Persistent and Transient Productive Inefficiency: A Maximum Simulated Likelihood Approach

Version: 20.5.2014

Massimo Filippini
Department of management, technology and economics,ETH ZurichandDepartment of Economics, University of Lugano, Switzerland

and

William Greene
Department of Economics
Stern School of Business
New York University
USA

Abstract

The productive efficiency of a firm can be seen as composed of two parts, one persistent and one transient. The received empirical literature on the measurement of productive efficiency has paid relatively little attention to the difference between these two components. Ahn, Good and Sickles (2000) suggested some approaches that pointed in this direction. The possibility was also raised in Greene (2004), who expressed some pessimism over the possibility of distinguishing the two empirically. Recently, Colombi et al. (2010), and Kumbhakar and Tsionas (2012), in a milestone extension of the stochastic frontier methodology have proposed a tractable model based on panel data the promises to provide separate estimates of the two components of efficiency. The approach developed in the original presentation proved very cumbersome actually to implement in practice. The latter study also suggested a partial solution, but stopped short of a full practical implementation of the original model. (The sensitivity of the Bayesian approach to the need to have informative priors over the efficiency distribution, which is the main object of estimation, also remains to be settled.) Finally, Kumbhakar, Lien and Hardaker (2012), in their survey of panel data models, suggested a method of moments estimator for the LR/SR model based on simple OLS. In this paper, we develop a practical full information maximum simulated likelihood estimator for the model. The approach is very effective and strikingly simple to apply, and uses all of the sample sistributional information to obtain the estimates. We alsoimplement the panel data counterpart of the JLMS (1982) estimator for technical or cost inefficiency. The technique is applied in a study of the cost efficiency of Swiss railways.

Keywords: Productive efficiency, Stochastic frontier analysis, Panel data, transient and persistent efficiency

JEL Classification:

1. Introduction

The productive efficiency of a firm can be seen as composed of two parts, one persistent and one transient. The persistent part is related to the presence of structural problems in the organization of the procuction process of a firm or the presende of systematikc shortralls in managerial capabilities. The transient part may be due to the presence of non-systematic management problems that can be solved in the short term. The received empirical litarature on the measurement of productive efficiency has paid relatively little attention to the different between these two components of productive efficiency.[1] Generally, the studies using stochastic frontier models for panel data do recognize that some econometric models produce indicators of efficiency that vary over time, whereas other econometric modelsprovidethe estimation of time invariant indicators of efficiency.[2] However, these studies generally do not address the possibility that the productive efficiency can be split into two parts, i.e. transient and persistent.

Some studies utilize several stochastic frontier models for panel data and comparethe ranking and the values of the estimated indicators of efficiency (e.g., Farsi et. al. (2005b), Abdulai and Tietje (2007), Faust and Baranzini (2014)). This comparison, usually performed by calculating a rank correlation coefficient of the index of efficiency, will ignore the possibility that the values obtained with the time invariant models reflect something different from those obtained from models with time varying inefficiency.

Ahn, Good and Sickles (2000) suggested an approach that pointed in the direction of distinguishing short-run and long-run efficiency levels. For this purpose, they proposed to use a stochastic frontier model with an autoregressive specification. The possibility was also raised in Greene (2004), who expressed some pessimism over the possibility of distunguishing empirically the persistent and transient part of the productive efficiency. Recently, Colombi et al. (2011, 2014) and Kumbhakar and Tsionas (2012), in a milestone extension of the stochastic frontier methodology, have proposed a tractable model based on panel data that promises to provide separate estimates of the two components of efficiency. As suggested by Kumbhakar and Tsionas (2012), we will call this four-way error component

[1] Books and surveys on the measurement of the level of productive efficiency do make the distinction between models that estimate time-varying inefficiency indicators and models that produce time invariant indicators. See for instance Kumbhakar and Lovell (2000) or Greene (2008). Measurement of the distinct parts has proved challenging.

[2] For a discussion of these models see Greene (2008).

stochastic frontier model the Generalized True Random Effects model (GTRE). The approached developed by Colombi et al. (2011, 2014) in the original presentation proved extremely cumbersome to actually implement in practice. Kumbhakar and Tsionas (2012) proposed a partial Bayesian Solution, but stopped short of a full practical implementation of the MLE. (The sensitivity of the Bayesian approach to the need to have informative priors over the main objects of estimation remains to be settled.) Finally, Kumbhakar, Lien and Hardaker (2012), in their survey of panel data models, proposed a method of moments estimator for the LR/ST models based on simple OLS.

The goal of this paper is to provide an alternative econometric approach for the estimation of the GTREM based on maximum simulated likelihood that allows the distinction between persistent and transient levels of efficiency. As we will discuss later in the paper, the advantage of this approach is in the transparency and effectiveness of the estimation procedure.

The paper is organized as follows. The next section presents a short overview of the most important stochastic frontier models for the estimation of the transient part of efficiency and of the persistent part of efficiency, while section 3 discusses the novel model/estimator based on maximumsimulated likelihood. Section 4 illustrates the application of this new stochastic frontier model using a public available data set on the cost of a sample of Swiss railway companies. The final section contains a summary and conclusion.

2. StochasticFrontier Models for the Estimation of the Persistent or the Transient Part of Productive Inefficiency

There are several different panel data stochastic frontier model (SFA) specifications that have been considered for the econometric estimation of one of the two components of the productive efficiency. Some will estimate the time invariant values of productive efficiency that tend to reflect the persistent part of the level of productive efficiency. Others estimate time varying values of productive efficiency that tend to capture the transient component. Thesereceived models do not provide the information if a firm is characterized by the presence of both parts of the productive inefficiency.

