Deregulation, Ownership, and Efficiency Change in Indian Banking: An Application of Stochastic Frontier Analysis*

Subal C. Kumbhakar
Department of Economics
State University of New York
Binghamton, NY 13902, USA
E-mail: kkar@binghamton.edu

and

Subrata Sarkar
Indira Gandhi Institute of Development Research
Gen. Vaidya Marg, Goregaon (East)
Mumbai 400 065, INDIA
E-mail: ssarkar@igidr.ac.in

October, 2004

Abstract

This chapter uses a Stochastic Cost Frontier Analysis to evaluate the efficiency of the Indian Banking System using panel data on public and private sector banks for the period 1986-2000. Econometric models that allow explanations of efficiency variations in terms of exogenous factors are used to analyze the time behavior of efficiency of the banking system, especially the changes in efficiency since the initiation of the reforms program in 1992. Ownership characteristics of banks are also incorporated into the analysis to examine if efficiency as well as efficiency changes have differed across ownership groups.

Our results indicate the presence of cost inefficiency in the Indian banking system, but there is a tendency for inefficiencies to decline over time. The results also indicate that cost inefficiency of banks has increased since the initiation of the reforms, though the reduction in inefficiencies over time continues albeit at a slower rate compared to that observed in the pre-deregulation period. We also find that private banks are generally more cost-efficient than public banks, but there are no significant differences in the impact of deregulation on the cost efficiency of these two bank groups. At the individual level, we find marked differences in the efficiency behavior of different banks with private banks exhibiting much more intra-group volatility in relative efficiency changes between the pre and post deregulation periods compared to that of public banks.

* We thank Rudran Sensarma and Arijit Ghosh for excellent research assistance. None, other than us, is responsible for remaining errors.

1. Introduction

Since the nationalization of the Imperial Bank of India (now the State Bank of India) in 1955, and until the decade of the 1990s, the banking system in India has been highly regulated, keeping in view its financial linkage with the rest of the economy, and to meet the social and economic objectives of development. Accordingly, there have been strict controls on interest rates, as well as stringent regulations relating to branch licensing, directed credit programs, and mergers. Over time, however, the banking system exhibited poor performance, and such under-performance was seen by many as the direct result of excessive regulations that were in place. Thus, starting in 1992, the Central Bank of India, i.e., the Reserve Bank of India (RBI)) initiated a number of liberalization measures to make the banking sector more productive and efficient. To a large extent such liberalization was influenced by the experience of developed countries, notably the U.S., where relaxation of regulations was found to be a major cause of productivity and efficiency increase (Morrison and Winston, 1995; Borenstein, 1992; Winston, 1993).

In this chapter we use the stochastic cost frontier analysis to evaluate the efficiency of the Indian banking system using panel data on public and private sector banks for the period 1986-2000. The long period of the panel, encompassing both the pre-deregulation period (1986-1992) and the post-deregulation period (1993-2000) enables us to examine econometrically if the time behavior of efficiency shows any structural break over these two periods thereby providing us a statistical way of judging the effect of the reforms on banking efficiency. In addition, the Indian banking system, that has commercial banks belonging to both public and private sectors, enables us to examine whether inter bank efficiency variations are related to ownership status, a possibility widely recognized in the literature. For both these analysis, we use stochastic cost frontier models that allows one not only to estimate producer specific efficiencies but also to explain their variations in terms of exogenous factors. Finally, since the frontier analysis gives efficiency estimates for each bank, the analysis enables us to identify those banks that are in need of relatively more policy intervention for increasing the efficiency of their operations.

Stochastic frontier analysis has been used by a number of studies in evaluating banking efficiency, for example by Berger and Mester (2001) with respect to U.S. banking; by Mendes and Reblo (1999) with respect to Portuguese banking; by Chaffai (1997) with respect to Tunisian banking; by Grifell-Tatje and Lovell (1996), and Kumbhakar et al. (2001) with respect to Spanish banking; and by Berg, et al. (1993), for Nordic countries, to name a few. With respect to India, there are quite a few studies which have looked into the efficiency and productivity of the Indian banking system. Subramanyam (1993) analyzed productivity growth in Indian banking, Bhattacharyya, Lovell, and Sahay (1996) examined the relative performance of commercial banks under three different kinds of ownership (private, public, and foreign) during the post-deregulation period, Bhattacharyya, Bhattacharyya and Kumbhakar(1997) analyzed productivity growth of and the effect of regulation on public sector banks over the period 1970-1991, while the study by Kumbhakar and Sarkar (2002) examined the relationship between deregulation, ownership and total factor productivity (TFP) growth of public and private sector banks over the period 1986-1996. However, almost all of these studies used the usual average response function in analyzing performance and productivity. In other words , these studies assumed all banks to be efficient so that the only deviation of output or cost from the maximum (minimum) attainable level, was due to purely random factors. However, previous empirical work (see Berger and Humphrey (1992) and the references therein) have demonstrated that there are often wide variations in the performance of individual banks within the banking industry. If this is case, then the stochastic frontier analysis framework is more suitable for analyzing banking performance, and accordingly we adopt this approach in our analysis.

The rest of the chapter is organized as follows. In Section 2 we present the basic framework of stochastic frontier analysis and discuss some models that are used to estimate efficiency. This is followed in Section 3 by an overview of the institutional structure and regulatory environment of Indian banking and the recent deregulation measures. Section 4 describes the data used and the econometric models estimated for the present study. Section 5 presents the empirical findings and Section 6 concludes.

2. Stochastic Frontier Analysis

Although the importance of efficient use of resources has long been recognized, the mainstream neoclassical paradigm in economics assumes that producers in an economy always operate efficiently. In reality, however, the producers are not always efficient. Two otherwise identical firms never produce the same output, and costs and profit are not the same. This difference in output, cost, and profit can be explained in terms of technical and allocative inefficiencies, and some unforeseen exogenous shocks. Given the resources (inputs), a producer is said to be technically inefficient if it fails to produce the maximum possible output. Similarly, a cost or profit maximizing producer is allocatively inefficient if it fails to allocate the inputs optimally, given input and output prices. Both inefficiencies are costly in the sense that cost (profit) is increased (decreased) due to these inefficiencies. Costs of these inefficiencies are also reflected in lower productivity of inputs. Alternatively, productivity growth will be lower in the presence of any one, or both, of these inefficiencies.

Inclusion of these inefficiencies into economic analysis is attractive for several reasons. First, it helps to identify which producers are inefficient and, if so, to what extent. By identifying the inefficient producers, policies designed to promote efficiency can be made more effective by directing the necessary help to those who are in the greatest need of assistance. Second, after identifying the presence of inefficiency, it is natural to examine factors responsible for inefficiency, i.e., identification of determinants of inefficiency. Once some explanatory factors are found, programs can be designed and support can be directed to the needy producers to achieve maximum effectiveness.

Empirical measurement of productive efficiency was first made by Farrell (1957), who showed how to define cost efficiency and decompose it into its technical and allocative components. He also provided an empirical application to U.S. agriculture, though he did not use econometric techniques. Stochastic frontier analysis had its origin in two papers, one by Meeusen and van den Broeck (June, 1977), and the other by Aigner, Lovell, and Schmidt (July, 1977) The stochastic frontier technique starts with a production technology that is specified as

$$y = f(x_1, \cdots, x_k; \beta) \times \exp\{v + u\} \tag{1}$$

where y denotes output, $x_1, \cdots, x_k$ are k inputs used to produce y, f is the production technology (black box) which converts inputs to output, and β is a technology parameter vector to be estimated. v is a random noise component, an exogenous shock unknown to the producer. It can be either positive (good luck, for example) or negative. If a producer is unable to produce the maximum possible output, given its input levels and the technology, it is said to be technically inefficient. Such inefficiency might arise due to factors such as, managerial errors arising from inertia and ignorance, poor quality of inputs, etc. Since a technically inefficient firm's output is always less than the maximum possible level determined by the stochastic production frontier (i.e., $f(x_1, \cdots, x_k; \beta)exp(v)$), given a specific input bundle, a one-sided term u ($u \leq 0$) is appended to (1) to capture technical inefficiency.

