Crawford School of Public Policy

Crawford School working papers

An Alternative Specification for Technical Efficiency Effects in a Stochastic Frontier Production Function

Crawford School Working Paper 1703
February 2017

Satya Paul
Centre for Economic Fragility and Governance, Amrita University, India

Sriram Shankar
ANU Centre for Social Research and Methods

Abstract

This paper proposes an alternative specification for technical efficiency effects in a stochastic production frontier model. The proposed specification is distribution free and thus eschews one-sided error term present in almost all the existing inefficiency effects models. The efficiency effects are represented by the standard normal cumulative distribution function of exogenous variables which ensures the efficiency scores to lie in a unit interval. An empirical exercise based on widely used Philippines rice farming data set illustrates the simplicity and usefulness of the proposed model.

THE AUSTRALIAN NATIONAL UNIVERSITY

Address for Correspondence:
Name   Dr Sriram Shankar
Position   Fellow
Address   ANU Centre for Social Research and Methods
Tel: +61 2 6125 2310
Email: sriram.shankar@anu.edu.au

Crawford School of Public Policy

College of Asia and the Pacific

The Australian National University

Canberra ACT 0200 Australia

www.anu.edu.au

The Crawford School of Public Policy is the Australian National University's public policy school, serving and influencing Australia, Asia and the Pacific through advanced policy research, graduate and executive education, and policy impact.

THE AUSTRALIAN NATIONAL UNIVERSITY

An Alternative Specification for Technical Efficiency Effects in a Stochastic Frontier Production Function

Satya Paul

Centre for Economic Fragility and Governance, Amrita University, India

and

Sriram Shankar[1]

Australian National University

Abstract

This paper proposes an alternative specification for technical efficiency effects in a stochastic production frontier model. The proposed specification is distribution free and thus eschews one-sided error term present in almost all the existing inefficiency effects models. The efficiency effects are represented by the standard normal cumulative distribution function of exogenous variables which ensures the efficiency scores to lie in a unit interval. An empirical exercise based on widely used Philippines rice farming data set illustrates the simplicity and usefulness of the proposed model.

[1] Corresponding author: 2.21 Beryl Rawson Building, ANU Centre for Social Research and Methods, Australian National University, email: Sriram.Shankar@anu.edu.au

1. Introduction

Since the seminal work of Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977), the literature on efficiency studies has proliferated. Among the wide variety of stochastic production frontier efficiency models, the most important ones are those that consider exogenous determinants of inefficiency effects in addition to estimating firm's efficiency. These inefficiency effects models are classified into two groups. The first group of models follow a two-step procedure in that the production frontier is first estimated and the technical inefficiency of each firm derived. These are subsequently regressed against a set of variables which are hypothesized to influence the firm's inefficiency. This approach has been adopted in a range of studies (e.g. Kalijaran, 1981; Pitt and Lee, 1981).

A problem with the two-stage procedure is the lack of consistency in assumptions about the distribution of inefficiencies. In the first stage, inefficiencies are assumed to be independently and identically distributed (i.i.d.) in order to estimate their values. However, in the second stage, estimated inefficiencies are assumed to be a function of a number of firm-specific factors, and hence are not identically distributed (Battese and Coelli, 1995). Another problem is that if the production function variables and efficiency effect variables are correlated, then the first step of the two-step procedure is plagued by omitted variable bias (Parmeter and Kumbhakar, 2014). Even if they are uncorrelated, the Monte Carlo simulations undertaken in Wang and Schmidt (2002) show that the second stage regression parameter estimates are likely to be biased downwards.

The second group of inefficiency effects models estimate the inefficiency scores and exogenous effects in one step. Amongst these models, the most popular are those of Kumbhakar, Ghosh and McGuckin (1991), Huang and Liu (1994) and Battese and Coelli (1995). In order to study the exogenous influence on inefficiency, these authors parameterize

the mean of pre-truncated distribution. These models are further complemented by Caudill and Ford (1993), Caudill, Ford and Gropper (1995) and Hadri (1999) by accounting for potential heteroscedasticity by parameterizing the variance of the pre-truncated distribution. Wang (2002) proposes a more general model that combines the two strands of one-step models mentioned above.

