Estimating Technical Efficiency of Crude Palm Oil in Malaysia

Mariah Binti Sabar[1] and Anton Abdulbasah Kamil[2]
[1]*School of Mathematical Sciences, Universiti Sains Malaysia, Penang 11800, Malaysia*
[2]*Faculty of Economics, Administrative and Social Sciences, Istanbul Gelisim University, Istanbul, Turkey*

Keywords: SFA, TE, Cobb-Douglas Functional Form, Translog Functional Form, CPO.

Abstract: The main purpose of this study is to apply parametric techniques in evaluating the technical efficiency (TE) of crude palm oil (CPO) production by the states in Malaysia. To achieve this, the parametric stochastic frontier analysis (SFA) approach was applied. This study involves a panel data consisting of 12 CPO producing states in Malaysia, over a 18 year time period from year 1999 to 2016. The output variable chosen was the annual CPO production and the input variables considered were plantation area, fruit mill capacity, labour and time variable. We found fruit mill capacity, labour and time as input variables that significantly affect the level of CPO output. Plantation area was proven to be statistically insignificant. Technical efficiency was found to be increasing over time. It was also found that the inefficiencies in the industry were mainly caused by 'pure' technical inefficiency rather than scale inefficiency. The overall mean TE of SFA is 0.79. Selangor is the top efficient state according to SFA. We concluded that the state of Malacca is overall the least efficient state due to their low ranking.

1 INTRODUCTION

Malaysia is one of the biggest palm oil producers in the world (Basiron, 2007). The country accounts for 44% of the world's exports of palm oil making the industry the fourth major revenue for the nation (M.P.C., 2014). The industry plays a huge role in the development of the country by reducing poverty rate from 50% in the 1960s, to less than 5% today. The success of the Malaysian palm oil industry, however, did not come without a price. From health campaign claiming the oil increased risk of heart diseases, alleged land grabs, deforestation and the extinction of the orang utan to the recent resolution by the European Parliament calling for the EU to phase out the use of palm oil in biodiesel that are allegedly produced in an unsustainable way, leading to deforestation.

With the continuous pressure and controversies surrounding the manufacturing of palm oil, it is only ideal that the Malaysian palm oil industry demonstrate sustainability by being more efficient in the usage of resources. Measuring efficiency is important not only to have a reliable record of the industry's progress, but also to be able to investigate the impact of any new and already existing implemented policies. Methods for estimating efficiency can be categorized into two, parametric approach and non-parametric approach. These approaches can either be deterministic or stochastic (Bogetoft et al., 2011). Among the various methods developed, parametric stochastic frontier analysis (SFA) is the most commonly used technique for estimating technical efficiency (Baten et al., 2009), (Hassan et al., 2012). The SFA technique involve mathematical programming and econometric methods, respectively (Coelli et al., 2005). To our knowledge, no study has yet used the most applied parametric SFA technique to find the efficiency of producing CPO by the states in Malaysia. The result could be an indicator to where each state stands in terms of producing CPO efficiently among the states in Malaysia. This can serve as a planning aid for management and policy makers to draw conclusion on existing and new regulations.

2 METHODOLOGY

Efficiency Measurement According to (Farrell, 1957), the efficiency of a firm could be looked at from two components; technical efficiency and allocative efficiency. Technical efficiency is the ability of a firm to produce the maximum amount of output from a given set of inputs. Meanwhile, allocative efficiency repre-

Sabar, M. and Kamil, A.
Estimating Technical Efficiency of Crude Palm Oil in Malaysia.
DOI: 10.5220/0012446800003848

In *Proceedings of the 3rd International Conference on Advanced Information Scientific Development (ICAISD 2023)*, pages 198-203
ISBN: 978-989-758-678-1

sents the firm's ability to use the optimal proportions of inputs given their respective prices and the production technology. This study focuses on technical efficiency (TE).

The following notations are used: $i, j = 1, ..., N$ the collection of decision making units (DMU), $t = 1, ..., T$ study period, $k, l = 1, ..., K$ number of inputs.

