Australian Journal of Agricultural Economics, Vol. 39, No. 2 (August 1995), pp 143-163

ANALYSIS OF THE TECHNICAL EFFICIENCY OVER TIME OF WEST JAVANESE RICE FARMS*

R. TREWIN*, L. WEIGUO*, ERWIDODO**
and S. BAHRI**
**Australian National University, Canberra, Australia,*
***Centre for Agro-Socioeconomic Research, Bogor, Indonesia*

The high costs and effectiveness of Indonesia's current mix of policies aimed at maintaining rice self-sufficiency, for example fertiliser and other subsidies, has been questioned. Consequently, attention has turned to developing policies to increase farmers yields through efficiencies, for example as a result of better extension. The main issue investigated in this article is whether existing yields can be significantly improved through increasing the efficiency of individual farmers. Production frontiers are estimated and technical efficiencies/inefficiencies predicted from panel data and for one year at a time. Different results are explained in relation to a number of aspects, especially the rate of adoption of new technologies, and policy actions are recommended. The robustness of the analysis is examined as conclusions obtained from past analyses have often been inconsistent.

Introduction

A key issue for Indonesian policy makers is how to maintain rice self-sufficiency, which was first achieved in 1984. The achievement of self-sufficiency was due to a mixture of policies promoting extensification (increases in the area harvested) and intensification (increases in yields). Extensification took place mainly as a result of increases in the area of land under irrigation. Intensification was the result of the introduction of high yielding seed varieties and the application of highly subsidised associated inputs such as water and fertilisers. These have been high cost policies and attention has turned towards developing more efficient ways of achieving self-sufficiency; for example through better use of existing irrigation infrastructure.

The focus in the analysis is on whether there are significant yield differences in Indonesian rice production due to differences in the effi-

Acknowledgements: Funding for the project in which this work was undertaken was provided by the Australian Centre for International Agricultural Research. The authors gratefully acknowledge assistance with concepts and analysis, without implications for the final content of the paper, given by Phil Kokic (Australian Bureau of Agricultural and Resource Economics), K. P. Kalirajan (ANU) and Ric Shand (ANU). Highly valuable comments were also provided by Tim Coelli and George Battese.

ciencies of farmers. Production frontiers are estimated and technical efficiencies/inefficiencies predicted from panel data, and for one year at a time. Differences in the analysis are explained in relation to a number of aspects, especially the rate of adoption of new technologies, and policy actions are recommended to improve farmers efficiencies. Efficiency of farmers over time will be specifically considered as a factor in any yield differences.

Past analysis of these issues has often resulted in inconsistent conclusions. The robustness of the conclusions obtained with regard to the choice of models, methods of analysis and form of applied data is also analysed.

In the next section, the models that can be used to determine possible explanations for any yield differences are discussed, along with alternative methods and forms of data for analysing the models. Models and methods are then applied to various sets of data from the Cimanuk Basin, Indonesia. A summary and conclusions, including the policy implications of the analysis, are discussed in the final section.

The Econometric Model and Analytic Methods

Frontiers and Envelopes

The production function is the technical relationship that transforms inputs into outputs and is the traditional way of representing farm operations. The choice of specific functional form is often important. For example, choices of functional form include the Cobb-Douglas (linear in logs) and various flexible functional forms such as the translog (see Kopp and Smith 1980 for discussion of the various forms). The choice of specific functional form is mainly an empirical issue although economic theory does impose some constraints.

In analysis of farmer efficiency/inefficiency it is not the average of observed relationships between farmers inputs and outputs that is of interest but the maximum possible output that is obtainable from a given combination of inputs — the frontier production function. A related concept, but one that will not always correspond to the frontier production function, involves the envelope encompassing all the input-output combinations of interest.

Similar concepts of a best-practices frontier (maximum output obtained with respect to the sample) and an absolute frontier (maximum output obtained with respect to all conceivable observations embodying the current technology) were introduced by Forsund *et al.* (1980). These frontiers are distinguished by Forsund *et al.* (1980) as being, respectively 'non-statistical' (no one-sided error distribution and typically 100%-efficient observation(s) on the envelope) or 'statistical', but Forsund *et al.* state that these frontiers would be expected to converge as the sample size grows.

However, there can be a time aspect to such concepts as well. Farms will never immediately adopt new technology as important activities such as water control and fertiliser application will have to be learnt to be undertaken efficiently on each farm (Squires and Tabor 1991). Other things being equal, yield differences will be apparent when comparing farm outcomes over time as a new technology is adopted. Moreover, a frontier estimated from a sample or even the population of farms in the early years of a new technology will invariably lie below that representing the situation when the technology has been completely adopted. However, it will approach this last frontier over time as the technology is completely adopted, given no new technology is introduced. Rather than a 'statistical' distinction being made between a best-practices and absolute frontier, here a distinction is being made in relation to all conceivable observations over the time during which full adoption takes place. An appropriate term for this concept might be the potential absolute frontier.