In the present setup, inputs are assumed to be given and the objective is to maximize output. Thus, the only inefficiency, if any, is technical. Since data are available only on output and input quantities, estimation of the unobserved inefficiency, u, for each producer from a sample of producers requires some special econometric techniques.

The question of resource allocation is not addressed in the above framework because inputs are assumed to be given. In reality, however, input allocation decisions also need to be made. Assuming that the objective of the producer is to minimize cost (of inputs), one can express the technology is terms of the dual cost function, viz.,

$$E = c(w_1, \ldots, w_k, y; \gamma)/CE \tag{2}$$

where E is actual cost, $c(\dot{)}$ is minimum cost function without any inefficiency, $w = (w_1, \cdots, w_k)$ are prices of inputs $x_1, \cdots, x_k$, y is output, and γ is the technology parameter vector (related to β in (1)). CE is the overall cost efficiency. Since actual cost is increased due to technical and allocative inefficiencies, $CE \leq 1$.

Allocative inefficiency arises when the producer fails to use inputs in such a way that the cost is minimized. In other words, some inputs are overused and some are underused. Such misallocation leads to an increase in costs. Similarly, compared to another producer who is technically efficient, the presence of technical inefficiency means that an inefficient producer has to use more of every input (which is going to increase cost) to produce a given level of output. This increase in cost due to technical and allocative inefficiencies is captured by the CE term. The reciprocal of CE can be used to measure the percent by which actual cost exceeds the minimum possible cost. The problem here is to (i) estimate the overall cost efficiency (CE), and (ii) then decompose it into technical and allocative efficiencies. Farrell (1957) showed that the overall cost efficiency (CE) is the product of technical and allocative efficiencies. The decomposition problem is mostly addressed when the technology is known. The problem is much harder in practice, because the task is to estimate an unobserved variable (CE) along with the production technology, and then decompose it into two components (see Kumbhakar (1997)).

While the cost-function approach is the dual of the production-function approach of modeling inefficiency, there are at least two advantages of using the cost-function approach. The first, is that while the cost-function approach can easily handle cases where producers produce multiple outputs, the production function approach to stochastic frontier analysis is done on the assumption of a single output. The assumption of single output is rather restrictive in modern day settings where a large number of firms produce multiple outputs. The second is, that while the cost-function approach, being an input oriented measure of efficiency, can make a distinction between variable inputs and quasi-fixed inputs (inputs fixed in the short run), the production-function approach, being an output oriented measure of efficiency, treats all inputs equally. However, the cost-function approach imposes a behavioral assumption on producers, i.e., producers minimize cost, while the production-function approach does not impose any such behavioral assumption explicitly (although implicitly one assumes output maximization, at least in a single output framework). However, in competitive environments in which input prices (rather than input quantities) are exogenous, and in which output is also demand driven and so can also be considered as exogenous, the cost-function approach may be more appropriate. Finally, the data requirements for the cost-function approach are higher compared to that for the production function approach. While the latter requires data only on output and inputs, the former requires data on total expenditure, outputs, and input prices. In addition, where a multiple-equations framework is used (see Kumbhakar and Lovell, (2000)), data on inputs or input-cost shares are also required. In this chapter we use the cost-function approach towards estimating and modeling inefficiency. Accordingly, we now provide the basic econometric framework for estimating these models.

As outlined above, the estimation of a single equation stochastic cost frontier assumes the existence of data on the prices of the inputs employed, the quantities of outputs produced, and the total expenditure made by each of the I producers. In this case, the estimable cost frontier can be expressed as

$$\ln E_i = \ln(c(y_i, w_i; \beta) \exp\{u_i\}) \quad i = 1, 2, \cdots, I \tag{3}$$

where $E_i = \sum_n w_{ni} x_{ni}$ is the actual cost incurred by producer i, $y_i = (y_{1i}, \cdots, y_{Mi}) \geq 0$ is the vector of outputs produced by producer i, $w_i = (w_{1i}, \cdots, w_{Ni}) > 0$ is the vector of input prices faced by the producer, $c(y_i, w_i; \beta)$ is the cost frontier common to all producers,

β is the vector of technology parameters to be estimated, and $u_i = \ln CI$ captures the percentage increase in cost due to inefficiency. Since actual cost is bounded below by the minimum cost $c(y_i, w_i; \beta)$, the random variable u_i is non-negative. Higher the value of u_i higher is the cost-inefficiency of the producer. Note that in this formulation the input vector x_i used by the producer i need not be observed. If this is indeed the case, then cost inefficiency cannot be decomposed into cost of technical inefficiency and cost of allocative inefficiency.

Given the above formulation, the cost efficiency (CE) of a producer i can be expressed as

$$CE_i = \frac{c(y_i, w_i; \beta)}{E_i} = \exp\{-u_i\} \tag{4}$$

which defines cost efficiency as the ratio of minimum possible cost to actual or observed cost. Since actual cost is greater than or equal to the minimum cost, it follows that the CE_i is always less than equal to 1 and equals 1 only when the producer is efficient, i.e, actual cost equals minimum attainable cost.

In equation 3 the cost frontier $c(y_i, w_i; \beta)$ is deterministic because that the entire excess of observed expenditure over minimum possible expenditure is assigned to cost inefficiency. However, sometimes cost increases can occur due to random exogenous shocks like weather, strikes, quality of inputs, etc. which are beyond the control of producers. In order to control for such exogenous factors, another random term is added to the cost function, and the model becomes:

$$\ln E_i = \ln c(y_i, w_i; \beta) + u_i + v_i \quad i = 1, 2, \cdots, I \tag{5}$$

Under this formulation

$$E_i = c(y_i, w_i; \beta) \exp\{v_i + u_i\} \tag{6}$$

and $c(y_i, w_i; \beta) \exp\{v_i\}$ is the stochastic frontier. The stochastic frontier consists of two components: a deterministic part $c(y_i, w_i; \beta)$ that is common to all producers, and producer-specific random part $\exp\{v_i\}$. We can calculate the producer specific efficiency exactly as before by:

$$CE_i = \frac{c(y_i, w_i; \beta) \exp\{v_i\}}{E_i} = \exp\{-u_i\} \tag{7}$$

and CE_i satisfies all the properties mentioned above.

Estimating equation (5) requires, (i) specification of a functional form for the deterministic kernel $c(y_i, w_I); \beta$, (ii) an assumption about the distribution of the random variable v_i, and (iii) an assumption about the distribution of the random variable u_i. Assumption relating to the random variable v_i is standard, namely that v_i is distributed as a normal variable with zero mean and finite variance. Empirical models tend to differ primarily in their assumption relating to the random variable u_i and in their specification of the deterministic kernel. The initial models specified either a half-normal distribution or an exponential distribution for u_i, while later models assumed a more general truncated normal distribution for u_i, with the truncation point occurring at zero to ensure non-negativity of u_i.

Given a particular specification for the random variables u_i and v_i, the Maximum Likelihood (ML) technique is used to estimate the unknown parameters. Subsequently, the producer-specific inefficiencies are estimated using the Jondrow et al. (1982) method. To illustrate the procedure, let us suppose that we make the following assumptions with respect to u_i and v_i (the analytics of the ML technique does not depend on the particular specification for the deterministic kernel).

(i) $v_i \sim iidN(0, \sigma_v^2)$

(ii) $u_i \sim iidN^+(0, \sigma_u^2)$[1]

(iii) v_i and u_i are distributed independently of each other, and of the regressors.