In this paper, we propose an alternative parametric efficiency effects model wherein the technical efficiency is represented by a cumulative density function of exogenous variables, ensuring the efficiency scores to lie in the unit interval. The existing inefficiency effects models contain two error terms and are estimated using maximum likelihood technique. One of the error terms is a symmetric random variable representing statistical noise (v) and the other is a one-sided random variable (u) which represents inefficiency. These models utilise a transformation proposed by Jondrow, Lovell, Materov and Schmidt (1982) (popularly known as JLMS estimator) in order to derive the technical inefficiency scores. Schmidt and Sickles (1984) have shown that the JLMS estimator is not consistent in the sense that the conditional mean or mode of the random variable representing inefficiency component (u) given the composite error (v-u) term, that is, $u|v-u$ never approaches u even when the number of cross-sectional units tends to infinity. However, if the panel data are used, the latter limitation can be overcome under certain other assumptions, some of which may be less realistic (Parmeter and Kumbhakar, 2014). The efficiency effects model/ specification proposed in this paper is much simpler and distribution free as it eschews one-sided error term present in almost all the existing inefficiency effect models. Our model contains only a statistical noise term (v), and its estimation can be done in a straight forward manner using non-linear least squares. Once the parameters are estimated, the efficiency scores are calculated directly.

Parmeter and Kumbhakar (2014) discuss about a distribution free inefficiency effects model which was first proposed in Simar, Lovell and van den Eeckaut (1994) and later explained in details in Wang and Schmidt (2002) and Alvarez, Amsler, Orea and Schmidt (2006). Parmeter, Wang and Kumbhakar (2016) non-parametrically estimate distribution free inefficiency effects using a partly linear model initially proposed by Robinson (1988). This model is similar to the one proposed by Deprins and Simar (1989a, 1989b) and extended by Deprins (1989). However, unlike the model proposed in this paper, these distribution free inefficiency effects models suffer from certain limitations which we shall discuss later in this paper.

The rest of the paper is organized as follows. Section 2 discusses the proposed model and estimation procedure. This is followed by a brief discussion of the limitations of existing distribution free 'inefficiency effects' specifications. Section 3 provides an empirical illustration of the proposed model using the Philippines data collected by the International Rice Research Institute for a panel of 43 rice farms observed over 8 years from 1990 to 1997 in the Tarlac region. Concluding remarks are made in Section 4.

2. Model

The specification of our stochastic production frontier efficiency effects model is as follows.

$$Y_{it} = \exp(x_{it}\beta + v_{it})H(z_{it}\gamma) \tag{1}$$

where the subscripts i and t respectively represent cross-sectional unit and time period, Y_{it} is the quantity of output, x_{it} represents a $(1 \times K)$ vector whose values are functions of input

quantities and time, β represents the corresponding coefficient vector $(K \times 1)$ and v_{it} represents the random noise. $H(z_{it}\gamma)$ represents the efficiency term and is required to lie between 0 and 1, that is, $0 \leq H(z_{it}\gamma) \leq 1$. Any cumulative distribution function (cdf) will satisfy this property. In this paper, we assume the efficiency term to take a probit functional form, that is, $H(z_{it}\gamma) = \Phi(z_{it}\gamma)$, where Φ is a standard normal cdf, the vector z_{it} contains a constant 1 and exogenous variables hypothesized to influence efficiency and γ is the corresponding coefficient vector.