Efficiencies / Inefficiencies

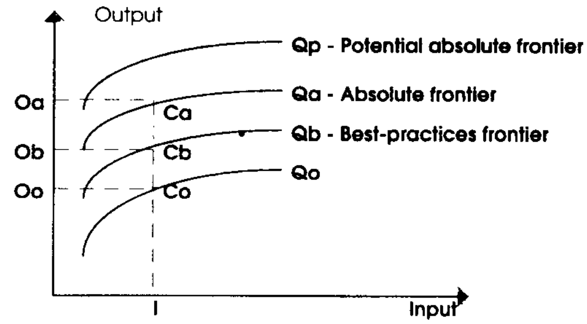
The distance a farm lies below its frontier measures the degree of technical inefficiency, thus it is a residual measure. The existence of technical inefficiency has been questioned. For example, Mueller (1974, p. 731) states '. . . little is known about the role of non-physical inputs, especially information and knowledge, which influence the firm's ability to use its available technology set fully. . . . Once all inputs are taken into account, measured productivity differences should disappear except for random disturbances.' This seems to be a question of what constitutes an appropriate input. In terms of the policy analysis, it is somewhat irrelevant whether extension advice, for example, improves the level of a 'non-core' input such as information and hence output under Mueller's view, or addresses inefficiencies due to a lack of information under a frontier function approach. There is also the question of whether all inputs can be taken into account, or measured, and thus avoid the need for a residual measure approach. The concept is represented diagrammatically in Figure 1, along with the other concepts already mentioned.

A production process is technically inefficient if maximum output is not produced from a given bundle of inputs. The complementary concept to technical inefficiency, that of technical efficiency, is measured by the ratio of the expected output to the maximum output, for example O_0/O_b in Figure 1; that is a comparison of output at points C_b and C_0, each with the same level of inputs but C_b lying on the best-practices frontier function Q_b (passing through a 100%-efficient sample point) whilst C_0 lies on Q_0 which represents a locus that is a neutral shift of the frontier Q_b and passes through the point C_0. The concept could be measured relative to other frontiers, for example the absolute frontier function Q_a lying above all sample points. Here, the ratio will be O_0/O_a or a comparison of output at points C_a on Q_a and C_0.

The potential absolute frontier is also represented in Figure 1. The potential absolute frontier, the maximum output obtained from all conceivable observations embodying the current technology (including over all time periods in which adoption takes place), is represented by Q_p which lies above Q_a. Over time, there would be a sequence of absolute frontier functions Q_a's (and associated levels of technical efficiency) moving up to the potential absolute frontier function Q_p. The concepts are analysed with data from the Cimanuk Basin in the next section.

FIGURE 1

Best-practices and (potential) Absolute Frontiers, and Measures of Inefficiency



Forms of frontiers and methodological approaches

One form of frontier function is the stochastic production frontier, developed independently by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977). Functions, being simplified representations of actual operations, require assumptions regarding the distribution of random errors before they can be used in empirical estimation. The usual assumption made in relation to average functions is a normal distribution which is symmetric and bell-shaped. The stochastic frontier (corresponding to the absolute frontier concept) incorporates two random components; a traditional random error component V_i, and a component U_i representing the degree of technical inefficiency. Various distributional assumptions have been made regarding this additional component, including the half-normal and truncated normal. In the deterministic frontier (often corresponding to the best-practices frontier concept), any variation in firm performance is relative to a single frontier and attributed purely to inefficiency. This ignores the possibility of variation due to specific factors not under a firm's control, such as the socioeconomic and physical environment, which are usually incorporated as random error.

The specific stochastic frontier models used in this paper was developed by Battese and Coelli (1992). This model can accommodate unbalanced panel data associated with a sample of N firms over T time periods and incorporates a simple exponential specification of time varying firm effects. The model is defined by

$$Y_{it} = f(X_{it};\, \beta)\exp(V_{it} - U_{it}) \qquad i = 1,\dots, N, \qquad t = 1, \dots, T,$$

and $$U_{it} = \eta_{it} U_i = \{\exp[-\eta\,(t - T)]\} U_i \qquad t\,\varepsilon\, I(i);\ i = 1, 2, \dots, N\,;$$

where Y_{it} represents the production for the ith firm in the tth time period; $f(X_{it};\, \beta)$ is a suitable function of a vector, X_{it}, of factor inputs (and firm specific variables), associated with the production of the ith firm in the tth time period; and a vector, β, of unknown parameters; the V_{it}'s are assumed to be independent and identically distributed $N(0,\sigma_v^2)$ random errors; the U_i's are assumed to be independent and identically distributed non-negative truncations of the $N(\mu, \sigma_u^2)$ distribution; z is an unknown scalar parameter; and $I(i)$ represents the set of T_i time periods among the T periods involved for which observations for the ith firm are obtained.

The parameterisation in which σ_v^2 and σ_u^2 are replaced with

$$\sigma_s^2 \equiv \sigma_v^2 + \sigma_u^2 \text{ and } \gamma \equiv \sigma_u^2/\sigma_s^2$$

is used as γ lies between 0 and 1 which provides a range that can be searched for a good starting point for an iterative maximisation of the likelihood.

The associated computer program, FRONTIER 2, was used to obtain maximum-likelihood estimates of the model parameters and predictors of the efficiencies of individual firms. The preferred model can be selected using generalised Likelihood Ratio Tests. Whether any form of stochastic production function is required can be tested via the significance of γ. Whether the model is time invariant can be tested via the significance of η. Whether the mean inefficiency effects are zero or not can be tested via the significance of μ.

However, the simple model has some limitations. For example, when used in conjunction with the full panel of data, the ordering of farms according to efficiency levels does not change from that obtained for the first year whereas estimating one year at a time results in considerable changes in orders. That is, $\eta > 0, = 0, < 0$ mean non-negative firm effects either decrease, remain constant, or increase as t increases.