Most of these frontier models using panel data are based on the fixed and the random effects models.For ourapplication, we will consider the estimation of a cost frontier. The first stochastic frontier modelthat specifically developedthe persistent part of inefficiency is Pitt and Lee (1981)(hereafter RE). They specified a model in which the inefficiency term u_i is assumed to be constant through time:

$$\begin{aligned} &\ln C_{it} = \alpha + \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it} + u_i, \\ &v_{it} \sim N[0,\sigma_v^2], \\ &u_i = |U_i|,\ U_i \sim N[0,\sigma_u^2], \\ &\varepsilon_{it} = v_{it} + u_i, \end{aligned} \qquad (1)$$

where $C_{it}$is total cost incurred by company $i$in year t,$\mathbf{x}_{it}$ is a vector of outputs and input prices in logs, $\boldsymbol{\beta}$ is the associated vector of parameters to be estimated; u_i, is a one-sided non-negative disturbance measuring the level of inefficiency, andv_{it}, is a symmetric disturbance representing the random noise. Usually v_{it} is assumed to be normally distributed, while the inefficiency term,$u_i$is assumed to follow a half-normal distribution.[3] This model, due to Pitt and Lee (1981), is a variant of the classical random effects model for panel data, where the individual effects are assumed to have a specific, non-normal distribution.

Several variants of the Pitt and Lee (1981) model have been proposed. to accommodate time variation in the inefficiency term. Most of these specify that the inefficiency term can be represented as a product of a deterministic function of time and the random effects, u_i,with the theone-sided non-negative disturbancesnow reflecting the transient effect of inefficiency. For instance, Kumbhakar (1990) specifies the inefficiency term as $u_{it} = [1+\exp(bt+ct^2)]^{-1}|U_i|$, Battese and Coelli (1992) as $u_{it} = \exp[-\eta(t-T)]|U_i|$, Battese and Coelli (1995) as $u_{it} = \exp[g(t,T,\mathbf{z}_{it})]|U_i|$ and Cuesta (2000) as $u_{it} = \exp[-\eta_i(t-T)]|U_i|$. We note at least some ambiguity in the interpretation of u_{it}in these contexts as time varying or time fixed inefficiency.

Schmidt and Sickles (1984) propose a model that estimate the persistent part of the inefficiency without specifying an explicit distribution of the inefficiency as in Pitt and Lee (1981). They propose to reinterpret the linear fixed effect model as:

$$\begin{aligned} &\ln C_{it} = \alpha_i + \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it}, \\ &v_{it} \sim N[0,\sigma_v^2], \end{aligned} \qquad (2)$$

[3] Other extensions of the basic frontier model have also considered exponential and truncated normal distributions for the inefficiency term. See for instance Battese and Coelli (1992).

where inefficiency $\hat{u}_i$ is computed from $\hat{u}_i = \left[\min_i\left(\hat{\alpha}_i\right)\right] + \hat{\alpha}_i$ where $\hat{\alpha}_i$ is the ithfixed effects estimate in the within groups fixed effects linear regression model. The model proposed by Schmidt and Sickles (1984) has been extended by Cornwell, Schmidt and Sickles (1990) in order to estimate a transient part of inefficiency without specifying an explicit distribution of the inefficiency. In this model, the transient part of the inefficiency is defined as $\alpha_{it} = \alpha_{0i} + \alpha_{1i}t + \alpha_{2i}t^2$. This interpretation would seem to remove the ambiguity in the interpretation of u_{it}. Since both the time invariant components, α_{0i} and the time varying components, $\alpha_{1i}t + \alpha_{2i}t^2$ are firm specific, this model could suggest a decomposition. It has two features that warrant further investigation: first, it leaves unaccounted for time invariant heterogeneity that is not inefficiency; second, it places a very strong assumption on the trajectory of time varying inefficiency – a firm specific quadratic function of time. (We do note, the authors did not make the distinction between permanent and transient inefficiency in their development. Cornwell, Schmidt and Sickles (1990) treated the estimator as time varying inefficiency.)

All these models that estimate a level of productive efficiency that varies over time are variants of the random or fixed effects models. Therefore, it is not completely clear if these models are really able to isolate the transient part of inefficiency, because all these inefficiency measures include a persistent partu_i or α_i. Moreover, in all these models any unobserved, time-invariant, individual-specific heterogeneity is captured by $\alpha_i$orα_{0i}, and therefore, considered as inefficiency. Consequently, these models tend generally to underestimate the level of efficiency.

Greene (2005a and 2005b) proposes two models based on the extension of the panel data version of the Aigner, Lovell and Schmidt (1977) half normal model by adding to the classical stochastic frontier model firm specific time-invariant effects. These models, called true fixed effects model (hereafter TFE) and true random effects model (hereafter TRE), include a term for time invariant unmeasured unobserved heterogeneity, a random noise term and a firm-specific inefficiency term. These models help to separate unobserved time-invariant effects from time-varying efficiency estimates. Therefore, the efficiency estimates obtained with these models provide information on the transient component of productive efficiency. The generic formulation of the model is

$$
\begin{aligned}
\ln C_{it} &= (\alpha + w_i) + \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it} + u_{it}, \\
v_{it} &\sim N[0, \sigma_v^2], \\
u_{it} &= |U_{it}|,\ U_{it} \sim N[0, \sigma_u^2].
\end{aligned}
$$

In the true fixed effects model, w_i represents time invariant heterogeneity that might be correlated with the included variables, $\mathbf{x}_{it}$. The estimator is simply pooled SFA with firm dummy variables added to the model to accommodate w_i. The characteristics of the TFE model have been examined in some recent studies, such as Chen et al. (2013) who develop an ML estimator based on likelihood function that applies after the 'within groups' transformation removes h_i. The TRE model is

$$
\begin{aligned}
\ln C_{it} &= \alpha_i + \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it} + u_{it}, \\
\alpha_i &= \alpha + w_i,\ \ w_i \sim N[0, \sigma_w^2], \\
v_{it} &\sim N[0, \sigma_v^2], \\
u_{it} &= |U_{it}|,\ U_{it} \sim N[0, \sigma_u^2], \\
\varepsilon_{it} &= v_{it} + u_{it}.
\end{aligned} \tag{3}
$$

The model is estimated bymaximum simulated likelihood. The term w_i is an *i.i.d.* random component in random-effects framework. The inefficiency term is assumed to be an *iid*random variable with a specific non-normal distribution (half-normal, exponential or truncated-normal distribution). This implies that the inefficiency is varying over time.

In the TRE and TFE settings, any time-invariant or persistent component of inefficiency is completely absorbed in the individual-specific constant term. Therefore, for example, in contexts characterized by certain sources of efficiency that result in time-invariant excess of inputs, the estimates of these models could be expected to provide relatively high levels of efficiency.

