Estimating a mixture of Stochastic Production Frontier - Technology gap and technical efficiency measures

Sawssen HAJJI DHAHRI *
GAFSA University (TUNISIA)
(sawssenh@gmail.com)

Received:09/03/2020 ; **Revised:**10/04/2020 ; **Accepted:**17/07/2020

Summary: This study aims to assess the ignorance impact of the technological heterogeneity on the technical efficiency measurement, using stochastic production frontier. We proposed a methodology based on a switching regression model in order to take account of the technology differences in the frontier specification. Indeed, we extend the switching model developed by Caudill (2003) for the panel case in two ways.First, we incorporate technical inefficiency in the model. Then, we consider that each firm has a specific probability to use one of the two technologies, in the industry. Next, we construct two technology homogenous groups. Finally, we estimate two production frontiers for each group of technologically homogeneous firms in order to estimate 'net' technical inefficiency scores.

This methodology was applied to a panel of French textile firms (1993-1997). The empirical results show that the technical efficiency is very low, 45% on average, under the homogeneous technology hypothesis.However, the average technical efficiency is about 80% with the extended switching regression model. Thus, if we ignore the technology heterogeneity, we underestimate the technical efficiency of approximately 30% on average

Keywords: Technology; technical efficiency; switching regression; French textile industry.
Jel Classification Codes : C01 ; C13 ; C51 ; E23.

I-Introduction:

Previous studies on stochastic frontier models consider that the same technology is shared by all the firms in the sample studies. The measurement of the inefficiency deduced could be highly sensitive to the firms which adopt technological innovations and have access to the new technologies in their production process. For instance, if the firms do not use the same technology, the "best practice" frontier can be represented by the firms which use the best technology. Moreover, for highly inefficient firms belonging to the same sample, the traditional measure of technical inefficiency will not only reflect the inputs misuse, but it could also reflect the differences in productivity due to differences in the technologies used.So, the traditional measure of technical inefficiency will be biased in this case. In terms of economic policy, the distinction between pure technical inefficiency and technology differences is important. The firms which are not using the recent technology should be penalized in terms of efficiency. In that case, they need to change their technology to improve their efficiency. If we have access to adequate informations dealing with the technology used by the firms in the data sets, the distinctions between technology differences and technical inefficiency differences is easy. Without such information, this distinction becomes more complicated.

In the literatures review, we find three main approaches to take account of the technology heterogeneity:

* Sawssen HAJJI DHAHRI sawssenh@gmail.com

* Caudill (2003) proposed the switching regression model to take account of possible differences in the technologies used by the firms. The model is very useful to detect differences in the technologies used, but it ignores possible differences in technical inefficiency of the firms studies. Thus, high differences in the technical inefficiency of the firms in the sample could be considered as differences in the technology used even if the firms are using the same technology. This model has been initially proposed by Gropper et al. (1991),and extended later by the same authors in order to estimate a time variant proportion$\lambda_t$in panel data. The problem is that this model suffers from the same limits that standard Switching model suffers from.

* Kalirajan et al. (1994, 1996) proposed a methodology based on a random coefficient model in order to take account of the technology differences in the frontier specification. The main idea of these authors is that there could be a non-neutral shift of the frontier. They consider a Cobb-Douglas random coefficient model specification in order to estimate the firm-specific technical efficiency for 48 machine building state enterprises in China. As the factor elasticity reflecting the firm specific technology differences, the best practice frontier is obtained by taking the maximum of these coefficients. This model is very interesting because it takes account of possible differences in the technologies used and of inefficiency differences, but it fails to decompose the total inefficiency measure obtained into technical inefficiency differences and into technology differences.

*Akahavein et al. (1997) proposed a model which takes account of possible differences in the efficiency and the technologies used. They estimate a linear profit frontier to allow each one to have a different potential of achieving maximal levels of profit. Moreover, technical and allocative inefficiency are incorporated in the model by adding an asymmetric component of inefficiency in the demand equation via the Hotelling lemma. The main limit of this model is that it could just be used for some linear dual functional forms for the technology. Moreover, there is no distinction between technology gap and differences in technical inefficiency.

