J. Int. Trade & Economic Development 8:4 419–436

Parametric and semi-parametric estimation of the effect of firm attributes on efficiency: the electricity generating industry in India

Madhu Khanna[†], *Kusum Mundra*[‡] *and Aman Ullah*[‡]

[†]University of Illinois, Urbana-Champaign, USA; [‡]University of California, Riverside, USA

Abstract

A stochastic frontier cost function is estimated using panel data for the electricity generating industry in India. The impact of distributional and functional form assumptions on technical inefficiency and the sources of inefficiency are investigated by using maximum likelihood, GLS and semi-parametric-GLS approaches and by incorporating firm-specific inefficiency effects in the cost function itself. Average inefficiency in the electricity generating industry in India is found to be high by all three methods. The estimate predicted by the maximum-likelihood approach is, however, lower than that predicted by the other two methods. This could be due to the distributional assumptions made under the maximum likelihood method. Public ownership and low capacity utilization are found to be significant determinants of inefficiency in the electricity generating industry in India.

Keywords

Stochastic frontier, maximum likelihood, semi-parametric, efficiency, panel data, electricity industry

1. INTRODUCTION

Privatization – the transfer of ownership from the public sector to the private sector – and modernization of production units have been key elements of industrial policy in several OECD economies in the 1980s. At present, these are major issues in the developing countries and in the economies under transition in Eastern Europe. The electricity industry, which has been traditionally owned and controlled by the state, is a prime target in public sector deregulation and privatization. A central argument for privatization across the world is that it increases productive efficiency and lowers costs as compared with public ownership. There is a large literature examining the efficient form of ownership of enterprises in general and of utilities in particular as well as on

Address for Correspondence
Madhu Khanna, Department of Agricultural and Consumer Economics, University of Illinois, Urbana-Champaign, IL 61801, USA. E-mail: khanna1@uiuc.edu

The Journal of International Trade & Economic Development
ISSN 0963-8199

the impact of age of equipment on efficiency. The evidence provided by these studies, however, is not conclusive.

In the case of privately-owned utilities, incentives for efficiency due to an ability to capitalize the gains from it are often confounded by the regulations imposed on the utilities. While Pollitt (1996) finds that privately-owned nuclear power plants in the UK are slightly more efficient than the publicly-owned plants, Bhattacharyya *et al.* (1995) and Fare *et al.* (1985) find that publicly-owned enterprises in the US are more efficient. Pollitt (1995) and Hjalmarsson and Veiderpass (1992) find no significant differences in efficiency between different types of ownership. Fung and Wan (1996) show that state-owned enterprises in China are less efficient than collective enterprises that have decentralized autonomous decision making and material incentives in the form of profit retention schemes and bonus systems. Several studies examining the impact of age on technical efficiency of firms find that older firms are less efficient than newer firms (Seitz, 1971; Jondrow *et al.*, 1982; Pitt and Lee, 1981; Sterner, 1990; Huang and Liu, 1994). On the other hand, Pillai and Srinivasan (1992) and Majumdar (1997) find older firms in the industrial sector in India to be more efficient. They attribute this to learning by doing and better labour-management relations among the older firms. Pollitt (1995), however, finds no significant learning effects among the electric utilities he studied.

A majority of these studies have used a non-stochastic specification of the objective function, which measures the mean value of the observations rather than the departure of the observations from the optimum values determined by the frontier. Although the idea of measuring productive efficiency of individual firms was originated by Farrell (1957), an econometric methodology for estimating stochastic frontier functions was introduced by Aigner *et al.* (1977) and Meeusen and van den Broeck (1977). This methodology has been used commonly to estimate the extent of inefficiency of electric utilities in the US (surveys by Bauer, 1990; Greene, 1993), but very seldom to examine the extent to which departures from the frontier can be systematically explained, particularly in countries other than the US (see the survey in Pollitt, 1995). Several studies (for example Hjalmarsson and Veiderpass, 1992; Seale, 1990; Sterner, 1990; Caves and Barton, 1990) analyse the sources of inefficiency by regressing the estimates of inefficiency on firm-specific characteristics. We believe that it is preferable to incorporate firm-specific variables in the frontier cost function because such variables may themselves influence the level of efficiency obtained. A few studies that do follow this approach (example, Coelli, 1996a; Battese and Coelli, 1995; Huang and Liu, 1994; Bhattacharyya *et al.*, 1995; Reifschneider and Stevenson, 1991) tend to make restrictive assumptions about functional forms and the distribution of the inefficiency.

This study has two objectives: first, to measure not only the extent of cost inefficiency, but also to examine whether inefficiency occurs randomly across firms or can be explained by firm-specific characteristics. We measure both time-varying and time-invariant inefficiency. We then hypothesize firm specific characteristics, such as alternative forms of ownership, public and private, age of plants and their capacity utilization levels, as possible sources of inefficiency. The impact of these hypothesized determinants of firm-specific inefficiency is examined by incorporating them directly in the specified cost function.

Second, this study assesses the impact of assumptions about the parametric form of the cost function and distribution of the stochastic departure from the frontier on the measurement of inefficiency and the sources of inefficiency. Results are obtained by following the conventional maximum likelihood (ML) approach, a generalized least squares (GLS) approach and a new semi-parametric-generalized least squares (SP-GLS) approach using panel data.

While incorporating the technical efficiency effects directly in the stochastic frontier specifications, Reifschneider and Stevenson (1991) make strong distributional assumptions by assuming that the inefficiency effects are the sum of a function of relevant explanatory variables and a truncated normal random error term, both of which must be non-negative. Huang and Liu (1994), Battese and Coelli (1995) and Coelli (1996a) make less restrictive distributional assumptions by assuming that the additive random error term is a truncation of a normal distribution with mean zero, whose point of truncation is dependent on firm characteristics, such that the inefficiency effects are non-negative. We apply this approach here.