2.1 Theoretical Stochastic Frontier Model

The model used was the production model for panel data proposed by (Battese and Coelli, 1992) expressed as:

$$lnyit = X_{it}\beta + (v_{it} - u_{it}, \quad (1)$$

$$u_{it} = u_i exp - \eta\left[-(t-T]\right., i = 1, ..., N, t = 1, ..., T \quad (2)$$

where y_{it} is the output of the i-th unit in the $t-th$ time period, x_{it} is a (K x 1) vector of transformation of the input quantities of the $i-th$ unit in the $t-th$ time period, β is a vector of unknown parameters to be estimated, v_{it} are random variables assumed to be independent and identically distributed $N(0, \sigma_v^2)$ and are independent of u_{it}, η is a unknown parameter to be estimated and u_i are non-negative random variables which are assumed to be independent and identically distributed as truncations at zero of the $N(0, \sigma_v^2)$ distribution and are assumed to represent the technical inefficiency in production.

The inefficiency model (2) can be in the form of a truncated normal distribution, half normal distribution or an exponential distribution (Hossain, 2013). However, in this study only the truncated normal or half-normal distributions were considered. (Battese and Corra, 1977) parameterized σv^2 and σ_u^2 by replacing them with:

$$\sigma^2 = \sigma_v^2 + \sigma_u^2 \quad (3)$$

$$\gamma = \sigma_u^2/\sigma^2 \quad (4)$$

Gamma (γ) is an unknown parameter that lies between zero and one. It explains the presence of the inefficiency component in the total error term (Coelli et al., 2005). The technical efficiency (TE) of the $i-th$ unit at the $t-th$ time period can be measured by:

$$TE_{it} = y_{it}/y_{it*} = exp(x_{it}\beta + v_{it} - u_{it})/exp(x_{it}\beta + v_{it}) = exp(-u_{it}) \quad (5)$$

where y_{it} is the observed output and $yit*$ is the corresponding stochastic frontier output.

The measurement of technical efficiency is the observed output of a unit relative to the output that potentially could be produced by a fully-efficient unit using the same amount of input (Coelli et al., 2005). The value can range between zero and one.

2.1.1 Application

Empirical Stochastic Frontier Model After the output and input variables involved were made clear, the functional form of translog production model (Battese and Coelli, 1992) was applied that can be defined as:

$$\begin{aligned} lnCPO_{it} &= \beta_0 + \beta_1 lnArea_{it} + \beta_2 lnMC_{it} \\ &+ \beta_3 lnLabour_{it} + \beta_4 t + 1/2[\beta_{11}(lnArea_{it})^2 + \\ &\beta_{22}(lnMC_{it})^2 + \beta_{33}(lnLabour_{it})^2 + \beta_{44}t^2 \\ &+ \beta_{12} lnArea_{it} * lnMC_{it} + \beta_{13} \\ & lnArea_{it} * lnLabour_{it} + \beta_{14} lnArea_{it} * t + \beta_{23} lnMC_{it} \\ &* lnLabour_{it} + \beta_{24} lnMC_{it} * t \\ &+ \beta_3 4 lnLabour_{it} * t + v_{it} - u_{it} \end{aligned} \quad (6)$$

where i = 1, 2, ..., 12 and t = 1, 2, ..., 18,

ln refers to the natural logarithm, CPO_{it} is the amount of crude palm oil production by the $i-th$ state at $t-th$ period,
$Area_{it}$ is the area under oil palm plantation in the $i-th$ state at $t-th$ period,
MC_{it} denotes the total fruit mill capacity available in the $i-th$ state at $t-th$ period,
$Labour_{it}$ is the number of plantation employee working in the $i-th$ state at the $t-th$ period, t is the study period from the value of 1 to 18 (year 1999 to 2016), β, v_{it} and uit are as defined in the previous section.