* Tsionas (2002) proposes a stochastic frontier model with random coefficients to separate technical inefficiency and technological differences. They consider a Cobb-Douglas cost function, and they use the Bayesian techniques to estimate this model. This model is very helpful because it takes account of possible differences in the technologies used and of efficiency differences. However, the authors consider that the intercept is constant (non-stochastic). In addition, there are many assumptions relative to the distribution of different errors terms.

In this paper, we extend the switching model developed by Caudill (2003) for the panel case in two ways. First, we incorporate technical inefficiency in the model. Next, we consider that each firm has a specific probability obtained by the model. Then, we construct two groups with homogeneous technology (this firm classification is based on the estimated probabilities obtained in the previous step). Finally, we estimate two production frontiers for each firms group in order to estimate what we call technical inefficiency net levels. This methodology is applied to a French panel of the textile industry (1993-1997). The preliminary results show that the technical efficiency is very low (nearly 45% on average) under the homogeneous technology assumption. However, the average technical efficiency is about 80% with the extended switching regression model. Thus, if we ignore the technology heterogeneity, we underestimate the technical efficiency of approximately 30% on average.

The paper is divided into four sections. In section 1, we derive the main characteristics of the "classic" switching regression model. In section 2, we develop the extended version of switching regression model. In section three, we give a brief description of the data used and an illustration of our empirical results.

II- The Switching Regression model

Before introducing the concept of technical efficiency in the switching regression model, it is important to present the main characteristics of the classic switching regression model. Quandt (1972) was the first author to apply a mixture of normal distributions approach in the economy field. This approach has already been equally processed by Battachaya since (1969), in the biology field. Quandt et Ramsey (1975) suggest that the switching regression models in an industry that is composed by two technologies are written as follows in the panel data case:

$y_{it1}=g_1(x_{it1};\beta_1^{'})+v_{it1}$ with a probability λ

i=1;2; ….; N , t=1;2; …;T (1)

$y_{it2}=g_2(x_{it2};\beta_2^{'})+v_{it2}$ with a probability (1-λ)

where

g_j represents a production function for the technology j (j=1;2).

y_{itj} represents the output vector produced by firm I at time t, when using thetechnology *j* (*j=1;2*).

x_{itj} is the input vector.

β_j is the production function parameters vector.

v_{itj} is the usual error term.

Therefore, we assume that each firm *i* has a probability λ to belong to the technology 1, and (1-λ) to belong the technology 2.

In the literatures review, we can find several estimation procedures of the switchingregression model. For instance, Day (1967) developed the method of moments. Hartley (1978) and Bead et al. (1991) proposed the EM procedure. Among all these methods, we will follow the EM algorithm that seems to be retrieved for the estimation of the Swtching models.

III-The switching regression model and the productive efficiency:

The classic Switching model ignores technical inefficiency. The error terms in (1) are symmetric and Gaussien. However, it takes account of possible differences in the technology used by each firm in the sample if λ≠ 0. We can improve this model by adding technical inefficiency component to the technology equations. One of the ways to do that is to follow the familiar Aigner, Lovell and Schmidt (1977) specification. For instance, we will add an asymmetric error term u_{it} ($u_{it} \geq 0$) which represents technical inefficiency component) to the production function specification in equation (1). Moreover, we will consider that each firm has its over probability to switch between the two technologies. In other words, we consider a switching parameter λ_i firm specific. Following these extensions, the switching model will be written:

$y_{it1}=g_1(x_{it1};\beta_1^{'})+v_{it1}-u_{it1}$ with a probability λ_i

(2)

$y_{it2}=g_2(x_{it2};\beta_2^{'})+v_{it2}-u_{it2}$ with a probability $(1-\lambda_i)$

This extended version of the switching model differs from the classical one (1) in two ways. First, it takes account of technical inefficiency in the sample (represented by the u_{it} terms). Second, each firm has its probability to use technology 1 or 2. In practice, if we consider that $\lambda_i=\lambda\forall$ i=1;2; ... ; N and u_{it}=0, we have model (1) as a particular case. Here we consider the assumption that the probability of switching is constant over time.

