011. In this case, the comparison highlights the fact that the long-run inefficiency component cannot be dropped. The G^2

[4] We used the Rice panel data set presented in Coelli *et al.* 2005, pp. 325-326. The data set is also available in the *R-package Frontier* developed by Coelli and Henningsen (2010).

[5] The data were collected by the International Rice Research Institute. Details of the survey are provided by Pandey *et al.* (1999).

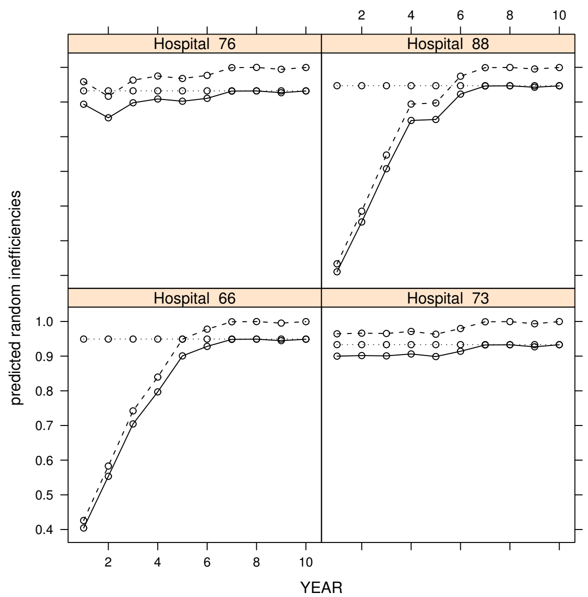

Fig. 1 Predicted hospitals' time-varying ineff. (dashed line), time-invariant ineff. (dotted line) and total ineff. (solid line) for the two most inefficient hospitals (66, 88) and the two most efficient ones (73, 76).

statistic to test model TFT against model TTT is clearly statistically significant, so model TTT does improve the fitting against model TFT. Finally, model FTT cannot be further simplified because both FFT and FTF must be rejected, as clearly emerges from comparing the values of the log-likelihood reported in Table 5. Last, notice that the omission of the time-dependent random-inefficiency component inflates the variance of the idiosyncratic random term (see Table 5). This implies that the presence of short-run sources of inefficiency is confounded with the random shocks.

Table 5 Skew-normal stochastic frontiers for the rice producers' data set

Model	Max. log-lik.	σ_b^2	$\sigma_{2u,1}^2$	σ_{1u}^2	σ_e^2
TTT	-49.707	0.0014	0.1783	0.0612	0.0170
TFT	-86.430	0	-	0.0721	0.0832
FTT	-49.810	-	0.1901	0.0655	0.0183
TTF	-52.350	0.0245	0.2178	-	0.0152
FFT	-86.431	-	-	0.0723	0.0832
FTF	-87.565	-	0.0606	-	0.0887
TFF	-88.605	0.1584	-	-	0.2901
FFF	-104.907	-	-	-	0.3283

Table 6 Log-likelihood ratio tests for the rice producers' data set

H_1	H_0	G^2	p-value	decision
TTT	TFT	75.716	0.000	Rejected
TTT	FTT	0.206	0.350	Accepted
TTT	TTF	5.286	0.011	Rejected

For the best-fitting model FTT, the estimated variances of the time-varying inefficiency components (i.e., $\sigma_{2u,t}^2$, $t = 1, 2, ...T$,) are displayed in Table 7. No evidence of a reduction in short-run efficiency is found in this case, differently from the hospitals' example.

Hence, it emerges from the analysis of firms' production of a standardized good (e.g., rice) with location concentrated in one region (i.e., high proximity between firms), that latent heterogeneity does not play an important role but that both long-run and short-run inefficiency components are important.

Table 7 Estimated $\sigma_{2u,t}^2$ in model FTT for the rice producers' data set

year	1990	1991	1992	1993	1994	1995	1996	1997
$\sigma_{2u,t}^2$	0.190	0.229	0.049	0.101	0.181	0.109	0.292	0.122

The last application covers the airport sector.[6] The sample is composed of $n = 38$ Italian airports and includes the years from 2005 to 2008. The output variable is given by *yearly number of aircraft movements*, and the inputs are given by the *number of runways*, the *total area* of the airports, the *number of check-in desks*, the *number of luggage claim lines*, and the *number of aircraft parking sites*. We take into account the possible impact of the factor *size of the airport*, using the EU classification. Hence, the airports are split into four groups with decreasing size: Great European Airports, National Airports, Domestic Airports, and Regional Airports.

[6] The features of the data set are presented in Malighetti *et al.* (2007).