Given these assumptions, the log-likelihood function for the sample of I producers can be written as (see Kumbhakar and Lovell (2000), pp. 140)

$$\ln L = constant - I \ln \sigma + \sum_i \ln \Phi(\frac{\epsilon_i \lambda}{\sigma}) - \frac{1}{2\sigma^2} \sum_i \epsilon_i^2. \tag{8}$$

where $\epsilon_i = u_i + v_i$, $\sigma^2 = (\sigma_u^2 + \sigma_v^2)$, $\lambda = \frac{\sigma_u}{\sigma_v}$ and $\Phi(.)$ is the standard normal cumulative distribution function. Battese and Corra (1977) suggested that the parameterization $\gamma = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2}$ be used in place of $\lambda = \frac{\sigma_u}{\sigma_v}$, because γ has a value between zero and one, while λ could be any non-negative value. Thus the parameterization γ is better suited in obtaining

[1] Subsequent studies have assumed the more general specification of $u_i \sim iidN^+(\mu, \sigma_u^2)$ for the one-sided error component, and modeled μ as a function of other variables.

the ML estimates as the search can be restricted within a known set. If we use the γ parameterization, the log-likelihood function is given by:

$$\ln L = constant - I \ln \sigma + \sum_i \ln[1 - \Phi(z_i)] - \frac{1}{2\sigma^2} \sum_i \epsilon_i^2. \tag{9}$$

where $z_i = \frac{\epsilon_i}{\sigma}\sqrt{\frac{\gamma}{1-\gamma}}$

The log likelihood function can then be maximized with respect to the unknown parameters to obtain the ML estimate. Note that $\gamma(\lambda) \to 0$ when either $\sigma_u^2 \to 0$ or $\sigma_v^2 \to \infty$ and $\gamma(\lambda) \to 1$ when either $\sigma_u^2 \to \infty$ or $\sigma_v^2 \to 0$. In the first case, the stochastic frontier collapses to the OLS cost frontier (i.e., average response function) with no inefficiency, while in the latter case, the stochastic frontier collapses to the deterministic frontier with no noise. Thus, log-likelihood test on $\gamma(\lambda) = 0$ can be done to judge the appropriateness of the stochastic frontier analysis vis-avis the OLS cost function approach of modeling efficiency.

Once the parameter estimates are obtained, the next step is to obtain the producer specific inefficiency estimates. Estimates of ϵ i.e. $\hat{\epsilon_i}$ are easily obtained from the residuals, viz., $\ln E_i - \ln c(y_i, w_i; \hat{\beta})$. However, this is a composite estimate of $u_i \hat{+} v_i$ from where we need to estimate on $\hat{u_i}$. It is obvious that $\hat{\epsilon_i}$ contains information about $\hat{u_i}$. Since the expected value of $\hat{v_i}$ is equal to zero, $\hat{u_i}$ is likely to be greater than zero when $\hat{\epsilon_i}$ is greater than zero. Accordingly, the conditional distribution of u_i given ϵ_i could be exploited to get estimates of producer specific inefficiency. This was first demonstrated, in the context of technical inefficiency, in a paper by Jondrow, Lovell, Materov, and Schmidt (1982) and since then this decomposition method of getting producer specific inefficiency has been known as the JLMS technique. The JLMS estimators of inefficiency are based on the conditional density of u_i given ϵ_i. Either the mean or the mode of this conditional distribution is used as point estimator of inefficiency. It can be shown (see Kumbhakar and Lovell (2000), pp. 141, for the exact derivation) that the mean estimator is

$$E(u_i \mid \epsilon_i) = \sigma_* [\frac{\phi(\frac{\epsilon_i \lambda}{\sigma})}{1 - \Phi(\frac{-\epsilon_i \lambda}{\sigma})} + (\frac{\epsilon_i \lambda}{\sigma})] \tag{10}$$

where $\sigma_*^2 = (\frac{\sigma_u^2 \sigma_v^2}{\sigma_u^2 + \sigma_v^2})$ and we have used the parameterization λ. Using these estimates of producer specific inefficiencies u_i, one can obtain producer specific cost efficiencies by plugging this value of u_i into equation (7), i.e,

$$\hat{CE_i} = \exp\{-E(\hat{u_i} \mid \epsilon_i)\}. \tag{11}$$

An alternative point estimator for producer specific inefficiency, first proposed by Battese and Coelli (1988), can also be obtained from

$$CE_i = E(\exp\{-u_i\} \mid \epsilon_i) = [\frac{1-\Phi(\sigma_* - \mu_{*i}/\sigma_*)}{1-\Phi(-\mu_{*i}/\sigma_*)}] \cdot \exp\{-\mu_{*i} + \frac{1}{2}\sigma_*^2\} \tag{12}$$

where $\mu_{*i} = \epsilon_i \sigma_u^2/\sigma^2$.

The two alternative estimators of cost inefficiency generally give different results since $\exp\{E(u_i \mid \epsilon_i)\} \neq E[\exp\{u_i\} \mid \epsilon_i]$. The Battese and Coelli point estimator is preferred to the JLMS estimator since the latter is only a first order approximation to the former. All inefficiency estimates computed in this chapter are based on the Battese and Coelli estimator given in equation (12).

Once we obtain estimates of inefficiencies for each producer at each time period, two natural questions arise. The first question is, what is the behaviour of the producer inefficiencies over time? Are they increasing, decreasing or constant? Such a question assumes importance especially in circumstances where policy interventions like deregulation, introduction of reforms, new entry, etc. take place at particular points in time. One could look at the time behaviour of inefficiencies to judge the impact of these events on producer performance. The second question is, what explains the variations in inefficiencies among producers and across time. The second question encompasses the first, but looking at the former is often done to get an aggregative idea.

Following Kumbhakar (1990), Battese and Coelli (1992) proposed a simple model that can be used to estimate the time behavior of inefficiencies . In their model, which was developed in a panel data context, the error term representing technical inefficiency was specified as:

$$u_{it} = \{\exp[-\eta(t-T)]\}u_i, \quad i = 1, 2, \cdots, N; \quad t = 1, 2, \cdots, T. \tag{13}$$

where the $u_{it} \sim N^+(\mu, \sigma^2)$, and η is a parameter to be estimated.

Under this specification, inefficiencies in periods prior to T depend on the parameter η. As $t \to T$ u_{it} approaches to u_T. Thus, inefficiency in period T can be viewed as the reference/benchmark point. Inefficiency prior to period T is the product of the terminal year's inefficiency and $\exp\{-\eta(t-T)\}$. If η is positive, then $\exp\{-\eta(t-T)\} = \exp\{\eta(T-t)\}$ is always greater than 1 and increases with the distance of the period t from the last period T. Thus when η is positive, inefficiencies *fall* over time. Conversely, when η is negative, inefficiencies increase over time. However, as Battese and Coelli note, the inefficiencies of different firms in any period t are equal to the *same* exponential function, $\exp\{-\eta(t-T)\}$, of the corresponding firm-specific inefficiency effects in the last period. Thus, the ordering of firms in terms of inefficiencies does not change during the period of analysis. This obviously rules out cases where relatively inefficient producers become more efficient over time and vice-versa. Though this is a restrictive feature of this model, we can interpret η as giving an "overall or average trend" in inefficiencies over the period of study across all producers, and thus this model provides a good starting point for modeling aggregative behavior. Battese and Coelli implement this model in their FRONTIER program under the option "Model 1."

With respect to the second question, since the main motivation for efficiency analysis to policy makers is to design policies to improve performance of producers – especially the inefficient ones – it is highly desirable to know whether or not there are factors that can explain inefficiency. Thus, the next stage of the analysis is to focus on the determinants (factors explaining) of inefficiencies.

The variables explaining inefficiency are usually neither inputs nor outputs of the production process, but which nonetheless exert an influence on producer performance. They are thought to characterize the environment in which production takes place, and so to influence the efficiency of production. Examples include the degree of competitiveness, input and output quality indicators, network characteristics, ownership form, changes in regulation, various management characteristics, and the like.

The next question is: How does one introduce these variables into the analysis? Two possible solutions are usually explored in the literature. The first solution is to include these variables in the production process as control variables. Using this interpretation, these variables influence the structure of the technology by which conventional inputs are converted into outputs, but not efficiency. To illustrate, let $x = (x_1, \cdots, x_N) \geq 0$ be an input vector to produce a scalar output y, $w = (w_1, \cdots, w_N)$ be the associated input-price vector, and let $z = (z_1, \cdots, x_M)$ be a vector of exogenous variables that influence the structure of the production process by which inputs are converted into outputs. Under this approach, since z is assumed to influence the production process and itself and hence the cost structure, it is included along with w in a stochastic frontier, which is written as

$$E = c(w, z, y; \gamma) \exp\{v + u\} \tag{14}$$

where $c(w, z, y; \gamma)$ is the deterministic kernel of the stochastic cost frontier $[c(w, z, y; \gamma) \exp\{v\}]$, $v \sim iidN(0, \sigma_v^2)$ captures the effect of random noise on the production process, $u > 0$ captures the effect of cost inefficiency, and the parameter γ to be estimated now includes cost parameters as well as environmental parameters. The above equation has exactly the same structure as a conventional stochastic cost frontier model discussed above and all the usual estimation techniques can be applied to estimate this model. Subsequently, cost inefficiency of individual producers in different time periods can be obtained by the application of the JLMS technique or the Battese and Coelli point estimator.

