Chapter 4

A Stochastic Frontier Production Function Incorporating a Model for Technical Inefficiency Effects

4.1 Introduction

Most theoretical stochastic frontier production functions do not explicitly formulate a model for the technical inefficiency effects. Empirical papers, in which the issue of the explanation of these inefficiency effects is raised, include Pitt and Lee (1981), Kalirajan (1981, 1982, 1989), Kalirajan and Flinn (1983) and Kalirajan and Shand (1989). These papers adopt a two-stage approach, in which the first stage involves the specification and estimation of the stochastic frontier production function and the prediction of the technical inefficiency effects of the firms involved. The second stage of the analysis involves the specification of a regression model for the predicted technical inefficiency effects of the firms in terms of various explanatory variables and an additive random error. The parameters of this second-stage inefficiency model have been generally estimated by using ordinary least-squares regression. Kalirajan (1981) specifies that the random errors in the second-stage model for technical inefficiency effects have half-normal distribution. In all these empirical studies, the methods of estimation of the parameters of the second-stage inefficiency model are based on assumptions which are clearly false, because the effects of estimation of the stochastic frontier production function are not accounted for.[1]

Pitt and Lee (1981) investigate the sources of technical inefficiency by specifying that firm intercepts in the stochastic frontier are a function of firm characteristics. The authors regress the estimated firm intercepts on the specified firm characteristics or incorporate the firm characteristics into the production frontier and jointly estimate the parameters involved.

[1] For example, many studies assume the technical inefficiency effects (usually denoted by U_i) are independently and identically distributed in the first-stage estimation. They then regress the predicted U_is upon firm-specific factors in a second stage. The specification of this second-stage model clearly conflicts with the assumption that the U_i are identically distributed.

More recently, models for the technical inefficiency effects in stochastic frontier production functions have been proposed in Kumbhakar, Ghosh and McGuckin (1991), Reifschneider and Stevenson (1991) and Huang and Liu (1994). Kumbhakar, Ghosh and McGuckin (1991) assume that the technical inefficiency effects are non-negative truncations of a normal distribution with mean, which is a linear function of exogenous factors whose coefficients are unknown, and an unknown variance. In addition, Kumbhakar, Ghosh and McGuckin (1991) consider allocative inefficiencies associated with the side conditions for profit maximisation not being exactly satisfied. In the application of their model to US dairy farms, they find that the technical inefficiency effects are significantly related to the level of education of the farmers and the size of their farming operations. Technical and allocative inefficiency effects are investigated in the context of a frontier production function of Zellner-Revankar (1969) type, which proves to be significantly different from the Cobb-Douglas model.

Reifschneider and Stevenson (1991) propose a model for the technical inefficiency effects of the stochastic frontier production function involving the sum of a non-negative function of relevant explanatory variables and a non-negative random variable, which is assumed to have half-normal, exponential or gamma distribution. This model is applied in the analysis of data on electricity generation in the US during three different time periods. The hypothesis, that the inclusion of the model for the technical inefficiency effects does not change the estimates of the frontier function parameters, is rejected in their study.

Huang and Liu (1994) consider a stochastic frontier production function in which the non-negative technical inefficiency effects are a linear function of variables involving firm characteristics. The additive random error of the model for the technical inefficiency effects is assumed to be the truncation of a normal distribution with mode zero, whose point of truncation is dependent on the firm characteristics, such that the technical inefficiency effects are non-negative. Hence the random errors are not required to be non-negative, as in the Reifschneider and Stevenson (1991) model. Huang and Liu (1994) apply their inefficiency frontier model in the analysis of cross-sectional data from the electronics industry in Taiwan and assume that the explanatory variables in the model for the technical inefficiency effects are a function of firm-specific variables and their interactions with the explanatory variables of the stochastic frontier. This makes their model a non-neutral shift of the traditional average response function, in that the marginal products of inputs and marginal rates of technical substitution depend on the firm-specific variables in the model for the technical inefficiency effects.

The model specified in the following section is a special case of the model of Kumbhakar, Ghosh and McGuckin (1991), which is extended to account for panel data. In doing so the model can account for both technical change in the stochastic frontier and time-varying technical inefficiency effects, along with other exogenous factors which influence the technical inefficiency effects. The model is applied in the analysis of farm-level data from an Indian village in Section 4.3. Some concluding comments are made in Section 4.4.

4.2 Model Specification

Consider the stochastic frontier production function for panel data, which is defined by equation (4.1),

$$Y_{it} = \exp(x_{it}\beta + V_{it} - U_{it}) \tag{4.1}$$

where Y_{it} denotes the production for the i-th firm at the t-th period of observation

(i=1,2,...,T; i = 1,2,...,N);

x_{it} is a $(1\times k)$ vector of values of known functions of inputs of production associated with the i-th firm at the t-th period of observation;

β is a $(k\times 1)$ vector of unknown parameters to be estimated;

the V_{it}s are assumed to be iid $N(0, \sigma_V^2)$ random errors, independently distributed of the U_{it}s which are non-negative random variables, associated with technical inefficiency of production;

the U_{it}s are assumed to be independently distributed, such that U_{it} is obtained by truncation (at zero) of the normal distribution with mean, $z_{it}\delta$, and variance, σ^2;

z_{it} is a $(1\times m)$ vector of firm-specific variables (and possibly input variables) which may vary over time; and

δ is an ($m \times 1$) vector of unknown coefficients of the explanatory variables for the technical inefficiency effects.

Although it is assumed that there are T time periods for which observations are available for at least one of the N firms involved, it is not necessary that all the firms are observed for all T periods.

Equation (4.1) specifies the stochastic frontier production function (e.g., of Cobb-Douglas or transcendental-logarithmic form) in terms of the original production values. However, the technical inefficiency effects, the U_{it}s, are assumed to be a function of a set of explanatory variables, the z_{it}s, and an unknown vector of coefficients, δ. The explanatory variables in the inefficiency model would be expected to include any variables which explain the extent to which the production observations fall short of the corresponding stochastic frontier production values, $\exp(x_{it}\beta + V_{it})$. The z_{it}-vectors may have the first element equal to one and include some firm- and time-specific variables. If the first z-variable has value one and the coefficients of all other z-variables are zero, then this case is similar to the models specified in Stevenson (1980) and Battese and Coelli (1988, 1992). If all elements of the δ-vector are equal to zero, then the technical inefficiency effects are not related to the z-variables and so the half-normal distribution originally specified in Aigner, Lovell and Schmidt (1977) is obtained. If interactions between firm-specific variables and input variables are included, then a non-neutral stochastic frontier model, similar to that proposed in Huang and Liu (1994), is obtained.

The technical inefficiency effect, U_{it}, in the stochastic frontier model (4.1) can be equivalently specified as,

$$U_{it} = z_{it}\delta + W_{it}, \tag{4.2}$$

where the random variable, W_{it}, is defined by the truncation of the normal distribution with zero mean and variance, σ^2, such that the point of truncation is $-z_{it}\delta$, i.e., $W_{it} \geq -z_{it}\delta$. These assumptions are consistent with U_{it} being a non-negative truncation of the $N(z_{it}\delta, \sigma^2)$-distribution.

The assumption that the U_{it}s are independently distributed for all $t = 1,2,...,T$, and $i = 1,2,...,N$, is obviously a simplifying, but restrictive, condition. Alternative models are required to account for possible correlated structures of the technical inefficiency effects over time.

It should be noted that the inefficiency frontier model (4.1)-(4.2) is not a generalisation of the Battese and Coelli (1992) model for time-varying technical inefficiency effects (see Chapter 3), even if they are time invariant. The Battese and Coelli (1992) model specifies that the technical inefficiency effects are the product of an exponential function of time and non-negative firm-specific random variables, i.e., $U_{it} = \{\exp[-\eta(t-T)]\}U_i$, where η is an unknown parameter and U_i is a non-negative truncation of the $N(\mu, \sigma^2)$-distribution. This model does not define the technical inefficiency effects in terms of additional explanatory variables. Further, the Battese and Coelli (1992) model implies particular correlated structures for the technical inefficiency effects over time for particular firms.

When the model in equation (4.1) is assumed, the technical efficiency of production for the i-th firm at the t-th observation is defined by

$$TE_{it} = \exp(-U_{it}) = \exp(-z_{it}\delta - W_{it}). \qquad (4.3)$$

If for two firms i and j, $z_{it}\delta + W_{it} > z_{jt}\delta + W_{jt}$, then it does not necessarily imply that the inefficiency effects for another time period, s, will have the same relationship, namely, $z_{is}\delta + W_{is} > z_{js}\delta + W_{js}$. Hence the same ordering of firms in terms of technical efficiency of production at one period of time does not necessarily apply for other time periods, as for the Battese and Coelli (1992) model.

The inefficiency frontier production function (4.1)-(4.2) differs from that of Reifschneider and Stevenson (1991) in that the W-random variables are not identically distributed, as in the latter paper. Reifschneider and Stevenson (1991) assume that the W-random variables in the model for the technical inefficiency effects are non-negative random variables which have half-normal, exponential or gamma distribution. In our model, the W-random variables could be negative if $z_{it}\delta > 0$ because W_{it} is not less than $-z_{it}\delta$, but they are independent truncations of the normal distribution with zero mean and variance, σ^2.

The technical inefficiency frontier model (4.1)-(4.2) is closely related to the models proposed by Kumbhakar, Ghosh and McGuckin (1991) and Huang and Liu (1994), in that it is an extension of these models to account for panel data and hence may be used to account for both technical change and time-varying technical inefficiencies. This extension for time-series data has the same distributional assumptions as if the cross-sectional dimension of the data was increased. However, for our panel-data model there would be particular interest in the behaviour of the technical efficiencies of production of the panel of firms over time.

The parameters of the model defined by (4.1) and (4.2) may be estimated by the method of maximum likelihood. The derivation of the likelihood function and its partial derivatives with respect to the parameters of the model are presented in Appendix 2. These functions are expressed in terms of the variance parameters $\sigma_S^2=\sigma_V^2+\sigma^2$ and $\gamma=\sigma^2/\sigma_S^2$, to facilitate obtaining the maximum-likelihood estimates.

4.3 Empirical Example

Data on paddy farmers from the Indian village of Aurepalle are considered for an empirical application of the stochastic frontier and technical inefficiency model discussed in the previous section. These data were collected by the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT). Battese and Coelli (1992) use data on 15 farmers over the ten-year period from 1975-76 to 1984-85. Because 21 observations were not available for some farmers in some of the years in the ten-year period, only 129 observations are used in that paper. Refer to Chapter 3 for further discussion of these data.

We have endeavoured to obtain data on farmer characteristics that may explain the level of the technical inefficiency of production. Information on the age and years of schooling of 14 of the 15 farmers are available. Hence the data in the present study comprise 14 farmers and a total of 125 observations. Information on other variables, such as the frequency of contacts with agricultural extension officers, access to credit and the use of high-yielding varieties, fertilisers, etc., were not readily available. While not providing a thorough analysis for practical policy purposes, the use of age, years of formal schooling and year of observation in the inefficiency model satisfactorily illustrate the methodology involved.

The stochastic frontier production function to be estimated is similar in structure to that considered in Chapter 3. It is defined by

$$\log(Y_{it}) = \beta_0 + \beta_1\log(Land_{it}) + \beta_2(IL_{it}/Land_{it}) + \beta_3\log(Labour_{it}) + \beta_4\log(Bullocks_{it}) + \beta_5\log(Costs_{it}) + \beta_6(Year_{it}) + V_{it} - U_{it} \quad (4.4)$$

where the technical inefficiency effects are assumed to be defined by,

$$U_{it} = \delta_0 + \delta_1(Age_{it}) + \delta_2(Schooling_{it}) + \delta_3(Year_{it}) + W_{it} \quad (4.5)$$

where

Age is the age of the primary decision maker in the farming operation;

Schooling refers to the number of years of formal schooling of the primary decision maker;

Year indicates the year of the observation involved;

the W_{it} are as defined in the previous section;

and all other variables are as defined in Chapter 3.

The variables in the production frontier (4.4) are those which are in the preferred model in Battese and Coelli (1992). However, the stochastic properties of that model are identical to the ordinary-least squares model, given the assumptions of the stochastic frontier model with time-varying technical inefficiency effects proposed in Battese and Coelli (1992). In this chapter, however, the technical inefficiency effects are assumed to be present in the stochastic frontier and be linearly related to age and education of the paddy farmers and the year of observation involved, such that an intercept parameter is included.

The inefficiency frontier model, defined by equations (4.4) and (4.5), account for both technical change and time-varying technical inefficiency effects. The *Year* variable in the stochastic frontier production function, defined by equation (4.4), accounts for Hicksian neutral technical change. However, the *Year* variable in the model for the technical inefficiency effects, defined by equation (4.5), specifies that the technical inefficiency effects may change linearly with respect to time. Given that the technical inefficiency effects are stochastic and have the specified distributional assumptions, the parameters associated with technical change and the time-varying technical inefficiencies are identified, in addition to the intercept parameters in the stochastic frontier and the model for the technical inefficiency effects.

Maximum-likelihood estimates of the parameters of the model, defined in equations (4.4) and (4.5), are obtained using the computer program, FRONTIER (see Coelli, 1994). This computer program is discussed in detail in Chapter 7. The parameter estimates are given in the second last column of Table 4.1, indicated by Model 1. The last column of Table 4.1 gives the maximum-likelihood estimates for the parameters of the preferred frontier model, to be discussed below, in which some parameters in the general model are specified to be zero.

The signs of the β-estimates are all as expected, with the exception of the negative estimate of the bullock-labour variable. Possible reasons for the parameter associated with bullock labour being negative are discussed in Saini (1979), Battese, Coelli and Colby (1989) and in Chapter 3 above. The positive coefficient of the proportion of land which is irrigated confirms the expected positive relationship between the proportion of irrigated land and total production.

The coefficients of the explanatory variables in the model for the technical inefficiency effects, defined by equation (4.5), are of particular interest to this study. The estimate for the coefficient associated with *Age* is positive, which indicates that the older paddy farmers are more technically inefficient than the younger ones. The estimate for the coefficient associated with *Schooling* is negative. This implies that the paddy farmers with greater years of schooling tend to be less technically inefficient. However, the relationship is very weak, because the coefficient is highly insignificant (by an asymptotic t-test). The negative coefficient of the *Year* variable suggests that the technical inefficiencies of production of the paddy farmers decline throughout the ten-year period.

The estimate for the variance parameter, $\gamma = \sigma^2/\sigma_S^2$, indicates that the variance, σ^2, associated with the inefficiency effects is about 95 percent of the total of the two variances.

Table 4.1

Maximum-likelihood Estimates for Parameters of Stochastic Frontier Production Functions and Inefficiency Models for Paddy Farmers in Aurepalle

Variable	Parameter	Model 1	Model 2
Stochastic Frontier			
Constant	β_0	2.86	3.01
		(0.60)*	(0.57)
log(Land)	β_1	0.37	0.37
		(0.12)	(0.13)
IL/Land	β_2	0.38	0.42
		(0.21)	(0.23)
log(Labour)	β_3	0.85	0.79
		(0.13)	(0.12)
log(Bullocks)	β_4	-0.33	-0.28
		(0.11)	(0.10)
log(Costs)	β_5	0.071	0.084
		(0.031)	(0.032)
Year	β_6	0.014	0
		(0.013)	
Inefficiency Model			
Constant	δ_0	-1.5	0
		(2.8)	
Age	δ_1	0.035	0.0154
		(0.034)	(0.0046)
Schooling	δ_2	-0.006	0
		(0.077)	
Year	δ_3	-0.57	-0.34
		(0.60)	(0.20)
Variance Parameters			
	σ_S^2	0.74	0.40
		(0.75)	(0.20)
	γ	0.952	0.922
		(0.047)	(0.048)
Log-likelihood Function		**-22.60**	**-23.06**

* Estimated standard errors are given in parentheses to two significant digits. The estimated coefficients are given to the corresponding numbers of digits behind the decimal places.