To estimate the model (2), we need to specify the technology used by imposing functional form for the function $g_1()$ and $g_2()$ in equation (2). Several functional firms have been proposed and need

in the empirical literature (Cobb-Douglas, Translog, …). We just need to choose one functional form to estimate the model. It is important to note here that it is not necessary to retain the same functional form for$g_1(\,)$ and $g_2(\,)$in equation (2). However, it is much easier to work with linear functional forms with respect to the β_i parameters.
Before developing the estimation procedures, we need additional assumption on the asymmetric error components, u, in the Switching model. We can consider the following assumption:
* The term u (representing the technical efficiency) is not stochastic.
* The case where the term u is stochastic.
Most of these assumptions are usually used in the panel data frontier modeling.

In this paper, we consider that the term u is stochastic. A particular advantage of this specification is that the frontier specification could also incorporate firm time-invariant attributes as explanatory variables.

The model under second consideration could be written as:

$$y_{it1} = x^{it1'}\beta_1' + \varepsilon_{it1} \qquad \text{with a probability } \lambda_i$$

$$y_{it2} = x_{it2}'\beta_2' + \varepsilon_{it2} \qquad \text{with a probability } (1-\lambda_i) \tag{4}$$

where

$$\varepsilon_{it1} = v_{it1} - u_{i1} \tag{5}$$

$$\varepsilon_{it2} = v_{it2} - u_{i2}$$

$u_{ij} \geq 0$where x' is the observations vector which is conform to the functional form retained for the technology used.

v_{itj} $(j = 1; 2)$ are assumed to be independed and identically distributed as Gaussien with zero mean and varianceσ_{vj}^2. The inefficiency components are treated as firm specific constant in this case.
We need a specific assumption on the distribution of the one sided error component in the model. Several distributions have been proposed in the frontier modeling literature (half normal, exponential, gamma …). We consider the half normal case in here.
To write the likelihood function of model (2), we must determine the density function of the error $\varepsilon_{itj} = v_{itj} - u_{ij}$ at first. If we suppose that the errors terms (v_{itj} and u_{ij}) are independent, this function will be written as follows (Aiger et al. (1977):

$$f_j(\varepsilon_{i.j}) = 2/\sigma_j * f_j(\varepsilon_{i.j}/\sigma_j)(1 - F_j(\varepsilon_{i.j}\delta_j/\sigma_j^*)) \qquad j = 1;2$$

with

$$\sigma_j^2 = (\sigma_{uj}^2 + \sigma_{vj}^2 / T) \qquad ; \qquad \delta_j = (T\sigma_{uj}^2 / \sigma_{vj}^2)^{1/2}$$

and

$$\varepsilon_{i.j} = 1/T\left[\sum_{t=1}^{T}(v_{itj} - u_{ij})\right]$$

F is the cumulative function of the standard normal distribution.

The likelihood function of the model is written as:

$$L = \prod_{i=1}^{N}\left[\left[\lambda_I f_1(\varepsilon_{i.j})\right] + (1-\lambda_I)\left[f_2(\varepsilon_{i.j})\right]\right]$$

The algorithm used previously needs to be changed according to the new formulation of the likelihood function. The new weights are formulated now by:

$$W_i^1 = \lambda_i(\varphi_i^1/\varphi_i)\left[1+(\sigma_i^{(n)}/\delta_1^{(n)})\left[(f(\varepsilon_{i.1}^{(n)}.\delta_1^{(n)}/\sigma_1^{(n)})/((1-F(\varepsilon_{i.1}^{(n)}.\delta_1^{(n)}/\sigma_1^{(n)})*(\varepsilon_{i.1}^{(n)}))\right]\right]$$

$$W_i^2 = \lambda_i(\varphi_i^2/\varphi_i)\left[1+(\sigma_2^{(n)}/\delta_2^{(n)})\left[(f(\varepsilon_{i.2}^{(n)}.\delta_2^{(n)}/\sigma_2^{(n)})/((1-F(\varepsilon_{i.2}^{(n)}.\delta_2^{(n)}/\sigma_2^{(n)})*(\varepsilon_{i.2}^{(n)}))\right]\right]$$

where

$$\varphi_i^1 = f_1(\varepsilon_{i.1}^{(n)})$$
$$\varphi_i^2 = f_2(\varepsilon_{i.2}^{(n)})$$
$$\varphi_i = \lambda_i \varphi_i^1 + (1-\lambda_i)\varphi_i^2$$
$$\varepsilon_{itj}^{(n)} = (y_{it} - x_{it}\hat{\beta}_j^{(n-1)}) \qquad ; j = 1;2$$