Taking logarithm on both sides of (1) we have

$$y_{it} = \ln(Y_{it}) = x_{it}\beta + \ln\left(H(z_{it}\gamma)\right) + v_{it} = g(x_{it}, z_{it}, \theta) + v_{it} \tag{2}$$

where $g(x_{it}, z_{it}, \theta) = x_{it}\beta + \ln\left(\Phi(z_{it}\gamma)\right)$ and $\theta = \left(\beta', \gamma'\right)'$. Equation (2) can be estimated by minimizing the following sum of squared errors with respect to parameter vector θ:

$$Q_N\left(\theta\right) = \sum_{i=1}^{N}\sum_{t=1}^{T}\left(y_{it} - g(x_{it}, z_{it}, \theta)\right)^2$$

The estimation can be done using the nonlinear least squares option available in any standard econometric package such as EViews.

Parmeter and Kumbhakar (2014) discuss a model possessing the scaling property, which can be estimated without making any distributional assumptions. Their model can be written as

$$y_{it} = x_{it}\beta + v_{it} - e^{z_{it}\gamma}u_{it} \tag{3}$$

where $e^{z_{it}\gamma}$ is the scaling function and u_{it} the basic distribution (such as half-normal or truncated normal). The conditional mean of y, given x and z is

$$E\left(y_{it} | x_{it}, z_{it}\right) = x_{it}\beta - e^{z_{it}\gamma}\mu \tag{4}$$

where $\mu = E(u_{it})$. The regression model given by (3) can be re-written as

$$y_{it} = x_{it}\beta - e^{z_{it}\gamma}\mu + v_{it} - e^{z_{it}\gamma}(u_{it} - \mu) = x_{it}\beta - e^{z_{it}\gamma}\mu + \varepsilon_{it} \tag{5}$$

where $\varepsilon_{it} = v_{it} - e^{z_{it}\gamma}(u_{it} - \mu)$ is independent but not identically distributed. Equation (5) can

be estimated with nonlinear least squares by minimizing $\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{it} - x_{it}\beta + e^{z_{it}\gamma}\mu)^2$.

Parmeter, Wang and Kumbhakar (2016) estimate the following partly linear regression model

initially proposed by Robinson (1988), which does not invoke the scaling property:

$$y_{it} = x_{it}\beta + v_{it} - u_{it} = x_{it}\beta - g(z_{it}\gamma) + v_{it} - (u_{it} - g(z_{it}\gamma)) = x_{it}\beta - g(z_{it}\gamma) + \varepsilon_{it} \tag{6}$$

where $\varepsilon_{it} = v_{it} - (u_{it} - g(z_{it}\gamma))$ and $E(u_{it}) = g(z_{it}\gamma) > 0$. In order to estimate β the following

equation is required.

$$y_{it} - E(y_{it}|z_{it}) = (x_{it} - E(x_{it}|z_{it}))\beta + \varepsilon_{it} \tag{7}$$

Since, $E(y_{it}|z_{it})$ and $E(x_{it}|z_{it})$ are unknown, in order to obtain a consistent estimate of β for

the partly linear model of Robinson (1988) the conditional means are replaced with their

nonparametric estimates.

Parmeter and Kumbhakar (2014) mention that the above two models, (5) and (6), suffer from

certain limitations. First, to avoid identification issues, z and x respectively in model (5) and

(6) cannot contain 1 to represent the intercept. Second, in the first model as ε depends on z

through $e^{z\gamma}$, x and z cannot contain common elements. However, Parmeter, Wang and

Kumbhakar (2016) show that $x - E(x|z)$ in (7) is uncorrelated with ε and so the correlation

between z and x is not an issue for them. Finally, as noted in Parmeter and Kumbhakar

(2014), it is possible to obtain negative estimates of $g(z)$ in model (6), which is inconsistent

with the notion that $g(z)$ represents average inefficiency. The model that we have proposed in this paper suffers from no such limitations.

3. An Empirical Illustration with Philippines Rice Data

We provide an empirical illustration of the proposed model of estimating technical efficiency using the Philippines data for a panel of 43 rice farms observed over 8 years from 1990 to 1997 in the Tarlac region. Rice is grown only in the rainy season, hence all the farms in the region get the same rainfall during any year. However, the rainfall varies across years. These data have been widely used in empirical studies such as Coelli, Rao, O'Donnell and Battese (2005) and Griffiths and Hajargasht (2016).