Generalised likelihood-ratio tests of null hypotheses that the technical inefficiency effects are absent or that they have simpler distributions are presented in Table 4.2. The second column of Table 4.2 gives the values of the logarithm of the likelihood when the restrictions specified by the null hypothesis in the first column are applied.

The null hypothesis that the technical inefficiency effects are absent from the model (i.e., H_0: $\gamma = \delta_0 = \ldots = \delta_3 = 0$) is rejected. The second null hypothesis considered in Table 4.2, H_0: $\gamma = 0$, specifies that the technical inefficiency effects are not stochastic. If the parameter, γ, is zero, then the variance of the technical inefficiency effects is zero and so the model reduces to a traditional mean response function in which the variables, age and schooling of the farmers, are included in the production function. However, if the γ-parameter is equal to zero, then the parameters, δ_0 and δ_3, are not identified, given that the production function involves an intercept parameter and year of observation. In this case the model reduces to a traditional average response function in which the constant term is β_0-δ_0, the coefficient of year of observation is β_6-δ_3 and the age and years of schooling of the farmers are explanatory variables along with the other variables specified in equation (4.4) If there are no random technical inefficiency effects in the model, then the parameters, δ_0 and δ_3, are not identified. However, the null hypothesis that the technical inefficiency effects are not random is rejected.

The null hypothesis that the technical inefficiency effects are not a linear function of the year of observation and the age and schooling of the farm operator, H_0: $\delta_1 = \delta_2 = \delta_3 = 0$, is also rejected. This indicates that the joint effect of these three explanatory variables on the levels of technical inefficiencies is significant, although the individual effects of one or more of the variables may not be statistically significant. However, the hypothesis that the technical inefficiency effects have no intercept parameter, H_0 : $\delta_0 = 0$, is not rejected.

Because the estimate for the intercept parameter in the model for the technical inefficiency effects is small relative to its estimated standard error, the model was re-estimated without this parameter. As expected, the estimates for the parameters in this model were little different from those obtained for the more general model, but the estimated coefficients of year of observation in the frontier and schooling in the

Table 4.2

Tests of Hypotheses for Parameters of the Stochastic Frontier and Inefficiency Model for Paddy Farmers in Aurepalle

Null Hypotheses	Log-likelihood Function	Test Statistic λ	Critical Value*	Decision
H_0: $\gamma = \delta_0 = \ldots = \delta_3 = 0$	-37.59	29.99	5.14-10.37	Reject H_0
H_0: $\gamma = 0$	-36.08	26.97	5.14-7.05	Reject H_o
H_0: $\delta_1 = \delta_2 = \delta_3 = 0$	-27.94	10.69	7.81	Reject H_0
H_0: $\delta_0 = 0$	-22.89	0.59	3.84	Accept H_0
H_0: $\beta_6 = \delta_0 = \delta_2 = 0$	-23.06	0.92	7.81	Accept H_0
Restrict: $\beta_6 = \delta_0 = \delta_2 = 0$				
H_0: γ=0	-36.25	26.39	2.71	Reject H_0
H_0: $\delta_1 = \delta_3 = 0$	-39.02	31.93	5.99	Reject H_0

* Critical values are calculated in the manner discussed in Chapter 3.

inefficiency model (β_6 and δ_2, respectively) were less than their estimated standard errors. In fact, the generalised likelihood-ratio statistic for testing the null hypothesis, H_0: $\beta_6 = \delta_0 = \delta_2 = 0$, is not significant and so we consider that the preferred stochastic frontier and inefficiency model has the three parameters, β_6, δ_0 and δ_2, equal to zero.

The maximum-likelihood estimates for the parameters of the preferred frontier model are presented in the last column of Table 4.1. All the parameter estimates for this model are considerably larger than their estimated standard errors. The generalised likelihood-ratio statistic for testing the null hypotheses of the absence of stochastic inefficiency effects, H_0: $\gamma = 0$, and of the absence of age and year effects in the model for the technical inefficiency effects, H_0: $\delta_1 = \delta_3 = 0$, in the preferred frontier model are highly significant (see Table 4.2).

The parameter estimates for the preferred stochastic frontier production function indicate that the elasticity of land is estimated to be 0.37. The estimated elasticity for labour, 0.79, is quite large. The elasticity for bullock labour is significantly less than zero. The estimated elasticity for other input costs is relatively small, 0.084, but is significantly different from zero. These estimates imply that the returns-to-scale parameter is estimated to be 0.965, with estimated standard error of 0.048. Thus the technology of the paddy farmers is such that the hypothesis of constant returns to scale would be accepted.

The technical inefficiency effects in the preferred model are significant, such that older farmers tend to have larger values of the technical inefficiency effects. However, the technical inefficiency effects for the paddy farmers tend to decrease over time.

The technical efficiencies of the paddy farmers in the different years involved are obtained using the predictor, presented in equation (A2.10) of Appendix 2. The parameters involved are estimated by their maximum-likelihood estimates. The predicted technical efficiencies obtained for the 14 paddy farmers involved are presented in Table 4.3.

The predicted technical efficiencies show considerable variability among the paddy farmers. The technical efficiencies of individual paddy farmers also vary up and down over time. Some farmers had the highest level of technical efficiency in one or more years, but had the lowest technical efficiency in at least one year, as well. For example, Farmer 1 had the highest technical efficiencies in the years 1975-76 and 1977-78, but also had the lowest technical efficiencies among the paddy farmers in 1978-79 and 1984-85. These values indicate that there is considerable variation in the levels of technical efficiencies over time for given paddy farmers, although there is a general decline in the technical inefficiencies of the paddy farmers over time. Given that the values of the explanatory variables in the model for the technical inefficiency effects (i.e., age of farmer and year of observation) change little from year to year, the variability in the technical efficiencies of the farmers in the panel is presumably largely due to random variations in the inefficiency model.

Table 4.3

Technical Efficiencies of Paddy Farmers in Aurepalle

Farmer	75-76	76-77	77-78	78-79	79-80	80-81	81-82	82-83	83-84	84-85
1	.887	.615	.928	.606	.856	.730	.733	.944	.839	.814
2	.724	.628	.898	.622	.853	.712	.727	.944	.835	.873
3	.518	.215	.835	.847	.908	.653	-	-	-	-
4	.540	.287	.751	.902	.777	.565	.744	.876	.872	.918
5	.460	.606	.886	.768	.837	.778	.904	.838	.918	.852
6	.730	.510	.922	.866	.794	.767	.761	.913	.884	.885
7	.505	.310	.914	.824	.759	.715	.763	.899	.797	-
8	.758	.465	.749	-	.873	.690	.906	.936	.899	.908
9	.623	.229	.792	.793	.820	.742	.763	.901	.904	.928
10	.664	.737	.875	.812	.913	.723	.928	.940	.898	.943
11	.718	.399	.819	.819	.868	.755	.885	.904	.861	.941
12	.569	.486	.886	.799	.764	.753	.915	.946	.914	.937
13	.420	.402	.888	.800	.825	.293	.570	-	-	-
14	-	.410	.884	.861	.896	-	-	-	-	-
mean	.624	.450	.859	.794	.839	.683	.800	.913	.875	.900

4.4 Conclusions

The results obtained in the empirical application of the proposed model for the stochastic frontier production function and technical inefficiency effects exhibit some interesting differences from those obtained in the application of the time-varying model for technical inefficiency effects presented in Battese and Coelli (1992) and discussed in Chapter 3. Given the specifications of the latter model, it is concluded that there are no technical inefficiencies of production, even though the analysis in Battese and Coelli (1992) involves essentially the same sample of paddy farmers as in this study. However, the Battese and Coelli (1992) model assumes that the technical inefficiency effects are the product of an exponential function of time and the (random) inefficiency effects for firms in the last period of the panel. The present model specifies that the technical inefficiency effects are a linear function of some firm-specific variables and time, together with an additive stochastic error which is assumed to be independent over time and among firms.

One possible reason for the differences in the results obtained in Chapter 4 relative to those in Chapter 3 (in particular the differences in the significance of the γ-parameter in the two analyses) could be that the model specification in Chapter 3 imposes a very rigid structure upon the pattern of the technical inefficiency effects. In that model the rankings of the firms in terms of technical inefficiency are assumed to not differ from one time period to the next and the technical inefficiency effects are also assumed to follow a particular expontial time pattern which is governed by a single parameter, η. This rigidity may be masking the existence of inefficiencies which only become apparent when the less rigid model specification in this chapter is considered. It appears that when the inefficiency effects of a particular firm are allowed to differ randomly between firms and time periods inefficiencies are observed which previously were not visible in the more restrictive panel data model considered in Chapter 3.[2]

The two models form Chapters 3 and 4 are clearly separate and so it is difficult to conclude which is the "best" model for the data involved. However, we do observe that the logarithm of the likelihood function for the data is greater under the assumptions of the above model than for the one proposed in Battese and Coelli (1992).

The next two chapters consider two additional applications of the stochastic frontier model specification proposed in this chapter. These applications involve the analysis of data on farmers from three different villages in India and an analysis of data on electricity generation by coal-fired power stations in Australia.

[2] An interesting issue arises from this discussion. When one is estimating a regular error-components panel data model and finds that the variance of the firm effect is insignificantly different from zero one would normally then revert to estimating an OLS regression. However, in the case of a stochastic frontier panel data model, if the variance parameter (γ) is found to be insignificantly different from zero one should not immediately assume that OLS is appropriate. One should first estimate a stochastic frontier model which does not assume the inefficiency effects are related across time and then conduct a hypothesis test to see if the γ-parameter is non-zero in that model.

Chapter 5

Identification of Factors Which Influence the Technical Inefficiency of Indian Farmers

5.1 Introduction

The measurement of the productive efficiency of a farm relative to other farms or to the "best practice" in an industry has long been of interest to agricultural economists. Much empirical work has centred on imperfect, partial measures of productivity, such as yield per hectare or output per unit of labour. Farrell (1957) suggested a method of measuring the technical efficiency of a firm in an industry by estimating the production function of a "fully-efficient firm" (i.e., a frontier production function). The technical efficiency of a farm may be defined as the ratio of its observed output to that output which could be produced by a fully-efficient firm, given the same input quantities.

Many subsequent papers have applied and extended Farrell's ideas. This literature may be roughly divided into two groups according to the method chosen to estimate the frontier production function, namely, mathematical programming versus econometric estimation. Debate continues over which approach is the most appropriate method to use. The answer often depends upon the application considered. The mathematical programming approach to frontier estimation is usually termed Data Envelopment Analysis (DEA).

The primary criticism of the DEA approach is that measurement errors can have a large influence upon the shape and positioning of the estimated frontier. Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977) independently proposed the stochastic frontier production function to account for the presence of measurement error in production in the specification and estimation of frontier production functions. Stochastic frontiers have two error terms, one to account for technical inefficiency of production and the other to account for other factors such as measurement error in the output variable, luck, weather, etc. and the combined effects of unobserved inputs on production. This favourable property of stochastic frontiers comes with a price, namely, that the functional form of the production function and the distributional assumptions of the two error terms, must be explicitly specified. Bauer (1990) and Greene (1993) present comprehensive reviews of the econometric estimation of frontiers.

In the agricultural economics literature the stochastic frontier (econometric) approach has generally been preferred (see Chapter 2). This is probably associated with a number of factors. The assumption that all deviations from the frontier are associated with inefficiency, as assumed in DEA, is difficult to accept, given the inherent variability of agricultural production, due to weather, fires, pests, diseases, etc. Furthermore, because many farms are small family-owned operations, the keeping of accurate records is not always a priority. Thus much available data on production are likely to be subject to measurement errors.

There have been many applications of frontier production functions to agricultural industries over the years. Some of these papers are reviewed in Chapter 2. Battese (1992) and Bravo-Ureta and Pinheiro (1993) also provide surveys of applications in agricultural economics, the latter giving particular attention to applications in developing countries. Bravo-Ureta and Pinheiro (1993) also draw attention to those applications which attempt to investigate the relationship between technical efficiencies and various socio-economic variables, such as age and level of education of the farmer, farm size, access to credit and utilisation of extension services. The identification of those factors which influence the level of technical efficiencies of farmers is, undoubtedly, a valuable exercise. The information provided may be of significant use to policy makers attempting to raise the average level of farmer efficiency. Most of the applications which seek to explain the differences in technical efficiencies of farmers use a two-stage approach. The first stage involves the estimation of a stochastic frontier production function and the prediction of farm-level technical inefficiency effects (or technical efficiencies). In the second stage, these predicted technical inefficiency effects (or technical efficiencies) are related to farmer-specific factors using ordinary least-squares regression. This approach appears to have been first used by Kalirajan (1981) and has since been used by a large number of agricultural economists, the most recent example of which may be found in Parikh and Shah (1994).

Recent papers by Kumbhakar, Ghosh and McGuckin (1991), Reifschneider and Stevenson (1991), Huang and Lui (1994) and Battese and Coelli (1995) specify stochastic frontiers and models for the technical inefficiency effects and simultaneously estimate all the parameters involved. The Battese and Coelli (1995) stochastic frontier, discussed in Chapter 4, is specified for panel data where the model for the technical inefficiency effects involves farmer-specific variables and year of observation. Battese and Coelli (1995) apply their model in the analysis of a small panel of ten years of data on fourteen paddy farmers from the village of Aurepalle in India. In this chapter a variant of the Battese and Coelli (1995) model is applied in the analysis of data for 34 farmers from this village and also in the analysis of data for farmers from two other Indian villages.

The method of simultaneous estimation of all parameters is preferred to the two-stage approach, referred to above, because the latter is not satisfactory on statistical grounds. There are inconsistencies in the assumptions regarding the distribution of the technical inefficiency effects in the two-stage approach. In the first stage, the technical inefficiency effects are usually assumed to be independently and *identically* distributed random variables. However, in the second stage, the predicted technical inefficiency effects are regressed upon a number of explanatory variables involving farmer- or farm-specific factors. The predicted technical inefficiency effects in this second equation are not independent and even their corresponding true values would only be identically distributed if the coefficients of the explanatory variables in the efficiency relationship were zero.

The remainder of this chapter consists of four sections. In Section 5.2, the data on the farmers from the three Indian villages are briefly described. In Section 5.3, the proposed stochastic frontier and inefficiency model is discussed. In Section 5.4, the empirical results are presented and several hypotheses are tested. In the final section some conclusions are made.

5.2 Panel Data on Indian Agriculture

During the decade from 1975-76 to 1984-85, the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) collected farm-level data on the agricultural operations of a sample of farmers in three different regions in India. These Village Level Studies (VLS) were designed to obtain reliable data on the broad agro-climatic sub-regions in the semi-arid tropics of India, in order to better understand traditional agriculture in the region, with a view to encouraging improved methods of agricultural production.

The three villages of Aurepalle, Kanzara and Shirapur were selected by ICRISAT for the in-depth study of the farming operations involved because they were considered broadly representative of the semi-arid tropics of India. These villages are located in the districts of Mahbubnagar, Akola and Sholapur, respectively, and are approximately 70 km south, 550 km north and 336 km west of the Headquarters of ICRISAT at Patancheru, near Hyderabad in the State of Andhra Pradesh. The three districts were selected because they represented the major soil types, rainfall and cropping patterns in the semi-arid tropics of India. Within each of the selected villages, farmers were stratified into small, medium and large farming operations. Samples of ten farmers were then selected from each of the three groupings in each of the three villages. The numbers of farmers involved in the three villages are 34, 33 and 35 for Aurepalle, Kanzara and Shirapur, respectively. These numbers exceed 30 because some farmers withdrew from the survey program and were replaced by other farmers from the appropriate size category. The total numbers of yearly observations involved in our analyses are 273, 289 and 268, for Aurepalle, Kanzara and Shirapur, respectively.