The most used functional forms are the Cobb-Douglas model and the transcendental logarithmic (trans-log) model. According to (Ferdushi, 2013), choosing the most appropriate model for our analysis is crucial as the functional form would significantly affect our results. Hence, to test whether the trans-log model above is the appropriate functional form for our model, the likelihood ratio test was conducted which would be explained in the next section. The time variable in the stochastic frontier model (6) was included to allow for Hicksian neutral technological change (Baten et al., 2009), while in the inefficiency model (2) the time variable is associated with the change in inefficiency as the time period increases (Coelli and Battese, 1996a). In model (6), the time-squared and the time interaction with each (log) input variable were considered to allow for non-monotonic technical change and non-neutral technical change respectively (See and Coelli, 2012). Hypothesis Test Several hypotheses would be tested to verify the validity of the results, to find the most appropriate functional form for the model and to select the distribution of the random variables assumed to represent the technical inefficiency (Ferdushi et al., 2011), (Mustapha, 2011). There are many different combinations and alternative models types to choose from. For the stochastic frontier model, the most common used are the Cobb-Douglas model or the trans-log model. For the inefficiency model, one can assume whether the inefficiencies follow a half-normal distribution or a truncated normal distribution. Since our data is a panel data, we also had to decide whether to assume time-varying or time invariant efficiencies. To solve this problem, a number of alternative models were estimated and then the likelihood ratio tests were carried out to select the most appropriate model (Coelli and Battese, 1996a).

We would be testing 4 hypotheses:

1. $H_0 : \gamma = 0$, testing the significance of the γ parameter is basically testing whether it is necessary to apply the stochastic frontier production function.

   From equation (4), we could see that if the null hypothesis is true, then the value of σ_u^2 would also be equal to zero meaning there is no technical inefficiency present. Thus, the uit term should be removed, turning the model into an ordinary linear regression model that could be solved using the ordinary least squares (OLS) method.

2. $H_0 : \beta_k l = 0 (k \leq l = 1, 2, 3, 4)$, the null hypothesis specifies that the coefficients of the squared input and the interaction between input variables of the stochastic frontier function are simultaneously zero. This means that the parameters $\beta_{11}, \beta_{22}, \beta_{33}, \beta_{44}, \beta_{12}, \beta_{13}, \beta_{14}, \beta_{23}, \beta_{24}, and \beta_{34}$ are restricted to the value of zero. If this is accepted, then the Cobb-Douglas functional form is more appropriate than the translog functional form.

3. $H_0 : \mu = 0$, this particular hypothesis is to test whether the distribution for the inefficiency is a half-normal distribution or a truncated normal distribution. The null hypothesis implies that the mean of the inefficiency distribution is equal to zero, making it a half-normal distribution which is a special case of the truncated normal distribution.

4. $H_0 : \eta = 0$, implies that the technical inefficiencies are time invariant.

As we can see from equation (2), if the null hypothesis $\eta = 0$ is accepted then it would mean that the technical inefficiencies are not affected by time.

All of these hypotheses were tested using the likelihood ratio test. The generalized likelihood ratio (LR) test statistic is defined by:

$$\begin{aligned} LR &= -2ln\,[L(H_0)/L(H_1)] \\ &= -2ln\,[L(H_0)] - ln\,[/L(H_1)] \end{aligned} \tag{7}$$

where $ln\,[L(H_0)]$ and $ln\,[L(H_1)]$ are the values of the log-likelihood function of the production frontier model under the null and the alternative hypotheses respectively. Under the null hypothesis, the LR statistic is assumed to be a Chi-square (or a mixed Chi-square) distribution with the degree of freedom equal to the number of restrictions involved (Coelli and Battese, 1996b). If the value of the LR test statistic exceeds the critical value, then the null hypothesis is rejected (Taymaz and Saatci, 1997).

3 RESULTS AND DISCUSSION

3.1 Maximum Likelihood Estimates of the Translog Stochastic Frontier Production Function

The maximum likelihood estimates for the parameters of the translog crude palm oil production model is shown in Table 1.

Table 1: Maximum likelihood estimates for the parameters of the translog production function.