Another frontier approach developed by Kokic *et al.* (1992) applies the robust regression technique of M-quantile regression (Breckling and Chambers 1988) to the function representing farm operations. Basically the technique is a generalisation of M-regression (Huber 1981) and weighs positive residuals by a factor 2p and negative residuals by a factor 2(1-p). For p close to 1, the M-quantile production function represents the average performance of efficient farms, and for p close to 0 the performance of inefficient farms. A measure of the jth farm's performance, p_j, with the desirable property of not being dependent on the level of inputs can be determined by the technique. Because this approach makes different assumptions to the stochastic frontier approach (for example, in relation to error distributions), similar conclusions will suggest robustness to these assumptions. Different conclusions should point to what may be the critical assumptions that require greater information and careful choice, for example whether or not an observation is truly an outlier.

Data Aspects

Frontier functions have been applied to both cross-sectional and to panel data such as that detailed later in the paper. The use of panel data has a number of advantages, for example, fewer distributional and independence assumptions are necessary. However, the use of panel data also introduces a number of complications such as the possibility of technological change.

Best-practices and absolute frontiers for a particular year could be estimated from the panel data, one year at a time. The potential absolute frontier could be estimated from the panel data as a whole so long as there was some modelling of the adoption rate of the new technology. If the best-practices and absolute frontiers vary randomly from year to year then this time variation could be used in the form of the pooled panel data to estimate a corresponding potential absolute frontier. In this case the potential absolute frontier would be conceptually like the meta-frontier (Pitt 1983). However, it would encompass individual time specific frontiers instead of technologically specific frontiers.

Application to Panel Data from the Cimanuk Basin

Cimanuk Basin Data

The data set used in this study was obtained from the Centre of Agro-Socioeconomic Research (CASER) and was collected from the rice production area of the Cimanuk River Basin, West Java, Indonesia. The rice production area of the Cimanuk River Basin is characterised by mainly irrigated rice farms set in an almost uniform climate. Six desa (villages) located in five kabupaten (the administrative unit between district and province level) were covered in the survey. These villages are listed amongst the dummy variables defined in the appendix listing all the variables used in the analysis.

The survey was conducted twice in 1977, collecting information in the wet season of 1975/76 and the dry season of 1976, and in addition in the wet season of 1976/77. In 1978 a similar survey was undertaken for the dry season of 1977. A follow up survey of the same farms/farmers for the 1981/82 and 1982 seasons was conducted in 1983. Altogether a panel of 171 respondents spread reasonably evenly over the six villages were continuously surveyed over six seasons.

One difficulty with panel data is that, being obtained from farms that have remained in the population over a period of time, they may not be representative of the population of farms at a particular point in time even though the initial sample was randomly selected. There may have been significant movements in the sample and population over time, resulting in the sample reflecting different characteristics to the population. To investigate this issue, panel sample estimates of various farm characteristics derived from the Cimanuk Basin data were compared with those obtained from other more general samples. Pingali *et al.* (1990) include farmer field data for 1980 and 1988 from a sample of 71 households in West Java, Indonesia; the same province as the Cimanuk Basin. Farm cost structure information for West Java is also available for 1982 from the Central Bureau of Statistics (CBS). These sets of data are compared in Table 1.

TABLE 1
Comparison of Farm Level Estimates

	Cimanuk Basin data					*CBS*	*Pingali et al.*[a]
	1976/77	1977	1981/82	1982	Average	1982	1980
Yield (kg/ha)	2,513	2,350	4,197	3,969	3,207	4,134	4,897 (20.7)
Seed (kg/ha)	40.6	37.2	42.2	37.8	39.4	35.4	
Nitrate (kg/ha)	220.0	192.7	268.1	250.0	232.7	193.3	235.8 (27.6)
Phosphate (kg/ha)	63.0	55.9	119.7	110.8	87.4	82.7	
Labour (8 hr days)	103.2	108.1	122.3	113.4	111.7		104.6 (29.5)

[a] Coefficients of variation in brackets

Compared with CBS data, CASER Cimanuk Basin panel data for 1982 suggest a smaller yield, and higher seed and fertiliser use, although these differences are within the realm of sampling errors. Compared with Pingali *et al.* (1990) data, CASER panel data show on average a smaller yield, roughly equal fertiliser and labour use in the Cimanuk Basin, although the Pingali *et al.* yields look high when compared to CBS data. Thus it would appear that the panel data is acceptably representative of the population of farms at particular points in time.

Stochastic Frontier Model Estimates for the Cimanuk Basin Panel Data

The earlier defined stochastic frontier approach, which has been applied previously to Cimanuk Basin panel data (Erwidodo 1990), was primarily used in the analysis. Total output per farm was the dependent variable and total quantity of seed, fertiliser, labour, farm size and a number of dummies (pesticide use, seed varieties, season, village) the independent variables. Initially, a hybrid form of the Cobb-Douglas stochastic production function was estimated. This form, additive in logs apart from the potassium fertiliser variable which was not logged, was used to overcome the difficulty caused by many farmers not using any potassium fertilisers.[1]

A general form of the hybrid model was estimated using FRONTIER 2 in conjunction with the panel data, and then various restricted forms

[1] Zeros in a Cobb-Douglas production function can be handled in a number of other ways, for example by adding the individual fertilisers (weighted or unweighted), or by converting the zeros to a small positive value or to unity. The appropriateness of the approach will depend on the need for separate fertiliser estimates and the structure of the untransformed data. Various approaches were applied and examined in the analysis.

were tested. Following the testing, the preferred model was one in which farm technical efficiency was time-invariant and the stochastic distribution had mean zero, both aspects assumed by Erwidodo (1990). Parameter estimates (see Table 2) were similar to those of Erwidodo (1990) apart from the hybrid parameter associated with potassium fertiliser.