A brief description of the agro-climatic conditions in the three districts involved is presented below. Walker and Ryan (1990) present a detailed discussion of the regions and the VLS data. Aurepalle is characterised by red soils of shallow-to-medium depth which generally have low water-retention capacities. Kanzara and Shirapur have black soils, which are deeper and have higher water-retention qualities than Aurepalle's red soils. The soils in Shirapur are regarded as better than the soils in Kanzara. Mean annual rainfalls over the ten-year period were 611 mm in Aurepalle, 629 mm in Shirapur and 850 mm in Kanzara, with year-to-year variation between 400 and 1200 mm. The majority of rain falls in the period from June to October. The predominant crops in the three villages are castor, sorghum and paddy in Aurepalle; cotton, pigeon pea and sorghum in Kanzara; and sorghum, chickpea, wheat and vegetables in

Shirapur. More details on the various input variables, and the age and education levels of the farmers, are in Table 5.1 , which is presented and discussed in Section 5.4.

5.3 The Stochastic Frontier and Inefficiency Model

The stochastic frontier production function which is specified for the farming operations in each village is

$$\log(Y_{it}) = \beta_0 + \beta_1\log(Land_{it}) + \beta_2(IL_{it}/Land_{it}) + \beta_3\log(Labour_{it})$$
$$+ \beta_4(HL_{it}/Labour_{it}) + \beta_5\log(Bullocks_{it}) + \beta_6\log(Costs_{it})$$
$$+ \beta_7(Year_{it}) + V_{it} - U_{it} \qquad (5.1)$$

where the technical inefficiency effects, U_{it}, are such that they are independently distributed and arise by truncation (at zero) of the normal distribution with variance, σ^2, and mean, μ_{it}, where μ_{it} is defined by

$$\mu_{it} = \delta_0 + \delta_1(Age_{it}) + \delta_2(Schooling_{it}) + \delta_3(Size_{it}) + \delta_4(Year_{it}) \qquad (5.2)$$

where *HL* is the quantity of hired labour; *Size* of the farming operation is proxied by the *Land* variable; and all other variables are as defined in Chapters 3 and 4.

The expected signs on the δ-parameters are not clear in all cases. The age of the farmers could be expected to have a positive or a negative effect upon the size of the inefficiency effects. The older farmers are likely to have had more farming experience and hence have less inefficiency. However, they are also likely to be more conservative and thus be less willing to adopt new practices, thereby perhaps having greater inefficiencies in agricultural production.

Schooling is expected to have a negative effect upon technical inefficiency effects. That is, we expect that a greater level of formal education will be associated with smaller values for the technical inefficiency effects.

The sign of the coefficient of the *Size* variable is expected to be negative. This expectation is partially based upon the likelihood that the farmers with smaller operations may have alternative income sources which are more important and hence put less effort into their farming operations compared to the larger farmers. It is also possible that the modified Cobb-Douglas form used in this analysis does not appropriately accommodate a range of scale economies and hence that some scale inefficiency may be included in the estimated technical inefficiencies of production.

The coefficient of year of observation in the model for the technical inefficiency effects is expected to be negative. This would imply that the levels of the technical inefficiency effects of farmers in the three villages tend to decrease over time. That is, farmers tend to become more technically efficient over time. This time-trend variable is expected to pick up the influence of factors which are not included in the inefficiency model which vary systematically through time. For example, it may reflect the influence of government agricultural extension programs over the sample period.

The stochastic frontier production function, defined by equation (5.1), is identical to those estimated in Chapters 3 and 4, except that the ratio of hired labour to total labour used, *HL/Labour*, is included to account for possible differences in the productivity's of hired and family labour in the farming operations in the three villages. This variable was not considered in the analyses of the Aurepalle paddy farmers in Chapters 3 and 4 because of earlier analyses which had indicated that hired and family labour were equally productive in the farming operations of the Aurepalle paddy farmers. The study in this chapter, however, considers both paddy and non-paddy farmers in Aurepalle, as well as farmers from two other villages, hence it is appropriate to include this variable in our models to allow hired and family labour to have differing productivity's.

The stochastic frontier production function, defined in equation (5.1), is a linearised approximation of a Cobb-Douglas production function in which the land and labour variables are linear combinations of irrigated and unirrigated land and hired and family labour, respectively. For more on this particular specification, see Battese, Coelli and Colby (1989), Battese and Coelli (1992) and the discussion in Chapter 3. A test of the hypothesis that hired and family labour are equally productive is obtained by testing the null hypothesis that the coefficient, β_4, of the labour-ratio variable, *HL/Labour*, is zero. This hypothesis is of particular interest in Indian agriculture, cf. Bardhan (1973). A similar test can be defined for the two different components of the land input.

As stated in Chapter 4, there is interest in testing the null hypothesis that the technical inefficiency effects are not stochastic, i.e., H_0: $\gamma = 0$, given the level of the inputs involved. Further, the null hypothesis that the technical inefficiency effects are not related to age or education of farmers, the size of their farming operations and the year of observation, is specified by H_0: $\delta_1 = ...= \delta_4 = 0$. Tests of these hypotheses are of interest in assessing the characteristics of the technical inefficiency effects for farmers in the three Indian villages involved.

5.4 Results and Discussion

A summary of the sample data on the different variables in the stochastic frontier and inefficiency model, defined by equations (5.1) and (5.2), is presented in Table 5.1. The sizes of the holdings are small relative to those seen in modern western agriculture. The average farm sizes vary from 4.29 ha in Aurepalle to 6.02 ha and 6.68 ha in Kanzara and Shirapur, respectively. The smaller holdings in Aurepalle could be attributed to the greater use of irrigation in Aurepalle (an average of 0.95 ha per farm in Aurepalle versus approximately 0.5 ha per farm in the other two villages). Labour use is higher in Aurepalle and Kanzara where paddy planting and cotton picking are labour-intensive activities. The use of bullock labour and costs of other inputs in Aurepalle and Kanzara are higher than in Shirapur. Much of this is due to the high input use required with the above two crops. The average age of farmers vary from 43.7 years in Kanzara to 53.9 years in Aurepalle, while average education levels are quite low, varying from about two years in Aurepalle to about four years in Kanzara.

5.4.1 Maximum-likelihood Estimates

The maximum-likelihood estimates for the parameters in the stochastic frontier and inefficiency model are presented in Table 5.2 for the three villages involved. The estimated δ-coefficients associated with the explanatory variables in the model for the technical inefficiency effects are worthy of particular discussion. We observe that age has a negative effect upon the technical inefficiency effects in Aurepalle and Kanzara. That is, the older farmers tend to have smaller technical inefficiencies (i.e., are more technically efficient) than younger farmers in Aurepalle and Kanzara, but the reverse is true in Shirapur. This mixture of signs is not unexpected, given the various effects that farmer age may have upon efficiency, as discussed in the Section 5.3. The result for Aurepalle differs from that reported in Chapter 4 in the analysis of Aurepalle paddy

Table 5.1

Summary Statistics for Variables in the Stochastic Frontier and Inefficiency Models for Farmers in Three Indian Villages

Variable	Sample Mean	Standard Deviation	Minimum Value	Maximum Value
Value of Output (Rupees)				
Aurepalle	3679.6	4559.2	10.15	18094
Kanzara	5231.3	7226.5	121.58	39168
Shirapur	3270.7	3482.7	22.00	26423
Land (hectares)				
Aurepalle	4.29	3.87	0.20	20.97
Kanzara	6.02	7.40	0.40	36.34
Shirapur	6.68	5.49	0.61	24.19
Irrigated Land (hectares)				
Aurepalle	0.95	1.41	0	7.09
Kanzara	0.51	1.22	0	9.79
Shirapur	0.64	1.07	0	4.96
Labour (hours)				
Aurepalle	2206.2	2744.1	26	12916
Kanzara	2578.5	3145.7	58	15814
Shirapur	1674.8	1576.9	40	11146
Hired Labour (hours)				
Aurepalle	1468.3	2349.6	0	11662
Kanzara	1841.2	2852.3	6	14130
Shirapur	719.1	768.4	24	4823
Bullock Labour (hours)				
Aurepalle	528.2	604.6	8	4316
Kanzara	570.6	765.1	12	3913
Shirapur	342.3	282.2	14	1240
Cost of Other Inputs (Rupees)				
Aurepalle	651.02	981.06	0	6205.0
Kanzara	628.96	978.49	0	5344.3
Shirapur	464.49	1038.00	0	6746.0
Age of Farmer (years)				
Aurepalle	53.9	12.6	26	90
Kanzara	43.7	9.6	23	67
Shirapur	48.2	10.2	24	72
Schooling of Farmer (years)				
Aurepalle	2.01	2.87	0	10
Kanzara	4.03	4.10	0	12
Shirapur	2.94	3.35	0	16

* Sample sizes are 273, 289 and 268 for Aurepalle, Kanzara and Shirapur, respectively.

farmers. However, the size of the farm is not considered as a factor in the inefficiency model in the Chapter 4 analysis and, furthermore, that study only involved those farmers who grew some amount of paddy while the present analysis involves all sample farmers in Aurepalle, including those who have no paddy fields.

The coefficient of *Schooling* is observed to be negative in Aurepalle and Shirapur, but positive in Kanzara. That is, in the villages of Aurepalle and Shirapur, farmers with greater years of formal education tend to be more technically efficient in agricultural production. The positive value obtained for Kanzara is unexpected, but could be due to the generally small numbers of years of formal schooling observed throughout the sample (see Table 5.1). We hypothesise that the result may have been different if a wider spread of education levels were observed.

The sign of the estimated coefficient of the *Size* variable in each village is negative, as expected. This indicates that farmers with larger farms tend to have smaller technical inefficiency effects than farmers with smaller operations. As discussed in Section 5.3, this result may be due to a number of factors, one of which could be some scale inefficiency being measured as technical inefficiency. We intend to investigate this issue by replacing the modified Cobb-Douglas functional form with a modified translog functional form in future work.

The coefficient of year of observation in the model for the technical inefficiency effects is also estimated to be negative in all three villages. This implies that the levels of the technical inefficiency effects of farmers in the three villages tend to decrease over time. That is, farmers tend to become more technically efficient over time. This time-trend variable may be picking up the influence of factors which are not included in the inefficiency model. For example, it may reflect the positive influence of government agricultural extension programs over the sample period.

Overall, the signs of the estimated δ-coefficients conform quite closely with our expectations. Only the coefficient of schooling in Kanzara has a sign which is contrary to our expectations. Note, however, that the ratio of this estimate to its estimated standard error (t-ratio) is only slightly larger than one in value, indicating that this

Table 5.2

Maximum-likelihood Estimates for Parameters of the Stochastic Frontier and Inefficiency Models for Three Indian Villages*

Variable	Parameter	Aurepalle	Kanzara	Shirapur
Stochastic Frontier				
Constant	β_0	-5.62	-4.90	-4.69
		(0.33)	(0.37)	(0.32)
Land	β_1	0.264	0.066	0.012
		(0.070)	(0.066)	(0.061)
IL/Land	β_2	0.093	0.083	-0.076
		(0.058)	(0.038)	(0.030)
Labour	β_3	1.212	0.785	0.905
		(0.076)	(0.079)	(0.060)
HL/Labour	β_4	-0.00047	-0.000019	0.00020
		(0.00012)	(0.000091)	(0.00040)
Bullocks	β_5	-0.430	-0.006	-0.086
		(0.056)	(0.060)	(0.060)
Costs	β_6	0.009	0.098	0.002
		(0.014)	(0.011)	(0.010)
Year	β_7	0.0279	-0.0182	0.016
		(0.0088)	(0.0081)	(0.012)
Inefficiency Model				
Constant	δ_0	-1.8	0.80	1.37
		(1.5)	(0.35)	(0.50)
Age	δ_1	-0.0150	-0.015	0.0133
		(0.0092)	(0.010)	(0.0099)
Schooling	δ_2	-0.064	0.039	-0.217
		(0.046)	(0.033)	(0.088)
Size	δ_3	-0.29	-0.083	-0.208
		(0.14)	(0.056)	(0.082)
Year	δ_4	-0.36	-0.077	-0.39
		(0.15)	(0.046)	(0.12)
Variance Parameters				
	σ_S^2	2.19	0.39	0.96
		(0.92)	(0.20)	(0.35)
	γ	0.9826	0.915	0.944
		(0.0069)	(0.040)	(0.023)
Log-likelihood Function		**-99.51**	**-80.29**	**-128.81**

* Estimated standard errors are given below the parameter estimates, correct to at least two significant digits. The parameter estimates are given correct to the corresponding number of digits behind the decimal places.

estimate may not be statistically significant. Also note that this t-ratio is the smallest among all the δ-estimates in any of the three villages.

The γ-parameter associated with the variances in the stochastic frontier is estimated to be greater than 0.9 in all of the three villages. Although this parameter cannot be interpreted as the proportion of the variance of the inefficiency effects relative to the sum of the variances of the inefficiency effects and the random variation, it indicates that the random component of the technical inefficiency effects do make a significant contribution in the analysis of agricultural production in the Indian villages involved.

The estimated coefficients of the stochastic frontier, defined by equation (5.1), reported in Table 5.2, have signs and sizes which generally conform with those obtained in past analyses of these data. The estimated coefficients of *Land* and *Labour* are positive for all of the three villages. The coefficient of *IL/Land* is expected to be positive, reflecting the higher productivity of irrigated land. However, for Shirapur the coefficient of the proportion of irrigated land is estimated to be negative and significantly different from zero. Further investigation is required to discern the basis for this result.

If the productivity of hired labour was lower than that for family labour, then the coefficient of *HL/Labour* would be negative. Negative estimates are obtained for Aurepalle and Kanzara, but for Shirapur the estimated coefficient is positive. However, the ratio of the estimated coefficient to the estimated standard error suggests that hired and family labour in Kanzara and Shirapur are equally productive.

The estimated coefficients of bullock labour are negative for all three villages, but only the estimate for Aurepalle is significantly different from zero. This negative influence is contrary to what one would expect, but conforms with earlier analyses, reported by Saini (1979), and Battese and Coelli (1992, 1995) and discussed in earlier chapters. A number of explanations have been suggested for this result, the most often quoted is, that the bullocks are often used for weed control and repairs of irrigation banks in poor seasons when the land is less water-logged. Thus the quantity of bullock labour may be acting as an inverse proxy for rainfall.

5.4.2 Tests of Hypotheses

Formal tests of hypotheses associated with the technical inefficiency effects are presented in Table 5.3. These tests of hypotheses involve the use of the generalised likelihood-ratio statistic, which has been discussed in Chapter 3. The generalised likelihood-ratio test is often preferred to the asymptotic t-test since the estimated standard errors can sometimes be unreliable when they are calculated as a by-product of the iterative procedure for ML estimation. Furthermore, the t-test can only be used when the null hypothesis involves a single restriction.

The first null hypothesis considered in Table 5.3, H_0: $\gamma = \delta_0 = \ldots = \delta_4 = 0$, specifies that the inefficiency effects in the frontier model are not stochastic (i.e., $\sigma^2=0$ and hence $\gamma=0$) and all the coefficients of the explanatory variables in the inefficiency model are zero. This hypothesis implies that the technical inefficiency effects are, in fact, zero or absent from the model, which, in turn, implies that the stochastic frontier model, defined by equation (5.1), is equivalent to the traditional average response function. This null hypothesis is clearly rejected by the data for all of the three villages involved. Thus the traditional average response function is not an adequate representation for the agricultural production in the three villages, given the specification of the stochastic frontier and inefficiency model, defined by equations (5.1) and (5.2).