The TRE can also suffer from the 'omitted variables bias', because the unobserved variablesmay be correlated with the regressors. (This motivates the FTE approach.) In order to solve this problem, Farsi et al. (2005b) suggest to use in the TRE an auxiliary equation introduced by Mundlak (1978). The use of this auxiliary equation allows considering the econometric problem of unobserved heterogeneity bias. The auxiliary equation is given by:

$$
\alpha_i = \alpha + \boldsymbol{\phi}'\overline{\mathbf{x}}_i + w_i,\ \ w_i \sim iid(0, \sigma_w^2) \tag{4}
$$

where $\bar{\mathbf{x}}_i$ is the vector of the firm means of all the time varying explanatory variables and $\boldsymbol{\phi}$ is the corresponding vector of coefficients. Equation (4) can be incorporated in the TRE.[4]

From this short review of some previous models, we can observe that there is an interest in an econometric model that allows at oncethe estimation of the persistent and transient parts of the productive efficiency. The recent papers on this topic mentioned in the introduction (Colombi (2010), Colombietal.(2011, 2014), Kumbhakar and Tsionas (2012), Kumbhakar, Lien and Hardaker (2012)), provide a theoretical platform on which to distinguish persistent from transient inefficiency. In what follows, we suggest a practical completion to the development by proposing a straightforward, transparent empirical estimation method.

3. Maximum Simulated Likelihood Estimation of the Generalized True Random Effects Model

The generic normal – half normal stochastic production frontier model is

$$\begin{aligned} y_{it} &= \alpha + \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it} - u_{it} \\ &= \alpha + \boldsymbol{\beta}'\mathbf{x}_{it} + \varepsilon_{it} \end{aligned} \tag{5}$$

where v_{it} is normally distributed with mean zero and variance σ_v^2 and$u_{it} = |U_{it}|$ where U_{it} is normally distributed with mean zero and variance σ_u^2. In the second line, ε_{it} has a two parameter skew normal distribution with parameters $\lambda = \sigma_u/\sigma_v$ and $\sigma = (\sigma_v^2 + \sigma_u^2)^{1/2}$. [See Azzalini (1985).] The log likelihood for this stochastic frontier model is

$$\begin{aligned} \log L(\alpha,\boldsymbol{\beta},\lambda,\sigma) &= \sum_{i=1}^{N} \left[\begin{array}{l} \log \dfrac{2}{\sigma} + \log \phi\left(\dfrac{y_{it} - \alpha - \boldsymbol{\beta}'\mathbf{x}_{it}}{\sigma} \right) \\ + \log \Phi\left(\dfrac{-(y_{it} - \alpha - \boldsymbol{\beta}'\mathbf{x}_{it})\lambda}{\sigma} \right) \end{array} \right] \\ &= \sum_{i=1}^{N} \left[\log \left\{ \frac{2}{\sigma} \phi\left(\frac{\varepsilon_{it}}{\sigma} \right) \Phi\left(\frac{-\varepsilon_{it}\lambda}{\sigma} \right) \right\} \right] \end{aligned} \tag{6}$$

where $\phi(z)$ is the standard normal density and $\Phi(z)$ is the standard normal cdf. [See Aigner, Lovell and Schmidt (1977).] The form in the second line displays the skew normal distribution of ε_{it}.[5]

[4] The Mundlak auxiliary equation has been proposed for a random effects linear regression model. This approach, based on normality and the linear model,might not strictly apply to stochastic frontier models estimated by ML, as these models possess an asymmetric composite error term ε_i. As the model captures the correlation between the individual effects and the explanatory variables at least partly, the resulting heterogeneity bias is expected to be minimal. The general approach has been used elsewhere in a variety of settings under the heading of 'correlated random effects models.' See, e.g., Wooldridge (2010).

The true random effects (TRE) model [Greene (2005)] adds a time invariant *random effect* to the normal-half normal stochastic frontier model;

$$y_{it} \quad = \alpha + w_i + \boldsymbol{\beta}'\mathbf{x}_{it} + v_{it} - u_{it} \tag{7}$$

where v_{it} and u_{it} are as defined earlier and the time invariant w_i is normally distributed with mean zero and variance σ_w^2. The random effect in (7) is understood to capture persistent firm level heterogeneity, not inefficiency. The log likelihood for the TRE model is formed as follows: Conditioned on w_i, the T observations for firm i are independent.[6] [See Butler and Moffitt (1982).] The conditional density is

$$f(\varepsilon_{i1},\ldots,\varepsilon_{iT}|w_i) = \prod_{t=1}^{T} \frac{2}{\sigma}\phi\left(\frac{(\varepsilon_{it}+w_i)}{\sigma}\right)\Phi\left(\frac{-(\varepsilon_{it}+w_i)\lambda}{\sigma}\right) \tag{8}$$

The unconditional density that will form the basis for MLE is obtained by integrating out w_i;

$$f(\varepsilon_{i1},\ldots,\varepsilon_{iT}) = \int_{-\infty}^{\infty} \left\{\prod_{t=1}^{T} \frac{2}{\sigma}\phi\left(\frac{(\varepsilon_{it}+w_i)}{\sigma}\right)\Phi\left(\frac{-(\varepsilon_{it}+w_i)\lambda}{\sigma}\right)\right\}\frac{1}{\sigma_w}\phi\left(\frac{w_i}{\sigma_w}\right)dw_i. \tag{9}$$

It is convenient to use a change of variable and write $w_i = \sigma_w W_i$ where W_i is normally distributed with mean zero and variance one. Combining terms, the log likelihood for the true random effects model is

$$\log L(\alpha,\boldsymbol{\beta},\lambda,\sigma,\sigma_w) = \sum_{i=1}^{N} \log\int_{-\infty}^{\infty} \left\{\prod_{t=1}^{T} \left[\begin{array}{l}\frac{2}{\sigma}\phi\left(\frac{y_{it}-\alpha-\boldsymbol{\beta}'\mathbf{x}_i-\sigma_w W_i}{\sigma}\right)\times \\ \Phi\left(\frac{-(y_i-\alpha-\boldsymbol{\beta}'\mathbf{x}_i-\sigma_w W_i)\lambda}{\sigma}\right)\end{array}\right]\right\}\phi(W_i)dW_i \tag{10}$$