The EM algorithm consists in calculating the expressions below in an iterative way until the convergence. At this step, the λ_i estimator is finally determined in the following way:

$$\hat{\beta}^1 = \left[x'\hat{W}^1 x'\right]^{-1}\left[x'\hat{W}^1 y\right]$$
$$\hat{\beta}^2 = \left[x'\hat{W}^2 x'\right]^{-1}\left[x'\hat{W}^2 y\right]$$
$$\hat{\sigma}_1^2 = (1/\varpi_1)(y - x\hat{\beta}^1)'\hat{W}^1(y - x\hat{\beta}^1)$$
$$\hat{\sigma}_1^2 = (1/\varpi_1)(y - x\hat{\beta}^1)'\hat{W}^2(y - x\hat{\beta}^1)$$
$$\hat{\lambda}_i = \sum_{t=1}^{T} \hat{W}_{it}^1 / T$$

where

$$\hat{W}^1 = diag\left[\hat{W}_{11}^1;\hat{W}_{21}^1;....;\hat{W}_{it}^1\right]$$
$$\hat{W}^2 = diag\left[\hat{W}_{11}^2;\hat{W}_{21}^2;....;\hat{W}_{it}^2\right]$$

$$\varpi_1 = \sum_{i=1}^{N}\sum_{t=1}^{T}\hat{W}_{it}^1$$
$$\varpi_2 = \sum_{i=1}^{N}\sum_{t=1}^{T}\hat{W}_{it}^2$$

At this stage, we remind that the aim of this model is to distinguish between the "pure" technical efficiency measures and the technology gaps. To do so, it is necessary to form firms groups that use a common technology on the estimated λ_i basis according to our model specification. In a first step, we will form these groups in the following way: If λ_i< 0.5 we consider that the firm *i* belongs to the first group while if λ_i> 0.5 it belongs to the second group bearing in mind that the firms which belong to the same group are using the same technology. So, we obtain what we call "pure" technical inefficiency measure. In second step, we estimate a specific production frontier for each firm group, and obtain the technical efficiency scores conditional on the groups formed below. So, we obtain what we call "pure" technical inefficiency measure. As technology differences are incorporated in the group's definition, the technical efficiency measure is a net measure not affected by technological differences. There exist several ways to obtain technical efficiency measures which could be used here, we consider the model proposed by Schmidt and Sickles (1984).

Moreover, we define the production frontier for group *j* (*j=1;2*), as follows:

$$y_{jit} = \alpha_j + x_{jit}\theta_j + v_{jit} - u_{ji} \qquad ; i = 1;2; \dots; N \qquad (6)$$
$$t=1;2; \dots;T$$

We can estimate this model by the maximum likelihood method or by the generalized least square GLS method.

Finally, we determine the efficiency scores according to the following expression:

$$\hat{u}_{gi} = Max\hat{\alpha}^*_{gi} - \hat{\alpha}^*_{gi} \qquad (7)$$

with the firm *i* use the technology g if we adopt the latter method.

IV- Empirical results:

The model proposed in section 3 is applied for the French textile industry. We have a panel data of 475 firms observed over the period 1993-1997[2]. We represent the technology by considering two production functions (Cobb-Douglas), Summary statistics of the French textile firms sample is given in table 1.

The objective is to derive estimators of technical efficiency and technology gap. We also propose an explanation of the gap observed in the technologies used.

We estimate a switching model by considering the following cases: technical efficiency is not stochastic and the efficiency is stochastic. At first, we estimate model (4)) and the $\lambda_i$by the method developed in section (3). In the next step, we group the firms according to the adopted technology (using the estimatorsλ_i). Then, we estimate the production frontier for each group to get the scores of the pure technical efficiency.

It is also interesting to estimate a production frontier that does not take into account technological differences.

