American Journal of Theoretical and Applied Statistics
2023; 12(6): 180-186
http://www.sciencepublishinggroup.com/j/ajtas
doi: 10.11648/j.ajtas.20231206.14
ISSN: 2326-8999 (Print); ISSN: 2326-9006 (Online)

Estimating the Determinants of Firm Innovation Inefficiency Through the Conditional Mean of Innovation Inefficiency Given a Composite Error

Vasilios Kanellopoulos

Department of Economics, University of Patras, Patras, Greece

Email address:
kanelb@upatras.gr

To cite this article:
Vasilios Kanellopoulos. Estimating the Determinants of Firm Innovation Inefficiency Through the Conditional Mean of Innovation Inefficiency Given a Composite Error. *American Journal of Theoretical and Applied Statistics s*. Vol. 12, No. 6, 2023, pp. 180-186. doi: 10.11648/j.ajtas.20231206.14

Received: October 17, 2023; **Accepted**: November 3, 2023; **Published**: November 11, 2023

Abstract: The present paper demonstrates that the estimations of the determinants of firm innovation inefficiency can be obtained through the conditional mean of innovation inefficiency given a composite error. We extract the estimations of the determinants of firm innovation inefficiency by replacing the true parameters in the equation of the conditional mean of innovation inefficiency given a composite error with Maximum Likelihood estimations from the Stochastic Frontier Approach. This is an alternative method for the estimation of the determinants of firm inefficiency besides those which are existent in the relevant literature. Based on statistical theory and algebra, we first present the case where innovation inefficiency is assumed to be distributed as a truncated normal with a nonzero constant mean. Second, we focus on the case where innovation inefficiency is assumed to be distributed as a truncated normal with a mean that varies across firms. There, we show that all the change in the error term of the Stochastic Frontier Knowledge Production Function originates from innovation inefficiency. The latter is modelled as having two components: a) a function of some firm-specific characteristics (variables) and b) random component. Then, we advance to the estimations of the determinants of firm innovation inefficiency via a generalized Stochastic Frontier Approach (generalized production frontier approach). Finally, we replace the true parameters in the equation of the conditional mean of innovation inefficiency given a composite error with Maximum Likelihood estimations from the generalized production frontier approach.

Keywords: Innovation Inefficiency, Firms, Conditional Mean Estimator, Stochastic Frontier Models

1. Introduction

This research will attempt to extract the estimations of the determinants of firm innovation inefficiency. This will demonstrate that it takes place through the conditional mean of innovation inefficiency given a composite error.

The measurement of the productive efficiency of the constructed frontiers is first introduced by Farrell M [1]. Its meaning refers to the fact that the firms and farmers will utilize best practices and techniques and will obtain a maximum feasible yield (Kalirajan K and Shand R [2], see also Alvarez R and Crespi G [3], Pestana Barros C and Dieke P [4]). By contrast, inefficiency refers to any deviation from the maximum feasible yield and reflects the firm's lack of use of the best practices.

Regarding the determinants of technical inefficiency, the most empirical studies follow a two-step procedure to estimate its determinants (see Alvarez R and Crespi G [3], Kalirajan K and Shand R [2], Page J [5], Pestana Barros C and Dieke P [4], Pitt M and Lee L [6]). Specifically, these studies estimate technical inefficiency from the production function. There, they do not consider that technical inefficiency is a function of other variables. In the second stage, technical inefficiency is regressed on a set of variables and firm characteristics. These explain differences in technical inefficiency among firms. However, this two-step procedure has some shortcomings. First, technical inefficiency is possible to be correlated with inputs creating endogeneity problems. Thus, inconsistent estimations of parameters and technical inefficiency take place (see

Kumbhakar S et al [7]). Another shortcoming is also described by Kumbhakar S et al [7] Kumbhakar S et al [7] (p. 280) argue that "the standard ordinary least squares results in the second step may not be appropriate since technical inefficiency-the dependent variable-is one-sided". In addition, they note that the meaning of the error term in the second step is vague. To face these problems, Kumbhakar S et al [7], Reifschneider D and Stevenson R [8] and Battese G and Coelli T [9] suggest a single-stage estimation procedure where the determinants of technical inefficiency are explicitly introduced in the model (see also Zeebari Z et al [10])[1].

In the literature of efficiency, most empirical studies deal with the technical and productive efficiency (see Battese G and Coelli T [9], De Borger B et al [11]). As a result, the first contribution of the present research to the existing literature is that it introduces the meaning of efficiency into innovation process. It succeeds it by assuming a Stochastic Frontier Knowledge Production Function. The latter deals with the efficient transformation of innovation and R&D inputs to desirable innovation output, such as patents or/and product and process innovation. Based on statistical theory and algebra, we first present the case where innovation inefficiency is assumed to be distributed as a truncated normal with a nonzero constant mean. After, we focus on the case where innovation inefficiency is assumed to be distributed as a truncated normal with a mean that varies across firms. There, we show that all the change in the error term of the Stochastic Frontier Knowledge Production Function originates from innovation inefficiency. The latter is modelled as having the following two components: a) a deterministic component which includes observed firm-specific characteristics and b) a random component. In addition, we advance to the estimations of the determinants of firm innovation inefficiency by using a generalized production frontier approach suggested by Kumbhakar S et al [7]. There, we maximize the log likelihood function by using differential calculus and first order conditions where necessary and obtain the estimations of variance and other parameters. Based on Jondrow J et al [12], we finally demonstrate that the estimations of the determinants of firm innovation inefficiency can be obtained through the conditional mean of innovation inefficiency given a composite error. We extract the estimations of the determinants of firm innovation inefficiency by replacing the true parameters in the equation of the conditional mean of innovation inefficiency given a composite error with Maximum Likelihood estimations of the parameters from the Stochastic Frontier Approach. This is the most important contribution of the present research and advances the relevant literature.