The integral does not exist in closed form, but it can be evaluated by simulation. The *simulated log likelihood* is

$$\log L_S(\alpha,\boldsymbol{\beta},\lambda,\sigma,\sigma_w) = \sum_{i=1}^{N} \log\frac{1}{R}\sum_{r=1}^{R} \left\{\prod_{t=1}^{T} \left[\begin{array}{l}\frac{2}{\sigma}\phi\left(\frac{y_{it}-\alpha-\boldsymbol{\beta}'\mathbf{x}_{it}-\sigma_w W_{ir}}{\sigma}\right)\times \\ \Phi\left(\frac{-(y_{it}-\alpha-\boldsymbol{\beta}'\mathbf{x}_{it}-\sigma_w W_{ir})\lambda}{\sigma}\right)\end{array}\right]\right\}. \tag{11}$$

---

[5]A stochastic cost frontier function will result from the simple change of $-u_{it}$ to $+u_{it}$ in (5). A few sign changes will also result in the log likelihood function. See, e.g., the survey in Greene (2008).

[6]There is no requirement that the number of observations, T_i, be the same for each firm. T_i is assumed here to be constant only for convenience to avoid another subscript in the presentation.

In the inner summation, W_{ir} is R simulated draws from the standard normal population. (In our applications, we use Halton sequences rather than pseudo-random numbers.) We rely on received results for properties of the MSLE. [See, e.g., Train (2003).] Derivatives for gradient based optimization and for computing the estimator of the asymptotic covariance matrix are also simulated. The model is otherwise conventional, and satisfies the regularity conditions that underlie familiar maximum likelihood estimation. Full results for the true random effects estimator appear in Greene (2004, 2005) and references therein.[7]

We now consider the 'Generalized True Random Effects' model (GTRE). The extension of the model adds to the TRE model above a time persistent counterpart to u_{it}in the time varying stochastic frontier. The two level stochastic frontier model is

$$y_{it} = \alpha + (w_i - h_i) + \boldsymbol{\beta}'\mathbf{x}_{it} + (v_{it} - u_{it}) \tag{12}$$

The random components v_{it}, u_{it} and $w_i$are as defined earlier while $h_i = |H_i|$ has a half normal distribution with underlying variance σ_h^2. The form in (12) might appear to include a four part disturbance with two time varying components and two time invariant components. [See, e.g., Colombi et al. (2011), Sec. 2, p. 4 where it is described as such.] By this view, identification and estimation would seem to be optimistic in the extreme. The crucial insight is that it is not a four part disturbance; it is a two part disturbance, one time varying, one time invariant, in which each of the two parts hasits own skew normal distribution rather than normal distribution. I.e., it is a random effects model with skew normal error components. Thus,

$\varepsilon_{it} = (v_{it} - u_{it})$ has the now familiar skew normal distribution with parameters σ and λ shown earlier while $\delta_i = (w_i - h_i)$ also has a skew normal distribution with parameters

$$\gamma = \sigma_h/\sigma_w \text{ and } \theta = (\sigma_w^2 + \sigma_h^2)^{1/2}. \tag{13}$$

The full unconditional log likelihood function for this model based on the joint distribution of $(\varepsilon_{i1},\ldots,\varepsilon_{iT},\delta_i)$ is derived by Colombi (2010) and Colombi et al. (2011). Before considering our preferred method of estimation, it is useful to show their result in detail. For the GTREM, the full log likelihood is

$$\log L(\alpha,\boldsymbol{\beta},\sigma_v,\sigma_u,\sigma_w,\sigma_h) = n(T+1)\log 2 + \sum_{i=1}^{N} \begin{bmatrix} \log \phi_T(\mathbf{e}_i \mid \mathbf{0}_T, \boldsymbol{\Sigma} + \mathbf{AVA}') + \\ \log \bar{\Phi}_{T+1}(\mathbf{Re}_i, \boldsymbol{\Lambda}) \end{bmatrix} \tag{14}$$

[7] The estimation strategy in (11) could, in principle, be applied to estimation of the stochastic frontier model in (6) by integrating u_{it}out of the conditionally normal linear regression model in (5). Maximization of (6) directly is extremely straightforward, however, and the MSLE would provide no improvement over direct MLE.

where

$$\mathbf{y}_i = \begin{pmatrix} y_{i1} \\ \vdots \\ y_{iT} \end{pmatrix},\ \mathbf{X}_\mathrm{i} = \begin{bmatrix} \mathbf{x}'_{i1} \\ \vdots \\ \mathbf{x}'_{iT} \end{bmatrix},\ (\mathbf{1}_T, \mathbf{0}_T) = \begin{pmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \end{pmatrix},\ \mathbf{I}_T = T \times T \text{ identity matrix, } \mathbf{A} = -[\mathbf{1}_T \mid \mathbf{I}_T],$$

$$\mathbf{e}_i = \mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta} - \mathbf{1}_T\alpha$$

$$\boldsymbol{\Sigma} = \sigma_v^2\mathbf{I}_T + \sigma_w^2\mathbf{1}_T\mathbf{1}'_T, \mathbf{AV} = \begin{bmatrix} \sigma_h^2\ \mathbf{R} & \mathbf{0}'_T\ \mathbf{A}\boldsymbol{\Sigma} \\ \mathbf{0}_T & \sigma_u^2\mathbf{I}_T \end{bmatrix},\ \boldsymbol{\Lambda} = \left[\mathbf{V}^{-1} + \mathbf{A}' \quad ^{-1} \quad \right]^{-1}, \qquad = \boldsymbol{\Lambda} \quad ' \quad ^{-1}$$

and ϕ_T ($\mathbf{z}|\boldsymbol{\mu},\boldsymbol{\Omega}$) denotes a T-variate normal density evaluated at $\mathbf{z}$ for random vector with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Omega}$ while $\overline{\Phi}_{T+1}(\boldsymbol{\tau},\boldsymbol{\Gamma})$ is the joint probability that a T+1-variate normal vector with mean vector $\boldsymbol{\tau}$ and covariance matrix $\boldsymbol{\Gamma}$ belongs to the nonnegative orthant.[8] Note that the full unconditional log likelihood involves the T-variate normal density and T+1 variate normal integrals in both of which the covariance matrix (which is, itself a complex function of the underlying parameters) is one of the objects of estimation. Maximization of log likelihoods of this sort is a notoriously challenging exercise.