The other solution is to associate variation in estimated efficiency with variation in the exogenous variables. The early papers (e.g., Pitt and Lee, 1981; and Kalirajan, 1981) implemented this approach in two stages. In the first stage a stochastic frontier equation was estimated (excluding the exogenous variables), typically by MLE under the usual distributional and independence assumptions, and the regression residuals were decomposed using the JLMS technique. In the second stage, these estimated inefficiencies were regressed on exogenous variables to explain/locate the source of inefficiency. Though this earlier approach was simple and intuitive it had a significant econometric problem, namely that in the first stage it was assumed that the inefficiency effects were independently and *identically* distributed to use the JLMS technique. However, the later assumption was clearly contradicted in the second stage in which the estimated efficiencies were assumed to have a functional relationship with the exogenous variables z_i.

Kumbhakar, Ghosh, and McGuckin (1991), and Reifcheneider and Setvenson (1991) first noted this inconsistency, which subsequently led to the development a series of models in which the inefficiency effects were specified as functions of the exogenous variables, and all the parameters of the stochastic frontier function as well as those of the inefficiency function was estimated together in a single MLE procedure.

In this chapter we use the second approach to explaining inefficiency. We use the model proposed by Battese and Coelli (1995), to determine the source of inefficiency. The Battese and Coelli model is similar in many respects to the models proposed by Huang and Liu (1994) and by Kumbhakar, Ghosh, and McGuckin (1991). Our preference for the use of the Battese and Coelli (1995) model is primarily because it is easily implementable since it is available as an option (Model "2") in the FRONTIER program. In brief, the model proposed by Battese and Coelli (1995), to explain production inefficiency within a panel data context, has the following structure:

$$\ln E_{it} = lnC(w_{it}, y_{it}; \beta) + v_{it} + u_{it} \tag{15}$$

$$u_{it} = \delta' z_{it} + \epsilon_{it} \tag{16}$$

where the first error term v_{it} represents the random noise in the production process of the ith firm in the tth period, and the second error term u_{it} captures the effect of technical inefficiency, which has a systematic component $\delta' z_{it}$ associated with the exogenous variables and a random component ϵ_{it}. The non-negativity requirement that $u_{it} = (\delta' z_{it} = \epsilon_{it}) \geq 0$ is modeled as $\epsilon_{it} \sim N(0, \sigma_\epsilon^2)$ with the distribution of ϵ_{it} being bounded below by the variable truncation point $-\delta' z_{it}$. Once the model is specified, the technological parameters and the inefficiency parameters are estimated by the MLE technique. Inefficiency estimates of individual producers at different time points are obtained as usual by the JLMS technique (see Battese and Coelli (1993) for the exact expression of the log-likelihood function, and the FRONTIER manual (Coelli, 1996) for calculating the producer and time specific inefficiencies.)

3. Institutional Structure and Regulatory Environment of Indian Banking

The banking system in India, like those in most developing economies, is characterized by the co-existence of different ownership groups, public and private, and within private, domestic and foreign. Public sector banks in India came into existence in several phases. In 1955, the Government of India (GOI) took over the ownership of the Imperial Bank of India and reconstituted it as the State Bank of India (SBI) under the State Bank of India Act of 1955. Later, in 1959, the State Bank of India (Subsidiary Banks) Act was passed enabling the SBI to take over seven banks of princely states as its associate banks. The

SBI and its associates were entrusted with the task of serving the banking needs of the hitherto neglected sectors. However, notwithstanding the progress made by these banks in terms of geographical coverage and credit expansion, it was felt that commercial bank credit was flowing mainly to the large and well established business houses, and sectors such as agriculture and small scale industries were being neglected. Thus in 1967, the policy of social control over banks was announced, and in 1969, fourteen of the largest private banks were nationalized under the Nationalization Act of 1969. A second phase of bank nationalization followed with six more private banks getting nationalized in 1980. The smaller private banks as well as the foreign banks were allowed to co-exist with the public sector banks, but their activities were highly restricted through entry regulation and strict branch licensing policies.

With the nationalization of the major commercial banks, a large number of regulatory measures were adopted by the RBI. Apart from changing the sectoral composition of credit, the RBI stipulated lending targets to priority sectors, provided refinancing facilities, set up credit guarantee schemes, and directed banks to open branches in rural and semi-urban areas to make banking accessible to all. The RBI also fixed maximum deposit rates on both savings and time deposits of all maturities and specified differential lending rates linked to borrowers' income and types of lending. The Lead Bank scheme was started for designing and implementing credit plans at the micro level. These measures led to the phenomenal growth of the banking system in general, and the public sector banks, in particular. By the early nineties, public sector banks accounted for nearly 90 percent of total deposits and advances, with the residual being almost equally split between private and foreign banks (Table 1).

However, by this time, the excessive focus on quantitative achievements had made many of the public sector banks unprofitable and under-capitalized by international standards. Many banks were earning less than reasonable rates of return, had low capital adequacy and high non-performing assets, and were providing poor quality customer service. Operating costs were increasing at a very high rate and the rapid growth in staff and promotions had diluted the quality of manpower.

In recognition of these growing illnesses, the RBI launched major banking sector reforms in 1991 aimed at creating a more profitable, efficient and sound banking system, based on the recommendations of the first Narasimham Committee on Financial Sector Reforms. The reforms sought to improve bank efficiency through entry deregulation, branch de-licensing, deregulation of interest rates, and mandating strong public sector banks to go to the capital market to raise funds up to 49 percent of their equity capital. The last move was primarily aimed towards generating market pressures on good public sector banks so that they became more efficient. The reforms also targeted to improve bank profitability through the gradual reduction of the Cash Reserve Ratio and the Statutory Liquidity Ratio, and to strengthen the banking system through the institution of the Bank of International Settlements (BIS) norm of a 8 percent Capital Adequacy Ratio, as well as stringent income recognition and provisioning norms (see Sarkar (1999) for an exhaustive review of the recent banking sector reforms).

These changes were aimed towards creating a competitive environment that, in the long-run, was expected to lead to substantial gains in efficiency, profitability, and productivity. Signs of increased competition has appeared in the Indian banking industry, as is evident in the decline of the four-bank asset concentration ratio from 0.49 in 1991-92 to 0.44 in 1994-95, by the growing presence of the private and foreign banks (Table 1), and in the appearance of service competition. The performance of public banks has also become more market-driven with growing emphasis put on profitability as an important benchmark for evaluating their performance by policy makers in the post-reforms era (MOF, 1993).

The existence of regulatory bottlenecks in the past, and their gradual liberalization in the recent years, provide us with a natural experiment that is well suited for studying the effects of deregulation on efficiency change. In addition, the Indian banking system which consists of commercial banks belonging to both public and private sectors, is particularly well suited for examining whether efficiency change could vary across ownership groups.

4. Data and the Econometric Model

4.1 Data

Multiple outputs and multiple inputs characterize the banking industry. This is independent of whether a value added, a user cost, or an asset approach is used. However, a considerable disagreement exists in the banking literature in defining what exactly a bank produces. Following Berger and Humphrey (1992), Grifell-Tatje and Lovell (1996) and Berg et al. (1993), we use the value added approach that treats both deposits and loans as outputs. Thus, our output vector (y) includes quantities (rupee value in 1980-81 prices) of deposits (DEP), loans and advances (AD), and investments (IN). Apart from these, we also include the number of branches (BRN) as an additional output. We do so keeping in mind the argument advanced in the literature that the number of branches can proxy for the quality and convenience that a bank offers to its customers (Grifell-Tatje and Lovell (1996), Berg et al. (1993). Indeed, branch expansion, especially in the rural and semi-urban areas, has been an important objective of the regulatory policy of RBI. Labor (L) and capital (K) are the two variable inputs. Total cost is the total operating cost (OPCOST) of a bank, less costs that are of a fixed nature like auditors' fees, lawyers' fees, etc. Price of labor (w_L) is obtained by dividing total expenses on labor by total number of employees. Similarly, price of capital (w_K) is obtained from $w_K =$ (total operating cost - total expenses on labor)/total fixed assets.