In this application, in contrast to the previous ones, we can accept the homoskedasticity hypothesis $\sigma^2_{2u,t} = \sigma^2_{2u,1}$, $t = 2, 3, 4$ on the variances of the time-varying inefficiency terms. Indeed, the log-likelihood ratio statistic to test these three constraints in the TTT model is $G^2 = 1.342$, with a p-value equal to 0.719.[7] Hence, different SF models are compared by assuming that the time-varying inefficiency components (when they are included in the model) have constant variance over time. Table 8 reports the results obtained by fitting the SF models under the constraints $\sigma^2_{2u,t} = \sigma^2_{2u,1}$, $t = 2, 3, 4$.

Table 8 Skew-normal stochastic frontiers for the airport data set

Model	Max. log-lik.	σ^2_b	$\sigma^2_{2u,1}$	σ^2_{1u}	σ^2_e
TTT	26.678	0.0493	0	0.3764	0.0159
TFT	26.677	0.0493	-	0.3764	0.0159
FTT	24.108	-	0	0.691	0.0126
TTF	25.699	0.1847	0	-	0.0159
FFT	24.107	-	-	0.6904	0.0162
FTF	-51.132	-	0	-	0.1474
TFF	25.699	0.4298	-	-	0.1261
FFF	-51.132	-	-	-	0.3387

Table 9 Log-likelihood ratio tests for the airport data set

H_1	H_0	G^2	p-value	Decision
TTT	TFT	0.002	0.482	Accepted
TTT	FTT	5.14	0.012	Rejected
TTT	TTF	1.958	0.081	Accepted

In this example, the relevance of the latent heterogeneity component emerges from the comparison of model TTT and model FTT, as shown in Table 9 ($G^2 = 5.14$, p-value $= 0.012$). The conclusion of rejecting model FTT is supported by the fact that the omission of the latent heterogeneity component inflates the estimate of σ^2_{1u}. The non-relevance of the time-invariant inefficiency is suggested by the comparison of model TTT with model TTF ($G^2 = 1.958$, $p = 0.081$). However, model TTF cannot be recommended because the omission of the time-invariant component inflates the estimates of the variance σ^2_b. The model TFT, when tested against TTT, gives $G^2 = 0.002$ and $p = 0.482$, so the time-dependent inefficiency can be dropped from TTT. Hence, this is a confirmation that when the panel data have a relatively

[7] Similar results are obtained when the same hypothesis is tested in the models FTF, TTF, and FTT.

short time horizon the impact of short-run inefficiency is limited, as correctly assumed in the previous contributions (e.g., Greene 2005).

5 Conclusions

We have presented a new SF model with four random components that allows us to disentangle latent heterogeneity from inefficiency, and to investigate the impact of long-run and short-run sources of inefficiency on firms' performances. Long-run inefficiency stays with the firm over time, while short-run inefficiency may vary in each period. We have applied the new model to analyze the efficiency of firms operating in three different sectors (hospitals, rice producers, and airports), and we have shown that the new model is particularly appropriate when firms are heterogeneous and when the panel is long. We have modeled short-run sources of inefficiency with independent random components. As Greene (2005, p. 29) underlines, "it seems reasonable to assert that if nothing else, there is some inertia (autocorrelation) in short run inefficiency." Hence, a possible generalization of our model is the inclusion of some dependency among the random components capturing the impact of short-run sources of inefficiency in a firm. It seems that this can be achieved by replacing the diagonal matrix $\boldsymbol{\Psi}$, introduced in Section 3.2, with an appropriate non-diagonal matrix. The specification of this non-diagonal matrix and the utility of the generalization needs further investigation.

Appendix

Proof of Proposition 1. In the TTT SF model presented in Eq. (1), the random components $b_i - u_{i0} - u_{it} + e_{it}$ can be written as the sum of the time-independent terms (i.e., $\varepsilon_i = b_i - u_{i0}$) and of the time-dependent terms (i.e., $\varepsilon_{it} = e_{it} - u_{it}$). According to our assumptions, these two terms are independent in probability and are given by the difference of a normal random variable and an independent left-truncated-at-zero normal random variable. It is well known (Kumbhakar and Lovell 2000) that ε_i has the following density:

$$f(\varepsilon_i) = 2\phi_1(\varepsilon_i, 0, \sigma_b^2 + \sigma_{1u}^2)\bar{\Phi}_1\left(\frac{-\sigma_{1u}^2}{\sigma_b^2 + \sigma_{1u}^2}\varepsilon_i, \sigma_{1u}^2\left(1 - \frac{\sigma_{1u}^2}{\sigma_b^2 + \sigma_{1u}^2}\right)\right).$$