Colombi (2010) notes that FIML estimation of the model is 'complex and time consuming.'[9] In the sequence of papers, Colombi (2010), Colombi et al. (2012), Kumbhakar, Lien and Hardaker (2012) and Kumbhakar and Tsionas(2012) have suggested other strategies, including a four step least squares method. The main point of this paper is that estimation of the model is neither complex nor time consuming. The extreme complexity of the log likelihood noted in Colombi is reduced by using simulation and exploiting the Butler and Moffitt (1982) formulation.

The obstacle to FIML estimation is the extreme complexity of multivariate normal integrals involving traces and determinants of matrices in the optimization. In fact, estimation of this extended model is no more complicated than the TRE model noted earlier. The GTRE model is simply a TRE model in which the time invariant effect has a skew normal distribution, rather than a normal distribution as assumed earlier. It is trivial to simulate draws from a skew normal as simply the sum of a normal minus (or plus) the absolute value of a normal draw. By the natural extension, the log likelihood function for the GTRE model is

---

[8]The authors allowed σ_u^2 to vary by period. This aspect can be added to the estimating equations by changing $\sigma_u^2\mathbf{I}$ to $\boldsymbol{\Psi}$ = diag($\sigma_{u,1}^2,\ldots,\sigma_{u,T}^2$) in the definition of $\mathbf{V}$. In the formulations below, the parameterization in terms of σ and λ would have to be replaced with the original parameterization in terms of σ_v and $\sigma_{u,t}$. Their results on the presence of this type of heteroscedasticity are mixed.

[9]In the original work, Colombi (2010) notes use of a self developed R routine named SNF-maxlik.

$$\log L(\alpha,\boldsymbol{\beta},\lambda,\sigma,\gamma,\theta) = \sum_{i=1}^{N} \log\int_{-\infty}^{\infty} \left\{\prod_{t=1}^{T} \left[\begin{array}{l}\frac{2}{\sigma}\phi\left(\frac{y_{it}-\alpha-\boldsymbol{\beta}'\mathbf{x}_i-\delta_i}{\sigma}\right)\times \\ \Phi\left(\frac{-(y_i-\alpha-\boldsymbol{\beta}'\mathbf{x}_i-\delta_i)\lambda}{\sigma}\right)\end{array}\right]\right\}\frac{2}{\theta}\phi\left(\frac{\delta_i}{\theta}\right)\Phi\left(\frac{-\gamma\delta_i}{\theta}\right)d\delta_i. \tag{15}$$

For practical purposes, it is more convenient to use the original parameterization. Recall,

$$\delta_i = \sigma_w W_i - \sigma_h |H_i|$$

where W_i and H_i are both normally distributed with mean zero and variance one. The usefulness of the parameterization is for the simulation, which is based on primitive draws from the standard normal populations – there are no additional parameters. Combining terms, the simulated log likelihood function for the GTRE model is

$$\log L_S(\alpha,\boldsymbol{\beta},\lambda,\sigma,\sigma_w,\sigma_h) = \sum_{i=1}^{N} \log\frac{1}{R}\sum_{r=1}^{R} \left\{\prod_{t=1}^{T} \left[\begin{array}{l}\frac{2}{\sigma}\phi\left(\frac{y_{it}-\alpha-\boldsymbol{\beta}'\mathbf{x}_i-(\sigma_w W_{ir}-\sigma_h|H_{ir}|)}{\sigma}\right)\times \\ \Phi\left(\frac{-(y_i-\alpha-\boldsymbol{\beta}'\mathbf{x}_i-(\sigma_w W_{ir}-\sigma_h|H_{ir}|)\lambda}{\sigma}\right)\end{array}\right]\right\} \tag{16}$$

This estimation problem is only slightly more difficult than that for the TRE model as it involves another parameter, σ_h. The simulation, itself, involves pairs of independent random draws from two standard normal populations. But, the optimization problem itself is essentially the same as the TRE. It is worth noting, the appearance that the MLE implied by (14) is 'direct' while that in (16) which involves simulation is 'approximate,' is a bit misleading. The full log likelihood in (14) involves T+1 dimensional integration of the normal distribution. This cannot be done 'directly.' Normal integrals of dimension 3 or larger require use of the GHK simulator, which approximates the same integral as in (16). The reason that maximizing (16) is so much faster than (14) is that (14) is a 'brute force' approach that does not make use of the greatly simplifying result that the actual integration needed to compute the term in the log likelihood involves integration over a single dimension, that of h_i. Experience with high dimensional integration using the GHK simulator suggests that because a large number of simulation points is needed to gain acceptable accuracy, the advantage of (16) over (14) should be substantial as T increases. In our application, T is 13.

Computing the technical efficiency uses a result from Colombi (2010) based on the moment generating function for the closed skew normal distribution;

$$\mathrm{E}[\exp(\mathbf{t}'\mathbf{u}_i) \mid \mathbf{e}_i] = \frac{\bar{\Phi}_{T+1}(\mathbf{Re}_i + \boldsymbol{\Lambda}\mathbf{t}, \boldsymbol{\Lambda})}{\bar{\Phi}_{T+1}(\mathbf{Re}_i, \boldsymbol{\Lambda})} \exp\left[\mathbf{t}'\mathbf{Re}_i + \tfrac{1}{2}\mathbf{t}'\boldsymbol{\Lambda}\mathbf{t}\right]$$

$$\mathbf{u}_i = \begin{bmatrix} h_i \\ u_{i1} \\ \vdots \\ u_{iT} \end{bmatrix}, \quad \mathbf{t} = \begin{bmatrix} -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ -1 \end{bmatrix} \tag{17}$$

(The residual, $\mathbf{e}_i$, is defined in (14).) This computation requires multivariate normal integration, which we do using the GHK simulator. Technical or cost inefficiency can be computed from the results in (17) using –log[E[exp($\mathbf{t}'\mathbf{u}_i$)|$\mathbf{e}_i$] element by element. Kumbhakar, Lien and Hardaker (2012) also suggested a total efficiency measure,

$$\textit{Overall efficiency}_{it} = \mathrm{E}[\exp(-h_i)|\mathbf{e}_i] \times \mathrm{E}[\exp(-u_{it})|\mathbf{e}_i] \tag{18}$$

which can also be computed from the results in (17).