TABLE 2
Maximum Likelihood Estimates of Stochastic Cobb-Douglas Production Function (Panel Data)

	Coefficients	t-ratio
Constant	4.97	26.35
ln KGS	0.16	5.71
ln KGN	0.13	7.00
ln KGP	0.07	6.06
ln LAB	0.23	7.93
ln LAND	0.43	13.78
D1	0.01	0.48
D2	0.14	2.62
D3	0.17	4.43
D4	0.04	2.03
D5	0.03	0.84
D6	–0.03	–0.67
D7	–0.03	–0.43
D8	–0.06	–1.04
D9	0.03	0.45
D10	0.09	1.45
$\sigma_s^2 \equiv \sigma_v^2 + \sigma_u^2$	0.13	16.08
$\gamma \equiv \sigma_u^2/\sigma_s^2$	0.13	2.36
Log-likelihood function = –367.60		
Chi-square test of one sided error (σ_u^2) = 6.00 with one degree of freedom		

Note: The dependent variable is GRKG in log form. V is iid normally distributed random errors, mean 0 and variance σ_v^2, and U is half-normally distributed random errors, mean 0 and variance σ_u^2.

The estimates of technical inefficiency were also similar to Erwidodo (1990) with a slightly higher, though still relatively small, mean value of 9.6 per cent (see Table 3).[2]

[2] A similar outcome was observed when zero observations for potassium fertiliser were replaced by unity, effectively resulting in zero entries in a log format, or when the individual fertilisers were summed.

TABLE 3
Frequency Distribution of Farmers Based on Level of Technical Inefficiency from Cobb-Douglas Production Frontier.

Technical Inefficiency (TI)	Number of Farms	Frequency Distribution (%)
≤ 5%	5	2.92
5% ≤ TI ≤ 10%	104	60.82
10% ≤ TI ≤ 15%	50	29.24
15% over	12	7.02
Mean	9.58	
Minimum	3.50	
Maximum	21.98	
Total number of farms	171	100.00

Robustness of the Estimates to Changes in Specification, Method and Data

A stochastic translog production frontier was estimated using FRONTIER 2 and the panel data as a test of the robustness to the choice of functional form. Fertiliser variables were aggregated in this model, the weights being determined from a regression of the fertiliser variables on yield. The form of this model encompasses the Cobb-Douglas form so a test of preference for one form over the other can be undertaken by analysing the significance of the cross terms in the translog form. The Likelihood Ratio Test of these terms suggested that the translog form is preferred.[3]

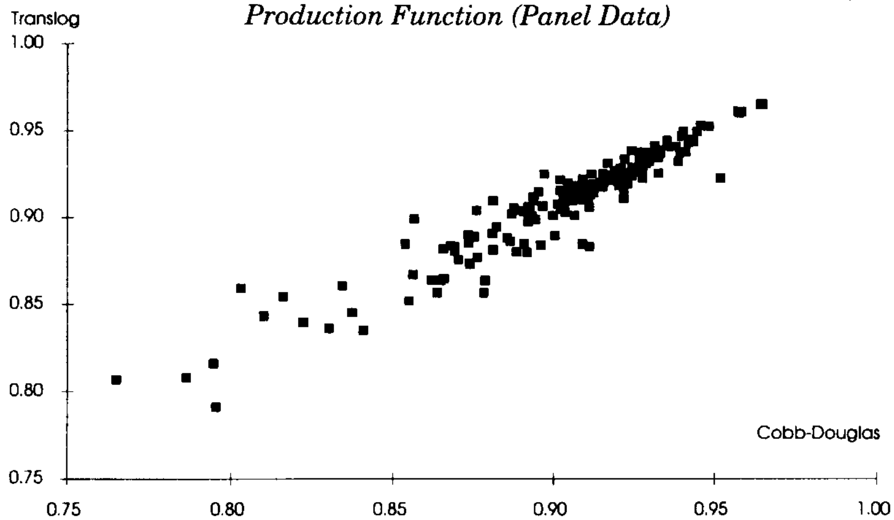
The preferred model was analysed to see if there were any marked differences in the key inefficiency estimates obtained previously. Again the preferred form was one in which farm technical efficiency was time-invariant and the stochastic distribution had mean zero. Multicollinearity was present in the model and the signs of the translog coefficients are difficult to interpret individually. However, the key focus of the analysis is on technical inefficiency. The estimates of technical inefficiency were slightly lower than the previous estimates, the mean value being 9.1 per cent.

The choice between these functional forms would not appear to make a significant difference to the key estimates. In fact, a mapping of individual farm technical inefficiency measures from the Cobb-Douglas model against those from the translog model closely follows a straight line 45 degrees from the origin (see Figure 2).[4]

[3] The Log Likelihood for the translog model was -340.56 compared to -367.60 for the Cobb-Douglas model giving a Likelihood Ratio Test value of 54.08 which is significant at the 1 per cent level.

[4] Squires and Tabor (1991), using a translog production function, found similar orders of technical inefficiency for rice production in West Java in 1983, as did Siregar (1987) using a Cobb-Douglas production function. In contrast, Esparon and Sturgess (1989), using a linear production function, found no technical inefficiencies for rice production in West Java in 1983.