The second null hypothesis in Table 5.3, H_0: $\gamma = 0$, specifies that the technical inefficiency effects in the frontier are non-stochastic. This null hypothesis is also strongly rejected for all three villages.

The third null hypothesis in Table 5.3, H_0: $\delta_0 = \ldots = \delta_4 = 0$, specifies that all the δ-parameters in the model for technical inefficiency effects in the stochastic frontier production function have value zero (and hence that the inefficiency effects have half-normal distribution). This hypothesis is also strongly rejected for all three villages.

Table 5.3

Tests of Hypotheses for Coefficients of the Explanatory Variables for the Technical Inefficiency Effects in Stochastic Frontier Production Functions for Three Indian Villages

Null Hypothesis	Log-likelihood Value	Test Statistic, λ	Critical Value	Decision
H_0: $\gamma = \delta_0 = \ldots = \delta_4 = 0$				
Aurepalle	-138.02	77.02	5.14-11.91	Reject H_0
Kanzara	-106.03	51.48	5.14-11.91	Reject H_0
Shirapur	-183.68	109.74	5.14-11.91	Reject H_0
H_0: $\gamma = 0$				
Aurepalle	-137.86	76.70	5.14-7.05	Reject H_0
Kanzara	-100.18	39.78	5.14-7.05	Reject H_0
Shirapur	-177.54	97.46	5.14-7.05	Reject H_0
H_0: $\delta_0 = \ldots = \delta_4 = 0$				
Aurepalle	-113.12	27.22	11.07	Reject H_0
Kanzara	-93.27	25.96	11.07	Reject H_0
Shirapur	-161.58	65.54	11.07	Reject H_0
H_0: $\delta_1 = \ldots = \delta_4 = 0$				
Aurepalle	-101.92	4.82	9.49	Accept H_0
Kanzara	-91.13	21.68	9.49	Reject H_0
Shirapur	-151.98	46.34	9.49	Reject H_0

The final null hypothesis considered in Table 5.3, H_0: $\delta_1 = \ldots = \delta_4 = 0$, specifies that all the coefficients of the explanatory variables in the inefficiency model are equal to zero (and hence that the technical inefficiency effects have truncated-normal distribution). This null hypothesis is rejected for the villages of Shirapur and Kanzara, but it is accepted for Aurepalle. Thus for Aurepalle, it could be concluded that the technical inefficiency effects are not significantly influenced by the age and education of the farmers, the size of the farming operation, and that they are not time-varying. Hence it appears that, given the specifications of the stochastic frontier and inefficiency model, defined by equations (5.1) and (5.2), the technical inefficiency effects for Aurepalle farmers can be regarded as independent and identically distributed random variables which arise from the truncation of a normal distribution with non-zero mean.

Thus, to summarise the tests of hypotheses in Table 5.3, it appears that there are significant technical inefficiencies in the agricultural production in the three villages considered in this study. In the villages of Shirapur and Kanzara, the explanatory variables (age, education, farm size and time) are observed to have a significant influence upon the technical inefficiency effects. In Aurepalle, however, these variables do not appear to have a significant influence. It could thus be concluded that there is considerable unexplained variation in the technical inefficiency effects in Aurepalle, suggesting the investigation of alternative explanatory variables, such as access to credit and extension advice, is particularly warranted in the case of Aurepalle.

Several tests of hypotheses regarding the β-parameters are also of interest. Generalised likelihood-ratio tests of the null hypothesis that the coefficient of the hired-labour ratio is zero are presented in Table 5.4 for the three villages. The null hypothesis, H_0: $\beta_4 = 0$, is rejected for farming operations in Aurepalle, but accepted for Kanzara and Shirapur. The conclusion that hired and family labour are not equally productive in Aurepalle may be associated with the labour-intensive operations required in paddy production, and the nature of the well-developed labour market in that region.

In our stochastic frontier production function, the cost of other inputs, such as fertiliser, manure and pesticides, is included as an explanatory variable. It has been suggested that this variable should not be used in a frontier production function, because it is a composite variable which contains the costs of various items which are likely to influence production in different ways. We maintain that the inclusion of this variable is preferable to its exclusion, on the grounds that it should reduce the degree of misspecification. Also considered in Table 5.4 is a test of the null hypothesis, H_0: $\beta_6 = 0$, which specifies that the coefficient of the cost of other inputs is zero. For Aurepalle and Shirapur, this null hypothesis is accepted, while for Kanzara it is strongly rejected. This result may be due in part to the importance of cotton production in Kanzara. The cotton plant is susceptible to a number of insect pests and so the regular use of pesticides in cotton production appears to be a highly significant factor in the agricultural production in Kanzara.

The final hypothesis considered in Table 5.4 relates to the question of technical change. This involves a test of the null hypothesis, H_0: $\beta_7 = 0$, that the coefficient of year of observation in the stochastic frontier is equal to zero. The test statistics indicate that the null hypothesis of no technical change is rejected in Aurepalle and Kanzara, but is

Table 5.4

Statistics for Tests of Hypotheses Involving Some Coefficients of the Stochastic Frontier Production Functions for Three Indian Villages

Null Hypothesis	Log-Likelihood Function	Test Statistic λ	Critical Value	Decision
H_0: $\beta_4 = 0$				
Aurepalle	-104.90	10.78	3.84	Reject H_0
Kanzara	-80.31	0.04	3.84	Accept H_0
Shirapur	-128.97	0.32	3.84	Accept H_0
H_0: $\beta_6 = 0$				
Aurepalle	-99.69	0.36	3.84	Accept H_0
Kanzara	-111.28	61.98	3.84	Reject H_0
Shirapur	-128.81	0.00	3.84	Accept H_0
H_0: $\beta_7 = 0$				
Aurepalle	-103.32	7.62	3.84	Reject H_0
Kanzara	-83.04	5.50	3.84	Reject H_0
Shirapur	-129.80	1.98	3.84	Accept H_0

accepted for Shirapur. We note that the coefficient of year of observation in the stochastic frontier, β_7, is positive for Aurepalle, but negative for Kanzara. The latter result is surprising and may merit further investigation. One possible reason why one may observe technical *regress* is the situation where intensive cropping practices reduce the nutrient content of the soil at a faster rate than fertiliser application replenishes it. A closer inspection of the farming practices in Kanzara is required before any conclusions can be made.

Finally, it is interesting to note that the conclusions of the generalised likelihood-ratio tests listed in Table 5.4 are the same as those that would have been made if asymptotic t-tests had been used. Thus, in this application, the standard errors of the ML estimators appear to be well estimated using the Davidon-Fletcher-Powell algorithm which is used in the program, FRONTIER.

5.4.3 Technical Efficiencies of Farmers

The technical efficiencies of farmers are predicted for each year in which they were observed, using the method proposed in Battese and Coelli (1993) and presented in Appendix 2. These predictions are derived from the estimated models presented in Table 5.2. The predicted technical efficiencies of the farmers in Aurepalle, Kanzara and Shirapur are presented in Tables 5.5, 5.6 and 5.7, respectively. Also presented in these tables are estimates for the mean technical efficiencies of each farmer (over the ten-year period) and the mean technical efficiencies for farmers in each of the years involved. The predicted technical efficiencies differ substantially within each village. They range from quite small values of less than 0.1 to values in excess of 0.9. The mean technical efficiencies of the farmers range from 0.353 for farmer 32 in Shirapur to 0.921 for farmer 28 in Kanzara. The mean technical efficiencies of the farmers in the three villages do not appear to differ substantially. They are 0.747 for Aurepalle, 0.738 for Kanzara and 0.711 for Shirapur.

To give a better indication of the distribution of the individual technical efficiencies, frequency distributions of the technical efficiencies are plotted for Aurepalle, Kanzara and Shirapur in Figures 5.1, 5.2 and 5.3, respectively. The plots are quite similar, with a thin tail in the left of the distribution, gradually rising to a maximum in the 0.8 to 0.9 interval, and then dropping sharply in the 0.9 to 1.0 interval. The fact that the mode of the distribution is not in this final interval offers support for the use of more general distributions (than the often considered half-normal distribution) for the technical inefficiency effects, such as the general truncated-normal distribution used in this study.

The annual mean technical efficiencies, which are presented in the bottom row of each of Tables 5.5, 5.6 and 5.7, are plotted in Figure 5.4. A general upward trend in the levels of mean technical efficiency is observed over the sample period in all three villages. The mean technical efficiencies in Shirapur tend to follow a rather smooth upward trend, in comparison with the more volatile results for Aurepalle and Kanzara. There is also a suggestion of a reduction in the variability of the mean technical efficiencies in the three villages towards the end of the ten-year period, relative to the greater divergence in the values in the earlier part of the sample period. This could reflect an improvement in the ability of the farmers to adjust their production methods to the year-to-year changes in the agro-climatic environments in the regions involved.

Table 5.5

Predicted Technical Efficiencies for Farmers in Aurepalle

Farm	75-76	76-77	77-78	78-79	79-80	80-81	81-82	82-83	83-84	84-85	Mean
1	-	-	-	-	-	.554	.590	.909	.764	.867	.737
2	-	-	-	-	-	.558	.573	-	.721	.351	.551
3	-	-	-	-	-	-	.323	.879	-	-	.601
4	-	-	-	-	-	.586	.790	.890	.805	.756	.765
5	.756	.772	.928	.587	.818	.700	.642	.918	.700	.550	.737
6	.745	.804	.908	.606	.825	.674	.651	.922	.702	.707	.754
7	.894	.837	-	.543	.664	.850	.388	.873	.826	.865	.749
8	.841	.154	.802	.800	.618	.582	.615	.846	.785	.847	.689
9	.767	.825	.472	-	.880	-	.664	.938	.709	.875	.766
10	.919	.749	.836	.887	.828	-	.607	.896	.905	.914	.838
11	.454	.599	.702	.795	.813	-	.475	.929	.681	.758	.689
12	.939	-	.811	.779	.680	.486	.304	.842	.045	.538	.603
13	.715	.778	.834	.834	.375	-	.604	.932	.563	.850	.721
14	.648	-	.809	.799	.835	.860	-	-	-	-	.790
15	.411	.372	.931	.750	.834	.758	-	-	-	-	.676
16	.705	.220	.826	.846	.908	.647	-	-	-	-	.692
17	.358	-	-	-	.487	.595	-	-	-	-	.480
18	.752	.452	.903	.890	.777	.799	.697	.869	.859	.851	.785
19	.665	.393	.662	.650	.704	.506	.676	.852	-	-	.638
20	.673	.365	.757	.906	.790	.588	.769	.843	.819	.874	.739
21	.620	.813	.888	.779	.825	.847	.890	.878	.905	.872	.832
22	.903	.452	.878	.879	.880	.456	.845	.837	.864	.874	.787
23	.890	.478	.800	.803	.707	.465	.649	-	-	-	.685
24	.875	.767	.933	.897	.847	.822	.805	.887	.848	.847	.853
25	.934	.231	.901	.869	.754	.583	.696	.716	.825	.690	.720
26	.654	.423	.930	.838	.764	.788	.827	.890	.749	-	.763
27	.833	.610	.802	-	.827	.653	.885	.920	.841	.847	.802
28	.748	.254	.785	.776	.781	.704	.702	.863	.823	.868	.730
29	.864	.765	.853	.800	.888	.826	.747	.829	.877	.887	.834
30	.807	.891	.913	.848	.926	.838	.932	.935	.874	.929	.889
31	.834	.505	.855	.857	.871	.728	.854	.859	.797	.905	.807
32	.694	.555	.895	.791	.741	.716	.881	.925	.869	.899	.796
33	.504	.463	.905	.822	.793	.312	.636	-	-	-	.634
34	-	.428	.894	.833	.844	-	-	-	-	-	.750
Mean	**.738**	**.554**	**.836**	**.795**	**.776**	**.660**	**.680**	**.880**	**.766**	**.801**	**.747**

Table 5.6

Predicted Technical Efficiencies for Farmers in Kanzara

Farm	75-76	76-77	77-78	78-79	79-80	80-81	81-82	82-83	83-84	84-85	Mean
1	-	-	.526	.558	.683	.378	.493	.774	-	-	.569
2	-	-	-	-	.596	.353	.737	.670	.690	-	.609
3	-	-	-	.552	.847	.596	-	-	.824	.790	.722
4	.832	.794	.598	.740	.729	.506	.881	.819	.896	.883	.768
5	.871	.750	.819	.309	.591	.440	.649	.875	.900	.885	.709
6	.916	.596	.653	.378	.614	.372	.738	.741	.883	.674	.657
7	.904	.460	.841	.602	.652	.458	.825	.817	.889	.852	.730
8	.856	-	-	.414	.425	.498	.530	.690	-	-	.569
9	.740	.523	.843	-	.669	.679	.915	.883	.675	.904	.759
10	.906	.844	.757	.602	.900	.640	.909	.773	-	-	.792
11	.919	.708	.735	.654	.843	.466	.585	.837	.947	.777	.747
12	.695	.365	.629	.687	.773	.754	.704	.860	.886	.879	.723
13	.847	-	-	-	-	-	-	-	.853	-	.850
14	.372	.880	.470	.132	.782	.617	-	.593	.897	.688	.603
15	.873	.809	.791	.565	.699	.625	.860	.866	.914	.820	.782
16	.739	.792	.415	.337	.804	.461	.606	-	.878	.908	.660
17	.702	-	-	-	-	.765	.597	.810	.826	.785	.748
18	.844	.793	.910	.819	.837	.639	.920	.924	.910	.851	.845
19	.867	.863	.605	.427	.249	.692	.534	.762	.660	.866	.652
20	.585	.908	.727	.830	.886	.551	.746	.793	.876	.767	.767
21	.768	.864	.431	.593	.706	.329	.783	.579	.896	.796	.674
22	.435	.654	.611	.686	.845	.464	.712	.759	.849	.847	.686
23	.863	.720	.479	.393	.709	.408	.740	.756	.721	.853	.664
24	.942	.848	.838	.891	.850	.635	.794	.811	.835	.851	.830
25	.854	.923	.855	.860	.823	.792	.867	.901	.932	.838	.864
26	.625	.553	.387	.452	-	-	-	-	-	-	.504
27	.805	.631	.606	.545	.783	.449	.733	.657	.812	.798	.682
28	.947	.934	.895	.867	.930	.901	.933	.944	.942	.918	.921
29	.754	.908	.808	.722	.780	.562	.842	.883	.883	.874	.802
30	.836	.777	.681	.402	.794	.458	.818	.773	.824	.850	.721
31	.903	.827	.653	.837	.756	.660	.902	.870	.881	.876	.817
32	.792	.815	.659	.626	.454	.855	.908	.925	.870	.862	.777
33	.856	.908	.872	.868	.898	.747	.902	.925	.939	.936	.885
Mean	**.795**	**.757**	**.682**	**.598**	**.730**	**.573**	**.764**	**.802**	**.855**	**.838**	**.738**