Variable	Parameter	Coefficient	Standard Error	t-ratio
Constant	β_0	19.21489***	3.11286	6.17274
Area	β_1	-0.00817	1.23538	-0.00661
MC	β_2	-2.83437***	1.02815	-2.75677
Labour	β_3	1.49079**	0.59092	2.52283
t	β_3	0.18653***	0.04569	4.08262
...	...	...	...	...
Variance Parameter				
Sigma-Squared	σ^2	0.05309**	0.02437	2.17910
Gamma	γ	0.71642***	0.12953	5.53094
Eta	η	0.04956***	0.01445	3.42982

Looking at the maximum likelihood estimates of the coefficient of the first order variables, it is clear that all the variables except plantation area significantly affect the level of crude palm oil production. Fruit mill capacity and time both yield coefficient that are highly statistically significant at 1% level of significance. The coefficient of time is estimated to be 0.187 meaning that as time increases by a year, then crude palm oil production would increase by 0.187 tonnes if the effects of all other predictors are held constant. It also implies that technical progress increases on average of 18.7% per year. Meanwhile, the coefficient of fruit mill capacity is - 48 2.834. The negative sign of the coefficient could possibly indicate that the current existing mills are not fully utilized to their full capacity. This could also suggest that smaller size fruit mills are more productive compared to the larger fruit mills because they are easier to manage and monitor. Labour yield a significant coefficient at 1.491 implying that the labour variable influences crude palm oil output positively. The value of the coefficient for plantation area is approximated at -0.008.

However, this value is proven to be statistically insignificant implying that plantation area does not affect the output level significantly. All of the second order variables are found to be insignificant. The coefficients of the product variables between plantation area with fruit mill capacity, fruit mill capacity with time and labour with time appear to be significant at the 10% level of significance. The other interactions between input variables were found to be insignificant to production.

The parameter of error σ^2 is estimated to be 0.053 with significance level at 5%. Since σ^2 is statistically significantly different from zero, we can say that the model is a good fit to our data set. The parameters γ and η are found to be significant at 1% level of significance. γ is estimated at 0.716, implying that 71.6% of the variation in deviation is caused by technical inefficiency whereas 28.4% is caused by the stochastic random error. This result shows that technical inefficiency is important in explaining the total variability within the production of crude palm oil. The parameter η is approximated to be 0.05. The positive value of η suggests that the technical inefficiency tends to decline over time. Thus, the technical efficiency increases over time.

3.2 Estimated Technical Efficiency of Production

Table 2 displays the readings of the estimated technical efficiency for the production of crude palm oil of each state for each year generated. The overall mean technical efficiency in the production of crude palm oil for the states in Malaysia from the year 1999 to 2016 is 0.792. This means that 79.2% of the potential output is achieved by the palm oil industry in Malaysia. However, this also shows that there exists technical inefficiency of around 20.8% that can be improved using the same amount of existing resources. The lowest reading of technical efficiency is 0.4 by the state of Malacca during 1999. On the other hand, the highest reading is 0.986 by Selangor in 2016. None of the states got 100% level in efficiency at any given year.

Table 2: Estimated technical efficiency of producing crude palm oil for the states in Malaysia from 1999 to 2016 by stochastic frontier analysis.

State	1999	2000	2001	...	2016	Mean
Selangor	0.968	0.970	0.971	...	0.986	0.978
Sarawak	0.931	0.934	0.937	...	0.969	0.952
Perak	0.892	0.897	0.902	...	0.952	0.926
N. Sembilan	0.891	0.896	0.901	...	0.951	0.925
Penang	0.759	0.770	0.779	...	0.888	0.831
Terengganu	0.726	0.737	0.748	...	0.871	0.806
Kedah	0.712	0.724	0.735	...	0.864	0.795
Sabah	0.590	0.606	0.620	...	0.797	0.702
Johor	0.581	0.596	0.612	...	0.791	0.695
Kelantan	0.554	0.570	0.586	...	0.775	0.674
Pahang	0.554	0.570	0.586	...	0.775	0.673
Malacca	0.400	0.418	0.436	...	0.674	0.544