FIGURE 2
Frontier Performance Measure — Cobb-Douglas versus Translog Production Function (Panel Data)



The M-quantile approach with a Cobb-Douglas functional form was then applied to the Cimanuk Basin data as a test of the robustness of the estimates to the chosen method.[5] Panel data were pooled in the analysis. This would seem appropriate given earlier evidence of time invariance. Sensitivity analysis was undertaken for the key parameters in the M-quantile approach to ensure its robustness to these parameters. Individual estimates of farm performance were compared with each other and with inefficiency measures obtained from the stochastic frontier approach. The M-quantile approach will always rank farms in terms of their performance even if the best and worse farm are not significantly different. Bootstrapping methods would be required to determine any significant differences. Key results of the M-quantile approach were:

- marked but random variation in individual farm performance measures over time
- marked differences in individual farm performance measures for the M-quantile approach and in inefficiency measures for the stochastic frontier approach. (In fact, a number of farms had upper extremes for one measure and lower extremes for the other, suggesting that the treatment of outliers could be critical to the measure obtained.)

The first result from M-quantile analysis that individual farm performance measures differed markedly from year to year, something unable to be ascertained from the way the model specified in FRONTIER 2 was applied,[6] suggests reconsideration of the earlier analysis in terms of the

[5] The M-quantile frontiers were estimated from a SAS program developed at ABARE by Phillip Kokic.

[6] FRONTIER 2 can be used to identify year to year differences if the user specifies that all the pooled observations are from the same time period (even though they are not), and then group the resulting technical inefficiency estimates into the different years.

potential absolute frontier concept. This is undertaken in the next subsection. The remainder of this subsection is devoted to some data diagnostic analysis on models for individual years, as suggested by the second result from the M-quantile analysis.

In regression analysis, some points can have greater influence than others. Regression data diagnostic analysis (see Besley *et al.* 1980) has been developed to ascertain which data points are influential in determining the estimated coefficients and so on by observing the response of these estimates to changes in the data.[7] This does not mean that these data points should be omitted. These points may be the only ones containing certain information which needs to be identified and judiciously used in the model development. Regression data diagnostics currently do not exist for stochastic frontier models. Because the majority of the annual models were average production functions estimated by ordinary least squares, the diagnostics applicable to such models were estimated as preliminary analysis of this issue.[8] There were 81 points with large DFBETAS and DFFITS. However, these points were not consistent over the years with 52 farms having only one outlying point across the six seasons and only one farm having more than three outlying points across the six seasons. Farms that had high levels of efficiency were usually associated with upper tail outliers in the DFFITS (a residual measure similar to efficiency measures), and vice versa. This situation appeared to apply uniformly when both M-quantile and frontier measures were available (see Tables 4a and b).[9]

The correspondence between the two measures and the appropriate data diagnostic was generally good. The points where the correspondence was not good between the two measures were not influential points. Whether the point is a true outlier and should therefore be excluded from the analysis is important as both approaches can treat these points as influential. Detailed analysis of the characteristics of the individual data points (see Seaver and Triantis 1989) would be required before a decision on the outlier status of an individual data point could be made.

Individual farm performance measures for the M-quantile approach were plotted against individual farm efficiency measures for the stochastic frontier approach for those years in which a frontier could be estimated

[7] Two basic component diagnostics are the diagonals of the least squares projection matrix (the 'hats') and the studentised residuals. The 'hat' matrix identifies points of high leverage that may be influential depending on the y values. Two diagnostic measures, DFBETAS and DFFITS, are respectively the scaled change in estimated coefficients and fit due to deleting an observation. Both of these are affected by the basic components, however, it is invariably necessary to consider a suite of diagnostic measures to obtain a full picture.

[8] Basic diagnostics for stochastic frontier models were derived by noting the effect on parameter estimates and forecasts of dropping each data point. These diagnostics were consistent with the preliminary analysis.

[9] Stochastic frontiers were not estimated for the middle four 'years' as the ordinary least squares models were preferred (γ accepted as equal to zero in testing).

TABLE 4a
Performance Measures and Outliers (1975/76 Wet Season)
(U=Upper tail, L=Lower tail)

ID	Efficiency Measures		HAT	DFFITS	DFBETAS				
	M-quantile	Frontier	Matrix		Intercept t	KGS	FERT	LAB	LAND
5	**1.00**	**0.92**		U	U	**L**	**L**		**U**
6	1.00	0.93	outlier	U	U	L	U		
16	0.02	0.70		L	L		U	U	L
28	0.00	0.74		L		U	L		
35	0.00	0.79	outlier	L	L				
37	1.00	0.91	outlier	U			L		
39	0.93	0.91				L			U
47	1.00	0.92			L		U		
95	1.00	0.93			L	U			L
101	1.00	0.92			U		L		U
106	0.03	0.77			L	U			L
114	0.00	0.74			L			U	L
118	1.00	0.93		U		U	U		
119	0.92	0.92	outlier	U					
141	0.08	0.74			L	U			L
142	0.11	0.77			L			U	
143	0.00	0.60		L		L	U		
144	0.07	0.73			U			L	U
145	0.02	0.72			U	L		L	U
151	0.95	0.91			U			L	U
162	0.57	0.87	outlier						
164	1.00	0.94		U			U	U	

TABLE 4b

Performance Measures and Outliers (1983 Dry Season)

(U=Upper tail, L=Lower tail)

ID	Efficiency Measures		HAT	DFFITS	DFBETAS				
	M-quantile	Frontier	Matrix		Intercept t	KGS	FERT	LAB	LAND
7	0.91	0.91			L		U	U	L
18	0.89	0.63		L	L				L
24	0.89	0.91			U		L		U
38	0.00	0.77			L	U			L
51	0.93	0.93	outlier	U	U			L	U
61	0.00	0.75			L			U	L
73	0.86	0.90			U				U
86	0.85	0.64		L			L	U	
110	0.00	0.63		L	U	L	L	U	U
113	0.44	0.91		U					U
125	0.31	0.92		U					
139	0.39	0.71		L	L			U	L
152	0.25	0.93		U		U	L	L	L
162	0.82	0.69		L	U	U	L		
163	0.75	0.94		U	L	U	U		L
167	0.47	0.93		U		L	U		
170	0.33	0.65		L	L	U			L
171	0.00	0.91				L		U	

Note: The SAS program was used to estimate the various diagnostic statistics for outliers. For a detailed definition of HAT matrix, DFFITS and DFBETAS, see Besley et al. (1980).