The data for the present study is obtained from various issues of the reports (i) *Financial Analysis of Banks*, (ii) *Performance Highlights of Public Sector Banks*, (iii) *Performance Highlights of Private Sector Banks*, and (iv) *Performance Highlights of Banks* published by the Indian Banks' Association. These publications report annual data from the profit and loss accounts and the balance sheets of all public and private banks operating in the Indian banking industry. The empirical model is estimated using data on 27 public banks and 23 private banks observed continuously over the years 1986 to 2000 [2]. In 2000, these 50 banks accounted for 99.6 percent of total branches and 89 percent of total deposits of the commercial banking sector (Table 1).

4.2 The Econometric Model

We use a translog specification of the cost frontier to estimate the efficiency of the individual banks. The translog function has been widely used in efficiency studies and can be viewed as a second order approximation of any unknown cost function. The translog cost function in the present case is:

[2] The year 2000 refers to the financial year beginning in April, 1999 and ending in March, 2000. Similarly, the year 1993 refers to the financial year beginning in April, 1992 and ending in March, 1993. We adopt this convention throughout the rest of the paper. Note that this convention is different than that in the paper by Kumbhakar and Sarkar (2002); and Sarkar, Sarkar, and Bhaumik (1998)

$$
\begin{aligned}
\ln E_{it} =& \alpha_0 + \sum_m \alpha_m \ln y_{mit} + \sum_j \beta_j \ln w_{jit} + \beta_t \, t \\
&+ \frac{1}{2}\Big\{ \sum_m \sum_l \alpha_{ml} \ln y_{mit} \ln y_{lit} + \sum_j \sum_k \beta_{jk} \ln w_{jit} \ln w_{kit} + \beta_{tt} \, t^2 \Big\} \\
&+ \sum_m \sum_j \alpha_{mj} \ln y_{mit} \ln w_{jit} \\
&+ \sum_m \alpha_{mt} \ln y_{mit} t + \sum_j \beta_{jt} \ln w_{jit} \, t + u_{it} + v_{it}
\end{aligned} \tag{17}
$$

where $i = 1, \ldots, I$ indexes banks and $t = 1, \ldots, T$ indexes time.

We impose the usual symmetry restrictions on the above cost function, viz., $\alpha_{ml} = \alpha_{lm}$, and $\beta_{jk} = \beta_{kj}$. Moreover, to ensure linear homogeneity of the cost function in w, the following restrictions are imposed

$$
\sum_j \beta_j = 1, \quad \sum_j \beta_{jk} = 0 \;\forall\; k, \quad \sum_j \alpha_{mj} = 0 \;\forall\; m, \quad \sum_j \beta_{jt} = 0. \tag{18}
$$

We estimate the above stochastic frontier models using the FRONTIER program developed by Coelli (1996). As outlined earlier, we estimate two alternative versions of the above model, namely the Battese and Coelli 1992 model for determining the time behaviour of efficiency, referred to as “Model 1” in the FORNTIER program, and the Battese and Coelli 1995 model for explaining inefficiency as a function of exogenous factors, referred to as “Model 2” in the FRONTIER program. Since our primary focus is on analyzing the effect of ownership and deregulation on bank efficiency, we include one ownership dummy namely PVT (for private banks), one deregulation dummy, DEREG (which equals one if year > 1992; and zero otherwise), and a time variable t, and their interactions, while estimating Model 2. We estimate these models using only public and private banks that were in existence prior to the initiation of the reforms. Foreign banks are not included in the analysis because many of the regulations relating to the percentage of priority sector lending, branch expansion, etc., have been quite different for foreign banks compared to the domestic banks. Since efficiency is a relative concept, we preferred to keep the comparison set to be as homogenous as possible. For similar reasons, we did not include the new banks (the entrants) in our analysis.

The year 1993 is taken as the beginning of the post-liberalization period keeping in view the fact that the first set of recommendations of the Narasimham Committee was started being implemented in January 1992. The norms of income recognition, asset classification, and loan-loss provisioning changed quite substantially in the post-deregulation period due to the institution of prudential regulations and adoption of the BIS norms. The response to these changes has varied between bank groups[3], as well as among banks within a particular group. Accordingly, in the econometric model that we estimate, the behavior of inefficiency prior and post 1993 would be of significant importance.

5. The Alternative Models and Results

Table 2 presents the means of the relevant variables that we use in our analysis. In 1986, public banks were, on an average, almost twenty times the size of private banks, operated eight times the number of branches as private banks and, had almost fifteen times the number of employees working in a typical private bank. However, by 2000, the relative size of the public banks had shrunk to about nine times the size of private banks, and the relative number of employees had reduced to twelve times that of a typical private bank, largely due to the higher growth rate experienced by private banks during this period. Public banks, though, continued to maintain their relative advantage in terms of number of branches, which is largely reflective of the conscious governmental policy of bringing banking services at the doorstep of every rural household. Looking into the sub-period of 1992-2000 i.e., the post liberalization era, Table 2 clearly reveals a much higher growth rate of private banks over public banks in terms of size (200 percent vis-a-vis 60 percent), branches (30 percent vis-a-vis 11 percent), and employees (20 percent vis-a-vis 1 percent). This reflects to a large extent the emphasis of the reforms measure to allow more freedom to private banks both in terms of their freedom of operation as well as geographical expansion, and the pressure on public banks to rationalize thier employee base through various labor reorganization procedures.

Table 2 also presents some simple indicators of productivity of public and private banks. Deposits, advances, and investments per employee of public banks were higher than those of private banks in 1985 and 1992, However, by 2000, this relative ranking had been reversed for each of the three indicators, with private banks showing a much higher growth rate (168 percent vis-a-vis 79 percent for deposits per employee, 153 percent vis-a-vis 51

[3] According to one estimate (MOF, 1993), the profits of the 28 public sector banks were reduced by 45 percent due to a switch to the new accounting norms.

percent for advances per employee, and 231 percent vis-vis 130 percent for investments per employee) during the post-liberalization period. The situation with respect to cost per employee is just the reverse, with private banks exhibiting a slightly lower cost per employee than public banks in the pre-liberalization period, and slightly higher figures in the post-liberalization period. The average salary per employee of private banks were slightly lower than public banks in the pre-liberalization period, and comparable to public banks in the post liberalization period.

Though Table 2 gives us an overall idea about the relative position of public and private banks in terms of some summary indicators, it is somewhat difficult to make judgments of efficiency using these indicators. In most situations where indicators of output are higher, so are the indicators of cost. To deduce efficiency propositions one needs a benchmark against which outputs and their associated costs can be judged. Similarly, one trend is visible in the summary data, namely that there appears to be a structural break and a role reversal between public and private banks during the post liberalization era. Whether this structural break and ownership difference is reflective of a general phenomenon or is driven by observations in a particular year or by a few banks needs to be evaluated in a statistical framework. This brings us to the estimation of the stochastic frontier models where costs are measured relative to the frontier level, and where efficiency can be explained in terms of exogenous variables like ownership and different time periods.

Table 3, column 1, presents the estimated parameters of the translog cost function, and the estimated parameters of the inefficiency function of Battese and Coelli Model 1, estimated using data on public and private banks for the period 1986-2000. Since the main focus of our analysis is on efficiency, we do not present a detailed discussion of the estimated cost function parameters. We only note that the estimated coefficients are theoretically consistent and fourteen out of the twenty-eight parameters of the translog cost functions are significant at the 5 percent level. We observe that the coefficients associated with the time variables t, t^2, and the interaction between t and the wage rate variable are negative and significant suggesting technical progress during the estimation period. Also, since the coefficient of t^2 is negative, the estimate parameters suggests an increasing rate of technical progress over time.