The next section presents the model and the last section describes the conclusions, the limitations of the present research and avenues for further research.

2. Model

2.1. Innovation Inefficiency Is Assumed to Be Distributed as a Truncated Normal with a Nonzero Constant Mean

We consider the following Stochastic Frontier Knowledge Production Function (KPF)[2]:

$$y_i = f(x_i; b), \tag{1}$$

Taking logarithms in the both sides of equation (1), we have:

$$y_i = a + bx_i + e_i, \tag{2}$$

where y_i is the maximum possible innovation output as a function of certain innovation inputs, $x_i = (x_1, x_2, \dots . x_N)'$ is the $N \times 1$ vector of innovation inputs, a is the constant term, $b = (b_1, b_2, \dots, b_N)'$ is the $N \times 1$ vector of coefficients which correspond to the innovation inputs, $e_i = (e_1, e_2, \dots., e_N)'$ is the $N \times 1$ vector of errors and $i = 1,2 \dots, N$ firms.

Here, we first assume that the error is:

$$e_i = u_i + v_i, \tag{3}$$

where $u_i \geq 0$. In addition, v_i and u_i which are distributed independently and identically represent the measurement and specification error and innovation inefficiency, respectively. We further assume that u and v are distributed as:

$$f(u) = \frac{1}{\left(1-F^*\left(-\frac{\mu}{\sigma_u}\right)\right)\sqrt{2\pi}\,\sigma_u} \exp\left[-\frac{1}{2} * \left(\frac{u-\mu}{\sigma_u}\right)^2\right], \tag{4}$$

for $u > 0$ and

$$g(v) = \left(\frac{1}{\sqrt{2\pi}\sigma_v}\right) \exp\left[-\frac{1}{2} * \left(\frac{v}{\sigma_v}\right)^2\right], \tag{5}$$

for all v. In other words, u is assumed to be distributed as a truncated normal with a nonzero constant mean (μ) and variance σ_u^2 and v is assumed to be distributed as a normal with zero mean and variance σ_v^2 (see Stevenson R [16]). The density function for $e = u + v$ is given by:

$$z(e) = \sigma^{-1} f^*\left(\frac{e-\mu}{\sigma}\right)\left[1 - F_1^*\left(-\frac{\mu}{\sigma\lambda} - \frac{e\lambda}{\sigma}\right)\right]\left[1 - F_2^*\left(-\frac{\mu}{\sigma_u}\right)\right]^{-1}, \tag{6}$$ [3]

where $\lambda = \sigma_u/\sigma_v$, $\sigma = (\sigma_u^2 + \sigma_v^2)^{1/2}$, $f^*(.)$ is the standard normal density function evaluated at $((e-\mu)/\sigma)$ and $F_1^*(.)$ and $F_2^*(.)$ are the standard normal distribution functions evaluated at $(-\frac{\mu}{\sigma\lambda} - \frac{e\lambda}{\sigma})$ and $\left(-\frac{\mu}{\sigma_u}\right)$, respectively. When $\mu = 0$, we take:

---

1 Here, technical inefficiency is assumed to be a function of some firm-specific characteristics and variables and a random component. Using this single-step estimation procedure, Kumbhakar S et al [7] give a more general specification to technical inefficiency.

2 For more details about knowledge production function at a firm level, see Pellegrino G and Piva M [13], Catozzella A and Vivarelli M [14] and Ramani S et al [15].

3 $f^*\left(\frac{e-\mu}{\sigma}\right) = dF^*(\frac{e-\mu}{\sigma})/d(\frac{e-\mu}{\sigma})$. Putting $\frac{e-\mu}{\sigma} = z$, we take $f^*(z) = d\Phi^*(z)/d(z)$. In addition, putting $-\frac{\mu}{\sigma\lambda} - \frac{e\lambda}{\sigma} = z_1$ and $-\frac{\mu}{\sigma_u} = z_2$, we take $\Phi_1^*(z_1) = P(-\infty < z_1 < z') = \int_{-\infty}^{z'} f(z_1) dz_1$ and $\Phi_2^*(z_2) = P(-\infty < z_2 < z'') = \int_{-\infty}^{z''} f(z_2) dz_2$.

$$z(e)|_{\mu=0} = \sigma^{-1} f^* \left(\frac{e}{\sigma}\right)\left[1 - F_1^*\left(-\frac{e\lambda}{\sigma}\right)\right], \quad (7)$$

This is the density function for e when the Aigner D et al [17] model takes place (see also Jondrow J et al [12]).