The stochastic frontier efficiency model that we estimate consists of a Cobb-Douglas Production frontier and an efficiency effects specification. The use of Cobb-Douglas function for the production frontier is in line with several other studies based on the Philippines data set (e.g. Griffiths and Hajargasht, 2016). For the production function, the output variable is tons of freshly threshed rice (*Y*). Input variables are: area planted for rice cultivation in hectares (*AREA*), person-days of hired and family labor (*LABOR*), kilograms of fertilizer (*FERTILIZER*) and *OTHER* which is a Laspeyres quantity index that combines inputs of seed, insecticides, herbicides, animals and tractors used during land preparation[2]. We also included year variable (*t*) which accounts for Hicksian neutral technological change. For the efficiency effects, the determinants (*z*) are: rainfall in mm. (*RAIN*), education of household head (*EDU*), age of the household head in years (*AGE*), and proportion of area classified as bantog (upland) fields which generally do not get flooded (*BANRAT*). Bantog

---

[2] Further details on *OTHER* are provided in Coelli, Rao, O'Donnell and Battese (2005).

fields are known to be better than those low land farms (called lubog) which are flood prone (see Pandey, Masicat, Velasco, and Villano, 1999). We expect that rainfall, education of household head and proportion of area under bantog are likely to enhance the technical efficiency in rice farming. However, it is difficult to predict *a priori* the sign of the effect of age of the household head on efficiency. If the younger people have better knowledge of farming techniques and management then the farms headed by young household heads are likely to be more technically efficient, other things remaining the same. On the other hand, if the experience gained over the years matters for farming, then farms managed by older household heads might be technically more efficient. Thus, the effect of age of household head on technical efficiency is an empirical issue.

The summary statistics of data presented in Table 1 reveal that the average size of land area under rice cultivation is 2.14 hectares. The minimum farm size is 0.20 hectare as against the maximum size of 7 hectares in the sample. The rainfall varies considerably across years. The percentage of area classified as bantog (upland) fields is 73%.

Table 1: Summary Statistics of Data

Variables	Mean	Std. Dev.	Minimum	Maximum
Output (in tonnes)	6.54	5.10	0.37	31.10
AREA (in hectares)	2.14	1.46	0.20	7.00
LABOR (person-days of hired and family labor)	108.34	77.19	8.00	437.00
FERTILIZER (in Kilograms)	189.23	169.80	10.00	1030.90
OTHER (Laspeyres quantity index)	125.35	158.24	1.46	1083.78
AGE (Age of Household head) (in years)	49.44	11.02	25.00	81.00
EDU (years of schooling of household head)	7.24	1.91	6.00	14.00
BANRAT (proportion of area classified as bantog (upland) fields	0.73	0.29	0.00	1.00
RAIN (rainfall in mm.)	1166.08	312.09	700.10	1618.80

 Note: The total number of observations on each of the above variables is 344.

The stochastic frontier efficiency model is specified as

$$\ln\left(Y_{it}\right) = \beta_0 + \beta_1 \ln(AREA_{it}) + \beta_2 \ln(LABOR_{it}) + \beta_3 \ln(FERTILIZER_{it}) + \beta_4 \ln(OTHER_{it})$$
$$+ \beta_5 t + \ln\left(\Phi\left(\gamma_0 + \gamma_1 AGE_{it} + \gamma_2 EDU_{it} + \gamma_3 BANRAT_{it} + \gamma_4 RAIN_{it}\right)\right) + v_{it}$$

$$(8)$$

The non-linear least squares (NLS) estimates of parameters of the model are obtained using

EVIEWS9 software package. These estimates are presented below.