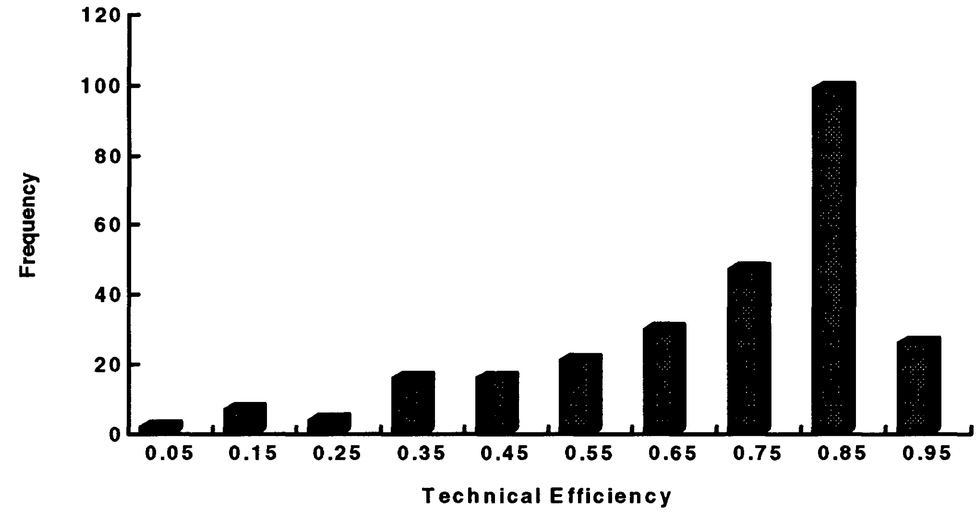
Table 5.7

Predicted Technical Efficiencies for Farmers in Shirapur

Farm	75-76	76-77	77-78	78-79	79-80	80-81	81-82	82-83	83-84	84-85	Mean
1	-	.613		.679	.715	.869	.800	.890	.910	.874	.775
2	-	.375	.670	.328	.181	-	-	-	-	-	.389
3	-	.749	.882	.727	.916	.867	.903	.712	.633	.392	.754
4	-	-	-	-	-	.707	.761	.802	.611	.821	.740
5	.568	.192	.340	.404	.608	.827	.721	.696	.599	-	.551
6	.352	.833	.811	.850	.885	.917	.770	.742	.463	.549	.717
7	.276	.739	.606	.781	.575	-	-	-	-	-	.595
8	.100	.298	.338	.764	.762	.637	.888	.900	.818	.877	.638
9	.022	-	-	-	.427	.099	.443	.556	.661	.468	.382
10	.361	.709	.523	.778	.629	.626	-	.806	.482	.450	.596
11	.390	.727	.496	.767	.872	.836	.897	.919	.896	.554	.735
12	.865	.859	.552	-	-	-	-	-	-	-	.759
13	.479	.737	.801	.789	.819	.839	.798	.567	.862	.880	.757
14	.345	.806	.454	.721	.721	.886	.722	.855	.760	-	.697
15	.180	.601	.885	.636	.936	.922	.903	.926	.765	.855	.761
16	.297	.445	.511	.346	.690	.700	.869	.900	-	-	.595
17	.316	.528	.743	.503	.685	.884	-	-	-	-	.610
18	.400	.688	.668	.586	.588	.847	.871	.892	.765	.877	.718
19	.178	.588	.745	.695	.843	.696	.864	.887	.893	.712	.710
20	.471	.882	.773	.845	.943	.910	.919	-	-	-	.820
21	.224	-	-	.464	-	.360	.778	.826	.864	.876	.628
22	.647	.756	.854	.787	.829	.859	.558	.891	.641	.912	.774
23	.152	-	-	-	.416	-	-	-	-	-	.284
24	.341	.718	.818	.780	.855	.848	.872	.876	.852	.859	.782
25	.700	.623	.828	.781	.928	.861	.905	.886	.804	.806	.812
26	.416	.700	.565	.731	.808	.717	.804	.838	.796	.867	.724
27	.776	.865	.926	.889	-	.599	.897	.905	.905	.460	.802
28	.735	.808	.855	.660	.769	.710	.901	.911	.893	.890	.813
29	.376	.813	.791	.849	.808	.833	.799	.891	.845	.834	.784
30	.892	.904	.812	.873	.888	-	-	-	-	-	.874
31	.932	.852	.827	-	-	-	-	-	-	-	.870
32	.353	-	-	-	-	-	-	-	-	-	.353
33	-	.195	.501	.523	.689	.768	-	-	-	-	.535
34	-	.713	.651	.530	.851	-	-	.830	.900	.867	.763
35	-	.892	.853	.863	.910	.883	.888	.933	.889	.893	.889
Mean	**.434**	**.674**	**.690**	**.687**	**.743**	**.760**	**.814**	**.833**	**.771**	**.753**	**.711**

Figure 5.3

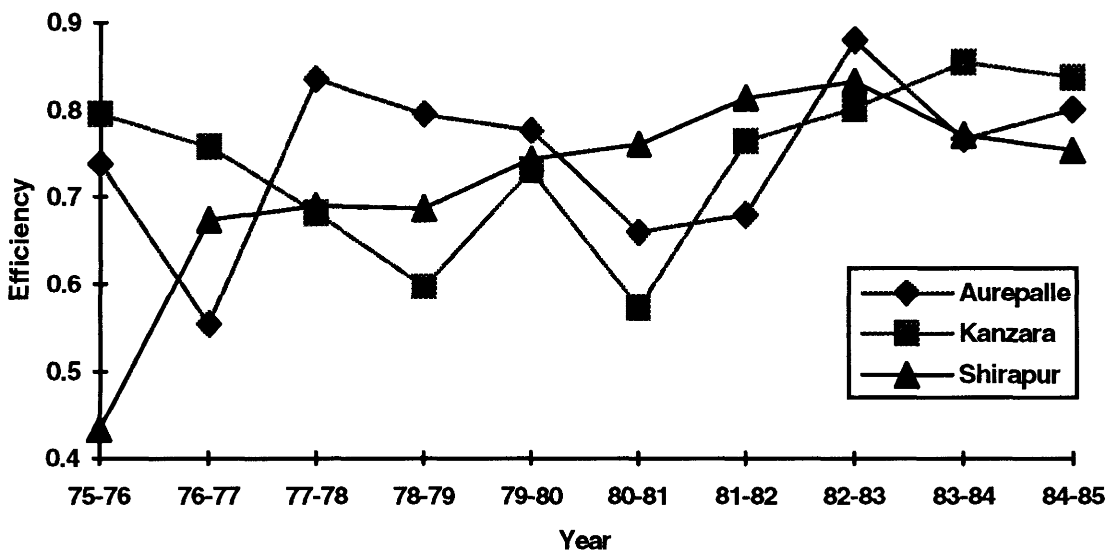
Frequency Distribution of Predicted Technical Efficiencies of Farmers in Shirapur



[Note that the numbers on the horizontal axis refer to the mid-point of the interval.]

Figure 5.4

Mean Technical Efficiencies of the Three Indian Villages



5.5 Conclusions

Stochastic frontier production functions and inefficiency models are estimated for each of three villages from diverse agro-climatic regions of the semi-arid tropics of India. The production frontiers involve the inputs of land, labour, bullock labour and cost of other inputs. The ratios of irrigated land to total land and hired labour to total labour are included in the functions to permit the productivity's of irrigated versus unirrigated land and hired versus family labour to differ. A time trend is used to proxy the influence of technical change. All estimates have the expected signs, with the exception of the coefficients of the ratio variables in the case of Shirapur and the coefficient of year of observation in the case of Kanzara. The results for Shirapur may be a consequence of there being no important labour-intensive irrigated crop grown in that village.

The model for the technical inefficiency effects in the production frontier includes the age and years of formal schooling of the farmer, size of the farm and the year of observation as explanatory variables. A number of tests of hypotheses are conducted to assess the relative influence of these factors and other random effects. The results indicate a significant random component in the technical inefficiency effects in all three villages and that the above four factors have a significant influence upon the size of the technical inefficiencies of farmers in Kanzara and Shirapur, but not in Aurepalle. Farm size and year of observation are estimated to be inversely related to the level of technical inefficiency in all villages. In two of the three villages, the effects of age and education of the farmers are found to be negatively related to the level of the technical inefficiency effects.

The technical efficiencies of each farmer, in each year that the farmer was surveyed, are predicted and tabulated. Technical efficiencies are observed to range from below 0.1 to above 0.9. The mean technical efficiencies for the three villages are estimated to be 0.747, 0.738 and 0.711 for Aurepalle, Kanzara and Shirapur, respectively. The mean level of technical efficiency follows an upward trend over the ten-year period in all three villages. The lowest annual mean technical efficiency was 0.434 in Shirapur during 1975-76 and the highest was 0.880 in Aurepalle during 1982-83.

The analyses reported in this chapter indicate that there are significant differences in the behaviour of value of output and technical inefficiencies of production in the different regions from which data were obtained in ICRISAT's Village Level Studies. Although this empirical study does not include discussion of various variables which might be important in modelling output and inefficiency effects, e.g., rainfall data, use of agricultural extension services and access to credit, it indicates the potential for more refined analysis, if such data were readily available. It is evident, that in order to be able to draw conclusions of significance for policy purposes, future studies need to be devised to obtain extensive data sets on relevant variables for production frontiers and models for technical inefficiency effects which are consistent with such policy orientations.

Chapter 6

Measurement and Sources of Technical Inefficiency in Australian Coal-Fired Electricity Generation

6.1 Introduction

The generation, transmission and distribution of electricity in Australia has traditionally been a Government enterprise. The distribution of electricity within each of the six States is conducted by State and/or Local Government Authorities, while the generation and transmission of electricity within each State is the sole domain of the State Governments.[1] This study is concerned with the measurement of the technical efficiency of coal-fired electricity generation in Australia. Over 75 per cent of electricity generated in Australia each year is produced by coal-fired power stations. Australia is fortunate to have large reserves of black and brown coal, which are inexpensive to extract, by world standards. The remainder of Australia's electricity is produced by hydro and gas-fired plants, with some smaller oil-fired plants used in remote districts.

Each State has traditionally been self-sufficient in its electricity needs, with interstate trade in electricity being quite rare until recent years. Transmission lines presently link the three States of New South Wales (NSW), Victoria and South Australia (SA) together, and the interstate transfer of electricity has steadily increased over the past few years. Proposals are presently being considered for the construction of transmission lines to permit the inclusion of two more States, namely Queensland and Tasmania, into this *inter-connected grid.*[2] At present, interstate sales of electricity, via the inter-connected grid, are being used primarily to allow the States to stagger the

---

[1]The one significant exception to this is the Snowy Mountains Hydro-Electric Scheme, which supplies some of the electricity consumed in the States of New South Wales and Victoria (and all of that consumed in the Australian Capital Territory).

[2]This would leave Western Australia (WA) as the only State without interstate trading opportunities. The inclusion of WA, however, is unlikely, given the vast distances between the population centres in WA and the remainder of Australia.

construction of new power plants, so as to reduce the amount of costly excess-generating capacity in each State.

The Federal and State Governments plan to use this inter-connected grid, not only for the above purpose, but also to introduce competition within the Australian electricity industry. They plan to establish a competitive national market in electricity. Part of this process will involve the separation and corporatisation of the generation, transmission and distribution divisions in each State, and, in certain States, the planned privatisation of some, and perhaps all, generating plants and distribution authorities. This process is at present well under way in Victoria, where one generating plant and a number of distribution authorities have been sold to the private sector.

The overall effect of the above changes will be to introduce competition to Australia's electricity generation industry for the first time in recent history. It is envisaged that the generation divisions of each State will be competing against each other, and that plants, or groups of plants, within a State will eventually be competing with each other in some cases. Given the prospect of this exposure to market forces, information on the relative efficiencies of each plant, both relative to other plants within a State, and relative to plants in other States, are of particular interest to the managers of the generation divisions within each State.

Econometric analyses of production and/or relative efficiency in the Australian electricity industry are few and far between. Only a few studies consider power plants within a particular State. Bateson and Swan (1989) estimate a cost function for power plants in NSW to measure scale economies and also to investigate the influence of capacity factor upon unit costs. Price et al. (1992) investigate the comparative productivity of NSW power plants using multilateral index numbers. There have also been a few comparisons of the relative productivity of the different State electricity utilities using the same methodology (e.g., Lawrence, Swan and Zeitsch, 1990). We were not able to identify, however, any plant-level analyses involving data from two or more States. It is this void in the literature which the present analysis hopes to fill.

The initial plan for this study involved the collection of physical and cost data on output and inputs for all major coal-fired power plants in Australia over a ten-year period. These data were to be used to estimate a variety of stochastic frontier production and cost functions, in order to investigate the structure of the technology and to estimate technical, allocative and overall economic efficiencies for each plant. The study proposal was readily accepted and funded by the Australian Electricity Supply Industry Research Board (AESIRB), which is the research arm of the Electricity Supply Association of Australia (ESAA). The collection of data, however, proved a difficult task. A number of State Authorities deemed some or all of the requested data too sensitive to release. This was most likely because of uncertainty regarding the final form of the competitive model planned for the Australian electricity industry. Thus, after three years of attempting to obtain these data, the present study is limited to an analysis of physical data only, from only three of the five States which have major coal-fired power plants.[3] Thus the present analysis involves the estimation of stochastic frontier production functions and the prediction of technical efficiencies from these estimated functions. The lack of cost data precludes the estimation of cost frontiers and the prediction of allocative or overall economic efficiencies.

This chapter is divided into five sections. Section 6.2 provides a brief review of literature on past analyses of electricity generation, involving both non-frontier and frontier methodologies. In Section 6.3 the data and model specification used in this study are detailed. Empirical results are presented and discussed in Section 6.4, and some brief concluding comments are made in the final section.

6.2 Literature

Many past analyses of efficiency in electricity generation involve the calculation of simple ratio measures, such as fuel efficiency (the ratio of power generated to the energy content of the fuel consumed) or labour productivity (power generated per employee). These measures can be very informative but can also be quite misleading because they consider only a single input in isolation. In this chapter we use the model outlined in Chapter 4 to obtain a measure of the relative efficiency of power stations which accounts for all the factors of production simultaneously.

[3]One State utility did provide all physical and cost data requested. Unfortunately the lack of cost data from other States meant that these data were not able to be utilised. It should also be noted that Tasmania does not have any large coal-fired plants.

We now provide a brief review of past studies which use econometric methods to model electric-power generation. This review begins with a discussion of analyses which use non-frontier econometric models, and progresses to describe some more recent studies which model power generation using frontier methods.

6.2.1 Non-frontier Analyses

This brief discussion deals almost entirely with analyses of electricity generation in the US and Europe. This is a consequence of the lack of Australian analyses, due primarily to the secular nature of the Australian electricity industry. A number of econometric analyses of electricity generation have used non-frontier econometric methods to investigate the structure of the production technology. Their primary interest is generally to investigate input-substitution possibilities, scale economies and technical change in electricity generation.

These studies may be divided into those which estimate the parameters of the production technology directly (e.g., Komiya, 1962 and Courville, 1974) and those which assume some form of behavioural assumption, such as cost minimisation or profit maximisation, and estimate a cost function, profit function, derived demand functions, or some combination thereof (e.g., Nerlove, 1963 and Christensen and Greene, 1976). An excellent survey of econometric analyses of electricity generation is provided by Cowing and Smith (1978) and hence we do not attempt that task here.

Cost minimisation appears to be the assumption most often made in econometric analyses of electricity. This is evident in the review paper by Cowing and Smith (1978), and is especially evident in the vast number of cost-function studies which have been published since 1978, such as the US studies by Stewart (1979), Gollop and Roberts (1983), and Atkinson and Halvorsen (1984), and the analyses of Bateson and Swan (1989) and Nemoto, Nakanishi and Madono (1993), involving the Australian and Japanese electricity industries, respectively. The popularity of this behavioural assumption is not surprising, given that a plant will normally have little say in what quantity of output it produces, and that electricity industries are generally highly regulated, to the extent that many are wholly government owned, as is the case in Australia. Given a behavioural assumption, such as cost minimisation , then the direct estimation of a production function suffers from simultaneous-equations bias, due to the endogeneity of the input levels. It is disappointing, therefore, that a lack of cost data prevents the estimation of the production technology from a cost perspective in this study. We therefore estimate the parameters of the production technology directly using a production function, and hope that the impact of any bias is not significant.[4]

6.2.2 Frontier Analyses

A number of studies apply frontier methodologies to a variety of electricity industries around the world. These studies involve the estimation of both production and cost functions, using both DEA and stochastic frontier approaches. Again, the vast majority of these studies are US applications. The following survey provides an indication of the breadth of analyses that have been conducted.