The GLS approach avoids the need to make any distributional assumptions about the inefficiency effects. However, like the ML estimation procedure it does require *a priori* assumptions about the functional form of the production/cost function (as in Schmidt and Sickels, 1984; Seale, 1990; Pitt and Lee, 1981). It is well known that any misspecification of the functional form may lead to inconsistent estimates of the parameters of the cost function as well as of inefficiency. In view of this, we propose a semi-parametric model with a composite error term that does not impose any *a priori* parametric functional form on the cost function or rely on distributional assumptions. A similar semi-parametric method has also been applied by Adams *et al.* (1998) but without incorporating the determinants of firm-specific inefficiency in the frontier function. These three methodologies are then applied to the electricity-generating sector in India using panel data for 1987–88 to 1990–91.

All three methods of estimation show that there is significant inefficiency among the power plants in the electricity generating sector in India, but the distributional assumptions about the error term, required for the ML approach, lead to different values of average inefficiency as compared with the other two methods. All three approaches show that public ownership and low capacity utilization significantly increase the inefficiency of plants. In Section 2 we present the model and estimation procedures. Section 3 describes the data used in this study. The results of the empirical analysis are in Section 4 and the conclusions are presented in Section 5.

2. THE STOCHASTIC COST FRONTIER WITH FIRM-SPECIFIC INEFFICIENCY

We assume that a power plant minimizes its costs of production to produce a given level of output. Its cost function can be expressed as follows with multiplicative disturbances:

$$C_{it} = C(P_{mit}, Y_{it}, D_i^c)\exp(\varepsilon_{it}); \quad t = 1, \ldots T; \quad i = 1,2, \ldots N;$$
$$m = \text{capital } (k), \text{ labour } (l), \text{ fuel } (f) \tag{1}$$

where C_{it} is the cost of production for the ith firm at time t, P_{mit} are the prices of the m inputs, Y_{it} is output and D_i^c is a dummy variable equal to one for coal-based plants and zero for oil/natural-gas based plants. The latter controls for the impact of fuel dependent technology on the cost of production. The variable C_{it} is the theoretical least cost of production and $\exp(\varepsilon_{it})$ is the stochastic error term. It represents the ratio of observed cost to its stochastic frontier cost for the same level of prices and output and measures the extent to which a firm's cost of production exceeds the least cost frontier at time t. Under the ML approach, we specify the term ε_{it} to be composed of a stochastic component v_{it} that represents random factors that affect cost and a non-negative component u_{it} which represents inefficiency for firm i as in Battese and Coelli (1995) and Coelli (1996a). We hypothesize that firm specific inefficiency, u_{it}, consists of a systematic component, $h(\mathbf{Z}_{it})$, where $\mathbf{Z}_{it}$ is a vector of characteristics hypothesized to be the sources of inefficiency. It also consists of a random error w_{it}, which denotes residual or unexplained inefficiency. We refer to this approach as ML I. Thus:

$$u_{it} = h(\mathbf{Z}_{it}) + w_{it} \tag{2}$$

In this study we investigate three factors that may lead to non-attainment of the cost function by a plant. These include its age, non-utilized capacity and its ownership. We also consider the case where firm-specific inefficiency is specified as: $u_{it} = h(\mathbf{Z}_{it}) + w_{it}$, with $\mathbf{Z}_i$ defined to be time-invariant. The results in this case (referred to as ML II) were not found to be much different as compared with those obtained with ML I. We estimate ML II, in order to make a more direct comparison between the ML approach and the other two approaches where we consider firm-specific inefficiency u_i to be time-invariant in order to distinguish it from stochastic inefficiency v_{it}. When both u and v vary over time, firm-specific efficiency cannot be identified under the GLS method (Pitt and Lee, 1981). Under the GLS and SP-GLS approaches, we therefore assume $u_i = h(\mathbf{Z}_i) + w_i$ for all t. Although not reported here, we also obtained ML estimates under this assumption and found that the estimate of average inefficiency was similar to that obtained using ML I and ML II (see Khanna *et al.*, 1998). The details of the three estimation procedures considered here, ML, GLS and SP-GLS are discussed below.

2.1 The Maximum Likelihood (ML) approach

To estimate the model in (1) by the ML method, functional forms of the cost function and of the systematic portion of the one-sided error term need to be specified. We assume that the cost function is a translogarithmic function of the independent variables. Assuming three inputs of production, capital, labour and fuel, and imposing price homogeneity on the cost function by normalizing input prices by the price of capital, P_{kit}, we specify the stochastic cost function as:

$$c_{it} = \ln \frac{C_{it}}{P_{kit}} = \alpha + \beta_1 \ln \frac{P_{fit}}{P_{kit}} + \frac{\beta_2}{2}\left[\ln \frac{P_{fit}}{P_{kit}}\right]^2 + \beta_3 \ln \frac{P_{lit}}{P_{kit}} + \frac{\beta_4}{2}\left[\ln \frac{P_{lit}}{P_{kit}}\right]^2 + \beta_5 \ln \frac{P_{fit}}{P_{kit}} \ln \frac{P_{lit}}{P_{kit}}$$
$$+ \beta_6 \ln Y_{it} + \beta_7[\ln Y_{it}]^2 + \beta_8 \ln Y_{it} \ln \frac{P_{fit}}{P_{kit}} + \beta_9 \ln Y_{it} \ln \frac{P_{lit}}{P_{kit}} + \beta_{10} D_i^c + \varepsilon_{it} \tag{3}$$

where $\varepsilon_{it} = u_{it} + v_{it}$. It is assumed that v_{it} is identically and independently distributed (iid) as $N(0, \sigma^2_v)$ and independent of u_{it}, which are non-negative and independently distributed random variables. It is further assumed that u_{it} is obtained by truncation at zero of the normal distribution with mean $h(\mathbf{Z}_{it})$ and variance σ^2. Furthermore, the systematic component of u_{it}, $h(\mathbf{Z}_{it})$ in (2), is taken as a linear function of a vector of firm specific variables $\mathbf{Z}_{it}$. A linear parametric form of $h(\mathbf{Z}_{it})$ is chosen because it provides a good fit to the data. Quadratic forms of the explanatory variables included in the vector $\mathbf{Z}_{it}$ were not significantly different from zero under any of the three estimation methods. They are therefore excluded and we specify u_{it} as follows:

$$u_{it} = \mathbf{Z}_{it}'\gamma + w_{it} \tag{4}$$

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 $-\mathbf{Z}_{it}'\gamma$, that is, $w_{it} \geq -\mathbf{Z}_{it}'\gamma$. This assumption is consistent with u_{it} being non-negative truncations of the $N(\mathbf{Z}_{it}'\gamma, \sigma^2)$ distribution. We rewrite the model in (3) by replacing ε_{it} by $\mathbf{Z}_{it}'\gamma + w_{it} + v_{it} = \mathbf{Z}_{it}'\gamma + e_{it}$, where $e_{it} = w_{it} + v_{it}$. The log-likelihood function for this model can then be written as

$$\ln L = -\frac{NT}{2}\ln(2\Pi) - \frac{NT\ln(\sigma_v^2 + \sigma^2)}{2} - \frac{1}{2}\sum_{i=1}^{N}\sum_{t=1}^{T}[e_{it}^2/(\sigma_v^2 + \sigma^2)]$$
$$- \sum_{i=1}^{N}\sum_{t=1}^{T}[\ln \Phi(d_{it}) - \ln \Phi(d_{it}^*)] \tag{5}$$

where $\Phi(x)$ is the standard normal cumulative distribution function evaluated at x, and

$$d_{it} = \frac{Z_{it}'\gamma}{\sigma}, \; d_{it}^* = \frac{\mu_{it}^*}{\sigma^*}, \; \mu_{it}^* = \frac{\sigma^{2*}}{\sigma_v^2\sigma^2}[\sigma_v^2 \mathbf{Z}_{it}'\gamma + \sigma^2\varepsilon_{it}], \; \sigma^{2*} = \frac{\sigma_v^2\sigma^2}{(\sigma^2 + \sigma_v^2)} \tag{6}$$

Further, the best predictor of the technical inefficiency of firm i, $\exp(u_{it})$, conditional on e_{it} is

$$E(\exp(u_{it})|e_{it}) = \left\{ \frac{\Phi[(\mu_{it}^*/\sigma_*) + \sigma_*]}{\Phi[(\mu_{it}^*/\sigma_*)]} \right\} \exp(\mu_{it}^* + \frac{1}{2}\sigma_*^2), \tag{7}$$

The expressions in (5) and (7) correspond to the expressions in Battese and Coelli (1993) for the case of production frontiers. These expressions also hold for $u_{it} = \mathbf{Z}_{it}'\gamma + w_{it}$ with $\mathbf{Z}_{it}$ replaced by $\mathbf{Z}_i$ in (6).

2.2 The Generalized Least Squares (GLS) method

The availability of panel data allows consistent estimation of firm-specific inefficiencies without making the distributional assumptions about the error structure required under the ML approach above. Here we consider

$$u_i = \mathbf{Z}_i'\gamma + w_i \tag{8}$$

and treat w_i, the unexplained portion of the one-sided error term u_i, as random effects with mean zero and variance σ^2 and independent of the error v_{it} which is *iid* $(0, \sigma^2{}_v)$. The functional form of the cost function, however, is assumed to be the same as that in (3). The estimation of the cost function in (3) can then be done using standard panel data estimation techniques such as the error com–ponents (random effects/fixed effects) models. Since D_i^c and $\mathbf{Z}_i$ include variables that are invariant over time, the fixed effects or 'within' estimator of w_i cannot be used. We therefore estimate (3) using random effects GLS estimation methods (Seale, 1990; Schmidt and Sickels, 1984). Given estimates of β and γ we recover estimates of w_i and u_i as:

$$\hat{w}_i = \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_v^2 + T\hat{\sigma}^2} \right) \sum_{t=1}^{T} \hat{e}_{it} \quad \text{and } \hat{u}_i^* = Z_i'\hat{\gamma} + \hat{w}_i \tag{9}$$

We use the fact that $u_i \geq 0$ to normalize the inefficiency (Seale, 1990) and define:

$$\hat{u}_i = \hat{u}_i^* - \min(\hat{u}_i^*) \tag{10}$$

The expected inefficiency of the ith firm relative to its stochastic frontier is measured as $\exp(\hat{u}_i)$.

2.3 Semi Parametric-Generalized Least Squares (SP-GLS) method

The SP-GLS estimator considered here avoids the specification of the cost function and relaxes the assumptions about the distribution of the error terms. We now take the logarithm of cost as being some unknown function of the logarithms of input prices and output, but continue to impose the condition of homogeneity in prices. The systematic component of the one-sided error term u_i is assumed to have a linear format as in (8). We write the cost function to be estimated as:

$$c_{it} = g(X_{it}) + \mathbf{Z}_i^{*\prime}\gamma_0 + w_i + v_{it} \tag{11}$$

where $c_{it} = \ln(C_{it}/P_{kit})$, $X_{it} = (\ln Y_{it}, \ln(P_{lit}/P_{kit}), \ln(P_{fit}/P_{kit}))'$, $\mathbf{Z}_i^{*\prime} = [D_i^c, Z_i']$, $\gamma_0 = [\beta_{10}, \gamma']'$, $g(X_{it})$ is an unknown function of X_{it}, and w_i and v_{it} are as in Section 2.2.