In the Stevenson R [16] model, the mean of e is:

$$E(e) = E(u+v), \quad (8)$$

or

$$E(e) = E(u) + E(v), \quad (9)$$

or

$$E(e) = E(u) + 0, \quad (10)$$

or

$$E(e) = E(u), \quad (11)$$

However,

$$E(e) = E(u) = \int_{-\infty}^{+\infty} u f(u) du, \quad (12)$$

or

$$E(e) = E(u) = \int_{-\infty}^{+\infty} u\left(\frac{1}{(1-F^*\left(-\frac{\mu}{\sigma_u}\right))\sqrt{2\pi}\,\sigma_u} \exp\left[-\frac{1}{2} * \left(\frac{u-\mu}{\sigma_u}\right)^2\right) du, \quad (13)^4$$

or

$$E(e) = E(u) = \frac{\mu\alpha}{2} + (\sigma_u\alpha/\sqrt{2\pi})\exp\left[-1/2 * (\frac{\mu}{\sigma_u})^2\right], \quad (14)$$

In addition, the variance of e is given by:

$$V(e) = V(u) + V(v), \quad (15)$$

or

$$V(e) = \sigma_u^2 + \sigma_v^2, \quad (16)$$

or

$$V(e) = E[(u - E(u))^2] + \sigma_v^2, \quad (17)$$

or

$$V(e) = E[(u - (\int_{-\infty}^{+\infty} u f(u) du))^2] + \sigma_v^2, \quad (18)$$

or

$$V(e) = E\left[\left(u - \left(\int_{-\infty}^{+\infty} u\left(\frac{1}{\left(1-F^*\left(-\frac{\mu}{\sigma_u}\right)\right)\sqrt{2\pi}\,\sigma_u} \exp\left[-\frac{1}{2} * \left(\frac{u-\mu}{\sigma_u}\right)^2\right)\right) du\right)\right)^2\right] + \sigma_v^2, \quad (19)$$

or

$$V(e) = \mu^2\left(\frac{\alpha}{2}\right)\left(1-\frac{\alpha}{2}\right) + \sigma_u^2\left(\frac{\alpha}{2}\right)\left(\frac{\pi-\alpha}{\pi}\right) + \sigma_v^2, \quad (20)$$

where $\alpha = [1 - F^*\left(-\frac{\mu}{\sigma_u}\right)]^{-1}$.

In the case of $\mu = 0$, the mean of e becomes:

$$E(e)|_{\mu=0} = \frac{0\alpha}{2} + (\sigma_u\alpha/\sqrt{2\pi})\exp\left[-1/2 * (\frac{0}{\sigma_u})^2\right], \quad (21)$$

or

$$E(e)|_{\mu=0} = (\sigma_u\alpha/\sqrt{2\pi})\exp[0], \quad (22)$$

or

$$E(e)|_{\mu=0} = (\sigma_u\alpha/\sqrt{2\pi}) * 1, \quad (23)$$

or

$$E(e)|_{\mu=0} = (\sigma_u\alpha/\sqrt{2\pi}), \quad (24)$$

or

$$E(e)|_{\mu=0} = (\sigma_u[1 - F^*\left(-\frac{0}{\sigma_u}\right)]^{-1})/\sqrt{2\pi}), \quad (25)$$

or

$$E(e)|_{\mu=0} = (\sigma_u[1 - F^*(0)]^{-1})/\sqrt{2\pi}), \quad (26)$$

or

$$E(e)|_{\mu=0} = (\sigma_u[1 - 0]^{-1})/\sqrt{2\pi}), \quad (27)$$

or

$$E(e)|_{\mu=0} = \sigma_u/\sqrt{2\pi}, \quad (28)$$

Similarly, when $\mu = 0$, the variance of e becomes:

$$V(e)|_{\mu=0} = V(u) + V(v) = 0\left(\frac{\alpha}{2}\right)\left(1-\frac{\alpha}{2}\right) + \sigma_u^2\left(\frac{\alpha}{2}\right)\left(\frac{\pi-\alpha}{\pi}\right) + \sigma_v^2, \quad (29)$$

or

$$V(e)|_{\mu=0} = \sigma_u^2\left(\frac{\alpha}{2}\right)\left(\frac{\pi-\alpha}{\pi}\right) + \sigma_v^2, \quad (30)$$

or

$$V(e)|_{\mu=0} = \sigma_u^2\left(\frac{[1-F^*\left(-\frac{0}{\sigma_u}\right)]^{-1}}{2}\right)\left(\frac{\pi-[(1-F^*\left(-\frac{0}{\sigma_u}\right))^{-1}]}{\pi}\right) + \sigma_v^2, \quad (31)$$

or

$$V(e)|_{\mu=0} = \sigma_u^2\left(\frac{[1-F^*(0)]^{-1}}{2}\right)\left(\frac{\pi-[(1-F^*(0))^{-1}]}{\pi}\right) + \sigma_v^2, \quad (32)$$

or

$$V(e)|_{\mu=0} = \sigma_u^2\left(\frac{[1-0]^{-1}}{2}\right)\left(\frac{\pi-[(1-0)^{-1}]}{\pi}\right) + \sigma_v^2, \quad (33)$$

or

$$V(e)|_{\mu=0} = \sigma_u^2\left(\frac{1}{2}\right)\left(\frac{\pi-1}{\pi}\right) + \sigma_v^2, \quad (34)$$