(see Figures 3a and 3b). As with similar analysis of pooled data, there were some marked differences in individual farm performance measures for the M-quantile approach and in inefficiency measures for the stochastic frontier approach, with a number of farms lying in the upper left corner of the figure, especially in the 1983 dry season. However, as already

FIGURE 3a
M-Quantile versus Frontier Performance Measure (Cobb-Douglas)
(1974 / 1976 West Season)

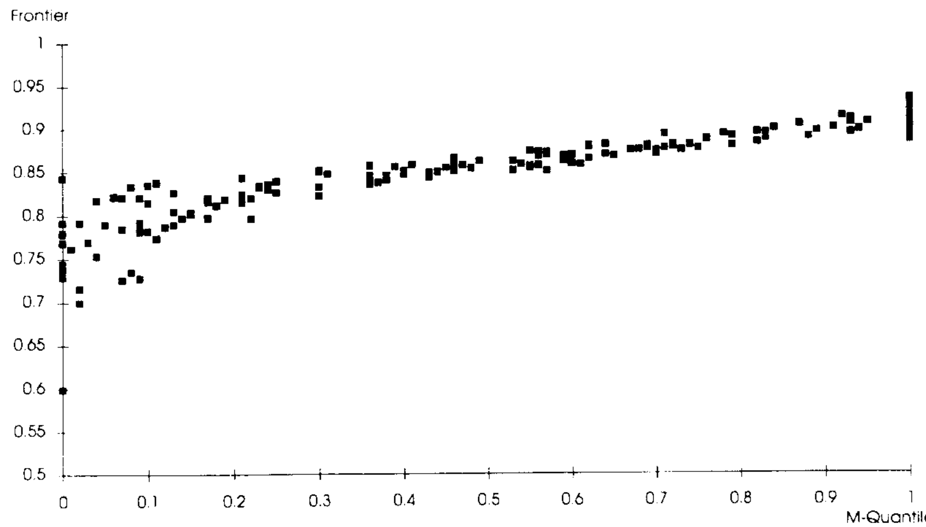
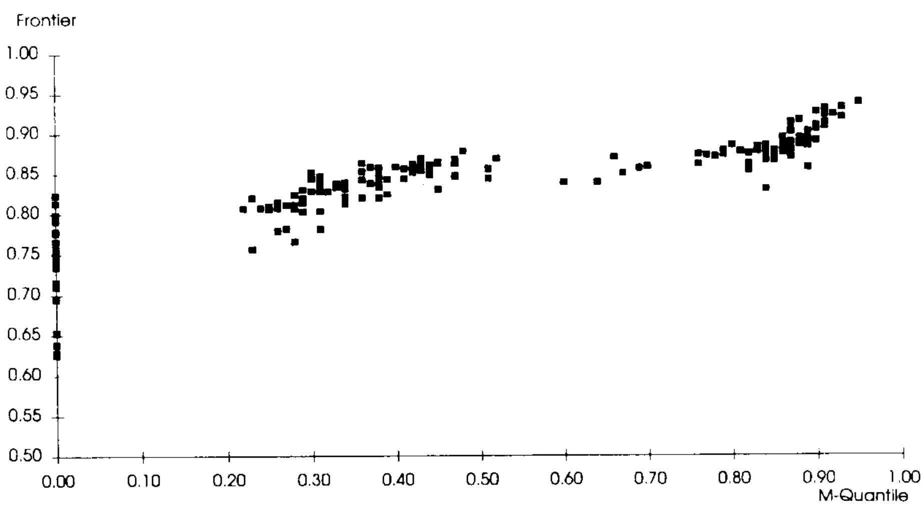


FIGURE 3b
M-Quantile versus Frontier Performance Measure (Cobb-Douglas)
(1983 Dry Season)



noted, these points were not influential points. Generally the situation for both the 1975/76 wet season and the 1983 dry season is that there is a good overall correspondence between the two sets of measures.

Frontier Model Estimates for the Cimanuk Basin Data by Individual Year

Initially, the M-quantile analysis was undertaken on an individual year basis for the Cobb-Douglas production function and estimates of the 95 percent M-quantiles produced. There were significant variations in these estimates over time,[10] confirming the need for individual year estimates.

Stochastic frontier estimates for individual years were undertaken for a Cobb-Douglas stochastic production function with half-normal errors. The Cobb-Douglas form was preferred over the translog on the basis of Likelihood Ratio Tests of additional cross terms in the translog form when compared on an individual year basis. This suggests the translog specification when applied to the panel data may have been preferred because its additional terms were picking up time variation, albeit at the cost of multicollinearity.

Farmers as a group appear to have been fully efficient during the dry seasons of 1976 and 1977 and the wet seasons of 1976-77 and 1981-82 (see Table 5). There was a consistent trend in the estimated coefficients for the fertiliser variable to increase year by year.[11]

Estimates of individual farm inefficiencies, where available, had relatively small mean values of around 15 to 16 per cent (see Table 6a and Table 6b).[12]

Analysing the data as a panel suggests small but significant degrees of technical inefficiency with no significant uniform trend over time. On analysing the panel data for a year at a time, it seems that farmers are in the main very efficient and that the frontier itself may vary each year (see for example Battese and Coelli 1992 on this last aspect), for example in relation to the efficiency of fertiliser application. These outcomes appear inconsistent. They would also seem to be inconsistent with prior information that no technological change (as distinct from the adoption of technology) has taken place over the period.