Stochastic frontier:

$$\ln\left(\hat{Y}_{it}\right) = -1.457 + 0.394\ln(AREA_{it}) + 0.317\ln(LABOR_{it}) + 0.268\ln(FERTILIZER_{it}) + 0.002\ln(OTHER_{it}) + 0.058t$$
$$(0.319) \quad (0.076) \qquad\qquad (0.074) \qquad\qquad\qquad (0.036) \qquad\qquad\qquad\qquad (0.019) \qquad\qquad\qquad (0.012)$$

Technical efficiency effects:

$$\hat{TE}_{it} = \Phi\left(-0.461 - 0.019AGE_{it} - 0.011EDU_{it} + 0.845BANRAT_{it} + 0.002RAIN_{it}\right)$$
$$(0.600) \quad (0.007) \qquad (0.036) \qquad\quad (0.325) \qquad\quad (0.001)$$

The values within parentheses are heteroscedasticity consistent standard errors (White, 1980). For overall significance F-stat $(10,333) = 243.04$ and number of observations used in the regression is 344.

The coefficients of *AREA*, *LABOR*, *FERTILIZER* and *OTHER* are the output elasticities in the production function. All the output elasticities are positive and statistically significant except the one with respect to *OTHER* which is statistically insignificant. In terms of magnitude of elasticities, the area under cultivation is the most important determinant of

output, followed by labor and fertilizer. The coefficient on the time variable (*t*) indicates that there is significant technological progress.

In the technical efficiency effects model, the coefficient of *BANRAT* and *RAIN* are positive and statistically significant implying that technical efficiency increases with rainfall and proportion of cultivated area classified as bantog (upland). The coefficient of *AGE* is negative and significant, implying that, *cetris paribus*, farms managed by younger household heads are more efficient than those managed by older household heads. The coefficient of *EDU* is negative but statistically insignificant. The null hypothesis that there are no efficiency effects ($H_0 : \gamma_0 = \gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 0$) is rejected at the 1% significance level by the Wald test (F-stat (5, 333) = 4.047).

The technical efficiency levels vary from 0.526 to 0.998 across farms and years (Table 2). The average efficiency level of farms has turned out to be 0.881. Figure 1 indicates that the estimated probability density function of technical efficiency is leptokurtic and skewed to the left. Table 3 presents estimates of technical efficiency scores for each farm averaged over the period 1990- 1997. We find that the top 5 technically efficient farms have an average score of 0.937. 92% of their farm area is upland (bantog) and they are managed by household heads whose average age is 37.5 years. In contrast, the lowest 5 farms have average efficiency score of 0.814, their upland farming area is only 31% of area cultivated, and they are managed by household heads with an average age of 51 years. These observations are in line with technical efficiency effects reported above.

It is also worth noting that the average efficiency level of farms during 1990-1993 is higher than the period 1994-1997 (Table 4). The data reveal that average rainfall during the first period was 1425.37 mm as against 906.77 mm during the second period. It appears that

relatively low average efficiency of farms observed in second period could be attributed to low rainfalls.