It was found that out of the 12 states, 7 states yielded mean technical efficiency above the overall average of 0.792. The most efficient state is the state of Selangor with a mean efficiency at 0.978. This implies that among all the states, Selangor is the most efficient in managing its resources to maximize production. It is clear that the least efficient state is the state of Malacca with mean efficiency reading of 0.544. The difference in score of the mean technical efficiency of Selangor and Malacca is a staggering 0.434. Meanwhile, the largest state in Malaysia, the state of Sarawak rank second with a yield mean efficiency score of 0.952. This is followed by Perak, Negeri Sembilan, Penang, Terengganu and Kedah with scores of 0.926, 0.925, 0.831, 0.806 and 0.795 respectively. The state of Sabah, which is the largest producer of crude palm oil between the states, ranked eighth following a mean efficiency score of 0.702. This indicates that Sabah can improve their output level by around 29.8% by fully utilizing their current available resources. After Sabah, the state of Johor, Kelantan and Pahang follow closely at 0.695, 0.674 and 0.673 respectively.

3.3 Selection of the Production Function and Hypotheses Testing

To determine the form of the production function, several hypothesis tests were carried out. The results are shown in Table 3:

According to (Coelli, 1995), if the null hypothesis involves $\gamma = 0$, then the asymptotic distribution requires a mixed Chi-square distribution. Thus, the critical value for the first null hypothesis is obtained from Table 1 of (Kodde and Palm, 1986). The null hypothesis is rejected since the value of the test statis-

Table 3: Generalized likelihood ratio test of hypothesis for the stochastic frontier production model.

Null Hypothesis	L-likelihood Function (H_0)	L-likelihood Function (H_1)	LR test Statistic	Critical Value	Decision
$H_0 : \gamma = 0$	87.1588	120.3780	66.4383	2.706*	Reject
$H_0 : \beta_{kl} = 0$	70.7475	125.4358	109.3765	18.307	Reject
$H_0 : \mu = 0$	125.3437	125.4358	0.1842	3.841	Accept
$H_0 : \eta = 0$	120.3780	125.3437	9.9314	3.841	Accept

tic exceeds the critical value. This result confirms that technical inefficiencies exist and are significant in explaining the performance in the production of crude palm oil by the states. The second null hypothesis $H0 : \beta kl = 0$ which specifies that the Cobb-Douglas production function is statistically more preferable than the translog production function is rejected. This indicates that the usage of translog production function is more appropriate for the data set. The third null hypothesis $H_0 : \mu = 0$ is accepted since the test statistic value did not exceed the critical value. We can conclude that the most suitable distribution for the inefficiency is the half-normal distribution. Finally, the null hypothesis $H_0 : \eta = 0$ implies that the technical inefficiencies are time invariant. This is rejected showing that time does significantly influence the technical inefficiencies in the production model. From the results of these hypothesis tests, we can conclude that the most preferable form of the production function for the data set is the translog stochastic frontier production function with the inefficiency assumed to follow a half-normal distribution and are time-variant.

4 CONCLUSION

This study set out to estimate the technical efficiency (TE) of producing crude palm oil (CPO) in Malaysia by applying the parametric stochastic frontier analysis (SFA) technique. The overall mean TE is 0.79. We found that fruit mill capacity, labour and time as input variables significantly affect the level of CPO output. Labour and time variables have positive relationship with the output level. On the other hand, fruit mill capacity was shown to have a negative relationship with the CPO production which could possibly indicate that the mills are not utilized to their full capacity. Plantation area was proven to be statistically insignificant in affecting output level. 71.6% of the variation in deviations were due to technical inefficiencies whereas 28.4% were cause by the stochastic random error. SFA estimated the state of Selangor to be the most efficient CPO producing state among our population and the state of Malacca to be the least efficient. Even though the average efficiency of the Malaysian CPO industry seems to be increasing gradually each year, there is still room for improvement. Inefficiencies could be reduced by managing existing resources better, utilization of idle capacity, operating at optimal scale and applying the ways of efficient states. The status of fruit mills in Malaysia needs to be looked at as it was discovered to have a negative relationship with output level. The existing mills possibly are not fully utilized. Future study should be done on the productivity of CPO production based on the size of fruit mills and whether smaller fruit mills are easier to manage and monitor. The productivity of the whole industry decreases each year due to technological change. Thus, investing in new technology is what needs to be done to encourage productivity growth in the industry. It is recommended that further study be done on identifying the factors influencing the TE of producing CPO in Malaysia preferably using the SFA (Coelli, 1995) model specification. The inclusion of environmental variables is highly suggested such as rainfall and temperature.