Taking the conditional expectation of (11) leads to:

$$E(c_{it}|X_{it}) = E(\mathbf{Z}_i^*|X_{it})'\gamma_0 + g(X_{it}) \tag{12}$$

Further subtracting (12) from (11) we get:

$$c_{it} - E(c_{it}|X_{it}) = (\mathbf{Z}_i^* - E(\mathbf{Z}_i^*|X_{it}))'\gamma_0 + w_i + v_{it} \tag{13}$$

The estimator of γ_0, γ_0^*, can now be obtained by applying the random effects GLS procedure of Section 2.2 on (13). However, this estimator will not be operational since $E(c_{it}|X_{it})$ and $E(\mathbf{Z}_i^*|X_{it})$ are not known. To make the estimator operational we, therefore, first estimate $E(c_{it}|X_{it})$ and $E(\mathbf{Z}_i^*|X_{it})$ by the non-parametric kernel method. The non-parametric kernel estimator of $E(c_{it}|X_{it})$ is

$$\hat{E}(c_{it}|X_{it}) = \frac{\sum_j \sum_{t'} c_{jt'} K((X_{jt'} - X_{it})/h)}{\sum_j \sum_{t'} K((X_{jt'} - X_{it})/h)} \tag{14}$$

where $j,i = 1,\dots N$; $t', t = 1,\dots T$; $K(.)$ is a kernel function and h is the window width. The estimator in (14) is the weighted average of the c_{it} values corresponding to those X_{jt} which are around X_{it}, the point at which the conditional mean is calculated. The kernel weight is chosen such that it gives a low weight to those observations that lie far from X_{it} and a high weight to the observations close to X_{it}. For our empirical analysis here we use the product of normal kernels, that is

$$K\left(\frac{X_{jt} - X_{it}}{h}\right) = \prod_{r=1}^{q} K\left(\frac{X_{rjt} - X_{rit}}{h}\right); K\left(\frac{X_{rjt} - X_{rit}}{h}\right) = (2\Pi)^{-1/2} \exp\left[-\frac{1}{2}\left(\frac{X_{rjt} - X_{rit}}{h}\right)^2\right] \tag{15}$$

and $q = 3$ is the number of regressors in X_{it}. The window width h of the kernel, which determines the size of the interval around X_{it} over which the observations are averaged, is determined by minimizing the asymptotic mean squared error of the estimator in (14). This is given by $h \propto n^{-1/(q+4)}$. For $q = 3$ in our empirical analysis we use $h = s_{xr}n^{-1/7}$ where s_{xr} is the standard deviation of the variable X_r. A cross-validated choice of h was also considered but it did not change the results. The non-parametric kernel estimator of $E(\mathbf{Z}_{it}|X_{it})$ can be similarly written by replacing c_{it} with $\mathbf{Z}_i$ in (14). For details on the kernel estimation technique and the choice of kernel and window width see Hardle (1990) and Pagan and Ullah (1999). The consistency and asymptotic normality properties of the operational SP-GLS estimator proposed above follow from Li and Ullah (1998). Once the estimate γ_0* is obtained, the estimates of w_i and of firm-specific inefficiencies are obtained as in (9) and the estimate of $g(X_{it})$ is obtained by using the non-parametric estimation technique of regressing $(c_{it} - \mathbf{Z}_i^{*\prime}\gamma_0^*)$ on X_{it} described in (14). For this we need to replace c_{it} with $(c_{it} - \mathbf{Z}_i^{*\prime}\gamma_0^*)$ in (14). Derivatives of this cost function at the sample mean and their t-ratios are obtained using the method described in Rilstone and Ullah (1989) and Pagan and Ullah (1999).

3. DATA DESCRIPTION

The empirical analysis is based on data for a sample of 66 power plants in India for the period 1987–88 to 1990–91[1]. These data were collected from the official records of the Central Electricity Authority of India, some of which are published in CEA (1992). These plants accounted for 59 per cent of the total thermal capacity in India in 1990–91. They included oil, gas and coal-based plants, with 87 per cent of plants being coal-based.

The dependent variable of our cost function is annual cost of electricity generated by a power plant. It is the sum of annual expenditures on the three inputs, capital, labour and fuel. The price of labour for a power plant in year t is determined by dividing annual expenditures on labour by the number of employees in that plant in year t. The annual fuel consumption of a plant is converted into tons of oil equivalent using information about the heating values of the fuel used by each of the plants. The price per ton of oil equivalent is then obtained by dividing annual expenditures on fuel by the tons of oil equivalent consumed annually. The expenditures on capital are obtained by adding up the annual interest payments, annual depreciation expense and the annual expenditures on repairs and maintenance reported by plants. This is divided by the total capacity of the plant to obtain the price per unit capacity. The output variable measures the annual net electricity generation by a plant in kilowatt-hours. Additionally, a dummy variable, equal to 1 if the plant is coal-based and equal to zero if it is gas/oil based is included to distinguish coal-based plants from gas/oil-based plants. The per unit capacity cost of constructing coal-based plants is higher than for oil/gas-based plants. Further, energy consumption per kilowatt hour of coal-based plants is higher than for oil/gas-based plants since the latter have a higher designed energy efficiency.

Three firm-specific variables, ownership of the power plant, age of the plant, and its non-utilized capacity factor are included to explain systematic variations in inefficiency across plants. The electricity industry in India includes plants belonging to three ownership groups. The majority of the plants are publicly-owned and belong either to the state governments and are controlled by their State Electricity Board (SEB) or to the central government. There are a few privately-owned enterprises that were allowed to retain ownership when the industry was nationalized in 1956. Of the plants included in our sample, 77 per cent belong to the SEBs, 17 per cent to the central government, and 6 per cent to the privately owned corporations. The effects of ownership are captured by creating dummy variables: *SEB*, equal to 1 if a plant belongs to an SEB and zero otherwise; *Center*, equal to 1 if a plant belongs to a central government corporation and zero otherwise; and *Private*, equal to 1 if a plant belongs to a privately-owned corporation and zero otherwise.