4. Empirical Analysis

In this section we illustrate an application of the new estimation approach of the four–way error component stochastic frontier model presented in the previous section that allows us to distinguish persistent and transient levels of productive efficiency. As discussed previously, following KumbhakarandTsionas(2012) we call this model "Generalized True Random Effects" model (hereafter GTRE), because is a model that nest two other models, i.e. pooled frontier model and the true random effects model. Further, in order to take into account a possible heterogeneity bias due to the correlation of the explanatory variables with the stochastic term, we propose the estimation of the GTRE with the Mundlak adjustment (hereafter MGTRE).

For comparison purposes we also estimate the stochastic cost frontier model using RE and TRE. We choose these two models because the first provide time-invariant inefficiency indicators, whereas the second estimate time variant inefficiency indicator.

The application of the GTREis based on an a data set and on a paper by Farsi et al. (2005) on the measurement of the cost efficiencyofa sample of Swiss railway companies.[10] A description of the model specification and on the variables used in this empirical application is available in Farsi et al. (2005).

The total cost of a railway company can be specified as a function of input prices and outputs. Moreover, as discussed in Farsi et. A. (2005a),in the cost model specification it is

[10] The data set is available at http://people.stern.nyu.edu/wgreene/Text/Edition7/TableF19-1.txt

possible to consider a number of output characteristics, which should take into account, at least partially, the railway companies' production environment.

The total cost can be can be written as:

$$TC = f\,(Y_1, Y_2, P_L , P_C , P_E , N, NS,\boldsymbol{d}_t) \qquad (19)$$

where TC is the total annual costs; Y_1 andY_2 are two outputs (numbers of passenger-kilometers and freight ton-kilometers); P_K , P_L and P_E are the prices of capital, labor and energyrespectively; N is the length of network , NS is the number of stopsand $\boldsymbol{d}_t$ is a vector of 12 year dummies from 1986 to 1997. With respect to the model used by Farsi et. al (2005a) we have included the number of stops in the model to quantify the environmental conditions more precisely. The cost function is concave,non-decreasing in input prices and output and linearly homogeneous in input prices.

To estimate the cost function in (19), a log-linear functional form is used.The cost function can be written as:

$$\ln\left(\frac{TC_{it}}{P_{E\,it}}\right) = \alpha_0 + \alpha_{y1}\ln Y_{1it} + \alpha_{y2}\ln Y_{2it} + \alpha_N \ln N_{it} + \alpha_{NS}\ln NS_{it} + \alpha_L \ln\left(\frac{P_{L\,it}}{P_{E_{it}}}\right) + \alpha_K \ln\left(\frac{P_{K\,it}}{P_{E_{it}}}\right) + \sum_{t=1986}^{t=1997}\alpha_t d_t + \varepsilon_{it}. \qquad (20)$$

Subscripts $i$and t denote the company and year respectively. The error term ε_{it} in (20) is composed of different independent parts depending on the econometric specification chosen as explained in table 1. The linear homogeneity restriction is imposed by normalizing the costs and input prices by the price of energy.

The data set used in this study, as discussed in more details in Farsi et. Al. (2005), is based on the financial reports of 50Swiss railway companies over the 13-year period from 1985 to 1997. Table 1 summarizes the four econometric specifications used in this empirical part of the paper.The estimation results for the cost frontier models using the four models discussed above are given in Table 2.

These results show that the output, output characteristics and input price coefficients are positive and highly significant across all models. The estimated coefficients are relatively similar across the different models. The only exceptions are the coefficients of the outputs in the Mundlak version of the GTREM. In this model these two coefficients are lower than in the other models.[11] These results suggest the presence of unobserved heterogeneity bias. The Mundlak version of the GTRE would be preferred to the other models on this basis. Moreover, on the basis of a likelihood ratio test we can reject that all Mundlak terms are equal zero (likelihood ratio test statistic is 44.62, higher than the $\chi^2_{(0.95;6)}$ value 12.592.

The coefficients of the year dummy positive and indicate that the total costs of railway companies increased over time.

Table 1: Econometric Specifications of the Stochastic Cost Frontier

	Model I **RE** (Pitt and Lee)	*Model II* **TRE**	*Model III* **GTRE**	*Model IV* **MGTRE**
Firm effects α_i	α	$N(\alpha,\sigma_w^2)$	$N(\alpha,\sigma_w^2)$	$\alpha_i = \boldsymbol{\phi}'\overline{\mathbf{x}}_i + w_i$ $w_i \sim N(\alpha,\sigma_w^2)$
Full random error ε_{it}	$\varepsilon_{it} = u_i + v_{it}$ $u_i \sim N^+(0,\sigma_u^2)$ $v_{it} \sim N(0,\sigma_v^2)$	$\varepsilon_{it} = w_i + u_{it} + v_{it}$ $u_{it} \sim N^+(0,\sigma_u^2)$ $v_{it} \sim N(0,\sigma_v^2)$ $w_{it} \sim N(0,\sigma_w^2)$	$\varepsilon_{it} = w_i + h_i + u_{it} + v_{it}$ $u_{it} \sim N^+(0,\sigma_u^2)$ $h_i \sim N^+(0,\sigma_h^2)$ $v_{it} \sim N(0,\sigma_v^2)$ $w_i \sim N(0,\sigma_w^2)$	$\varepsilon_{it} = w_i + h_i + u_{it} + v_{it}$ $u_{it} \sim N^+(0,\sigma_u^2)$ $h_i \sim N^+(0,\sigma_h^2)$ $v_{it} \sim N(0,\sigma_v^2)$ $w_i \sim N(0,\sigma_w^2)$
Persistent Inefficiency Estimator	$E(u_i \mid \boldsymbol{\varepsilon}_{i1},\ldots\boldsymbol{\varepsilon}_{iT})$	None	$E(h_i \mid \mathbf{y}_i)$	$E(h_i \mid \mathbf{y}_i)$
Transient Inefficiency Estimator	None	$E(u_{it} \mid \varepsilon_{it})$	$E(u_{it} \mid \mathbf{y}_i)$	$E(u_{it} \mid \mathbf{y}_i)$

Since total costs and all the continuous explanatory variables are in logarithms, the estimated coefficients can be interpreted as cost elasticities. For instance, the output coefficients suggest that the increase in cost due to a one percent increase in the number of stops, keeping all other explanatory variables constant, varies between 0.1 to 0.2 percent. The coefficient of network length suggest that the increase in cost due to a one percent extension in the network keeping all other explanatory variablesconstant is approximately 0.4 percent.