4 Following integration by parts, we take: $\int_{-\infty}^{+\infty} u f(u) du = \int_{-\infty}^{+\infty} u (f(u))' du = u(f(u))' - \int_{-\infty}^{+\infty} u' f(u) du = u(f(u))' - \int_{-\infty}^{+\infty} f(u) du$.

or

$$V(e)|_{\mu=0} = \frac{\sigma_u^2}{2} * \left(\frac{\pi-1}{\pi}\right) + \sigma_v^2, \quad (35)$$

2.2. Innovation Inefficiency Is Assumed to Be Distributed as a Truncated Normal with a Mean That Varies Across Firms

In this point, we assume that innovation inefficiency is distributed as a truncated normal with a mean which varies across firms because of variation in firm-specific factors like input quality, education as well as other characteristics (Kumbhakar S et al [7]). There, innovation inefficiency is made up of two components. The first which constitutes the deterministic component is a vector of exogenous variables and the second is the unobserved random component. Thus, innovation inefficiency can be modelled as[5]:

$$u' = k\beta + \varepsilon, \quad (36)$$

where u' is innovation inefficiency, $k = (k_1, k_2, \ldots\ldots, k_M)'$ is the $M \times 1$ vector of exogenous variables. This vector includes firm-specific factors like input, education and other characteristics which affect innovation inefficiency despite the fact that they do not enter into the stochastic frontier knowledge production function. Finally, $\beta = (\beta_1, \beta_2, \ldots\ldots \beta_M)'$ is the $M \times 1$ vector of coefficients which corresponds to the vector of exogenous variables, while ε is the unobserved random component.

Then, the density function for innovation inefficiency is:

$$f(u') = \frac{1}{(1-F^*\left(-\frac{k\beta}{\sigma_u'}\right))\sqrt{2\pi}\,\sigma_u'} \exp[-\frac{1}{2} * \left(\frac{u'-k\beta}{\sigma_u'}\right)^2], \quad (37)$$

for $u' > 0$ and

$$g(v) = \left(\frac{1}{\sqrt{2\pi}\sigma_v}\right) \exp[-\frac{1}{2} * \left(\frac{v}{\sigma_v}\right)^2], \quad (38)$$

for all v.

Similarly, the joint density function of $e' = u' + v$ is given by:

$$z(e') = (\sigma^{-1})' f^*\left(\frac{e'-k\beta}{\sigma'}\right)\left[1 - F_1^*\left(-\frac{k\beta}{\sigma'\lambda'} - \frac{e'\lambda'}{\sigma'}\right)\right][1 - F_2^*\left(-\frac{k\beta}{\sigma_u'}\right)]^{-1} \quad (39)$$

where $\lambda' = \sigma_u'/\sigma_v$, $\sigma' = ((\sigma_u^2)' + \sigma_v^2)^{1/2}$, $f^*(.)$ is the standard normal density function evaluated at $((e' - k\beta)/\sigma')$ and $F_1^*(.)$ and $F_2^*(.)$ are the standard normal distribution functions evaluated at $(-\frac{k\beta}{\sigma'\lambda'} - \frac{e'\lambda'}{\sigma'})$ and $\left(-\frac{k\beta}{\sigma_u'}\right)$, respectively.

Then, the mean of e' is:

$$E(e') = E(u' + v), \quad (40)$$

or

$$E(e') = E(u') + E(v), \quad (41)$$

or

$$E(e') = E(u') + 0, \quad (42)$$

or

$$E(e') = E(u'), \quad (43)$$

However,

$$E(e') = E(u') = \int_{-\infty}^{+\infty} u' f(u') du', \quad (44)$$

or

$$E(e') = E(u') = \int_{-\infty}^{+\infty} u'\left(\frac{1}{(1-F^*\left(-\frac{k\beta}{\sigma_u'}\right))\sqrt{2\pi}\,\sigma_u'} \exp[-\frac{1}{2} * \left(\frac{u'-k\beta}{\sigma_u'}\right)^2\right) du', \quad (45)$$

or

$$E(e') = E(u') = \frac{k\beta\alpha'}{2} + (\sigma_u'\alpha'/\sqrt{2\pi})\exp[-1/2 * (\frac{k\beta}{\sigma_u'})^2], \quad (46)$$