We begin by estimation of the switching regression model with technical efficiency. Table 2 illustrates the parameters of this switching regression model (4). We note that all the parameters are significant. With a Cobb-Douglas specification, we obtain 273λ_i< 0.5 and 203$\lambda_i \geq$0.5 .

After having formatted the two homogeneous groups, we estimate one frontier for each group. The results are presented in table 3 and 4. At first, we note that all the parameters are significant. Thus, we notice that the French textile sector is capital-intensive. Moreover, if we compare the estimated elasticities of group 1 and 2, it’s clear that there are technological differences.

Now, we can estimate the “net” technical efficiency according to expression (7). The statistics of these efficiency scores for the two groups are presented in tables 5, and 7. The average value of technical efficiency in the French textile industry is 76% for group 1 and 75% for group 2 respectively. However, if we consider that u is stochastic, the technical efficiency for group 1 and 2 is 70% and 85% on average.

[1] data Source :Coface SCRL.

It is also important to get the traditional measure of technical efficiency under the assumption that all the firms in the sample are sharing the same technology. We use the Schmidt and Sickles (1994) model to obtain the technical efficiency scores. The results of these estimations are presented in table 6 below.

The average technical efficiency score is fairly low. It is about 52% over the studied period. According to these analyses, the French textile firms can increase their production by 80% (on average) using the same technology. This result seems to be exaggerated. On the other hand, there are some firms which can survive in the competition of the textile sector with a technical inefficiency more than 80%.

If we compare the results of table 6 and the results of table 5, we notice that the "pure" technical efficiency scores obtained for group1 and group 2 are higher than those obtained under the assumption that the production frontier is common to all the firms.

Thus, when we take account of the heterogeneity of the technology, we also notice that the minimal values of the technical efficiency scores have improved for groups 1 and 2 (71% and 27% respectively).

Moreover, if take account of the technological heterogeneity, we notice that more than 80% of firms in the sample have a productive efficiency over 60%. Nevertheless, if we adopt common frontier, this score would be 18%.

V-Conclusion:

In this paper, we have extended the switching regression model to incorporate technical inefficiency. We also derived efficiency measures taking into account the heterogeneity of the technology. The application of this model into the French Textile Industry case has shown that there is a technological heterogeneity. the "net" technical efficiency scores for the two groups is 76% for group 1 and 75% for group 2 respectively. However, if we consider that the frontier is stochastic, the technical efficiency for group 1 and 2 is 70% and 85% on average.

In this sector and that the technical efficiency obtained from the traditional stochastic frontier modeling overestimates technical inefficiency in this industry by about 28% on average. If we ignore the technology heterogeneity, the French textile firms can increase their production by 80%, on average, using the same technology; this score is not ''net'' technical inefficiency. Indeed, if take account of the technological heterogeneity, we notice that more than 80% of firms in the sample have a productive efficiency over 60%.

- Appendices:

Table (1): Statistics of the variables

	Mean	Standard deviation
CA	$3.39\ (10)^7$	$5.131\ (10)^7$
Capital	9125580	$1.677\ (10)^7$
Labor	54	92

Table (2): Production Specification

Cobb-Douglas	Regime 1	Regime 2
Log(K)	0.77	0.68

	(99.7)	(88.9)
Log(L)	0.25 (37.6)	0.012 (2.7)
Constant	3.67 (37.6)	4.72 (44.7)

Table (3): Production Function (Group1)
N=135; T=4

Variable	Estimator coefficient
Constant	*4.3* *(19.01)*
Log(K)	*0.72* *(42.4)*
Log(L)	*0.28* *(11.7)*

Table (4): Production Function (Group2)
N=340; T=4

Variable	Estimator coefficient
Constant	*5.23* *(31.78)*
Log(K)	*0.73* *(65.7)*
Log(L)	*0.12* *(10.8)*

Table (5) : “pure” technical efficiency measure

	Average	Min	Max
EFFG1	85%	71%	100%
EFFG2	70%	27%	100%

EFFGi : is the technical efficiency of group i

Table 6: efficiency measure (Common frontier)

	Average	Min	Max

EFF*	52%	15%	100%
EFFG1**	53%	16%	100%
EFFG2**	52%	15%	100%

** EFF: designate the measure of technical efficiency for the entire sample*
*** EFFGj: designate the measure of technical efficiency (deduced from the commune frontier)for the firms of group j (j=1;2).*