[11] The values of the coefficients of the MGTRE are very close to the coefficients obtained using a classical fixed effects model.

Further, the coefficient of number of stops suggests that the increase in cost due to a one percent increase in the number of stops, keeping all other explanatory variables constant, varies between 0.02 to 0.1 percent.

Table 3 provides descriptive statistics for the cost efficiency estimates for the 50 railway companies obtained from the econometric estimation of the six models. The estimation results of the new cost frontier models (GTRE and MGTRE) provide estimates the persistent (GTRE-P, MGTRE-P)as well the transient component of cost efficiency (GTRE-T, MGTRE-T). The RE model produce values of the cost efficiency that are time-invariant and therefore should reflect the persistent part of the cost efficiency and the TRE produce values that are time-varying and therefore should reflect the transient part of the cost efficiency.

The values reported in table 3 show that the estimated average values of the persistent efficiency varies from 55% in the REM to 74% in the TGTRE and 78% in the TMGTRE. The estimated average values of the transient efficiency in the TRE, the TGTRE and in the TMGTREis approximately 94%.Table 4 provides the correlations between the estimated levels of cost efficiency obtained from the different model specifications.

Generally, the value of the correlation coefficients between the values of the transient cost efficiency obtained with TRE andGTRE is relatively high, 0.85. But, the correlation between the values of the persistent cost efficiency obtained with RE and GTRE is relatively low, 0.15. This result suggeststhat the result obtained with the RE model is not measuring the persistent efficiency of the firmscorrectly. As Greene (2005) suggested earlier, the reason could be that in this model all unobserved time invariant heterogeneityis captured by the individual effect that is also used to compute the level of efficiency.

Table 2: Swiss Railways, Estimation Results
(Asymptotic t-ratiosin parentheses. Mundlak terms not shown.)

	RE		TRE		GTRE		MGTRE	
LNY1	0.193	***	0.162	***	0.155	***	0.106	***
	(6.242)		(31.018)		(27.634)		(3.479)	
LNY2	0.021	***	0.027	***	0.021	***	0.016	***
	(7.936)		(19.776)		(14.823)		(5.491)	
LNN	0.424	***	0.392	***	0.417	***	0.422	***
	(10.798)		(49.913)		(49.926)		(11.836)	
LNPL	0.539	***	0.551	***	0.541	***	0.556	***
	(18.588)		(29.185)		(27.974)		(18.788)	
LNPE	0.151	***	0.134	***	0.149	***	0.131	***
	(4.428)		(6.701)		(7.245)		(3.915)	
LNSTOPS	0.115	***	0.026	***	0.050	***	0.113	***
	(2.597)		(2.955)		(4.901)		(2.231)	
YEAR86	0.011		0.017		0.016		0.019	
	(0.280)		(0.600)		(0.551)		(0.611)	
YEAR87	0.014		0.024		0.024		0.031	
	(0.412)		(1.028)		(0.985)		(1.122)	
YEAR88	0.029		0.043		0.043		0.050	
	(0.691)		(1.315)		(1.253)		(1.194)	
YEAR89	0.05332		0.068	*	0.068	*	0.075	*
	(1.159)		(1.987)		(1.913)		(1.828)	
YEAR90	0.071	**	0.088	**	0.089	***	0.096	***
	(2.077)		(2.824)		(2.913)		(2.873)	
YEAR91	0.080	***	0.102	***	0.103	***	0.111	***
	(2.872)		(4.381)		(4.417)		(4.111)	
YEAR92	0.094	***	0.116	***	0.116	***	0.122	***
	(2.862)		(5.191)		(5.081)		(4.161)	
YEAR93	0.081	***	0.105	***	0.103	***	0.110	***
	(2.601)		(4.202)		(4.08)		(3.685)	
YEAR94	0.064	*	0.087	**	0.084	***	0.092	***
	(1.708)		(2.884)		(2.816)		(2.639)	
YEAR95	0.048		0.064	**	0.060	**	0.063	**
	(1.512)		(2.749)		(2.528)		(2.238)	
YEAR96	0.032		0.052	***	0.046	**	0.045	*
	(1.267)		(2.336)		(2.043)		(1.907)	
YEAR97	0.032		0.048	*	0.043		0.044	
	(1.045)		(1.858)		(1.66)		(1.571)	
α	-7.727	***	-9.507	***	-6.242	***	-3.687	***
	(-11.676)		(-36.704)		(-23.684)		(-10.284)	
σ_w	-		0.353	***	0.522	***	0.365	***
			(88.250)		(90.153)		(90.639)	
λ	0.820	***	1.635	***	0.095	***	0.094	***
	(6.296)		(7.712)		(30.305)		(25.475)	
σ	11.692	*	0.099	***	1.611	***	1.561	***
	(1.857)		(24.750)		(8.197)		(6.823)	
σ_h	-		-		0.664	***	0.851	***
					(12.008)		(13.506)	
Log likelihood	598.644		591.559		599.231		621.544	

Table 3: Cost Efficiency Scores

Variable	Mean	Std.Dev.	Minimum	Maximum
RE	0.548	0.208	0.126	0.986
TRE	0.935	0.036	0.716	0.988
TGTRE	0.939	0.031	0.733	1.000
PGTRE	0.747	0.038	0.738	1.000
TMGTRE	0.941	0.030	0.737	1.000
PMGTRE	0.784	0.032	0.775	1.000

Table 4: Correlation Coefficients

	RE	TRE	TGTRE	PGTRE	TMGTRE	PMGTRE
RE	1	-0.11	-0.06	0.15	-0.06	0.15
TRE	-0.11	1	0.85	-0.19	0.84	-0.19
TGTRE	-0.06	0.85	1	0.29	1.00	0.28
PGTRE	0.15	-0.19	0.29	1	0.29	1.00
TMGTRE	-0.06	0.84	1.00	0.29	1	0.29
PMGTRE	-0.15	-0.19	0.28	1.00	0.29	1

5. Conclusions

In the measurement of the level of productive efficiency of a firm, it is possible to distinguish between persistent and transient levels of efficiency. Empirical studies on efficiency measurement have paid relatively little attention to the distinction between these two components in estimates of productive efficiency. Recently, Colombi et al. (2011, 2014), Kumbhakar and Tsionas (2012), and Kumbhakar, Lien ahd Hardaker (2012) have proposed some econometric approaches to provide separate estimates of the two components of efficiency. However, the approaches are relatively cumbersome or are based on a multistep manipulation of OLS that is not completely satisfactory from an econometric point of view.