The potential absolute frontier concept enables a consistent interpretation of the analytic evidence. It suggests that in most years farmers are a homogenous group in terms of efficiency and that over time they move as a group, although not always in a smooth fashion, towards a higher stable frontier as they adopt the new technology as a group. It is also suggested by the analysis that the simple exponential specification of the

[10] For example, the Kruskal-Wallis test, which is a non-parametric version of a one-way analysis of variance, was highly significant.

[11] The imposition of constant returns to scale as suggested by the estimates made little difference.

[12] The 1983 results were consistent with results obtained by Squires and Tabor (1991) in the same year and region.

TABLE 5

Maximum Likelihood Estimates for Parameters for Stochastic Cobb-Douglas Production Function (Half-normal)

	1975/76 Wet	1976 Dry (OLS)	1976/77 Wet (OLS)	1977 Dry (OLS)	1982/83 Wet (OLS)	1983 Dry
ln KGS	0.07	0.13	0.12	0.17	0.18	0.10
	(1.20)	(2.37)	(2.17)	(2.72)	(2.83)	(1.53)
ln FERT	0.08	0.14	0.17	0.17	0.26	0.28
	(3.08)	(4.03)	(6.11)	(5.39)	(4.60)	(5.57)
ln LAB	0.19	0.29	0.35	0.20	0.03	0.18
	(2.90)	(3.90)	(6.22)	2.96	(0.42)	(2.80)
ln LAND	0.68	0.40	0.34	0.44	0.59	0.55
	(8.00)	(6.06)	(6.29)	(6.13)	(8.04)	(7.29)
Constant	6.20	4.44	4.46	3.41	6.24	6.17
	(12.43)	(5.70)	(7.16)	(5.78)	(10.61)	(11.88)
R2 based on OLS	0.92	0.87	0.91	0.88	0.94	0.94
Bruesch-Pagan-Godfrey test based on OLS						
χ^2 test with 13 d.f.	10.65	27.04	31.39	25.53	18.01	23.14
Frontier 2 diagnostics						
$\sigma_s^2 \equiv \sigma_v^2 + \sigma_u^2$	0.10 (2.91)					0.11 (3.08)
$\gamma \equiv \sigma_u^2/\sigma_s^2$	0.49 (1.44)					0.50 (1.56)
Log-likelihood function	-9.72					-17.73
Chi-square test of one sided error with 1 d.f.	0.48					0.47

Notes: 1 The dependent variable is GRKG in log form. Dummy coefficients are not reported. Figures in parentheses are t values.
2 The critical value for $\chi^2(13)_{0.05}$ is 22.4 and for $\chi^2(1)_{0.05}$ is 3.84.

way a firm's production varies over time may not be flexible enough to appropriately represent the actual situation taking place in the application.

TABLE 6a

Frequency distribution of farmers based on the level of technical inefficiency from Cobb-Douglas production frontier (1975/76 wet season)

Technical inefficiency (TI)	Number of Farms	Frequency Distribution (%)
≤ 5%	0	0
5% < TI ≤ 10%	27	15.79
10% < TI ≤ 15%	72	42.11
15% ≤ TI ≤20%	41	23.98
20% < TI ≤ 25%	22	12.87
25% over	9	5.25
Mean	15.12	
Minimum	6.32	
Maximum	40.05	
Total Number of Farms	171	100.00

TABLE 6b

Frequency distribution of farmers based on the level of technical inefficiency from Cobb-Douglas production frontier (1983 dry season)

Technical inefficiency (TI)	Number of Farms	Frequency Distribution (%)
≤ 5%	0	0
5% < TI ≤ 10%	18	10.53
10% < TI ≤ 15%	73	42.69
15% ≤ TI ≤20%	54	31.58
20% < TI ≤ 25%	18	10.53
25% over	8	4.67
Mean	15.88	
Minimum	6.07	
Maximum	37.42	
Total Number of Farms	171	100.00

Summary and Conclusion

The main purpose of this paper was to ascertain whether there were significant yield differences in Indonesian rice production due to differences in the efficiency of farmers, giving specific consideration to changes over time and to the robustness of estimates. This was investigated by analysing panel data from the Cimanuk Basin in West Java.

Analysis using a stochastic frontier approach on panel data indicated that there were relatively small but significant degrees of technical inefficiencies but with no significant trend over time. This outcome appeared at odds with analysis of technical inefficiencies using a stochastic frontier approach a year at a time, which suggested that most inefficiencies were small and that the frontier itself may vary over time. These outcomes were reconciled by introducing a concept that distinguished the potential absolute frontiers achieved from using fully adopted technologies.

The robustness of the analysis was investigated by comparing a Cobb-Douglas and a translog model specification as well as stochastic frontier and M-quantile regression approaches, and by considering the influence of individual data points. There appeared little difference between the measures obtained from either model specification. The use of different methods did not result in major variations in efficiency measures but did illustrate the importance of considering the analysis on a year by year basis. Differences were investigated in terms of the influential data points identified from regression data diagnostics and it was suggested that the treatment of the influential data points was fundamental to the farm efficiency or performance measure obtained.