Plants owned by the SEBs are managed by politically appointed members of the SEBs and their objectives may not always be cost minimization. Owing to a below cost pricing policy of the state governments, these plants are accumulating large losses and require subsidies to cover their costs of production. Financial dependence on the government has increased political interference in the management of these plants and eroded their autonomy. It has also reduced the availability of funds for maintenance of equipment. Chronic power shortages and subsidies provided by the government virtually guarantee that all plants are operated, irrespective of their efficiency. There is thus no reward for efficient production practices or pressure to reduce costs.

Plants owned by the central government have greater autonomy in making operational decisions than the plants owned by the SEBs although they too are subject to some bureaucratic interference. These corporations have to raise debt-capital from the open market and are not guaranteed government subsidies. Their institutional structure therefore provides them with incentives for managerial efficiency as compared with those owned by SEBs. On the other hand, privately-owned corporations are completely autonomous in their managerial decisions. They finance their capital investments through both debt and equity capital. The management is appointed by, and therefore accountable to, the shareholders.

The age of a plant could affect its productive efficiency in several ways. An older age of plants may reflect older technology embodied in the plant and greater wear and tear of equipment, which lowers its productive efficiency. On the other hand, better adaptation of production conditions in the plant with older equipment as well as high costs of learning about new technology could result in older plants being more efficient. Plants in our sample varied between 5 and 43 years with the average age being 18 years in 1990–91. Older plants in the electricity industry in India were more likely to have equipment that was imported from abroad rather than manufactured domestically. This equipment had difficulty adapting to the quality of coal available in India. Newer equipment is manufactured domestically and, although it is more suited for production conditions in India, it is more expensive and has suffered from inadequate availability of spare parts and technical assistance. In the presence of these confounding factors, the *a priori* impact of age on productive efficiency in the electricity-generating sector in India is ambiguous. In ML I, the age of the plant (*Age*) is defined as the number of years that it had been in operation over the period 1987–91. In ML II, we measure the age of the plant by the number of years it had been in operation in 1991 (measured by the difference between 1991 and the year in which the plant began production). Two joint variables, *Age* times *Private* and *Age* times *Center* are included to capture the differential effects of age on efficiency depending on ownership.

The capacity utilization factor is the ratio of actual output produced to the maximum output of electricity that a plant could produce if it were to operate continually at maximum capacity. In theory, capacity utilization is strongly correlated with productive efficiency. Low capacity utilization implies high frequency of shutdowns. This indicates higher expenditures on repairs and maintenance, which adds to the cost of capital. Frequent shut down and start up of equipment, is also fuel intensive and contributes to low fuel efficiency, relative to plants being operated continuously. The average capacity utilization of plants in the electricity-generating sector in India has been rather low with the average being only 47 per cent over the period 1987/88–1990/91. We capture its effect by the variable non-utilized capacity factor measured by 1-{net electricity generated/(total kilowatt installed capacity $\times$ 8760 hours)}. Since the non-utilized capacity factor was found to be strongly correlated with the output variable in the cost function, we use its average value over the four-year period for each plant in estimating firm-specific inefficiency even in the ML I case.

4. RESULTS

The results of estimating the stochastic cost frontier function using the ML, GLS and SP-GLS methods are presented in Table 1. ML results were obtained using FRONTIER 4.1 (see Coelli, 1996b). The last six variables in Table 1 represent the deterministic components of the inefficiency error term. Under the ML approach, the specification of firm-specific inefficiency in (4) includes an intercept term that can be identified separately from the intercept of the cost function. This is, however, not possible under the other two approaches where the estimation method is such that this intercept is subsumed within a common intercept term for the estimated function. In the case of the ML approach, a likelihood ratio test is performed to test the null hypothesis that the coefficients of the variables explaining inefficiency are zero. The estimated value of the test statistic with 8 degrees of freedom is 115.8 in ML I and 83.6 in ML II. In both cases, this leads us to reject at the 1 per

Table 1 Estimated parameters of the stochastic cost function

Independent variables	ML I	ML II	GLS	SP-GLS
Intercept	–0.57	–0.15	–0.55	0.622
	(0.72)	(0.97)	(0.84)	(0.031)
Price of labour	–0.26	–0.39	–0.79	0.17
	(0.20)	(0.22)*	(0.20)	(0.36)
Price of fuel	0.20	–0.18	0.11	0.46
	(0.20)	(0.24)	(0.17)	(0.39)
Output	0.78	0.76	0.32	0.70
	(0.11)***	(0.14)***	(0.15)**	(0.28)**
Price of labour squared	0.39	0.022	0.017	–
	(0.15)**	(0.016)	(0.017)	
Price of labour × Price of fuel	–0.013	0.022	–0.010	–
	(0.03)	(0.033)	(0.028)	
Price of fuel squared	0.053	0.038	0.029	–
	(0.019)**	(0.021)*	(0.017)*	
Output squared	–0.0018	–0.015	0.017	–
	(0.0086)	(0.014)	(0.009)**	
Output × Price of fuel	0.012	0.020	0.024	–
	(0.0157)	(0.017)	(0.016)	
Output × Price of Labour	0.014	0.017	0.015	–
	(0.012)	(0.014)	(0.015)	
Coal plants	0.25	0.27	0.32	0.34
	(0.05)	(0.05)***	(0.07)***	(0.11)***
Intercept of inefficiency function	–1.96	–2.08	–	–
	(0.21)***	(0.55)***		
Age of plant	0.0005	0.0042	–0.0078	–0.0041
	(0.0031)	(0.0031)	(0.0037)**	(0.0042)
Age × Private Ownership	0.039	0.035	0.013	0.0082
	(0.009)***	(0.018)**	(0.007)*	(0.0073)
Age × Central Ownership	0.0052	0.003	0.012	–0.0051
	(0.0041)	(0.004)	(0.006)**	(0.0061)
SEB ownership	1.73	2.02	0.49	0.53
	(0.17)***	(0.53)***	(0.19)**	(0.25)**
Central ownership	1.66	1.94	0.38	0.51
	(0.18)***	(0.52)***	(0.21)*	(0.29)*
Non-utilized capacity factor	1.10	0.66	0.45	0.49
	(0.16)***	(0.14)***	(0.18)**	(0.24)**
Corr(c_{it}, $\hat{c}_{it}$)	0.957	0.962	0.983	0.975
σ^2	0.035	0.035	0.048	0.057
σ^2_v	0.009	0.009	0.009	0.017
$\lambda = \sigma^2/\sigma^2_v$	3.95	3.95	6.8	3.35

Estimated standard errors are reported in parenthesis to two significant digits. The estimated coefficients are given to the corresponding numbers of digits behind the decimal places; *** indicates statistically significant at 1%; ** indicates statistically significant at 5%; * indicates statistically significant at 10%

cent confidence level the hypothesis that the inclusion of an inefficiency function is not warranted.