In this paper, we propose to estimate the two components of productive efficiency using a full information maximum simulated likelihood estimator. The extreme complexity of the log likelihood noted in Colombi et. Al. (2011) is reduced by exploiting Butler and Moffitt's (1982) formulation in the simulation. The approach is then applied with success in the estimation of a cost frontier function for a sample of Swiss railways.

From the methodological point of view we show that this method is relatively straightforward and effective to apply. Further, we show that the transient and the persistent parts of productive efficiency are relatively different in absolute value and not highly correlated. These indicators measure different things. We find that the efficiency indicators obtained with the TRE are highly correlated with the transient efficiency indicators obtained with the GTRE. Therefore, the TRE model tend to estimate the transient part of efficiency. The contribution of the GTRE is to decompose further the time persistent effect built into the TREM. Finally, the indicators obtained with RE model are not correlated with the persistent efficiency indicator of the GTRE. The classical RE model appears not to be measuring the level of persistent efficiency of the firms in the sample. This result may be due to the fact that in this model all unobserved time invariant variables, notably time invariant heterogeneity, are captured by the individual effect that is used to compute the level of efficiency.

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*(3), 393–416.

Ahn, S. C., & Sickles, R. C. (2000). Estimation of long-run inefficiency levels: a dynamic frontier approach. *Econometric Reviews*, *19*(4), 461–492.

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.

A. Azzalini (1985). A class of distributions which includes the normal ones. *Scand.J.Statist.* 12, 171-178.

Baranzini, A., & Faust, A.K. (2014). Water Supply: The cost structure of water utilities in Switzerland. *Journal of Productivity Analysis*, Forthcoming.

Battese, G. E., & Coelli, T. J. (1992). Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India. In T. R. G. Jr & C. A. K. Lovell (Eds.), *International Applications of Productivity and Efficiency Analysis* (pp. 149–165). Springer Netherlands.

Battese, G. E., & Coelli, T. J. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics, 20*(2), 325–332.

Butler, J., and R. Moffitt, 1982. A Computationally Efficient Quadrature Procedure for the One Factor Multinomial Probit Model,*Econometrica*, 50, 761-764.

Chen, Y., Schmidt, P. And P. Wang, (2014). Consistent Estimates of the Fixed Effects Stochastic Frontier Model, Journal of Econometrics, Forthcoming.

Colombi, R. (2010). A Skew Normal Stochastic Frontier Model for Panel Data. Proceedings of the 45-th Scientific Meeting of the Italian Statistical Society

Colombi, R., Martini, G., & Vittadini, G. (2011). A stochastic frontier model with short-run and long-run inefficiency random effects. *Department of Economics and Technology Management, Universita Di Bergamo, Italy*.

Cornes, R. (1992). *Duality and Modern Economics*. CUP Archive.

Cornwell, C., Schmidt, P., & Sickles, R. C. (1990). Production frontiers with cross-sectional and time-series variation in efficiency levels. *Journal of Econometrics*, *46*(1–2), 185–200.

Cuesta, R. A. (2000). A Production Model With Firm-Specific Temporal Variation in Technical Inefficiency: With Application to Spanish Dairy Farms. *Journal of Productivity Analysis*, *13*(2), 139–158.

Farsi, M., Filippini, M., & Greene, W. (2005 a). Efficiency Measurement in Network Industries: Application to the Swiss Railway Companies. *Journal of Regulatory Economics*, *28*(1), 69–90.

Farsi, M., Filippini, M., & Kuenzle, M. (2005 b). Unobserved heterogeneity in stochastic cost frontier models: an application to Swiss nursing homes. *Applied Economics*, *37*(18), 2127–2141.

Greene, W. (2004). Distinguishing between heterogeneity and inefficiency: stochastic frontier analysis of the World Health Organization's panel data on national health care systems. *Health Economics*, *13*(10), 959–980. doi:10.1002/hec.938

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

Greene, W. (2005 b). Fixed and Random Effects in Stochastic Frontier Models. *Journal of Productivity Analysis*, *23*(1), 7–32.

Greene, W. (2008). The econometric approach to efficiency analysis. In *Fried HO, Lovell CAK, Shelton SS (eds) The measurement of productivity efficiency and productivity growth* (pp. 92–250). Oxford University Press.

Jondrow, J., Knox Lovell, C. A., 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.

Kumbhakar, S. C. (1990). Production frontiers, panel data, and time-varying technical inefficiency. *Journal of Econometrics*, *46*(1–2), 201–211.

Kumbhakar, S. C., Lien, G., & Hardaker, J. B. (2012). Technical efficiency in competing panel data models: a study of Norwegian grain farming. *Journal of Productivity Analysis*, *41*(2), 321–337.

Kumbhakar, S. C., & Lovell, C. A. K. (2003). *Stochastic Frontier Analysis*. Cambridge University Press.

Kumbhakar,& Tsionas, E. G., S. C. (2012). Firm Heterogeneity, Persistent and transient Technical Inefficiency: A generalized True Random Effects model. *Journal of Applied Econometrics*, *29*(1), 110–132.

Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica*, *46*(1), 69–85.

Pitt, M. M., & Lee, L.-F. (1981). The measurement and sources of technical inefficiency in the Indonesian weaving industry. *Journal of Development Economics*, *9*(1), 43–64.

Schmidt, P., & Sickles, R. C. (1984). Production Frontiers and Panel Data. *Journal of Business & Economic Statistics*, *2*(4), 367–374.

Train, K., 2003. *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge

Wooldridge, J. M. (2010). Correlated random effects models with unbalanced data. *MIMEO*, *University of Michigan*.