One of the earliest applications of frontier methods to electricity generation is an analysis of 181 steam-electric plants by Seitz (1971), which involved the estimation of a frontier production function, using linear programming, and the calculation of technical, allocative and overall efficiency measures. A second-stage regression of the technical efficiency measures upon a number of firm-specific factors (including number of units and unit size) was conducted and found evidence of significant relationships.

Several papers, written in the late 1970s, use data from the electricity industry to illustrate advances in stochastic frontier methodologies. These include analyses by Schmidt and Lovell (1979, 1980) involving 111 US steam-electric plants to illustrate extensions of the stochastic frontier model to allow for allocative inefficiency, and the studies of Stevenson (1980) and Greene (1980b) which use US data to investigate more general distributions for the inefficiency effects and more flexible functional forms, respectively.

Kopp and Smith (1980) estimate stochastic frontier production functions for 43 US coal-fired electric power plants. They consider three alternative functional forms, three estimation methods, and divide their data into two capital-vintage groups, finding that all three factors have an influence upon the measures of mean technical efficiency obtained.

---

[4] It should be noted that even if cost data were available, an estimated cost function may suffer from specification error, due to government regulations resulting in other than cost-minimising behaviour.

A number of papers apply DEA methods to US electricity data during the 1980s. The Färe, Grosskopf and Logan (1983, 1984) studies consider the efficiency of Illinois electric utilities and the relative efficiency of public- and privately-owned utilities, respectively. The analyses by Färe, Grosskopf and Pasurka (1986, 1989) consider the effects of environmental regulation upon relative efficiency.

The two themes of the influence of ownership and pollution controls upon efficiency are also prevalent in the more recent literature. Bernstein, Feldman and Schinnar (1990) use DEA methods to investigate the effects of pollution controls in US plants, while Hausman and Neufeld (1991) use DEA to investigate the influence of ownership in the US industry. Hammond (1992) uses a stochastic frontier cost function to look at the same issue in the UK. The results of these last two studies, and those of Färe, Grosskopf and Logan (1984) are of particular interest to policy makers in Australia at present. Hausman and Neufeld (1991) and Färe, Grosskopf and Logan (1984) find public plants more efficient than privately owned plants (with the difference in the latter study not being significant), while Hammond (1992) finds the converse to be true in the UK. Thus, the evidence is not conclusive in either direction at this stage.

6.3 Data and Model Specification

6.3.1 Data

The sample data used in this study comprise annual measures of output, inputs and a variety of other variables, from each of 13 Australian coal-fired power plants. Of these 13 plants, six are from NSW, four are from Victoria and three are from Western Australia (WA). The NSW data were obtained from annual reports[5] while the data for the other two States were obtained from a combination of annual reports and the direct assistance of employees of the respective State electricity commissions of Western Australia (SECWA) and Victoria (SECV).[6] Annual data were obtained for each financial year from 1981-82 to 1990-91, with some exceptions. Some observations were not available in certain years because a few plants did not begin operating until after 1981-82, and, in some cases, because output and labour figures were missing

[5] The NSW data were collected by Michael Plumb from Sydney University, and were kindly made available for use in this study.

[6] Thanks are due to Joy Johnson and Nenad Ninkov of SECWA and Michael Freeman from SECV.

from the data provided. Because of these omissions, the final set of data involved 114 observations.

A number of choices needed to be made when deciding upon the exact definition of each measure to be used. When considering an output measure, a choice had to be made between electricity *generated* and electricity *sent out*. Engineers generally prefer the first measure, but from an economic point of view, the amount of power sent out is the measure of useful output. Hence the amount of electricity sent out is used as the measure of output in this study.[7] The capital measure used is the name-plate capacity (in mw). This is not the preferred choice. A measure of the overall capital investment, adjusted for depreciation and embodied technical change, would have been preferred, but unfortunately these data were not always available.[8] The labour input used is simply the number of employees. A labour measure which accounts for differences in hours worked and degree of training would have been preferred, but such a measure was not possible because of data limitations.

The selection of a measure of the fuel input was the most difficult decision of all. The most obvious choice was tonnes of coal burned, given that only coal-fired power plants are considered. However, this measure suffers from a number of problems. The main problem is that coal quality varies substantially from one State to another, and also, from one coal mine to another within a State. The most visible difference in coal quality is between the low quality brown coal used in Victorian plants and the higher quality black coals used in other States. The average energy contents, in megajoules per kilogram, vary from 7.4 for brown coal at Yallourn in Victoria to 26.4 for black coal at Munmorah in NSW, which is different by a factor of over 350 per cent. The fuel issue is further complicated by the additional use of fuel oil, briquettes and natural gas in varying quantities in different plants, generally to assist with re-starting a unit after a period of down-time. It was thus decided that the best way to avoid these problems was to convert all fuels into terrajoule equivalents and to aggregate the resulting figures. This approach, however, is not without problems. The main

[7]This measure of annual output in kwh does not account for the required distribution of production through each day, nor throughout the year. The formulation of output measures to account for this deficiency is beyond the scope of this study. For further discussion of this issue, see Cowing and Smith (1978).

[8]See Swan (1990) for an example of the type of capital measure that can be constructed when the necessary data are available.

difficulty is that one terrajoule of brown coal is not the same input as one terrajoule of black coal, because of the extra volume that must be handled and burned in the case of the lower quality brown coal. For further discussion of these and other problems associated with using aggregative measures of fuel energy, see Berndt (1978).

The 13 plants in the sample represent approximately 50 per cent of Australia's total generating capacity. The sample means, standard deviations, minima and maxima of the variables used in this study are presented in Table 6.1.[9] The average capacity of plants in the sample is 1270.5 megawatts (mw) and the average unit size is 345.18 mw, indicating an average of three to four generating units per plant. The largest plant has a total of 2640 mw of generating capacity, comprising four 660 mw units, and the smallest plant has 120 mw, involving four 30 mw units. The ages of the plants vary significantly from newly commissioned plants to one that was 32 years old in 1990-91. The average age of the plants was 12.34 years (over the ten years considered). Coal is the most important component of fuel in each plant. Oil is used in small quantities for unit re-starts in all plants, with the exception of two Victorian plants which use briquettes instead. The gas figure in Table 6.1 is due, almost entirely, to the Kwinana plant which had two of its units converted to allow either gas or coal to be burned.

6.3.2 Model Specification

A translog stochastic frontier production function is specified for the Australian electricity generation industry. The output of a plant is assumed to be a function of the three inputs of capital, labour and fuel; technical change is permitted to be non-neutral; and the stochastic frontier is assumed to have the properties of the model specified in Battese and Coelli (1993, 1995) and discussed in Chapter 4. That is, the stochastic frontier production function is assumed to be described by:

$$\begin{aligned}\log(Q_{it}) = {} & \beta_0 + \beta_1\log(K_{it}) + \beta_2\log(L_{it}) + \beta_3\log(F_{it}) + \beta_4[\log(K_{it})]^2 + \beta_5[\log(L_{it})]^2 \\ & + \beta_6[\log(F_{it})]^2 + \beta_7\log(K_{it})\log(L_{it}) + \beta_8\log(K_{it})\log(F_{it}) + \beta_9\log(L_{it})\log(F_{it}) + \\ & \beta_{10}\log(K_{it})t + \beta_{11}\log(L_{it})t + \beta_{12}\log(F_{it})t + \beta_{13}t + \beta_{14}t^2 + V_{it}-U_{it}, \\ & i=1,2,\ldots,N;\ t=1,2,\ldots,T, \qquad (6.1)\end{aligned}$$

[9] A full listing of the data is provided in Appendix 3.

Table 6.1

Summary Statistics for Observations on 13 Coal-fired Electricity Generating Plants in Australia during 1981-82 to 1990-91

Variable	Sample Mean	Standard Deviation	Minimum Value	Maximum Value
Output (gwh)	5502.6	3744.9	65.1	15406.0
Capacity (mw)	1270.5	720.6	120.0	2640.0
Labour (persons)	639.79	329.39	41.00	1787.00
Fuel (terrajoules)	62925.0	39519.0	1282.0	156370.0
Capacity factor (%)	46.31	16.47	6.19	85.16
Age of units (average in yrs)	12.34	8.48	0.00	32.00
Size of units (average in mw)	345.18	227.65	30.00	660.00
Coal (1,000t)	4342.8	4427.7	60.6	18030.0
Gas (terrajoules)	1636.6	6977.0	0.0	40610.0
Oil (1,000t)	8.76	13.25	0.00	90.30
Briquettes (1,000t)	13.54	37.76	0.00	216.30

where Q_{it} represents the electricity sent out (in kwh) by the i-th plant in the t-th year;

K_{it} represents the capacity of the plant (in mw);

L_{it} represents labour (number of employees);

F_{it} represents fuel usage (in terrajoules);

t is a time trend;

the β_j are unknown parameters to be estimated;

the V_{it}s are iid $N(0,\sigma_v^2)$ random errors, which are assumed to be independently distributed of the U_{it}s;

the U_{it}s are non-negative random variables associated with technical inefficiency, which are assumed to be independently distributed, such that the distribution of U_{it} is obtained by truncation at zero of the normal distribution with mean, μ_{it}, and variance, σ^2, where μ_{it} is defined by

$$\mu_{it} = \delta_0 + \delta_1 C_{it} + \delta_2 A_{it} + \delta_3 S_{it} + \delta_4 B_{it} \tag{6.2}$$

and C_{it} represents capacity factor;[10]

A_{it} represents the average age of installed units;

S_{it} represents the average size of installed units (in mw);

B_{it} is a dummy variable which takes the value 1 when brown coal is used and 0 when black coal is used; and

the δ_j are unknown parameters to be estimated.[11]

As discussed in Chapter 4, we replace σ_v^2 and σ^2 with $\sigma_s^2=\sigma^2+\sigma_v^2$ and $\gamma=\sigma^2/(\sigma_v^2+\sigma^2)$. This transformation has advantages in the estimation process, where γ can be searched between zero and one to obtain a suitable starting value for an iterative maximisation process. Values of the 22 unknown parameters in the above stochastic frontier and inefficiency model are simultaneously estimated by the method of maximum likelihood.

It should be noted, that the four firm-specific factors included in the analysis are not the only factors which could possibly influence the degree of technical inefficiency of plants. A variety of management factors, such as the experience of managers and the degree of bureaucratic and/or union constraints upon management could also be expected to have an influence upon the technical efficiency of a plant. Data on such variables have not yet been collected.

The above model specification permits certain firm-specific factors to shift the mean of the technical inefficiency effects. It is possible that the firm-specific factors considered in the study may not have a significant influence upon the degree of technical inefficiency of electricity generating plants. This hypothesis, along with a number of

---

[10] Capacity factor, in this study, is defined as the ratio of the actual power sent out to the amount of power that theoretically could be sent out if all units produced to their name-plate ratings for 100% of the time with no down-time and no wastage.

[11] Note that a time trend was included in the inefficiency model in an earlier analysis. This was included to account for the possibility of technical efficiency change through time resulting from influences which had not been already accounted for in the model. This trend variable was found to be insignificant and hence has been omitted from the models discussed in this Chapter.

other hypotheses relating to the inefficiency effects, are tested in the following section. Furthermore, we recall from Chapter 2, that the translog functional form, assumed in the above specification, is a more general representation of the production structure than is often assumed in empirical analyses of production, where simpler forms, such as the Cobb-Douglas, have been more prevalent. The translog form permits more general substitution, scale and technical change possibilities than simpler forms, such as the Cobb-Douglas, but at the expense of needing to estimate substantially more parameters. If the production technology is suitably represented by a simpler form, then the estimation of unnecessary parameters will result in inefficient estimates. Hence, a number of hypotheses, regarding restrictions upon this functional form are considered in the following section.

6.4 Results and Discussion

6.4.1 Maximum-likelihood Estimates

The maximum-likelihood estimates of the parameters of the translog stochastic frontier and inefficiency model, defined by equations (6.1) and (6.2), are obtained using the computer program, FRONTIER Version 4.1, discussed in detail in Chapter 7. These estimates are presented in the first column of Table 6.2. Asymptotic standard errors are presented in parentheses below each estimate. The ratios of the estimated coefficients to their corresponding standard errors (t-ratios) provide an indication of the statistical significance of the coefficients. Only three of the estimates of the coefficients associated with the production inputs and technical change have t-ratios larger than 1.96 in absolute value, suggesting that very few of them are significantly different from zero at the five per cent level. Furthermore, only eight of the β-coefficients have t-ratios larger than one in absolute value. This may suggest that the model is a fairly poor fit. The consideration of these t-ratios, however, can be misleading on two counts. First, the sizes of these tests will not be equal to five percent when more than one test is conducted in sequence; and the second, and probably the most important point, is that multicollinearity, resulting from the inclusion of second order terms, may be contributing to the high standard errors observed. If

Table 6.2

Maximum-likelihood Estimates of the Stochastic Frontier and Inefficiency Model for Electricity Generation in Australia

Coefficient	Parameter	Translog	Translog (neutral technical change)	Cobb-Douglas
Stochastic Frontier				
Intercept	β_0	-1.17	-1.70	-3.16
		(0.98)	(0.73)	(0.15)
log(Capital)	β_1	-0.71	-0.80	0.112
		(0.29)	(0.25)	(0.041)
log(Labour)	β_2	-0.46	-0.41	0.017
		(0.26)	(0.22)	(0.027)
log(Fuel)	β_3	1.62	1.59	0.982
		(0.23)	(0.21)	(0.037)
$[\log(\text{Capital})]^2$	β_4	-0.102	-0.058	
		(0.058)	(0.052)	
$[\log(\text{Labour})]^2$	β_5	0.010	0.065	
		(0.056)	(0.051)	
$[\log(\text{Fuel})]^2$	β_6	-0.018	0.028	
		(0.068)	(0.059)	
log(Capital)log(Labour)	β_7	0.25	0.26	
		(0.11)	(0.10)	
log(Capital)log(Fuel)	β_8	0.06	0.01	
		(0.13)	(0.12)	
log(Labour)log(Fuel)	β_9	-0.13	-0.210	
		(0.13)	(0.099)	
log(Capital)t	β_{10}	0.0082		
		(0.0069)		
log(Labour)t	β_{11}	-0.0092		
		(0.0078)		
log(Fuel)t	β_{12}	0.0032		
		(0.0067)		
t	β_{13}	-0.027	0.004	0.0068
		(0.035)	(0.011)	(0.0029)
t^2	β_{14}	-0.00007	0.00029	
		(0.00096)	(0.00088)	
Variance Parameters				
	σ^2	0.00234	0.00300	0.0049
		(0.00042)	(0.00038)	(0.0015)
	γ	0.162	0.234	0.40
		(0.064)	(0.067)	(0.22)
Inefficiency Model				
Intercept	δ_0	0.14	-0.009	-0.33
		(0.13)	(0.071)	(0.15)
Capacity factor	δ_1	-0.00593	-0.00567	-0.0036
		(0.00022)	(0.00077)	(0.0015)
Age of units	δ_2	0.0100	0.0146	0.019
		(0.0018)	(0.0024)	(0.0044)
Size of units	δ_3	0.000106	0.000174	0.00038
		(0.000085)	(0.000029)	(0.00024)
Brown coal	δ_4	0.114	0.189	0.304
		(0.023)	(0.039)	(0.060)
Log-likelihood Function		**185.16**	**182.01**	**159.50**

this is the case, the consideration of these individual t-ratios may lead to the omission of some important coefficients, resulting in misspecification of the model.