A fuller analysis would include more flexible model specifications. For example, consideration should be given to models representing shifts in technology to see if these more general specifications would give greater insight into the efficiency of individual farms over time. Moreover, the yield differences could be viewed as being due to different technologies even when the basic method, for example the use of high yielding varieties, is the same. This is because the specific implementation of technology, for example the application of fertiliser (broadcasting, briquette, deep siting, etc) is evolving. In such cases a specification with time varying parameters could be more appropriate. Influential data points in the estimated frontiers should also be assessed to determine whether or not they are true outliers and thus should be excluded from the analysis (see Seaver and Triantis 1989). Once robust measures of efficiency have been obtained, then possible explanations could be analysed, especially those that can be influenced by policies. The greater robustness of measures obtained by the above approach may lead to the identification of more significant and stable explanations of the inefficiencies than have been observed in the past.

Key policy options that have been used or considered as vehicles for addressing inefficiencies include input subsidies, infrastructure investment, extension advice and research. The most likely policy for dealing with differences between farmers' yields is extension policy. To the extent that reality is represented in the results from analyses of the distribution of farmer performances relative to an absolute frontier and the absolute frontiers relative to the potential absolute frontier, then useful information about policy options is provided. Traditional single frontier analysis that shows all farmers are close to the frontier, and implies that extension is less important (see Esparon and Sturgess 1989), needs to be cognizant of the extent to which the restricted sample nature of such analysis (for example, representation of time periods) may temper such conclusions. Possible gains in group performance outside of the sample information (for example, over time as adoption takes place) are not counted. The results of this analysis suggests that improved extension should be considered from an overall perspective (for example, the timing or form of fertiliser applications) rather than being aimed at individual groups of farms (for example, villages or specific sized farms), as no individual group stands out as being more inefficient than any other.[13]

References

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

Battese, G. E. and T. J. Coelli (1992), 'Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India', *Journal of Productivity Analysis* 3, 153-169.

Belsley, D. A., E. Kuh and R. E. Welsch (1980), *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, New York: Wiley.

Breckling, J. and R. Chambers (1988), 'M-quantiles', *Biometrika* 75, 761-71.

Erwidodo (1990), *'Panel data analysis on farm level efficiency, input demand and output supply of rice farming in West Java, Indonesia'*, Unpublished PhD thesis, Michigan State University, East Lansing, Michigan.

Esparon, N. M. and N. H. Sturgess (1989), 'The measurement of technical efficiency using frontier production functions of rice farms in West Java', *Bulletin of Indonesian Economic Studies* 25(3), 99-119.

Forsund, F. R., C. A. K. Lovell and P. Schmidt (1980), 'A survey of frontier production functions and of their relationship to efficiency measurement', *J. Econometrics* 13, 5-25.

Huber, P. J. (1981), *Robust Statistics*, New York: Wiley.

Kokic, P., J. Breckling, S. Beare and R. Chambers (1992), *'A measure of production performance'*, Unpublished.

Kopp, R. J. and V. K. Smith (1980), 'Frontier production function estimates for steam electric generation: a comparative analysis', *Southern Economic Journal* 47, 1049-1059.

Meeusen, W. and J. van den Broeck (1977), 'Efficiency estimation from Cobb-Douglas production functions with composed error', *International Economic Review* 18(2), 435-444.

Mueller, J. (1974), 'On sources of measured technical efficiency: the impact of information', *American Journal of Agricultural Economics*, 56, 730-38.

[13] This last aspect, also evident from analysis by Squires and Tabor (1991), may be a result of the intensive forms of extension that Indonesia has undertaken at various times or of irrigation and other systems resulting in farmers undertaking very similar forms of production.

Pingali, P.L., P. F. Moya and L. E. Velasco (1990), '*The post-green revolution blues in Asian rice production — the diminished gap between experiment station and farmer yields*', IRRI Social Science Division Papers 90-01, Los Banos, Laguna.

Pitt, M. M. (1983), 'Farm-level fertilizer demand in Java: a meta-production function approach', *American Journal of Agricultural Economics*, 65, 502-08 .

Seaver, B. L. and K. P. Triatis (1989), 'The implications of using messy data to estimate production-frontier based technical efficiency measures', *J. Business Economic Statistics* 7, 49-59.

Siregar, M. (1987), 'Effects of some selected variables on rice farmers' technical efficiency', *Jurnal Agro Ekonomi*, Vol. 6.

Squires, D. and S. Tabor (1991), 'Technical efficiency and future production gains in Indonesian agriculture', *The Developing Economies* 29-3, 258-70.

APPENDIX

Definition of Variables

GRKG: Gross output of rice (kg)

KGS: Seed (kg)

KGN: Urea (kg)

KGP: Phosphate (TSP, kg)

FERT: KGN plus KGP

LAB: Total labour (including family and hired labour, hours)

LAND: Cultivated farm size (hectares)

D1: dummy variable of pesticide use, 1 if farmer uses pesticides and 0 otherwise

D2: dummy high yielding varieties (HYV), 1 if HYV and 0 otherwise

D3: dummy mixed varieties (MV), 1 if MV and 0 otherwise

D4: dummy variable of season, 1 if wet season and 0 otherwise

D5: dummy variable of farm size, 1 if farm size greater than 0.5 ha and 0 otherwise

D6: dummy village, 1 if desa Lanjan kabupaten Indramayu and 0 otherwise

D7: dummy village, 1 if desa Gunung Wangi kabupaten Majalengka and 0 otherwise

D8: dummy village, 1 if desa Malausma kabupaten Majalengka and 0 otherwise

D9: dummy village, 1 if desa Sukaambit kabupaten Sumedang and 0 otherwise

D10: dummy village, 1 if desa Ciwangi kabupaten Garut and 0 otherwise