Figure 1: Estimated Probability Density Function of Technical Efficiency
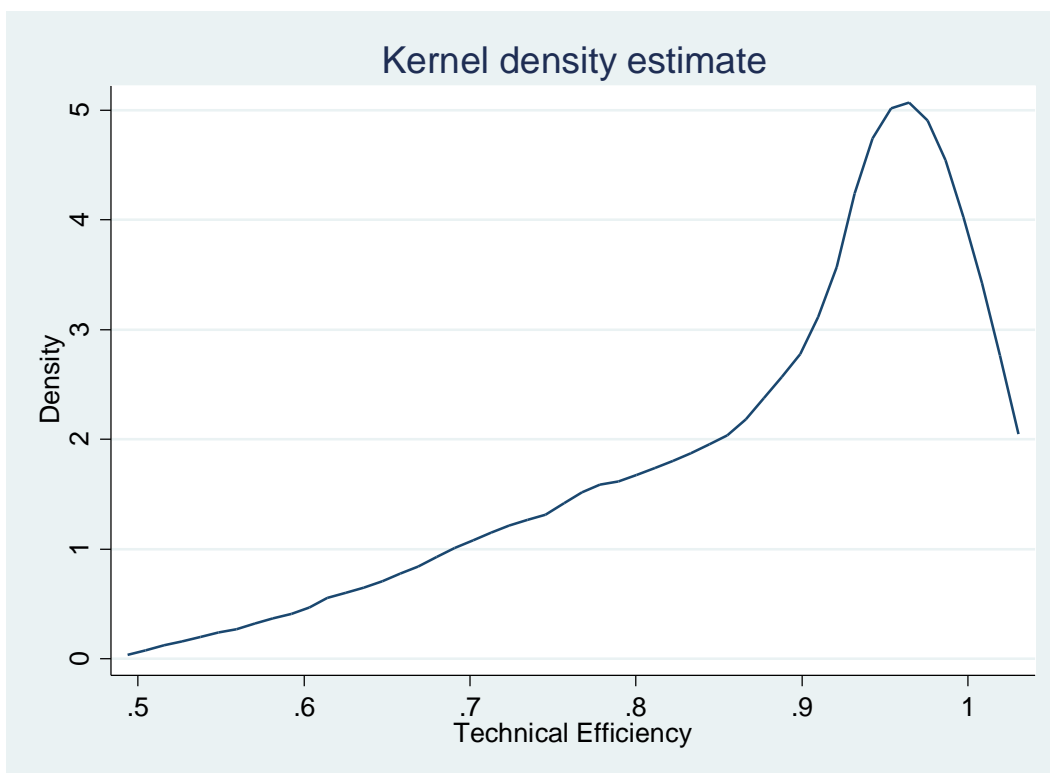

Table 2: Summary Statistics of Estimated Technical Efficiency

Mean	0.881
Maximum	0.998
Minimum	0.526
Std. Dev	0.115
Skewness	-1.021
Kurtosis	3.105
Observations	344

Table 3: Farm-Wise Estimates of Mean Technical Efficiency

Farm	Mean Technical Efficiency	Ranking
1	0.935	3
2	0.931	4
3	0.853	34
4	0.860	32
5	0.878	27
6	0.909	14
7	0.924	6
8	0.904	15
9	0.858	33
10	0.860	30
11	0.823	39
12	0.827	38
13	0.911	11
14	0.896	18
15	0.909	12
16	0.881	21
17	0.909	13
18	0.914	9
19	0.878	26
20	0.880	23
21	0.860	31
22	0.829	37
23	0.803	43
24	0.812	42
25	0.897	17
26	0.949	1
27	0.871	29
28	0.880	25
29	0.814	41
30	0.923	7
31	0.945	2
32	0.887	20
33	0.900	16
34	0.820	40
35	0.917	8
36	0.839	36
37	0.914	10
38	0.928	5
39	0.845	35

40	0.892	19
41	0.880	24
42	0.881	22
43	0.876	28
Overall Mean	0.881	

Table 4: Year-Wise Mean Technical Efficiency

1990	0.986
1991	0.962
1992	0.921
1993	0.988
1994	0.826
1995	0.914
1996	0.682
1997	0.772
1990-1993	0.964
1994-1997	0.798
1990-1997	0.881

4. Conclusions

This paper has proposed a specification for estimating technical efficiency effects in the stochastic production frontier model. The proposed specification eschews one-sided error term present in almost[3] all the existing inefficiency effects models. The efficiency effects are represented by a standard normal cumulative distribution function of exogenous variables, ensuring the efficiency scores to lie between zero and one. The efficiency effects model is distribution free. The efficiency scores are obtained directly once parameters of the model are

---

[3] Even though the model (equation 5) as proposed in Parmeter and Kumbhakar (2014) requires no distributional assumptions for the inefficiency term, it does invoke the scaling property in which the inefficiency term is initially assumed to have a basic distribution such as half or truncated normal distribution. Further, Parmeter, Wang and Kumbhakar (2016) make no distributional assumptions concerning the inefficiency term but the estimation is performed in a non-parametric framework.

estimated using non-linear least squares. An empirical exercise based on the Philippines data set of rice farms illustrates the simplicity and usefulness of the proposed methodology.