The bottom part of the table presents the parameters that can be used to judge the suitability of using the stochastic frontier model. Under the present formulation, testing for the presence of bank-specific inefficiency, and hence the necessity of using the frontier model, translates into testing the composite hypothesis $H_0 : \gamma = \mu = \eta = 0$[4]. The test is done using the usual likelihood ratio (LR) test, but the test statistic has a mixed chi-squared distribution (Battese and Coelli, 1996) and the critical value for a given level of significance, is lower than that reported in the usual chi-squared tables. The value of the LR statistic, along with its degrees of freedom are reported on the last two rows of Table 3. At the 1 percent level of significance, the critical value of the (usual) chi-squared distribution (with 3 degrees of freedom) is 11.341. The value of the test statistic in our case is much larger than this value (as noted above, the critical value is lower than 11.341) suggesting that our analysis overwhelmingly rejects the null hypothesis. Thus the standard average response function is not adequate for analyzing the cost behavior of banks and a frontier model is required. Also note that the null hypothesis $H_0 : \gamma = 1$ is rejected at the 1% level of significance (the associated test statistic, which is asymptotically normally distributed, has value of -9.03), implying that a stochastic frontier specification fits the data better than a deterministic frontier. Thus the estimated parameters imply that the performance of banks are better analyzed within a stochastic frontier framework.

The estimated value of the parameter η is positive and significant at the 1 percent level. Recall that a positive value of η implies that inefficiencies of producers decrease over time. Thus our estimate suggests that the cost efficiency of Indian banks has improved during the estimation period. This is true for every bank since the parameter η in the Battese and Coelli (1992) model, is bank invariant. The calculated mean efficiencies[5] for each year are reported in Table 4 and represented in Figure 1. The means are reported for all banks as a group and separately for public and private banks. According to these estimates, the Indian banking system exhibits significant inefficiency, with the mean efficiency score varying from 69 percent in 1986 to 75 percent in 2000. It can also be seen that the mean efficiency score of public banks as a group is lower than that for private banks, the relative efficiency of the former being about 90 percent of the latter group. The mean efficiency

[4] Note that in this model testing only the null hypothesis of H_0:γ=0 does not imply the absence of inefficiency as γ=0 is consistent with the presence of bank-invariant inefficiency.

[5] Whenever we calculate mean efficiency scores, these are the simple means of the efficiency scores of the individual banks.

of the banking system as a whole, as well as of each group, show an increasing trend, and in fact an exactly similar trend, with private banks being more efficient than public banks in every year. These observations, however, could be reflective of the specification of the Battese and Coelli (1992) model.

Table 5 presents the ordering of the banks in terms of their efficiency scores for the year 2000. Recall that under the Battese and Coelli (1992) model, the ordering of the banks is constrained to be the same in every year, though their efficiency scores can vary over the years. In our sample there are 50 banks, of which 23 are private, and 27 are public. Private banks appear to be much more efficient than public banks, with the top 25 spots being occupied by 16 private sector banks. The inter-bank variation in efficiency scores is much wider than suggested by our earlier means based analysis. Within private banks, the efficiency scores of the top two private banks namely the Dhanlakshmi Bank Ltd. and the Lord Krishna Bank Ltd. (with efficiency scores of 98 percent and 97 percent respectively) are much higher compared to that of the lowest two private banks namely the Punjab Cooperative Bank Ltd. and the United Western Bank Ltd. (with efficiency scores of 71 percent and 65 percent respectively). Within public banks, the variation in efficiency score is still sharper, with the Allahabad Bank being the most efficient public bank (efficiency score of 97 percent) and the Canara Bank the least efficient of the group (with efficiency score of 58 percent). The ordering shows that though there are some public banks that compare very well with private banks in terms of their cost efficiency, a majority of private banks come out to be more efficient than public banks. This shows that our earlier conclusion of the relative efficiency of private banks as a group (based on mean efficiency) was reflective of a general characteristic and not dictated by a handful of very efficient private banks.

As outlined earlier, though the Battese and Coelli (1992) model is useful in obtaining an overview of the efficiency of the banking system, the model has two restrictive features, namely, (i) the time behavior of efficiency is constrained to be a smooth monotonic function over the entire estimation period, and (ii) the ordering of banks is constrained to be the same for every year. Thus it is difficult to use this model in situations where exogenous shocks like reforms and deregulation are expected, and indeed targeted, to alter the operational efficiency of banks. Also since different banks or bank groups can react differently to the reforms and deregulation measures, the rankings of banks in terms of their efficiency can be reasonably expected to change.

One way to adapt the Battese and Coelli (1992) model to handle such situations is to estimate the model separately for different time periods and different ownership groups (public and private). Columns 2 and 3 present the estimation results of the Battese and Coelli Model 1 for the pre and the post liberalization periods respectively, while columns 4 and 5 present the results for public and private banks (over the entire estimation period) respectively. The estimated values of η reported in columns 2 and 3 suggest efficiency to be increasing over both the sub-periods, though the level of significance of these estimates is now reduced for both sub-periods. The estimated values of η reported in columns 4 and 5 suggest that while efficiency of private banks have increased over the estimation period, the efficiency of public banks has not changed significantly over time. However, the latter two models are subject to the same restrictions as outlined earlier. Again, one could in principle estimate two models (one for the pre and one for the post liberalization period) for each of the ownership groups, but then one starts reducing the degree of freedom quickly under this approach. Also, even assuming that one does estimate many such alternative models, comparisons of efficiency across periods and across groups are not strictly valid because one is allowing the benchmark (the frontier cost function) to change, which amounts to measuring efficiency changes with an elastic scale. The Battese and Coelli (1995) model that we outlined earlier overcomes many of these problems and is very well suited for analyzing the issues of impact of deregulation and ownership, and it is to this model that we now focus our attention.

Tables 6a and 6b present the estimated coefficients and the associated t-ratios respectively, of the translog cost function based on Battese and Coelli Model 2. Four alternative models are estimated for different specifications of the inefficiency function, i.e. the function used to explain bank inefficiencies in terms of exogenous variables (referred to as the Z variables). The estimated parameters of the inefficiency functions are presented in Table 7. Like Model 1, the LR statistics for all the alternative models reconfirm that the stochastic frontier specification is the appropriate framework for analyzing bank performance; the LR statistic being significant at the 1 percent level for each of the four models[6]. Note that,

[6] In Battese and Coelli Model 2, the null hypothesis of the absence of bank specific inefficiencies translates into H_0:γ=δ_0=$\delta_1 \cdots \delta_m$=0, where δ's are the parameters associated with the Z variables.

the estimated value of γ is much lower in Model 2 than what we estimated in Model 1. This implies that the exogenous variables in the inefficiency function are able to explain a substantial part of the unconditional variance of the one-sided error term[7]. We briefly note that the estimated coefficients of the translog cost function are theoretically consistent and about half of the twenty-eight parameters are significant at the 5 percent level in each of the four models. We also observe that similar to Model 1, the coefficients associated with the time variables t, t^2, and the interaction between t and the wage rate and the advances variables are negative and significant in most models suggesting technical progress during the estimation period.

Lets now look at the parameters of the inefficiency function reported in Table 7. Model A is the simplest, in which inefficiency is modeled in terms of a private ownership dummy. This model suggests, that if we look at the estimation period as a whole, then private banks as a group come out to be more efficient than public banks, a statistical confirmation of the result that we found based on the simple Battese and Coelli Model 1. Model B incorporates an additional time trend variable and allows this trend to be different for public and private banks. Both coefficients are negative, almost similar in magnitude, but are statistically insignificant. This suggests, that for the estimation period as a whole, we do not find any significant change in efficiency over time. Models C and D brings in the effect of deregulation into the analysis. Model C, is again a simple model, allowing only one time intercept shift in the inefficiency function in the deregulation period. The estimated parameter associated with the deregulation dummy is positive and significant, indicating that, on a average (across years), the cost-efficiency of Indian banks has declined (i.e., inefficiency has increased) in the post-deregulation period. The coefficient associated with the interaction of the deregulation dummy and the private dummy is negative, suggesting that decline in efficiency of private banks has been less, but this coefficient is not significant even at the 20 percent level of significance.