Analogously, the densities of ε_{it} $(t = 1, 2, ..., T)$ are:

$$f(\varepsilon_{it}) = 2\phi_1(\varepsilon_{it}, 0, \sigma_e^2 + \sigma_{2u,t}^2)\bar{\Phi}_1\left(\frac{-\sigma_{2u,t}^2}{\sigma_e^2 + \sigma_{2u,t}^2}\varepsilon_{it}, \sigma_{2u,t}^2\left(1 - \frac{\sigma_{2u,t}^2}{\sigma_e^2 + \sigma_{2u,t}^2}\right)\right).$$

The previous two densities are $(1, 1)$ closed-skew normal densities, and so the random components $b_i - u_{i0} - u_{it} + e_{it}$ of the vector $\mathbf{1}_T b_i + \boldsymbol{A}\boldsymbol{u}_i + \boldsymbol{e}_i = -\boldsymbol{A}(\varepsilon_i, \varepsilon_{i1}, ..., \varepsilon_{iT})'$ are sums of two closed-skew normal random errors.

Let $\boldsymbol{\Upsilon}$ be the diagonal matrix with the ratios $\frac{\sigma_{2u,t}^2}{\sigma_e^2+\sigma_{2u,t}^2}$ on the main diagonal. From Theorem (3) of Gonzáles-Farías *et al.* (2004), it follows that the $T+1$ independent random variables $\varepsilon_i, \varepsilon_{it}$ $(t = 1, 2, ..., T)$ have a joint $(T+1, T+1)$ closed-skew normal density function with parameters $\boldsymbol{\nu}_0 = \mathbf{0}$, $\boldsymbol{\mu}_0 = \mathbf{0}$, and

$$\boldsymbol{\Gamma}_0 = \begin{bmatrix} \sigma_{1u}^2 + \sigma_b^2 & \mathbf{0}_T' \\ \mathbf{0}_T & \boldsymbol{\Psi} + \sigma_e^2\boldsymbol{I}_T \end{bmatrix}, \ \boldsymbol{D}_0 = \begin{bmatrix} \frac{-\sigma_{1u}^2}{\sigma_b^2+\sigma_{1u}^2} & \mathbf{0}_T' \\ \mathbf{0}_T & -\boldsymbol{\Upsilon} \end{bmatrix},$$

$$\boldsymbol{\Delta}_0 = \begin{bmatrix} \sigma_{1u}^2(1 - \frac{\sigma_{1u}^2}{\sigma_b^2+\sigma_{1u}^2}) & \mathbf{0}_T' \\ \mathbf{0}_T & \boldsymbol{\Psi} - \boldsymbol{\Upsilon} \end{bmatrix}.$$

Finally, from the previous result and from Theorem (1) of Gonzáles-Farías *et al.* (2004), it follows that the T dimensional random vector $\mathbf{1}_T b_i + \boldsymbol{A}\boldsymbol{u}_i + \boldsymbol{e}_i = -\boldsymbol{A}(\varepsilon_i, \varepsilon_{i1}, ..., \varepsilon_{iT})'$ with components $b_i - u_{i0} - u_{it} + e_{it} = \varepsilon_i + \varepsilon_{it}$ has a $(T, T+1)$ closed-skew normal distribution with parameters $\boldsymbol{\nu} = \mathbf{0}$, $\boldsymbol{\mu}_1 = \mathbf{0}$, $\boldsymbol{\Gamma} = \boldsymbol{A}\boldsymbol{\Gamma}_0\boldsymbol{A}' = \boldsymbol{\Sigma} + \boldsymbol{A}\boldsymbol{V}\boldsymbol{A}'$, $\boldsymbol{D} = \boldsymbol{D}_0\boldsymbol{\Gamma}_0\boldsymbol{A}'\boldsymbol{\Gamma}^{-1} = \boldsymbol{R}$, $\boldsymbol{\Delta} = \Delta_0 + \boldsymbol{D}_0\boldsymbol{\Gamma}_0\boldsymbol{D}_1' - \boldsymbol{D}_0\boldsymbol{\Gamma}_0\boldsymbol{A}'\boldsymbol{\Gamma}^{-1}\boldsymbol{A}\boldsymbol{\Gamma}_0\boldsymbol{D}_0' = \boldsymbol{\Lambda}$. Because $\boldsymbol{\mu}$ is a location parameter and $\boldsymbol{y}_i = \mathbf{1}_T\beta_0 + \boldsymbol{X}_i\boldsymbol{\beta} + \mathbf{1}_T b_i + \boldsymbol{A}\boldsymbol{u}_i + \boldsymbol{e}_i$, the statement of the Proposition follows.