6.4.2 Tests of Hypotheses

A more appropriate testing procedure is to simultaneously test the significance of groups of coefficients. In this study the generalised likelihood-ratio test is used. This involves the calculation of

$$\lambda=-2\{\log[L(H_0)]-\log[L(H_1)]\}, \tag{6.3}$$

where $L(H_0)$ and $L(H_1)$ are the values of the likelihood function under the null and alternative hypotheses, respectively. This λ-statistic has asymptotic chi-square distribution, with degrees of freedom equal to the difference between the number of parameters involved in H_0 and H_1.[12]

The first hypothesis test considered is a test of the null hypothesis of Hicks-neutral technical change. Technical change is Hicks-neutral if the coefficients of the interactions between the logarithms of the inputs and the time trend are all zero. This is the first null hypothesis considered in Table 6.3, i.e., H_0: $\beta_{10}=\beta_{11}=\beta_{12}=0$. The maximum-likelihood estimates of the parameters of the model defined by equations (6.1) and (6.2), with these restrictions imposed, are listed in the second column of Table 6.2. As can be seen from the results in Table 6.3, the value of the likelihood-ratio statistic is calculated to be 6.30, which is less than 7.81 (the 5% critical value from the χ_3^2-distribution). Hence the null hypothesis of Hicks-neutral technical change is not rejected, implying that technical change has not favoured the use of one particular input over another in this industry.

The second null hypothesis considered in Table 6.3 specifies that there has not been any technical change over the sample period. This test involves a test of the restrictions that all the coefficients associated with the time trend are equal to zero, i.e., testing H_0: $\beta_{10}=\beta_{11}=...=\beta_{14}=0$. The results for this test, listed in Table 6.3, show a

[12] As noted in Chapter 3, this statistic has a distribution which is a mixture of chi-square distributions when the null hypothesis specifies that the γ-parameter is zero.

Table 6.3

Tests of Hypotheses of Parameters of the Stochastic Frontier and Inefficiency Model for Electricity Generation in Australia

Null Hypothesis	Log-likelihood Value	Test Statistic λ	Critical Value	Decision
H_0: $\beta_{10}=\beta_{11}=\beta_{12}=0$	182.01	6.30	7.81	Accept H_0
H_0: $\beta_{10}=\beta_{11}=...=\beta_{14}=0$	177.27	15.78	11.07	Reject H_0
H_0: $\beta_4=\beta_5=...=\beta_{14}=0$	159.50	51.32	18.31	Reject H_0
H_0: $\gamma=\delta_0=\delta_1=...=\delta_4=0$	154.34	61.64	5.14-11.91	Reject H_0
H_0: $\delta_1=\delta_2=\delta_3=\delta_4=0$	155.97	58.38	9.49	Reject H_0

test statistic of 15.78 which exceeds the χ_5^2-critical value of 11.07, resulting in a rejection of the null hypothesis of no technical change.[13]

The third null hypothesis that is considered in Table 6.3 is that the Cobb-Douglas production frontier with neutral technical change is an adequate representation of the data. This null hypothesis is specified by H_0: $\beta_4=\beta_5=...=\beta_{12}=\beta_{14}=0$. The maximum-likelihood estimates of this Cobb-Douglas model are listed in the last column of Table 6.2. The value of the log-likelihood function has reduced substantially to 159.50. This provides a generalised likelihood-ratio test statistic of 51.32 which exceeds the χ_{10}^2-critical value of 18.31 by a large amount. Thus we confidently reject the Cobb-Douglas form, given the specification of the translog frontier model. It therefore appears that the extra effort involved in estimating and analysing the translog form is warranted in this instance. Furthermore, we note that the rejection of the translog functional form because of the small t-ratios associated with the individual β-coefficients would have involved poor statistical inference.

We now turn our attention to the estimates of the coefficients associated with the technical inefficiency effects of the model specification. We note that, in column 1 of

---

[13] This hypothesis test, and all other hypothesis tests listed in Table 6.3 are conducted with the unrestricted translog as the model under the null hypothesis. Since we have seen that the hypothesis of Hicks-neutral technical change is not rejected, it could be argued that the translog with Hicks-neutral technical change should be used as the model for subsequent hypothesis tests. These tests were also conducted, but have not been reported because none of the conclusions differ from those in Table 6.3.

Table 6.2, the estimated coefficients associated with the γ- and the δ-parameters (except δ_0) are much larger than their corresponding standard errors. Thus the indications are that these terms are significant additions to the model. However, even though multicollinearity is unlikely to have as large an influence upon these coefficients, as was the case with the β-coefficients, we conduct two generalised likelihood-ratio tests to confirm our observations.

First, we consider a test of the null hypothesis that the technical inefficiency effects are absent in this industry. The omission of U_{it} is equivalent to imposing the restrictions specified in the null hypothesis, H_0: $\gamma=\delta_0=\delta_1=...=\delta_4=0$. The relevant information for the test of this hypothesis test are in Table 6.3. The generalised likelihood-ratio statistic is calculated to be 61.64, which is substantially larger than the critical value range of 5.14 to 11.91. Thus we reject the null hypothesis of no technical inefficiency effects, given the specifications of the stochastic frontier and inefficiency model.

One question of particular interest to this study, is whether the four firm-specific factors, considered in the inefficiency model, have a significant influence upon the degree of technical inefficiency associated with the plants. Thus a test of the null hypothesis that H_0: $\delta_1=\delta_2=\delta_3=\delta_4=0$ is conducted. The results of this test are listed in the final row of Table 6.3. The generalised likelihood-ratio statistic is calculated to be 58.38, which is much larger than the χ_4^2-critical value of 9.49. Hence the null hypothesis that these four factors do not have an influence upon technical inefficiency is also rejected in this case.

6.4.3 Economic Plausibility of the Results

Production Structure

When the conclusions of the above five tests of hypotheses are considered together, the preferred model appears to be that, defined by equations (6.1) and (6.2), with Hicks-neutral technical change imposed. The estimates of the parameters of this model are given in the second column of Table 6.2. Due to the complexity of the translog form, the economic plausibility of the estimated coefficients is not easy to assess without first calculating some more easily interpreted estimates. Thus, the estimated values of the production elasticities of the three inputs, evaluated at the sample means,

Table 6.4

Key Estimates Derived From the Translog Frontier and Inefficiency Model With Hicks-Neutral Technical Change

Description	Estimate
Capital elasticity	0.170
	(0.038)
Labour elasticity	-0.022
	(0.034)
Fuel elasticity	0.932
	(0.040)
Returns-to-scale parameter	1.080
	(0.022)
Technical change	0.0074
	(0.0026)

are given in Table 6.4 (above). Also given in Table 6.4 are the estimates for the returns-to-scale parameter and the annual percentage change in production due to technical change. Approximate standard errors of these estimates are listed in parentheses under each estimate.

The estimated elasticities have the expected positive signs, except for labour, but the estimate is not significantly different from zero at the 20% level using an asymptotic t-test. The production elasticity for capital is 0.170 and that for fuel is 0.932. These results are not unlike those seen in many past analyses of electricity production. For example, in their analysis of 111 privately owned steam-electric generating plants in the US, Schmidt and Lovell (1979) obtained production elasticities which are not significantly different from zero for labour and also obtain a value close to one for the elasticity of fuel. Kopp and Smith (1980) conducted preliminary analyses of 43 private and public coal-fired plants in the US, using the three inputs of capital, labour and fuel, and decided to omit the labour input from the reported analysis completely. Kopp and Smith (1980, p.1053) argued that "capital and fuel appear to be the most important inputs to the production technology" and that labour "appears to bear a direct relationship to the scale of the plant". Thus it appears that the elasticity estimates in Table 6.4 are similar to those obtained in other studies and that, in particular, the insignificant labour elasticity is not unusual in electricity generation.

The estimated returns-to-scale parameter of 1.080 indicates mildly increasing returns to scale. This value is significantly different from zero, according to an asymptotic t-test. This result again does not conflict with the findings in previous studies. If anything, the scale elasticity obtained here is slightly smaller than those reported in previous studies. For example, Kopp and Smith (1980) report values ranging from 1.142 to 2.131 for a variety of estimation methods and capital vintages. Given that many of these past studies are based upon data a decade or more before the data used in this study (e.g., Kopp and Smith use data from 1969 to 1973), it is not surprising that unexhausted scale economies diminish as the plants being analysed become progressively larger. This observation conforms with the conclusions of Christensen and Greene (1976) who investigate differences in scale economies in US electric power generation between 1955 and 1970 using a translog cost function.

The final estimate listed in Table 6.4 is a measure of technical change. The value of 0.0074, indicates that the industry has experienced a rate of technical progress over the sample period of approximately 0.74 per cent per year. This indicates that a hypothetical plant could produce 7.4 per cent more output in 1991 than could be produced in 1981, using the same levels of inputs. This estimate of technical progress is found to be significantly different from zero using an asymptotic t-test. The statistical significance of the technical change estimate conforms with the earlier finding in this study, of the coefficients associated with the time trend being a significant addition to the model, using a generalised likelihood-ratio test (see the second hypothesis test in Table 6.3).

Technical Inefficiency Effects

The maximum-likelihood estimates of the variance ratio parameter, γ, and the δ-parameters for the preferred model are listed towards the bottom of the second column of Table 6.2. All of these estimates have t-ratios which are larger than 1.96 in absolute value, with the exception of δ_0. These significant t-ratios are not surprising, given the conclusions of the likelihood-ratio tests above. The interpretation of the γ-estimate of 0.234 is not as clear in this model specification as it is in the half-normal stochastic frontier (i.e., the model where all the δ_i are zero), where it could be shown to be a simple function of the ratio of the variance of the inefficiency error term to the sum of the variances of the two error terms. For the model specification used in this study, it may be loosely interpreted as an indication of the amount of *unexplained* variation in the technical inefficiency effects, relative to the sum of this value and the variance of the random error, V_{it}.

The signs of the δ-parameters need to be considered carefully. The negative sign of the estimated coefficient of capacity factor indicates that an increase in capacity factor results in a decrease in the value of the technical inefficiency effect and hence an increase in technical efficiency. This conforms with the expectation that a plant which is permitted to utilise more of its capacity is likely to appear to be more technically efficient using the measures defined in this study. The estimated coefficient associated with the age of the generating units at a plant is observed to be positive. Thus, the older plants tend to have greater levels of technical inefficiency relative to the newer plants. This also conforms with what one would expect, given that the capital measure used in this study is simply name-plate capacity, and hence that no allowances have been made for the effects of embodied technical change in this capital measure. The positive sign on the estimate of the coefficient of the size of units in a plant is somewhat surprising. It was expected that technical inefficiency would decrease as the size of the generating units increase, because of labour savings, etc. One possible explanation for this unexpected sign is that the plants with smaller unit sizes are more flexible in their ability to adjust to unexpected demand variation.[14] The positive sign on the estimated coefficient of the dummy variable associated with the use of brown coal is consistent with expectations. The plants which are using this lower quality brown coal must handle larger volumes of coal than the black coal plants. Hence it is not surprising that this contributes to the level of the technical inefficiency effects for these plants.

Technical Efficiencies of Plants

The technical efficiencies of each plant in each year can be predicted from the estimated model. Given the stochastic frontier and inefficiency model defined by equations (6.1) and (6.2), the technical efficiency of production of the i-th plant in the t-th year is defined by

[14] Tom Cowing suggested this interpretation in a personal communication.

$$TE_{it} = \exp(-U_{it}). \quad (6.4)$$

This is predicted using the conditional expectation of $\exp(-U_{it})$, given the value of $E_{it}=V_{it}-U_{it}$. This expression is presented in Appendix 2 for the model involved in Chapter 4.

The technical efficiency predictions for our preferred model, calculated by the FRONTIER program, are listed in Table 6.5. They are also plotted in Figures 6.1a and 6.1b. These range in value from 0.618 for Morwell in 1990-91 to 1.000 for Vales Point B in 1982-83 and 1983-84. The mean of technical efficiencies in this industry is calculated to be 0.925. This suggests that, on average, plants produce 92.5 per cent of the output that could be potentially produced with the same bundle of inputs by a technically efficient plant. This figure is comparable to estimates of mean technical efficiency reported in other studies of electricity generation. For example, Kopp and Smith (1980) report estimated mean technical efficiencies of 0.846 and 0.954 from the estimation of stochastic frontier models for two different capital vintages. The estimates of mean technical efficiency for electricity generation, reported in the literature, tend to be larger, on average, than those reported for many other industries. For example, mean technical efficiencies reported in analyses of agricultural industries are usually in the region of 0.6 to 0.7.[15] The higher levels of mean technical efficiency in electricity generation are most likely a consequence of the size of the plants and hence the resources that they have available to ensure that they are always aware of, and using, the latest advances in technology.

It should be noted, however, that the measures of technical efficiency, reported in this paper, are calculated relative to a frontier that has been estimated using a sample of firms taken from the Australian industry only. If we were to estimate a frontier using data taken from electricity industries from a number of countries, it is conceivable that the mean technical efficiency of these Australian firms may be lower, relative to this "international best-practice frontier".

[15]Refer to Battese (1992) for a survey of applications of frontier production functions to agricultural industries.

Table 6.5

Technical Efficiencies for 13 Power Plants in Australia, 1981-82 to 1990-91

Plant	81-82	82-83	83-84	84-85	85-86	86-87	87-88	88-89	89-90	90-91	MEAN
Loy Yang A			0.804	0.981	0.979	0.984	0.995		0.996	0.995	**0.962**
Hazelwood	0.905	0.845	0.843	0.852	0.826	0.828	0.821		0.777	0.817	**0.835**
Yallourn W	0.988	0.992	0.987	0.984	0.972	0.977	0.957		0.957	0.973	**0.976**
Morwell	0.887	0.911	0.886	0.813	0.772	0.747	0.725		0.646	0.618	**0.778**
Bayswater					0.996	0.996	0.995	0.996	0.997	0.997	**0.996**
Eraring		0.997	0.996	0.995	0.992	0.992	0.993	0.994	0.992	0.994	**0.994**
Liddell		0.989	0.992	0.991	0.948	0.922	0.919	0.888	0.888	0.860	**0.933**
Munmorah		0.990	0.982	0.920	0.937	0.906	0.878	0.863	0.857	0.788	**0.902**
Vales Point B		1.000	1.000	0.997	0.996	0.994	0.993	0.985	0.994	0.991	**0.994**
Wallerang C		0.989	0.983	0.992	0.996	0.955	0.994	0.990	0.989	0.931	**0.98**
Bunbury		0.877	0.849	0.761	0.716	0.710	0.663	0.716	0.706	0.681	**0.742**
Muja	0.996	0.997	0.996	0.996	0.994	0.995	0.993	0.996	0.997	0.997	**0.996**
Kwinana	0.971	0.933	0.947	0.951	0.968	0.979	0.983	0.967	0.957	0.935	**0.959**
MEAN	**0.950**	**0.956**	**0.939**	**0.936**	**0.930**	**0.922**	**0.916**	**0.933**	**0.904**	**0.891**	**0.925**

The means of the technical efficiencies of each plant are presented in the last column of Table 6.5 and are graphed in Figure 6.2. These range in value from 0.742 for Bunbury to 0.996 for Bayswater and Muja. It is interesting to note that the four least efficient plants, appear to also be the oldest and smallest plants in the sample.[16] These plants have also had their capacity factors reduced over the sample period. The addition of extra capacity in the system has resulted in an excess of available capacity because of a lower than expected expansion of demand for electricity over the sample period. Hence a number of the hypothesised factors appear to be contributing to the low mean technical efficiencies of these four plants.

The means of the predicted technical efficiencies in each of the ten years during the sample period, presented in the bottom row of Table 6.5, are reproduced in Figure 6.3. A gradual decline is observed in this plot, with mean technical efficiencies declining by approximately 6 per cent over the sample period. This figure should not be considered in isolation. Recall that a value of 7.4 per cent is estimated for the effect of technical change. Thus the combined effect of these two influences could be a small overall increase in productivity for the average plant. One should also note, that the four plants which have the lowest technical efficiencies and which have also experienced the

[16] This observation can be made when the technical efficiencies listed in Table 6.5 are compared with the data used in estimation which is listed in Appendix 3.