In order to compare the estimated coefficients across the three models, we calculate the partial derivatives of cost with respect to each of the variables that appears jointly or in quadratic form in the parametric models. These partial derivatives are evaluated (as in Rilstone and Ullah, 1989) at the sample means of the variables. These estimates are reported in Table 2. The coefficients for the relative prices of labour and fuel and for output using all three approaches are rather close. They show that the cost function is monotonically increasing in all input prices and in output. While output is statistically significant in all models, fuel price is significant only in the parametric models. All three estimation methods show that coal plants have significantly higher costs of production than gas/oil fired plants.

Among the variables hypothesized to explain the systematic portion of inefficiency, we find that both state and central ownership have positive and statistically significant effects on the costs of production under all three approaches (Table 1). This indicates that publicly-owned plants are more inefficient than privately-owned plants. The partial derivatives of the cost function with respect to the three ownership categories show that under all three methods, the coefficient for central ownership is lower than that for state ownership and higher than that for private ownership (Table 2). This suggests that even among plants under public ownership, managerial autonomy and appropriate incentives for cost minimization can reduce technical inefficiency to some extent.

The coefficients for the age variables estimated in the parametric models suggest that age had a positive impact on inefficiency among plants owned by the central sector and the private sector. This implies that it had a negative impact on the efficiency of plants owned by the SEBs. This indicates either the

Table 2 Partial derivatives of the cost function

	ML I	ML II	GLS	SP-GLS
Price of labour	0.08	0.08	0.10	0.17
	(0.46)	(0.49)	(0.31)	(0.36)
Price of fuel	0.62	0.62	0.66	0.46
	(0.27)**	(0.28)**	(0.25)***	(0.39)
Output	0.90	0.88	0.80	0.70
	(0.01)***	(0.02)***	(0.02)***	(0.277)**
Age	0.0037	0.0065	–0.0051	–0.0043
	(0.005)	(0.006)	(0.0046)	(0.0041)
SEB ownership[a]	1.73	2.02	0.482	0.52
	(0.17)***	(0.53)***	(0.189)**	(0.25)**
Centre ownership[a]	1.66	1.95	0.39	0.51
	(0.18)***	(0.52)***	(0.20)*	(0.51)
Private ownership[a]	0.039	0.038	0.0054	0.004
	(0.010)***	(0.018)***	(0.0094)	(0.009)

[a]coefficients estimated for a one-year old plant. Standard errors in parenthesis; *** indicates statistically significant at 1%; ** indicates statistically significant at 5%; * indicates statistically significant at 10%

presence of learning by doing among the plants owned by the SEBs or the ability of the central and private sectors to make better equipment purchase decisions so that their newer equipment was substantially more efficient than their older equipment. The net impact of age on efficiency after purging the effects of ownership on efficiency is insignificant in all three models as shown in Table 2. The three models also show that the fraction of unutilized capacity has a positive and significant impact on technical inefficiency. The significance of capacity utilization in determining efficiency is consistent with other studies such as Pollitt (1995, 1996), Singh (1991) and Reifschneider and Stevenson (1991).

In order to see the goodness of fit of the three models we estimated the correlation between the observed and the predicted value of the dependent variable of the cost function, $c_{it} = \ln(C_{it}/P_{Kit})$. We find that all three methods provide good fits to the data. The ML methods had the lowest correlation coefficient of 96 per cent. The correlation coefficient of the GLS model (98.3 per cent) was somewhat higher than that of the SP-GLS method (97.5 per cent). This indicates that the assumption of a translog cost function in the ML and GLS models was a reasonably good approximation to the unknown form of the cost function for our sample of electricity generating plants.

4.1 Estimates of technical efficiency

We now analyse the magnitudes of the firm specific technical inefficiencies calculated using the different methods discussed above. The average inefficiency of plants in the sample is 48 per cent and 43 per cent under the ML I and ML II methods respectively (Table 3). Average inefficiency increased slightly over the four years, from 46 per cent to 52 per cent under ML I and from 42 per cent to 47 per cent under ML II. The average inefficiency of power plants is much higher under the GLS method (97 per cent) and the SP-GLS method (99 per cent). Under both the GLS and SP-GLS methods, the least efficient plant was at least 300 per cent as inefficient as a plant on the frontier. The ability of power plants in the electricity industry in India to continue operating despite such low efficiencies is possible because of the lack of com-

Table 3 Summary results for inefficiency estimates

	ML I	ML II	GLS	SP-GLS
Mean inefficiency	1.48	1.43	1.97	1.99
Standard Error	0.02	0.03	0.05	0.05
Median	1.41	1.36	1.95	1.99
Skewness	0.91	0.44	0.10	0.22
Range	1.67	0.95	2.06	2.42
Maximum	2.68	1.97	3.06	3.42

petition in the sector, the excess demand for electricity at existing prices and the subsidies provided by the state governments. The shortage of the electricity supply in India implies that all available capacity is operated irrespective of its costs of production.