The variance of e' is

$$V(e') = V(u') + V(v), \quad (47)$$

or

$$V(e') = (\sigma_u^2)' + \sigma_v^2, \quad (48)$$

or

$$V(e') = E[(u' - E(u'))^2] + \sigma_v^2, \quad (49)$$

or

$$V(e') = E[(u' - \left(\int_{-\infty}^{+\infty} u' f(u') du'\right))^2] + \sigma_v^2, \quad (50)$$

or

$$V(e') = E[(u' - \left(\int_{-\infty}^{+\infty} u'\left(\frac{1}{\left(1-F^*\left(-\frac{k\beta}{\sigma_u'}\right)\right)\sqrt{2\pi}\,\sigma_u'} \exp\left[-\frac{1}{2} * \left(\frac{u'-k\beta}{\sigma_u'}\right)^2\right)\right) du'\right))^2] + \sigma_v^2, \quad (51)$$

or

$$V(e') = (k\beta)^2\left(\frac{\alpha'}{2}\right)\left(1 - \frac{\alpha'}{2}\right) + (\sigma_u^2)'\left(\frac{\alpha'}{2}\right)\left(\frac{\pi-\alpha'}{\pi}\right) + \sigma_v^2, \quad (52)$$

where $\alpha' = [1 - F^*\left(-\frac{k\beta}{\sigma_u'}\right)]^{-1}$. Thus, (46) becomes:

$$E(e') = E(u') = \frac{k\beta\left[1-F^*\left(-\frac{k\beta}{\sigma_u'}\right)\right]^{-1}}{2} + (\sigma_u'([1 - F^*\left(-\frac{k\beta}{\sigma_u'}\right)]^{-1})/\sqrt{2\pi})\exp[-1/2 * (\frac{k\beta}{\sigma_u'})^2], \quad (53)$$

Furthermore, (52) becomes:

5 In other words, $u' \sim N(k\beta, (\sigma_u^2)')$.

$$V(e') = (k\beta)^2 \left(\frac{\left(\left[1-F^*\left(-\frac{k\beta}{\sigma_u'}\right)\right]^{-1}\right)}{2}\right)\left(1-\frac{\left(\left[1-F^*\left(-\frac{k\beta}{\sigma_u'}\right)\right]^{-1}\right)}{2}\right) + (\sigma_u^2)'\left(\frac{\left(\left[1-F^*\left(-\frac{k\beta}{\sigma_u'}\right)\right]^{-1}\right)}{2}\right)\left(\frac{\pi-\left(\left[1-F^*\left(-\frac{k\beta}{\sigma_u'}\right)\right]^{-1}\right)}{\pi}\right) + \sigma_v^2, \quad (54)$$

2.3. The Estimation of the Determinants of Firm Innovation Inefficiency: A Generalized Production Frontier Approach

In this point, we form the following logged likelihood function (see Stevenson R [16]) in order to estimate the determinants of firm innovation inefficiency through a generalized production frontier approach (Kumbhakar S et al [7]):

$$\ln L(y|\beta,\lambda,\sigma^2,k\beta) = -\frac{n}{2}*ln\sigma^2 - \frac{n}{2}*\ln 2\pi - \frac{1}{2\sigma^2}\sum_{i=1}^{N}((y_i-b'x_i)-k\beta)^2 + \sum_{i=1}^{N}\ln[1-F_1^*(\sigma^{-1}(-(y_i-b'x_i)\lambda))] - n\ln[1-F_2^*((-k\beta/\sigma)*(\lambda^{-2}+1)^{\frac{1}{2}})], \quad (55)$$

Taking partial derivatives, we have:

$$\frac{\partial lnL}{\partial b} = \left(-\frac{1}{2\sigma^2}\right)2\sum_{i=1}^{N}((y_i-b'x_i)-k\beta)(-x_i) + \sum_{i=1}^{N}\frac{1}{1-F_1^*}*f_1^*\,(-x_i\lambda\sigma^{-1}), \quad (56)$$

or

$$\frac{\partial lnL}{\partial b} = \left(-\frac{1}{\sigma^2}\right)\sum_{i=1}^{N}((y_i-b'x_i)-k\beta)(-x_i) + \sum_{i=1}^{N}\frac{1}{1-F_1^*}*f_1^*\,(-x_i\lambda\sigma^{-1}), \quad (57)$$

or

$$\frac{\partial lnL}{\partial b} = \left(-\frac{1}{\sigma^2}\right)\sum_{i=1}^{N}((y_i-b'x_i)-k\beta)(-x_i) + \sum_{i=1}^{N}\frac{1}{1-F_1^*}*f_1^*\left(-x_i\lambda(\frac{1}{\sigma})\right), \quad (58)$$

or

$$\frac{\partial lnL}{\partial b} = \left(-\frac{1}{\sigma^2}\right)\sum_{i=1}^{N}((y_i-b'x_i)-k\beta)(-x_i) + \sum_{i=1}^{N}\frac{1}{1-F_1^*}*f_1^*\left(-x_i(\frac{\lambda}{\sigma})\right), \quad (59)$$

or

$$\frac{\partial lnL}{\partial b} = \left(\frac{1}{\sigma^2}\right)\sum_{i=1}^{N}((y_i-b'x_i)-k\beta)x_i - (\lambda/\sigma)\sum_{i=1}^{N}\frac{f_1^*}{1-F_1^*}x_i, \quad (60)$$