Figure 6.1a

Technical Efficiencies for 13 Power Plants in Australia, 1981-82 to 1990-91

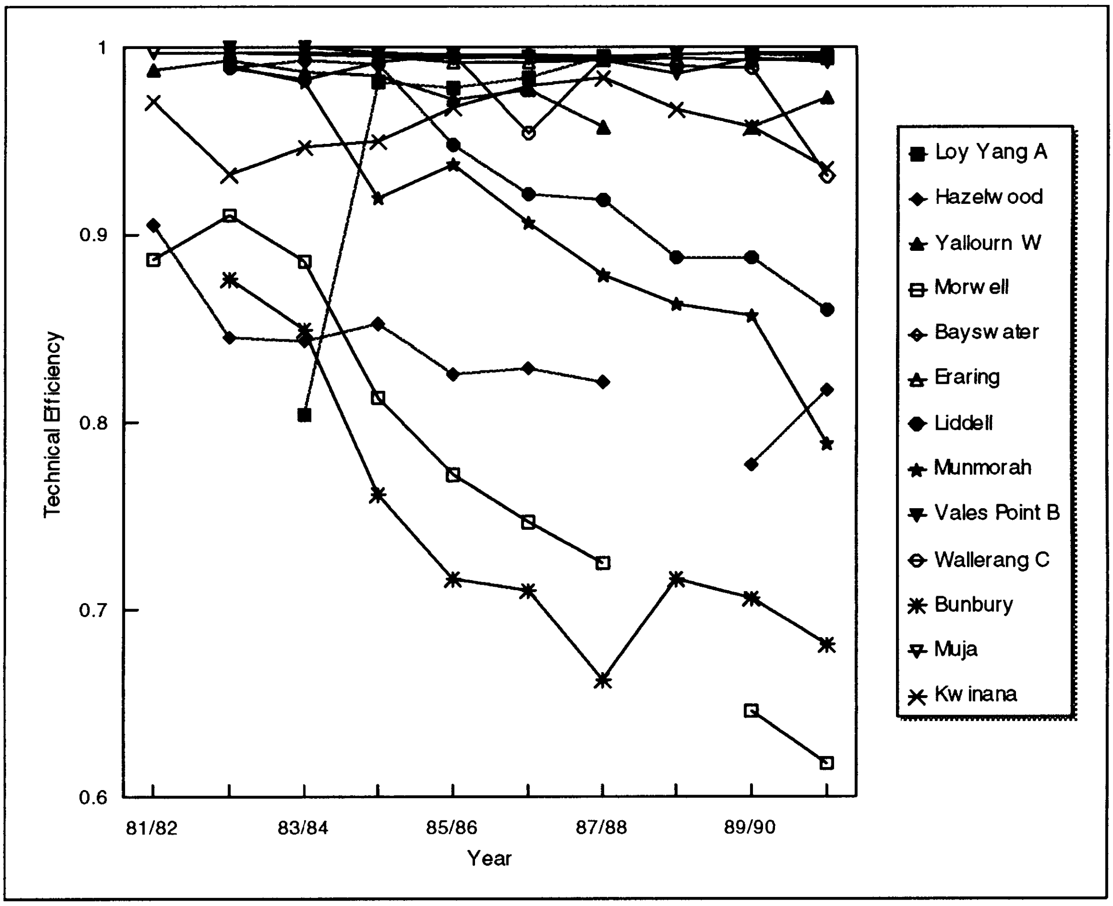


Figure 6.1b

Technical Efficiencies for 13 Power Plants in Australia, 1981-82 to 1990-91

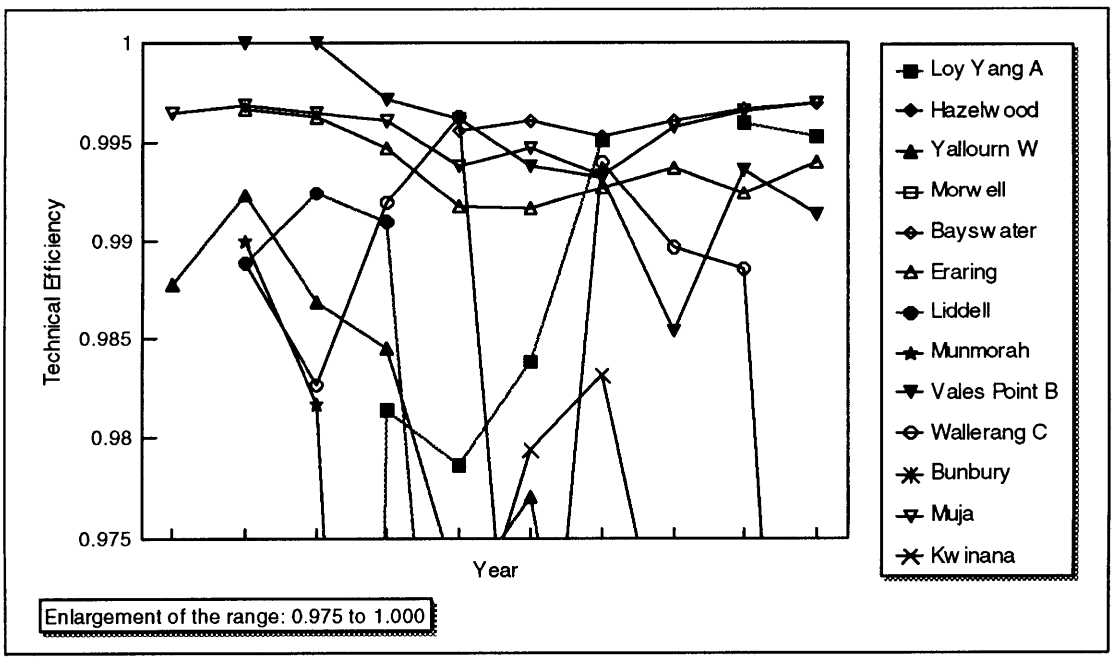
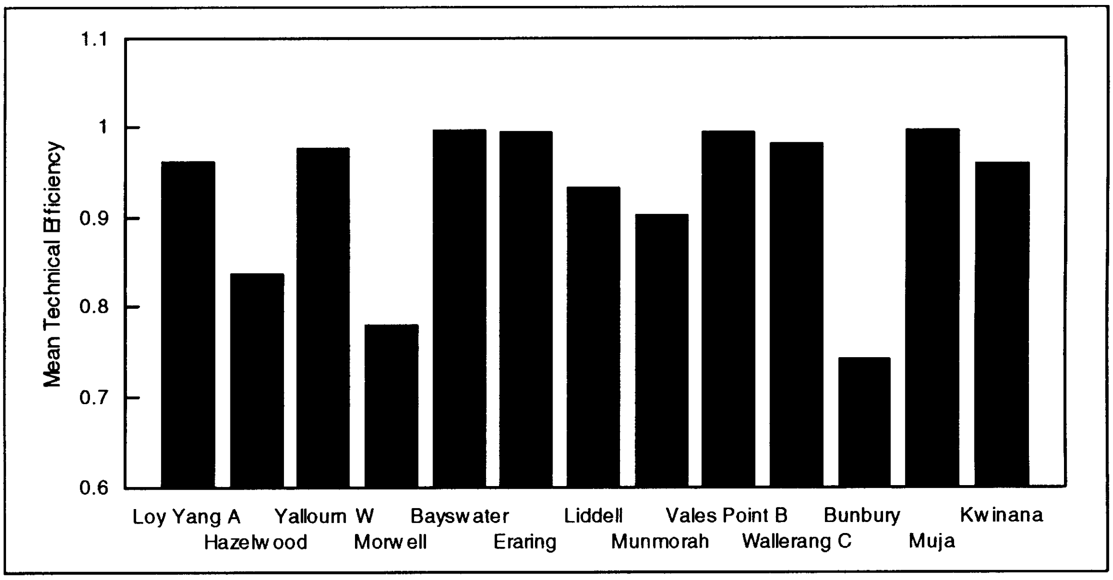


Figure 6.2

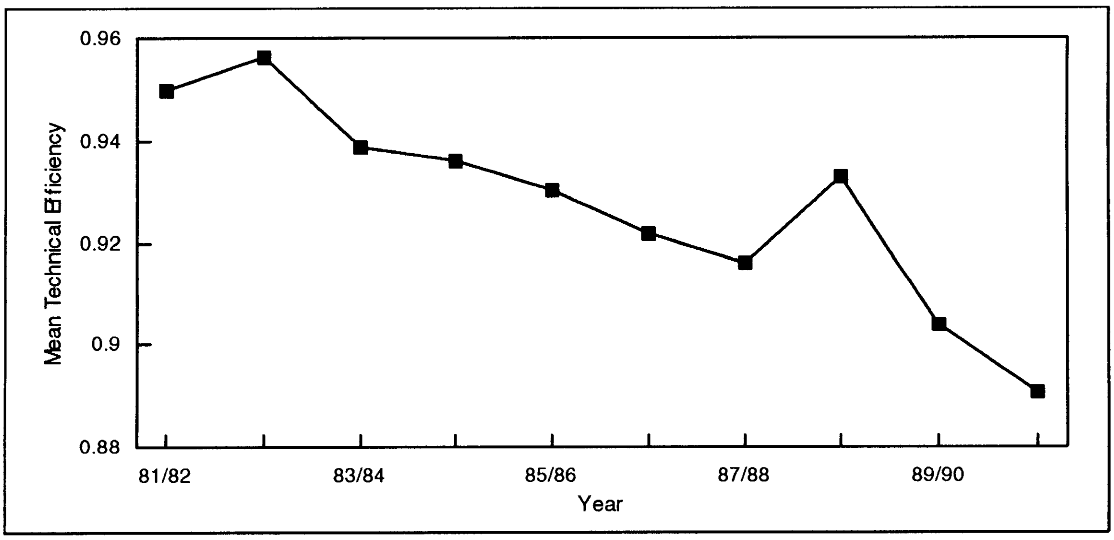
Mean Technical Efficiencies for 13 Power Plants in Australia, 1981-82 to 1990-91*



* Means are for the years 1981-82 to 1990-91, with some observations missing in some years.

Figure 6.3

Annual Mean Technical Efficiencies of 13 Power Plants in Australia, 1981-82 to 1990-91



greatest decline in technical efficiencies (see Figures 6.1a and 6.1b) are the smallest plants in the sample. Hence, if the technical efficiencies were weighted by the scale of the plant, the resulting measure of the average decline in technical efficiency would be much smaller than that indicated by Figure 6.3.

6.4.4 A Comparison with the Two-stage Approach

The 19 parameters of the stochastic frontier and inefficiency model, defined by equations (6.1) and (6.2), with Hicks-neutral technical change imposed, are estimated simultaneously in the above analysis. Given that the majority of past analyses of the determinants of technical inefficiency have estimated the parameters of similar models in two stages, we also estimate the above model in this way for comparative purposes. The first stage involves the ML estimation of the parameters of the stochastic frontier model, defined by equation (6.1), assuming that the U_{it} are independently and identically distributed as truncations at zero of an $N(\delta_0,\sigma^2)$ distribution. The predicted technical efficiencies, $\exp(-U_{it})$, are obtained from this model. The second stage of the estimation process involves the regression of the negative of the logarithms of the predictions of the technical efficiencies from the estimated first-stage model upon the four firm-specific factors.[17] This involves the estimation of

$$-\log(\hat{TE}_{it}) = \alpha_0 + \alpha_1 C_{it} + \alpha_2 A_{it} + \alpha_3 S_{it} + \alpha_4 B_{it} + W_{it} \tag{6.5}$$

where $\hat{TE}_{it}$ is the technical efficiency prediction from the first-stage model and the W_{it} is an error term, such that $-W_{it} < \alpha_0 + \alpha_1 C_{it} + \alpha_2 A_{it} + \alpha_3 S_{it} + \alpha_4 B_{it}$. The range of W_{it} must be limited in this way to ensure that the values of $-\log(\hat{TE}_{it})$ do not become negative and hence that $\hat{TE}_{it}$ does not exceed one. Ordinary least-squares (OLS) estimation of the parameters of equation (6.5) is unlikely to be optimal, because the non-normality and bounded range of the error term, W_{it}, is not taken into account. A more suitable estimation method would involve recognising that this is a limited dependent variable model, and to consider estimation using an approach such as that proposed by Tobin (1958). However, since none of the predicted efficiencies from the

[17] The negative of the logarithms of the technical efficiency prediction, produced by the FRONTIER program, is chosen as the dependent variable in this regression so that the estimated coefficients of the firm-specific factors would have similar interpretations to those in the single-stage model.

first-stage estimation were exactly equal to one, this method could not be implemented. Hence OLS estimation is used in this instance. The OLS estimates of the parameters in the model, defined by equation (6.5), are listed in Table 6.6. The signs of the estimated coefficients are the same (as those obtained in the single-stage estimation) in the case of capacity factor and coal quality, but they differ for age and size. The t-ratios, however are all less than 1.96 in value and a joint test of the significance of the four regressors yields an F-value of 0.882 which also is insignificant. Thus the message from this two-stage procedure is that the four firm-specific factors do not explain any of the variation in technical efficiency. This result is in direct contrast to those obtained from the single-stage procedure discussed earlier. This may be viewed as support for the single-stage procedure, as it is argued that the simultaneous estimation of all parameters is more efficient than the two-stage estimation procedure.

6.5 Conclusions

The primary contribution of this study is a substantial addition to the stock of knowledge regarding the structure of production and relative efficiencies in coal-fired electricity generation plants in Australia. Since there have been no previous econometric analyses of electricity generation using plant-level data from different States in Australia, and also that there have not been any past analyses of relative efficiency, other than those using simple partial measures (such as fuel conversion ratios), this analysis has broken new ground in this industry.

The main conclusions are that, when compared with the translog production frontier, the Cobb-Douglas functional form is not an adequate representation of the production technology in the electricity generation industry in Australia. The industry appears to be characterised by Hicks-neutral technical progress and mildly increasing returns to scale. The mean level of technical efficiency is estimated to be 0.925. The mean technical efficiencies of the plants vary from 0.742 to 0.996, with capacity factor, age and size of generating units in plants, and coal quality found to have a significant influence upon technical inefficiency of generation. The mean level of technical efficiency appears to decline over the sample period. This is likely to be due in part to the increase in excess capacity in the industry during this time.

Table 6.6

OLS Estimates of the Second-stage Regression of Technical Inefficiency Effects

Regressor	Estimate
Constant(α_0)	0.083
	(0.022)
Capacity factor	-0.00050
	(0.00030)
Average age of unit	-0.00064
	(0.00069)
Average size of unit	-0.000016
	(0.000025)
Coal quality	0.007
	(0.022)

This study also makes two secondary contributions. The first is as an illustration of the single-stage estimation of stochastic frontiers which incorporate a model for the technical inefficiency effects which has been used in only a few studies to date. The second contribution is the humble beginnings of a data base containing information on electricity generating plants in Australia. This data base will hopefully grow over time with additional States providing data on their electricity generating plants. It is hoped that cost information can also be made available in the near future. With the planned partial de-regulation of the electricity generating industry, the public reporting of statistics on power plants will be very important for monitoring purposes. Data on the variables we tried to obtain for this study will most likely form part of the minimum reporting requirements.

A number possible extensions to this work could be considered. These include: comparing the results when tonnes of coal is used as the fuel measure; estimating separate functions for brown coal and black coal plants; and analysis involving extra data from other States and/or the inclusion of cost data. Negotiations are also currently under way with the funding body (the AESIRB) regarding a new project which will include data on plants from the US and Europe in the analysis so as to assess the performance of Australian plants relative to international best practice. That study is also likely to involve the use of both stochastic frontier and DEA methods so as to investigate the robustness of results to choice of methodology.