Figures 1–4 present the kernel density of inefficiencies under the ML, GLS and the SP-GLS method. For the kernel density estimation method, see Hardle (1990) and Pagan and Ullah (1999). The inefficiency distributions obtained using the GLS and SP-GLS methods are relatively more symmetric, with the mean and the median values being almost identical, as compared to those obtained using the ML methods. The inefficiency distributions estimated using the ML method are relatively more skewed towards the higher values. The ML methods also predict a lower magnitude of average inefficiency among the sample plants and a smaller range of the inefficiency distribution as compared with the other two approaches. A comparison of the distribution of inefficiencies estimated by the alternative models indicates that the ML method predicts a larger percentage of plants having lower inefficiency as compared with the other two methods. This difference in the distribution of ML inefficiency compared with the other two methods may be due to the distributional assumption imposed by the ML approach.

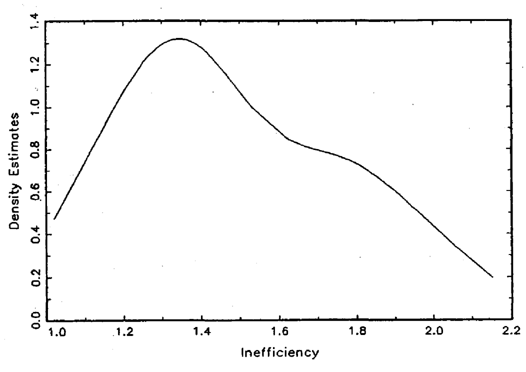

Figure 1 ML 1 Method

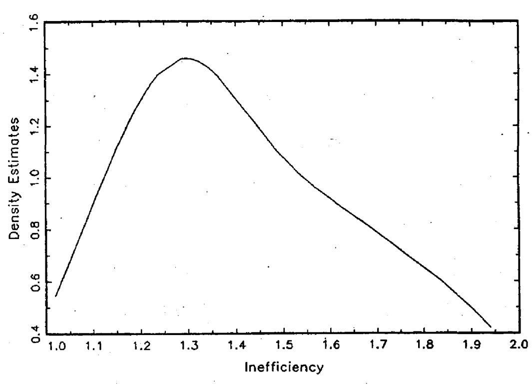

Figure 2 ML II Method

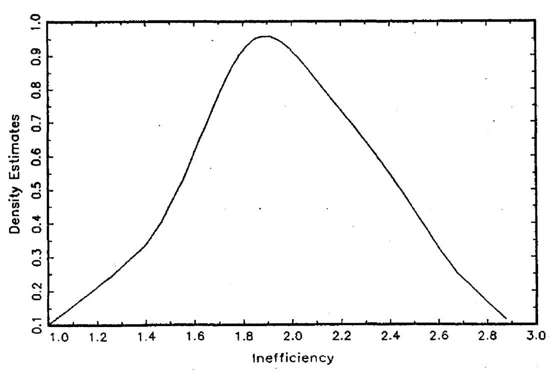

Figure 3 GLS Method

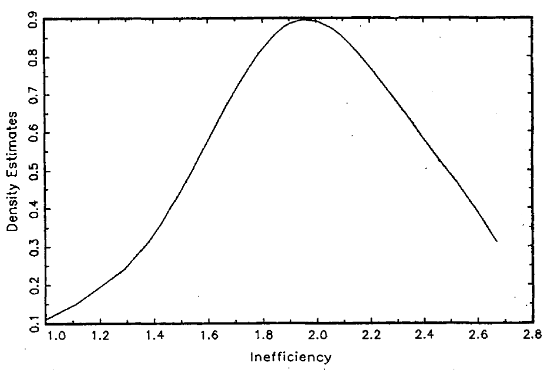

Figure 4 SP Method

We compare the magnitude of each plant's inefficiency as estimated by the three methods by calculating the Pearson's and the Spearman's Rank Correlation. The correlation between the plant level inefficiencies estimated by all three approaches is significant at the 1 per cent level (two-tailed test) (Table 4). Both Pearson's and the Spearman's Rank Correlation between the inefficiencies estimated using SP-GLS and GLS were higher than those between the SP-GLS and ML methods. The magnitudes of the plant level inefficiencies estimated by GLS were, however, more closely correlated with those estimated using ML relative to those estimated using SP-GLS.

5. CONCLUSIONS

In this paper we use panel data to estimate frontier cost functions while relaxing the parametric and distributional assumptions typically associated with frontier function estimation to varying degrees. Firm-specific attributes that could lead to systematic departures from the frontier are incorporated in the cost function. A new SP-GLS approach is proposed here and its estimates are compared with those obtained using ML and GLS approaches. These three approaches are applied to estimate and explain the level of cost-inefficiency among electricity generating power plants in India with differences in ownership structures, age and capacity utilization levels.

Our empirical analysis shows that publicly-owned plants are more inefficient than privately-owned plants and that capacity utilization is a significant determinant of inefficiency in the electricity generating industry in India. The age of plants by itself did not contribute to plant inefficiency. All three methods of estimation provide a good fit to the data indicating that the assumption of a translog cost function in the ML and GLS models was a reasonably good approximation to the unknown form of the cost function for our sample of electricity generating plants. The average inefficiency predicted by the ML approach is lower than that predicted by the GLS and the SP-GLS methods which could be due to the distributional assumptions and error structure assumed under the ML method.

Table 4 Pearson's and Spearman's correlation of firm-specific inefficiencies

	GLS	SP-GLS	ML I	ML II
GLS	1	0.65[b]	0.82[b]	0.86[b]
SP-GLS	0.73 [a]	1	0.49[b]	0.56[b]
ML I	0.77 [a]	0.50 [a]	1	0.99[b]
ML II	0.83 [a]	0.58 [a]	0.99 [a]	1

[a]Pearson's correlation coefficient; [b]Spearman's correlation coefficient. All correlations are significant at the 1% level (two-tailed).

ACKNOWLEDGEMENTS

The authors are thankful to the referees for their helpful and constructive comments. They are also thankful for the comments by the participants of seminars at the Washington State University, University of Guelph, McGill University and the South East Asian Econometric Society Meeting at New Delhi, India. The authors gratefully acknowledge the research support provided by the Campus Research Board, University of Illinois, Urbana-Champaign and by the Academic Senate, University of California, Riverside.