In addition,

$$\frac{\partial lnL}{\partial k\beta} = \left(-\frac{1}{2\sigma^2}\right)2\sum_{i=1}^{N}((y_i-b'x_i)-k\beta)(-1) + \sum_{i=1}^{N}\frac{1}{1-F_1^*}*f_1^*\left(-\sigma^{-1}\left(\frac{1}{\lambda}\right)\right) - n\frac{1}{1-F_2^*}*f_2^*(-(1/\sigma)*(\lambda^{-2}+1)^{\frac{1}{2}}), \quad (61)$$

or

$$\frac{\partial lnL}{\partial k\beta} = \left(-\frac{1}{\sigma^2}\right)\sum_{i=1}^{N}((y_i-b'x_i)-k\beta)(-1) + \sum_{i=1}^{N}\frac{1}{1-F_1^*}*f_1^*\left(-\left(\frac{\sigma^{-1}}{\lambda}\right)\right) - n\frac{1}{1-F_2^*}*f_2^*(-\left(\frac{1}{\sigma}\right)*(\lambda^{-2}+1)^{\frac{1}{2}}), \quad (62)$$

or

$$\frac{\partial lnL}{\partial k\beta} = \left(\frac{1}{\sigma^2}\right)\sum_{i=1}^{N}((y_i-b'x_i)-k\beta) + \sum_{i=1}^{N}\frac{1}{1-F_1^*}*f_1^*\left(-(1/\sigma)(\frac{1}{\lambda})\right) - n\frac{1}{1-F_2^*}*f_2^*(-\left(\frac{(\lambda^{-2}+1)^{\frac{1}{2}}}{\sigma}\right)), \quad (63)$$

or

$$\frac{\partial lnL}{\partial k\beta} = \left(\frac{1}{\sigma^2}\right)\sum_{i=1}^{N}((y_i-b'x_i)-k\beta) + \sum_{i=1}^{N}\frac{1}{1-F_1^*}*f_1^*\left(-\frac{1}{\sigma\lambda}\right) + n\frac{1}{1-F_2^*}*f_2^*\left(\frac{(\lambda^{-2}+1)^{\frac{1}{2}}}{\sigma}\right), \quad (64)$$

or

$$\frac{\partial lnL}{\partial k\beta} = \left(\frac{1}{\sigma^2}\right)\sum_{i=1}^{N}((y_i-b'x_i)-k\beta) - \left(\frac{1}{\sigma\lambda}\right)\sum_{i=1}^{N}\frac{f_1^*}{1-F_1^*} + \left(\frac{n(\lambda^{-2}+1)^{\frac{1}{2}}}{\sigma}\right)\frac{f_2^*}{1-F_2^*}, \quad (65)$$

In turn,

$$\frac{\partial lnL}{\partial \lambda} = \sum_{i=1}^{N}\frac{1}{1-F_1^*}f_1^*[(-\sigma^{-1}k\beta\lambda^{-2}) - (\sigma^{-1}(y_i-b'x_i))] - n\frac{1}{1-F_2^*}f_2^*[\left(-\frac{k\beta}{\sigma}\right)(-2)\lambda^{-3}(\lambda^{-2}+1)^{\frac{1}{2}}], \quad (66)$$

or

$$\frac{\partial lnL}{\partial \lambda} = \sum_{i=1}^{N}\frac{1}{1-F_1^*}f_1^*\left[\left(-\sigma^{-1}k\beta\left(\frac{1}{\lambda^2}\right)\right) - \left(\left(\frac{1}{\sigma}\right)(y_i-b'x_i)\right)\right] - 2n\frac{1}{1-F_2^*}f_2^*[\frac{k\beta}{\sigma}\lambda^{-3}(\lambda^{-2}+1)^{\frac{1}{2}}], \quad (67)$$

or

$$\frac{\partial lnL}{\partial \lambda} = \sum_{i=1}^{N}\frac{1}{1-F_1^*}f_1^*\left[\left(-\left(\frac{1}{\sigma}\right)*\left(\frac{k\beta}{\lambda^2}\right)\right) - ((y_i-b'x_i)/\sigma)\right] - 2n\frac{1}{1-F_2^*}f_2^*[\frac{k\beta}{\sigma}(\frac{1}{\lambda^3})(\lambda^{-2}+1)^{\frac{1}{2}}], \quad (68)$$

or

$$\frac{\partial lnL}{\partial \lambda} = \sum_{i=1}^{N}\frac{1}{1-F_1^*}f_1^*\left[\left(-\left(\frac{k\beta}{\sigma\lambda^2}\right)\right) - ((y_i-b'x_i)/\sigma)\right] - 2n\frac{1}{1-F_2^*}f_2^*[(\frac{k\beta}{\sigma\lambda^3})(\lambda^{-2}+1)^{\frac{1}{2}}], \quad (69)$$

or

$$\frac{\partial lnL}{\partial \lambda} = \sum_{i=1}^{N}\frac{f_1^*}{1-F_1^*}\left[\left(-\left(\frac{k\beta}{\sigma\lambda^2}\right)\right) - ((y_i-b'x_i)/\sigma)\right] - 2n\frac{f_2^*}{1-F_2^*}[(\frac{k\beta}{\sigma\lambda^3})(\lambda^{-2}+1)^{\frac{1}{2}}], \quad (70)$$

Finally,

$$\frac{\partial lnL}{\partial \sigma^2} = -\frac{n}{\sigma^2} + \frac{1}{\sigma^4}*\sum_{i=1}^{N}((y_i-b'x_i)-k\beta)^2 + \sum_{i=1}^{N}\frac{f_1^*}{1-F_1^*}\left[\left(-\frac{k\beta}{\sigma^3\lambda}\right) -\right.$$

$$\frac{(y_i-b'x_i)\lambda}{\sigma^3}\Big] + n\frac{f_2^*}{1-F_2^*}\Big[\Big(\frac{k\beta}{\sigma^3}\Big) * (\lambda^{-2}+1)^{\frac{1}{2}}\Big], \quad (71)$$