Model D is the most comprehensive model. It allows ownership effects, time effects, and deregulation effects. In addition, the deregulation effects and the time effects are allowed to vary across ownership groups by incorporating suitable interaction terms.

[7] Referring to our discussion at the end of Section 2, one can notice that in Battese and Coelli Model 2, the parameter $\gamma = \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2+\sigma_u^2}$.

Looking at the pre-deregulation period one observes that the efficiency of banks has tended to increase over time, the coefficient associated with the time variable being negative and significant at the 5 percent level. Also, private banks appear to be more efficient than public banks. The effect of deregulation is captured in terms of the four variables (i) *dereg*, which gives the change in the intercept of the inefficiency function in the post-deregulation period, (ii) *dereg*t*, which gives the change in the slope of the inefficiency function with respect to the time variable, (iii) *dereg*pvt* and (iv) *dereg*t*pvt* which measure the differences in effects of *dereg* and *dereg*t*, respectively between private and public banks. We first observe that neither of the two variables *dereg*pvt* and *dereg*t*pvt* are significant at the 5 percent level (the associated p-values are 0.38 and 0.40 respectively), implying that there are no ownership effects of deregulation and accordingly there are no significant changes in the relative efficiency of public and private banks in the post-liberalization period. The coefficient associated with *dereg*t* variable is positive and significant suggesting that the efficiency of the Indian banks has exhibited a declining trend in the post-deregulation period *relative* to that in the pre-deregulation period. The coefficient is, however, of lower magnitude than that associated with the t variable. This implies, that the total effect (i.e, the sum of t and *dereg*t*) is still negative, so that the efficiency of Indian banks has continued to increase over the years in the post-deregulation period, albeit at a much slower rate compared to that in the pre-deregulation period. Finally, as observed earlier, since there are no significant ownership effects of deregulation on banking efficiency, and private banks were, on the average, found to be more efficient than public banks in the pre-deregulation period, the relative ranking has been preserved in the post-deregulation years.

Table 8 reports the annual average (over banks) efficiency scores for the banking system as a whole, as well as for public and private banks, implied by the inefficiency function based on Model D. Figure 2 gives a graphical representation of the figures in Table 8. A number of observations can be made from the table. First, in conformity with the estimated parameters for model D, the efficiency scores for both public and private banks show an increasing trend from the year 1986 till the year 1992, the end of the pre-deregulation period. In the year 1993, there is a drop in efficiency of both bank groups, and since then a gradual increase. However, the rate of increase in efficiency over the years, is less than the rate observed in the pre-deregulation period. By the year 2000, the last period of the sample, the efficiency level of the banking system as a whole, as well as of the public sector banks were at the same level as in the year 1988, while for private banks the efficiency level was comparable to that in the year 1989. Efficiency levels in 2000, in general, were lower compared to the level reached at the end of the pre-deregulation period, namely the year 1992. Second, the efficiency of private banks are higher than that of public banks, as a group, and this ranking is maintained during the entire estimation period. Finally, the estimated average efficiencies based on Battese and Coelli Model 2 are much higher compared to those based on Model 1, and this is true for all models, A through D. This result is not surprising because Model 2 takes the determinants of inefficiency into account explicitly.

Table 9 presents the ranking of banks in terms of their average efficiency scores. Three rankings are given, one based on the entire estimation period (1986 to 2000), the next based on the pre-liberalization period (1986 to 1992), and finally the last based on the post-liberalization period (1993 to 2000). Under each of the three rankings, private banks appear to be more efficient than public banks. Unlike Model 1, no public bank appears in the top 25 rankings (for the full estimation period). This could be partly a reflection of the parametric nature of the inefficiency function fitted in Model 2 (the private dummy being negative and significant). However, the relative rankings of the public banks within its own group is quite robust between Model 1 and Model 2, with Allahabad bank and the Dena bank appearing within the top three spots, and the State bank of Saurashtra, the Syndicate bank, and the Canara bank being the three least cost efficient public banks. Similarly, within private banks, the five cost efficient banks in terms of Model 1, also appear in the first six cost efficient banks in terms of Model 2. The Spearman's rank correlation coefficient between the two rankings is 0.73.

More interesting observations can be made by looking at the change in rankings of banks between the pre and the post deregulation periods. Private banks display a much wider variation in their ranks compared to that of public banks. While the Spearman's rank correlation coefficient for public banks is 0.66, it is only 0.34 for private banks. Thus, within private banks, there are significant inter-bank changes in rankings, with some banks leap-frogging over others in terms of relative efficiency. The South Indian Bank Ltd., the Dhanalakshmi Bank Ltd., and the Lord Krishna Bank Ltd., exhibit substantial relative efficiency gains, while the City Union Bank Ltd., the Benaras State Bank Ltd., the Punjab Cooperative Bank Ltd., and the Bareilly Bank Ltd., exhibit significant loss in relative efficiency. Compared to private banks, the ranking of public banks are much more stable over the two periods, with only the Oriental Bank and the Corporation Bank showing marked improvement in their rankings, and the Indian Bank showing significant decline in rankings.

The greater volatility of private banks in terms of their efficiency rankings between the pre and the post liberalization period is perhaps natural. The limits to branch and size expansion and other restrictions that existed prior to the reforms could have constrained private banks in a manner in which the efficiency differentials among banks could not materialize. This is particularly likely to be relevant when there are scale economies and previous studies (Kumbhakar and Sarkar, 2002) do indicate the presence of such scale economies in Indian banking. Under such circumstance, removal of constraints is likely to have different effects on different private banks. This explanation also fits quite well with the relatively lower volatility in the rankings of public banks all of which were allowed to expand unfettered during the pre liberalization years.

Figures 3(a) to 3(j) give the time series plot of the cost-efficiency (based on Model D) of some public and private banks over the estimation period. It is easily observed that the efficiency behavior is quite different for different banks. While all banks show a decline in cost-efficiency in the years surrounding the initiation of the reforms, their subsequent behavior are markedly varied. The rankings and the efficiency plots at the bank level show the usefulness of stochastic frontier analysis in identifying those banks which are in need of greater assistance so that appropriate policy actions can be taken.

6. Conclusion

This chapter used stochastic frontier analysis to evaluate the efficiency of public and private sector banks in India over the period 1986 to 2000. An important purpose of the analysis was to illustrate how the stochastic frontier approach could be used to explain the variations in inefficiency in terms of exogenous factors that could be of use to policy makers. In Indian banking this translated into examining the effect of ownership and especially the recent deregulation measures in affecting bank performance.

Our results indicate that Indian banks, on average, do exhibit the presence of cost inefficiency in their operations, though there is a tendency for inefficiencies to decline over time. The results indicate that recent deregulation has led to an increase in the cost inefficiency of the Indian banks and a fall in the rate of inefficiency reduction. This phenomenon of increase in cost inefficiency subsequent to liberalization has been observed in many other banking studies as well, perhaps because liberalization brings with it significant changes in technology, procedures, and practices, all of which shifts the cost-frontier inwards, but individual banks are slow to respond and reorient themselves to these changes.

Our results also indicate the presence of some ownership effects; private banks, on average, are generally more cost-efficient than public banks. However, we do not find any significant differences in the impact of deregulation on the cost efficiency of these two ownership groups. At the individual level, we do find marked differences in the efficiency behavior of different banks and their response to the deregulation measures. Private banks show much more intra-group volatility in relative efficiency changes between the pre and post deregulation periods compared to that of public banks. This is perhaps because pre-deregulation constraints were more binding on private banks than on public banks. The ordering of the banks in terms of their rankings in the pre and the post deregulation periods can help us identify those banks which are in need of greater assistance and monitoring.