NOTE

1 This refers to a financial year, which begins on April 1 in India.

REFERENCES

Adams, R. M., Semenick Alam, I.M. and Sickles, R.C. (1998) 'Time variation in banking efficiency: do banks converge to the frontier?' Manuscript, Rice University, Houston, Texas.

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

Battese, G.E. and Coelli, T.J. (1993) 'A stochastic frontier production function incorporating a model for technical inefficiency effects'. Working Papers in Econometrics and Applied Statistics, No. 69, Department of Econometrics, University of New England, Armidale.

Battese, G.E. and Coelli, T.J. (1995) 'A model for technical inefficiency effects in a stochastic frontier production function for panel data'. *Empirical Economics* 20, 325–332.

Bauer, P.W. (1990) 'Recent developments in the estimations of frontiers'. *Journal of Econometrics* 46, 39–56.

Bhattacharyya, A., Harris, T.R., Narayanan, R. and Raffiee, K. (1995) 'Specification and estimation of the effect of ownership on the economic efficiency of the water utilities'. *Regional Science and Urban Economics* 25, 759–784.

Caves, R.E. and Barton, D.R. (1990) *Efficiency in U.S. Manufacturing Industries*. Cambridge, Massachusetts.

CEA (1992) 'Performance review of thermal power stations, 1990–91'. Central Electricity Authority, Government of India, Ministry of Energy, Department of Power, New Delhi.

Coelli, T.J. (1996a) 'Measurement and sources of technical inefficiency in Australian coal-fired electricity generation'. CEPA Working Paper, University of New England, Armidale.

Coelli, T.J. (1996b) '*A Guide to FRONTIER Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation*'. CEPA working paper 96/07, University of New England, Armidale.

Fare, R., Grosskopf, S. and Logan, J. (1985) 'The relative performance of publicly-owned and privately-owned electric utilities'. *Journal of Public Economics* 26, 89–106.

Farrell, M.J. (1957) 'The measurement of productivity efficiency'. *Journal of Royal Statistics Society*, series A, 120, 253–81.

Fung, M.K. and Wan, K.K. (1996) 'Ownership and efficiency differentials in Chinese industry: Further evidence from data envelopment analysis'. *Applied Economics Letters* 3, 475–482.

Greene, W. (1993) 'The econometric approach to efficiency analysis'. In: H.O. Fried, C.A.K. Lovell, and S.S. Schmidt (eds) *The Measurement of Productive Efficiency*. Oxford: Oxford University Press.

Hardle, W. (1990) *Applied Nonparametric Regression*. Cambridge: Cambridge University Press.

Hjalmarsson, L. and Veiderpass, A. (1992) 'Efficiency and ownership in Swedish electricity retail distribution'. *Journal of Productivity Analysis* 3, 7–23.

Huang, C. and Liu, J.-T. (1994) 'Analysis of a non-neutral stochastic frontier production function'. *Journal of Productivity Analysis* 5, 171–80.

Jondrow, J., Lovell, C.A.K., Materov, I.S. and Schmidt, P. (1982) 'On the estimation of technical inefficiency in the stochastic frontier production function models'. *Journal of Econometrics* 19, 233–38.

Khanna, M., Mundra, K. and Ullah, A. (1998) '*Parametric and Semi-Parametric Estimation of the Effect of Firm Attributes on Efficiency: The Electricity Generating Sector in India*' Working Paper, Department of Economics, University of Riverside, California.

Li, Q. and Ullah, A. (1998) 'Estimating partially linear panel data models with one-way error components errors'. *Econometric Reviews* 17, 145–66.

Majumdar, S.K. (1997) 'The impact of size and age on firm-level performance: some evidence from India,' *Review of Industrial Organization* 12, 231–41.

Meeusen, W. and van den Broeck, J. (1977) 'Efficiency estimation from Cobb-Douglas production function with composed error'. *International Economic Review* 18, 435–44.

Pagan, A. and Ullah, A. (1999) *Non Parametric Econometrics*. New York: Cambridge University Press.

Pillai, P.M. and Srinivasan, J. (1992) 'The age and productivity of machine tools in India'. In: *Indian Industrialization*, A. Ghosh, K.K. Subramanian, M. Eapen, H.A. Drabu (eds) Delhi: Oxford University Press.

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

Pollitt, M.G. (1996) 'Ownership and efficiency in nuclear power production'. *Oxford Economic Papers* 48, 342–60.

Pollitt, M.G. (1995) *Ownership and Performance in Electric Utilities: The International Evidence on Privatization and Efficiency*. Oxford: Oxford University Press.

Reifschneider, D. and Stevenson, R. (1991) 'Systematic departures from the frontier: a framework for the analysis of firm inefficiency'. *International Economic Review* 32 (3), 715–23.

Rilstone, P. and Ullah, A. (1989) 'Nonparametric estimation of response coefficients'. *Communications in Statistics, Theory and Methods* 18(7), 2615–27.

Schmidt, P. and Sickles, R.C. (1984) 'Production frontiers and panel data'. *Journal of Business & Economic Statistics* 2(4), 367–74.

Seale, J.L., Jr. (1990) 'Estimating stochastic frontier systems with unbalanced panel data: the case of floor tile manufactories in Egypt'. *Journal of Applied Econometrics* 5, 59–74.

Seitz, W.D. (1971) 'Productivity efficiency in the steam-electric generating industry'. *Journal of Political Economy* 79, 878–86.

Singh, J. (1991) 'Plant size and technical efficiency in the Indian thermal power industry'. *Indian Economic Review*, 26(2), 239–52.

Sterner, T. (1990) 'Ownership, technology, and efficiency: an empirical study of co-operatives, multinationals, and domestic enterprises in the Mexican cement industry'. *Journal of Comparative Economics* 14, 286–300.