Taking first order conditions, we have:

$$\frac{\partial lnL}{\partial \lambda} = 0, \quad (72)$$

or

$$\frac{\partial lnL}{\partial \lambda} = \sum_{i=1}^{N} \frac{f_1^*}{1-F_1^*}\Big[\Big(-\Big(\frac{k\beta}{\sigma\lambda^2}\Big)\Big) - \big((y_i - b'x_i)/\sigma\big)\Big] - 2n\frac{f_2^*}{1-F_2^*}\Big[\Big(\frac{k\beta}{\sigma\lambda^3}\Big)(\lambda^{-2}+1)^{\frac{1}{2}}\Big] = 0, \quad (73)$$

Putting (73) into (71), we take:

$$\frac{\partial lnL}{\partial \sigma^2} = -\frac{n}{\sigma^2} + \frac{1}{\sigma^4} * \sum_{i=1}^{N}((y_i - b'x_i) - k\beta)^2, \quad (74)$$

Taking again first order conditions, we have:

$$\frac{\partial lnL}{\partial \sigma^2} = 0, \quad (75)$$

$$\frac{\partial lnL}{\partial \sigma^2} = -\frac{n}{\sigma^2} + \frac{1}{\sigma^4} * \sum_{i=1}^{N}((y_i - b'x_i) - k\beta)^2 = 0, \quad (76)$$

or

$$\frac{1}{\sigma^4} * \sum_{i=1}^{N}((y_i - b'x_i) - k\beta)^2 = \frac{n}{\sigma^2}, \quad (77)$$

or

$$n\sigma^4 = \sigma^2 \sum_{i=1}^{N}((y_i - b'x_i) - k\beta)^2, \quad (78)$$

Dividing both sides of (78) by σ^2, we take:

$$n * (\sigma^4/\sigma^2) = [\sigma^2 \sum_{i=1}^{N}((y_i - b'x_i) - k\beta)^2]/\sigma^2, \quad (79)$$

or

$$\widehat{\sigma^2} = \frac{1}{n} * \sum_{i=1}^{N}((y_i - b'x_i) - k\beta)^2), \quad (80)$$

Equation (80) gives the Maximum Likelihood (ML) estimator and the solution to the problem. Advancing to the appropriate transformations, we also extract the estimations of β which is the vector of coefficients of the determinants of innovation inefficiency (k).

2.4. The Estimation of the Determinants of firm Innovation Inefficiency Through the Conditional Mean of Innovation Inefficiency Given a Composite Error

Here, we assume that innovation inefficiency is distributed as a truncated normal with a mean that varies across firms. Considering that $u_i \sim N(k\beta, \sigma_u^2)$ and $v_i \sim N(0, \sigma_v^2)$, the conditional density function of u_i given e_i, $f(u_i|e_i)$, is $N_+(k\beta_{*i}, \sigma_*^2)$, where:

$$k\beta_{*i} = (-e_i\sigma_u^2)/\sigma^2, \quad (81)$$

and

$$\sigma_*^2 = (\sigma_v^2\sigma_u^2)/\sigma^2, \quad (82)$$

as Jondrow J et al [12] and Kumbhakar S et al [18] show.

In order to estimate the innovation inefficiency at the firm level, we can employ the expected value (mean) or the mode of innovation inefficiency conditional on the realization of composed error of the model $e_i = u_i + v_i$ (see Kumbhakar S et al [18])[6]. Then, we take:

$$\widehat{u_i} = E[u_i|e_i] = k\beta_{*i} + (\sigma_* f(-k\beta_{*i}/\sigma_*))/(1 - F(-k\beta_{*i}/\sigma_*)), \quad (83)$$

where f and F are the standard normal probability density function and cumulative density function, respectively. In turn, we obtain the individual estimates by replacing the true parameters in (83) with Maximum Likelihood estimations from the Stochastic Frontier Approach (Kumbhakar S et al [18]). Then, (83) can be written as:

$$E[u_i|e_i] = k\widehat{\beta_{*i}} + (\widehat{\sigma_*} f(-k\widehat{\beta_{*i}}/\widehat{\sigma_*}))/(1 - F(-k\widehat{\beta_{*i}}/\widehat{\sigma_*})), \quad (84)$$

Here, we have the estimation of variance as well as the estimation of β_{*i} that is the vector of coefficients of k. Thus, we obtain the estimations of the determinants of innovation inefficiency through the conditional mean of innovation inefficiency given a composite error.

3. Conclusions

The present paper demonstrates that the determinants of firm innovation inefficiency can be estimated through the conditional mean of firm innovation inefficiency given a composite error. We replace the true parameters in the equation of the conditional mean of innovation inefficiency given a composite error with the estimations of the parameters from Maximum Likelihood when we apply Stochastic Frontier Analysis. Here, innovation inefficiency has two components: a) a function of some firm-specific characteristics and b) a random component. In this case, innovation inefficiency is distributed as a truncated normal with a mean that varies across firms. This is the main contribution of the present research to the existing literature, and this constitutes an alternative method for the estimation of the determinants of firm inefficiency besides those which are already existent in the relevant literature (see Kumbhakar S et al [7]).