REFERENCES

Basiron, Y. (2007). Palm oil production through sustainable plantations. *European Journal of Lipid Science and Technology*, 109(4):289–295.

Baten, M., Kamil, A., and Mohammad, A. (2009). Modeling technical inefficiencies effects in a stochastic frontier production function for panel data. *African Journal of Agricultural Research*, 4(12):1374–1382.

Battese, G. and Coelli, T. (1992). Frontier production functions, technical efficiency and panel data: With application to paddy farmers in india. *Journal of Productivity Analysis*, 3(1-2):153–169.

Battese, G. and Corra, G. (1977). Estimation of a production frontier model: with application to the pastoral zone of eastern australia. *Australian Journal of Agricultural Economics*, 21(3):169–179.

Bogetoft, P., Otto, L., and Boles, J. (2011). Benchmarking with dea, sfa and r. In *Proceedings of the Annual Meeting (Western Farm Economics Association*, New York. Springer.

Coelli, T. (1995). Estimators and hypothesis tests for a stochastic frontier function: a monte carlo analysis. *Journal of Productivity Analysis*, 6(3):247–268.

Coelli, T. and Battese, G. (1996a). Identification of factors which influence the technical inefficiency of indian farmers. *Australian Journal of Agricultural and Resource Economics*, 40(2):103–128.

Coelli, T. and Battese, G. (1996b). Identification of factors which influence the technical inefficiency of indian farmers. *Australian Journal of Agricultural Economics*, 40(2):103–128.

Coelli, T., Rao, D., O'Donnell, C., and Battese, G. (2005). *An Introduction to Efficiency and Productivity Analysis*. Springer Science & Business Media, New York.

Farrell, M. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society. Series A (General*, 120(3):253–290.

Ferdushi, K. (2013). Stochastic metafrontier production model, flexible risk efficiency and its properties: with application to rice croppings systems in bangladesh.

Ferdushi, K., Baten, M., Kamil, A., and Mustafa, A. (2011). Wage augmented stochastic frontier model with truncated normal distribution. *International Journal of Physical Sciences*, 6(14):3288–3295.

Hassan, M., Kamil, A., Mustafa, A., and Baten, M. (2012). Estimating stock market technical efficiency for truncated normal distribution: Evidence from dhaka stock exchange. *Trends in Applied Sciences Research*, 7(7):532–540.

Hossain, M. (2013). Improved data envelopment analysis efficiency with statistical distributions: A role of environment impact.

Kodde, D. and Palm, F. (1986). Wald criteria for jointly testing equality and inequality restrictions. *Econometrica: Journal of the Econometric Society*, 54(5):1243–1248.

M.P.C. (2014). *Reducing unnecessary regulatory burdens on business: Growing oil palm*. Malaysia Productivity Corporation, Selangor, Malaysia.

Mustapha, N. (2011). Technical efficiency for rubber smallholders under risda's supervisory system using stochastic frontier analysis. *Journal of Sustainability Science and Management*, 6(1):156–168.

See, K. and Coelli, T. (2012). An analysis of factors that influence the technical efficiency of malaysian thermal power plants. *Energy Economics*, 34(3):677– 685.

Taymaz, E. and Saatci, G. (1997). Technical change and efficiency in turkish manufacturing industries. *Journal of Productivity Analysis*, 8(4):461–475.