Finally, a few comments on the limitations of our analysis and scope for future research. First, in this analysis, we have looked only at cost efficiency. A bank could appear to be cost inefficient but may be revenue efficient. It is particularly relevant for service industries like banking, where because of difficulty in controlling for output quality, banks which provide more services may appear to be cost inefficient. But if such services are value enhancing, customers may be willing to pay for better quality, which should show up in these banks earning higher revenue compared to others. Thus cost efficiency has to be combined with revenue efficiency for effective policy intervention. Clearly, a bank, which is neither cost nor revenue efficient, is in need of help. Second, we have included branches as outputs. We have done so to capture the convenience and quality of service that banks offer to its customers. In our case, this inclusion has also been motivated by the regulator's objective of social banking so vigorously followed in India. However, it can be argued that including branches as outputs preempts the possibility of detecting inefficiency that can arise due to having a suboptimal number of branches. Since public sector banks in particular have a large number of branches, this could bias the (efficiency) estimates in favour of public banks. Therefore, examining the inefficiencies by excluding branches as outputs is worth exploring in the future. Finally, we have not considered directly the effect of new private banks as well as foreign banks on the efficiencies of the public and the old private banks, though we have incorporated the indirect effect as encapsulated in the proxy variable for time. Since efficiency is a relative concept, it would be instructive to include the new banks into the analysis to examine the effect of their entry on the relative standing of the public and old private banks.

References

Aigner, D.J., Lovell, C.A.K., and Schmidt, P. (1977), Formulation and Estimation of Stochastic Frontier Production Function Models, *Journal of Econometrics*, 6:1 (July), 21-37.

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, 325-32.

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 Economics, University of New England, Armidale.

Battese, G.E. and Coelli, T.J. (1992), Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India, *Journal of Productivity Analysis* 3:1/2 (June), 153-69.

Battese, G.E. and Coelli, T.J. (1988), Prediction of Firm-Level Technical Efficiencies with a Generalized Frontier Production Function and Panel Data, , *Journal of Econometrics*, 38, 387-99.

Battese, G.E. and Corra G.S. (1977), Estimation of a Production Frontier Model: With Application to the Pastoral Zone off Eastern Australia, *Australian Journal of Agricultural Economics*, 21:3, 169-79.

Berg, S.A., Forsund, F.R., Hjalmarsson, L. and Suominen, M. (1993), Banking Efficiency in the Nordic Countries, *Journal of Banking and Finance*, 17, 371-388.

Berger, A.N. and Mester, L.J. (2001), Explaining the Dramatic Changes in Performance of U.S. Banks: Technological Change, Deregulation, and Dynamic Changes in Competition, *Working Paper*, University of Pennsylvania.

Berger, A.N. and Humphrey, D.B. (1992), Measurement and Efficiency Issues in Commercial Banking, in *Output Measurement in the Service Sectors, Studies in Income and Wealth*, Z. Griliches (ed.), Vol. 56, The University of Chicago Press, Chicago, IL, for National Bureau of Economic Research, Chicago, 245-296.

Bhattacharyya, A., Bhattacharyya, A. and Kumbhakar, S.C. (1997), Changes in Economic Regime and Productivity Growth: A Study of Indian Public Sector Banks, *Journal of Comparative Economics*, 25, 196-219.

Bhattacharyya, A., Lovell, C.A. Knox and Pankaj, S. (1996), The Impact of Liberalization on the Productive Efficiency of Indian Commercial Banks, *European Journal of Operational Research*, 21.

Borenstein, S. (1992), The Evolution of U.S. Airline Competition, *Journal of Economic Perspectives*, 6, 45-73.

Chaffai, M.E. (1997), Productivity and Efficiency Performances of the Tunisian Banking Industry before and after the Economic Liberalization Program: An Econometric Study Using Panel Data, in *Models for Economic Policy Evaluation Theory and Practice: An International Experience. Research in Human Capital and Development*, Dahel, R. and Sirageldin, I. (eds.), Vol. 11, Part B, JAI Press, 335-350.

Coelli, T.J. (1996), A Guide to FRONTIER Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation, Centre for Efficiency and Productivity Analysis, University of New England, Armidale.

Farrell, M.J. (1957), The Measurement of Productive Efficiency, *Journal of the Royal Statistical Society*, Series A, General, 120, Part 3, 253-81.

Grifell-Tatje, E. and Lovell, C.A.K. (1996), Deregulation and Productivity Decline: The Case of Spanish Savings Banks, *European Economic Review*, 40, 1281-1303.

Huang, C. J., and J.-T. Liu (1994), Estimation of a Non-Neutral Stochastic Frontier Production Function, *Journal of Productivity Analysis*, 5, 171-80.

Indian Banks' Association, *Performance Highlights of Public Sector Banks*, Mumbai, INDIA, various issues.

Indian Banks' Association, *Performance Highlights of Private Sector Banks*, Mumbai, INDIA, various issues.

Indian Banks' Association, *Performance Highlights of Banks*, Mumbai, INDIA, various issues.

Indian Banks' Association, *Financial Analysis of Banks*, Mumbai, INDIA, various issues.

Jondrow, J., Lovell, C.A.K., Materov, I.S., and Schmidt, P. (1982), On the Estimation of Technical Efficiency in the Stochastic Frontier Production Function Model, *Journal of Econometrics*, 19:2/3 (August), 233-38.

Kalirajan, K.P. (1981), An Econometric Analysis of Yield Variability in Paddy Production, , *Canadian Journal of Agricultural Economics*, 29, 283-94.

Kumbhakar, S.C. (1997), Modeling Allocative Inefficiency in a Translog Cost Function and Cost Share Equations: An Exact Relationship, *Journal of Econometrics*, 76, 351-356.

Kumbhakar, S.C. (1990), Production Frontiers, Panel Data and Time-Varying Technical Inefficiency, *Journal of Econometrics*, 46, 201-212.

Kumbhakar, S. and Sarkar, S. (2002), Deregulation, Ownership, and Productivity Growth in the Banking Industry: Evidence from India, *Journal of Money Credit and Banking*, forthcoming.

Kumbhakar, S.C., Lozano-Vivas, A., Lovell, C.A.K. and Hasan, I. (2001), The Effects of Deregulation on the Performance of Financial Institutions: The Case of Spanish Savings Banks, *Journal of Money, Credit and Banking*, 33, 101-120.

Kumbhakar, S.C. and Lovell, C.A.K (2000), *Stochastic Frontier Analysis*, Cambridge University Press, U.K.

Kumbhakar, S.C., Ghosh, S., and McGuckin, J.T. (1991), A Generalized Production Frontier Approach for Estimating Determinants of Inefficiency in US Dairy Farms, *Journal of Business and Economic Statistics*, 9:3 (July), 279-86.

Mendes, V. and Reblo, J. (1999), Productive Efficiency, Technological Change and Productivity in Portuguese Banking, *Applied Financial Economics*, 9, 513-521.

Ministry of Finance (MOF), Government of India, Public Sector Commercial Banks and Financial Sector Reform: Rebuilding for a Better Future. New Delhi: Government of India, 1993.

Morrison, S. and Winston, C. (1995), The Evolution of the Airline Industry, *Brookings Institution*.

Meeusen, W. and Broeck, J. van den (1977), Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error, *International Economic Review*, 18:2 (June), 435-44.

Pitt, M. and Lee, L.F. (1981), The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry, *Journal of Development Economics*, 9, 43-64.

Reifschneider, D. and Stevenson, R. (1991), Systematic Departures from the Frontier:

A Framework for the Analysis of Firm Inefficiency, *International Economic Review*, 32:3 (August), 715-23.

Sarkar, J. (1999), India's Banking Sector: Current Status, Emerging Challenges and Policy Imperatives in a Globalized Environment, in *Financial Sector in India*, James Hanson and Sanjay Kathuria (eds.), Oxford University Press.

Sarkar, J., Sarkar, S. and Bhaumik, S. (1998), Does Ownership Always Matter? Evidence from the Indian Banking Industry, *Journal of Comparative Economics*, 26, 262-281.

Subramanyam, G. (1993), Productivity Growth in India's Public Sector Banks: 1970-1989, *Journal of Quantitative Economics*, 9, 209-223.

Winston, C. (1993), Economic Deregulation: Days of Reckoning for Microeconomists, *Journal of Economic Literature*, 31, 1263-1289.