A limitation of the current study relates to the fact that the conditional mean estimator shrinks the innovation inefficiency towards its mean (Wang W and Schmidt P [19], see also Zeebari Z et al [10]). This leads to a different distribution compared to the distribution of the innovation inefficiency. Furthermore, Zeebari Z et al [10: p. 79] argue that "the mean and mode are not fully representative characteristics of the conditional distribution of the inefficiency, especially if each unit is observed once". Therefore, conditional mean estimator leads to an inconsistent estimator of the innovation inefficiency (see Zeebari Z et al [10]). This creates the need for further development of inefficiency and stochastic frontier models

6 The conditional mean of u_i given e_i gives a point prediction of u_i (Kumbhakar S et al [18]).

that deal with the endogeneity bias and the inconsistent estimators of the innovation inefficiency. Discussing the limitations of the present research, we should bear in mind that the distribution of the conditional density function of innovation inefficiency differs from the distribution of the conditional mean estimator. This creates issues related to the inaccurate estimations and underestimations of innovation inefficiency (Badunenko O et al [20], Zeebari Z et al [10]). In order to face these problems, the future research should focus on the models that combine Stochastic Frontier Models and Data Envelopment Analysis (see Andor M et al [21], Tsionas M [22]).

Conflicts of Interest

The authors declare no conflicts of interest.

ORCID

https://orcid.org/0000-0002-9473-5731

References

[1] Farrell M J (1957) The measurement of productive efficiency. *Journal of the Royal Statistical Society, Series A* 120(3): 253-290.

[2] Kalirajan K P and Shand R T (1985) Types of education and agricultural productivity: A quantitative analysis of Tamil Nadu. *Journal of Development Studies* 21(2): 232-243.

[3] Alvarez R and Crespi G (2003) Determinants of technical efficiency in small firms. *Small Business Economics* 20(3): 233-244.

[4] Pestana Barros C and Dieke P U C (2008) Technical efficiency of African hotels. *International Journal of Hospitality Management* 27(3): 438-447.

[5] Page J M (1980) Technical efficiency and economic performance: Some evidence from Ghana. *Oxford Economic Papers* 32(2): 319-339.

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

[7] Kumbhakar S, Ghosh S and McGuckin J T (1991) A generalized production frontier approach for estimating determinants of inefficiency in U. S. dairy farms. *Journal of Business & Economics Statistics* 9(3): 279-286.

[8] Reifschneider D and Stevenson R (1991) Systematic departures from the frontier: A framework for the analysis of firm inefficiency. *International Economic Review* 32(3): 715-723.

[9] Battese G and Coelli T (1995) A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics* 20(2): 325-332.

[10] Zeebari Z, Mansson C Sjolander P Soderberg M (2023) Regularized conditional estimators of unit inefficiency in stochastic frontier analysis, with application to electricity distribution market. *Journal of Productivity Analysis* 59(1): 79-97.

[11] De Borger B, Kerstens K Moesen W and Vanneste J (1994) Explaining differences in productive efficiency: An application to Belgian municipalities. *Public Choice* 80(3-4): 339-358.

[12] Jondrow J, Lovell K C A Materov I S and Schmidt P (1982) On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics* 19(2-3): 233-238.

[13] Pellegrino G and Piva M (2020) Innovation, industry and firm age: are there new knowledge production functions? *Eurasian Business Review* 10(1): 65-95.

[14] Catozzella A and Vivarelli M (2007) Beyond the Knowledge Production Function: The role of R&D in a multi-faceted innovative process. Jena Economic Research Paper No. 2007-087.

[15] Ramani S, El-Aroui M-A and Carrere M (2008) On estimating a knowledge production function at the firm and sector level using patent statistics. *Research Policy* 37(9): 1568-1578.

[16] Stevenson R (1980) Likelihood functions for generalized stochastic frontier estimation. *Journal of Econometrics* 13(1): 57-66.

[17] Aigner D, Lovell K C A and Schmidt P (1977) Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics* 6(1): 21-37.

[18] Kumbhakar S, Parmeter C and Zelenyuk V (2020) Stochastic Frontier Analysis: Foundations and Advances II. In S. Ray R. Chambers and S. Kumbhakar (eds.) *Handbook of Production Economics* (pp. 1-40). Singapore: Springer.

[19] Wang W S and Schmidt P (2009) On the distribution of estimated technical efficiency in stochastic frontier models. *Journal of Econometrics* 148(1): 36-45.

[20] Badunenko O, Henderson D and Kumbhakar S (2012) When, where and how to perform efficiency estimation. *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 175(4): 863-892.

[21] Andor M, Parmeter C and Sommer S (2019) Combining uncertainty with uncertainty to get certainty? Efficiency analysis for regulation purposes. *European Journal of Operational Research* 274(1): 240-252.

[22] Tsionas M G (2021) Optimal combinations of stochastic frontier and data envelopment analysis models. *European Journal of Operational Research* 294(2): 790-800.