Referrals and references:

1 Aigner, D., Lovell, C.and Schmidt,P. (1977). **Formulation and estimation of stochastic frontier production function models**.Journal of econometrics. 6, pp 39-56.
2 Akhavein, J., Swamy, P., Taubman, S. andSingamsetti, R.(1997). **A general method of deriving the inefficiencies of banks from a profit function**.Journal of Productivity Analysis. 8, pp 71-93.
3 Battesse, G.andCoelli, T. (992). **Frontier production functions, technical efficiency and panel data with application to paddy farmers in India**. Journal of Productivity Analysis. 3, pp 153-169.
4 Battese, G., Rao, P. (2002).**Technology gap, efficiency, and a stochastic Metafrontier function**. International Journal of Business and Economics. 1, pp 87-93.
5 Beard, T., Caudill, S. and Gropper, D. (1991). **Finite Mixture Estimation of Multiproduct Cost Functions**.Review of Economics and Statistics. 73, pp 654-664.
6 Beard, T., Caudill, S. and Gropper, D. (1997). **The diffusion of production processes in the U.S banking industry : A finite mixture approach**. Journal of Banking and Finance. 21, pp 721-740.
7 Caudill S.(2003).**Estimating a mixture of stochastic frontier regression models via the em algorithm: A multiproduct cost function application.** Empirical Economics. 28, pp 581-598.
8Chaffaï,M.and Plane, P. (2014). **Some recent developments on the measurement of productive performance: application to the Moroccan Garment Sector**. Economie et Développement. 22, pp 91-107.
9 Greene, W. (2005).**Reconsidering heterogeneity in panel data estimators of the stochastic frontier model**. Journal of Econometrics. 126, pp 269-303.
10 Gropper D., Caudill, S. and Beard, T.(1999). **Estimating multiproduct cost functions over time using a mixture of normal**. Journal of Productivity Analysis. 11, pp 201-218.
11Havenner, A. and Swamy, P. (1981). **A random coefficient approach to seasonal adjustment of economic time series**. Journal of Econometrics. 15, pp 177-209.
12Hartley, M. (1978). **Estimating mixtures of normal distributions and switching regression, Comment**. Journal of American Statistical Association. 73(364), pp 738-741.
13Kalirajan, K. et Obwana, M.(1994). **Frontier production: the stochastics coefficient approach**. Oxford Bulleletin of Economics and Statistic. 56, pp 87-95.
14Kim Y., Schmidt P. (2000).**A Review and empirical comparison of Bayesian and classical approaches to inference on efficiency levels in stochastic frontier models with panel data**. Journal of Productivity Analysis. 14,pp 91-118.
15Konstantinos, G., Tran, K., Tzouvelekas, V. (2003).**Predicting technical efficiency in stochastic production frontier models in the presence of misspecification: a Monte-Carlo analysis**. Applied Economics. 35, pp 153-161.
16 Kumbhakar, S., Parmeter, C. andTsionas, E. (2013).**A zero inefficiency stochastic frontier model**. Journal of Econometrics. 172, pp 66-76.

17Quandt, R.(1972).**A new approach to estimating switching regressions**. Journal of American Statistical Association. 67, pp 306-310.
18Quandt, R. (1978).**Estimating mixtures of normal distributions and switching regressions**. Journal of American Statistical Association. 73, pp 730-738.
19 Wallace,M. and Hennessy T. (2016). **Technical efficiency and technology heterogeneity of beef farms: a latent class stochastic frontier approach**. Working paper, University of Warwick, England.
20Schmidt, P. andSickels, R. (1984). **Production frontiers and panel data**. Journal of business and economic statistics. 2 (4), pp 368-374.
21Tsionas, G. (2002). **Stochastic frontier models with random coefficients**. Journal of Applied Econometrics. 17, pp 127-147.

How to cite this article by the APA method:

Sawssen HAJJI DHAHRI (2020), **Estimating a mixture of Stochastic Production Frontier-Technology gap and technical efficiency measures**, Algerian review of economic development, Volume 07 (Number 02), Algeria: Kasdi Marbah University Ouargla, pp. 321-330.