The research presented in this paper may be extended as follows. We have represented the efficiency effects using the cdf of a standard normal distribution function. It would be interesting to see the sensitivity of results based on other cumulative distribution functions such as logistic and student-t. Further, it would be worthwhile to empirically compare inefficiency effects based on existing distributions such as half normal, exponential, truncated normal and gamma with models of efficiency effects represented by a cumulative distribution function.

References

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

Alvarez, A., Amsler, C., Orea, L. and Schmidt, P. (2006), "Interpreting and testing the scaling property in models where inefficiency depends on firm characteristics," *Journal of Productivity Analysis*, 25, 201–212.

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

Caudill, S.B. and Ford, J.M. (1993), "Biases in frontier estimation due to heteroscedasticity," *Economic Letters*, 41, 17–20.

Caudill, S.B., Ford, J.M. and Gropper, D.M. (1995), "Frontier estimation and firm-specific inefficiency measures in the presence of heteroscedasticity," *Journal of Business & Economic Statistics*, 13, 105–111.

Coelli, T.J., Rao. D.S.P., O'Donnell C.J. and Battese G.E. (2005), *An Introduction to efficiency and productivity analysis* (2nd ed.), Springer, New York.

Deprins, D. (1989), Estimation de frontieres de Production et Mesures de l'Efficacite Technique, Louvain-la-Neuve, Belgium: CIACO.

Deprins, P. and Simar, L. (1989a), "Estimating technical efficiencies with corrections for environmental conditions with an application to railway companies," *Annals of Public and Cooperative Economics*, 60, 81–102.

Deprins, P. and Simar, L. (1989b), "Estimation de frontieres deterministes avec factuers exogenes d'inefficacite," Annales d'Economie et de Statistique 14, 117–150.

Griffiths, W. and Hajargasht, G. (2016), "Some models for stochastic frontiers with endogenity," *Journal of Econometrics*, 190, 341-348.

Hadri, K. (1999), "Estimation of a doubly heteroskedastic stochastic frontier cost function," *Journal of Business & Economic Statistics,* 17, 359–363.

Huang, C.J. and Liu, J.T. (1994), "Estimation of a non-neutral stochastic frontier production function," *Journal of Productivity Analysis*, 5, 171–180.

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

Kalirajan, K. (1981), "An econometric analysis of yield variability in paddy production," *Canadian Journal of Agricultural Economics*, 29, 283-294.

Kumbhakar, S.C., 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 & Economic Statistics*, 9, 279–286.

Meeusen, W. and van den Broeck, J. (1977), "Efficiency estimation from Cobb-Douglas production functions with composed error," *International Economic Review*, 18, 435-444.

Pandey, S.P., Masicat, P., Velasco, L. and Villano, R. (1999), "Risk analysis of rainfed rice production systems in Tarlac central luzon," *Experimental Agriculture*, 35, 225-237.

Parmeter, C.F. and Kumbhakar, S.C. (2014), "Efficiency analysis: A primer on recent advances", *Foundations and Trends(R) in Econometrics, Now publishers,7, 191-385.*

Parmeter, C. F., Wang, H. J. and Kumbhakar, S.C. (2016), "Nonparametric estimation of the determinants of inefficiency," *Journal of Productivity Analysis*, DOI: 10.1007/s11123-016-0479-x

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

Robinson, P.M. (1988), "Root-N-consistent semiparametric regression," *Econometrica*, 56, 931–954.

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

Simar, L., Lovell, C.A.K. and van den Eeckaut, P. (1994), "Stochastic frontiers incorporating exogenous influences on efficiency", *Discussion Papers No. 9403*, Institut de Statistique, Universite de Louvain.

Wang, H.J. (2002), "Heteroscedasticity and Non-Monotonic Efficiency Effects of a Stochastic Frontier Model," *Journal of Productivity Analysis*, 18, 241-253.

Wang, H.J. and Schmidt, P. (2002), "One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels," *Journal of Productivity Analysis*, 18,129–144.

White, H. (1980), "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity," *Econometrica*, 48, 817–838.