Proof of Proposition 3. To prove b) note that:

$$f(\boldsymbol{u}_i|\boldsymbol{y}_i) = \frac{\phi_T(\boldsymbol{y}_i, \mathbf{1}_T\beta_0 + \boldsymbol{X}_i\boldsymbol{\beta}, \boldsymbol{\Sigma} + \boldsymbol{A}\boldsymbol{V}\boldsymbol{A}')\phi_q(\boldsymbol{u}, \boldsymbol{R}\boldsymbol{r}_i, \boldsymbol{\Lambda})2^q}{f(\boldsymbol{y}_i)} \tag{9}$$

and use the result of Proposition 1 about $f(\boldsymbol{y}_i)$. Noting that $f(\boldsymbol{u}_i|\boldsymbol{y}_i)$ is the density of a left-truncated-at-zero multi-normal random variable, d) follows from Lemma 13.6.1 of Dominguez-Molina *et al.* (2004). To prove a) we observe that:

$$\begin{aligned} f(b_i|\boldsymbol{y}_i) = & \quad \phi(b_i, \sigma_b^2\mathbf{1}'\boldsymbol{\Delta}\boldsymbol{r}_i, \tilde{\sigma}_b^2) \times \\ & \times \frac{\int_0^\infty ... \int_0^\infty \phi_q(u, \boldsymbol{R}\boldsymbol{r}_i - \boldsymbol{R}\mathbf{1}_T\sigma_b^2\tilde{\sigma}_b^{-2}\tilde{b}_i, \tilde{\boldsymbol{\Lambda}})du_0...du_T}{\int_0^\infty ... \int_0^\infty \phi_q(\boldsymbol{u}, \boldsymbol{R}\boldsymbol{r}_i, \boldsymbol{\Lambda})du_0...du_T} \end{aligned} \tag{10}$$

Point a) of the proposition follows immediately. Observing that $f(b_i|\boldsymbol{y}_i)$ is a closed-skew normal density, point c) follows from the result (3) on the moment-generating function of a closed-skew normal random variable.

References

1. Arellano-Valle E, Bolfarine H, Lachos H (2005) Skew-normal linear mixed models. J. Data. Sci. **3**: 415-438
2. Arellano-Valle E, Azzalini A (2006) On the unification of families of skew-normal distributions. Scand. J. Stat. **33**: 561-574
3. Berta P, Callea G, Martini G, Vittadini G (2009) The effects of upcoding, cream skimming and readmissions on the Italian hospitals efficiency: a population-based investigation. Econ. Mod. **27**: 812-821
4. Coelli T, Henningsen A (2010) Frontier: Stochastic Frontier Analysis. R package version 0.996-6 http://www.R-project.org
5. Coelli T, Rao D, O'Donnell C, Battese G (2005) An introduction to efficiency and productivity analysis. Springer, New York
6. Dominguez-Molina A, Gonzales-Farias G, Ramos-Quiroga R (2004) Skew normality in stochastic frontier analysis. In: Genton, M. (ed) Skew Elliptical Distributions and their Applications. Chapman and Hall CRC, London
7. Genz A, Bretz F (2009) Computation of multivariate normal and t probabilities. Springer, New York
8. Gonzáles-Farías G, Domínguez Molina A, Gupta A (2004) Additive properties of skew normal random vectors. J. Stat. Plan. Infer. **126**: 521-534
9. Greene W (2005) Fixed and random effects in stochastic frontier models. J. Productivity Analysis. **23**: 7-32
10. Kumbhakar S, Lovell K (2000) Stochastic frontier analysis. Cambridge University Press, Cambridge
11. Kumbhakar S, Heshmati A (1995) Efficiency measurement in Swedish dairy farms: An application of rotating panel data, 1976-1988. American J. Agricultural Econ., **77**: 660-674
12. Lin T, Lee C (2005) Estimation and prediction in linear mixed models with skew-normal random effects for longitudinal data. Statist.Med. **27**: 1490-1507
13. Malighetti P, Martini G, Paleari S, Redondi R (2007) An empirical investigation on the efficiency, capacity and ownership of Italian airports. Rivista di Politica Economica. **47**: 157-188
14. Pandey S, Masicat P, Velasco L, Villano R (1999) Risk analysis of the rainfed rice production system in Tarlac, Central Luzon, Philippines. Experimental Agriculture, 35: 225-237
15. Parke W (1986) Pseudo maximum likelihood estimation: the asymptotic distribution. Ann.Stat. **14**: 355-357
16. Pitt M, Lee L (1981) Measurement of sources of technical inefficiency in the Indonesian weaving industry. J.Dev.Econ. **9**: 43-64
17. R Development Core Team (2009) R: A language and environment for statistical computing. R Foundation for Statistical Computing ISBN 3-900051-07-0 http://www.R-project.org
18. Silvapulle M, Sen P (2005) Constrained Statistical Inference. Wiley, Hoboken, NJ