CHAPTER IV

DATA BASE AND MEASUREMENT OF VARIABLE

A major issue that one encounters while analysing a banking firm is defining and measuring the inputs and outputs. Since banking is a typical service oriented industry wherein a number of products are jointly produced, measuring services provided by banks poses significant difficulties. However, it is very important to define inputs and outputs at the onset of any empirical study into the banking industry as results of the analysis may very substantially according to the choice of inputs and outputs and measurement of these variables. For example, the magnitude of measured level of efficiency and productivity depends to a large extent on the assumptions made in relation to structure of production technology, choice of input and output variables, construction of prices for different inputs and so on.

Present chapter discusses the measurement of relevant inputs and outputs and other variables used in the study and nature and sources of data. The chapter is organised into two sections. Section 4.1 discusses various theoretical issues related with modelling a banking firm while section 4.2 explains the sources of data and choice of sample.

SECTION I

4.1 MODELING A BANKING FIRM

In modeling the production structure of a multi product firm such as bank, one deals with the problem of appropriate specification of inputs and outputs. Despite the large body of literature on performance measurement of banks, there has been no consensus on how to define the inputs and outputs of a multi product banking firm. The two most widely recognized approaches in this regard are intermediation approach and production approach. The main difference between the two approaches is with regard to the treatment of deposits.
The production approach, introduced by Benston (1965), views the financial firms as typical production units which use real inputs (manpower and real assets) to produce financial services. These financial services include both deposits and earning assets, measured in terms of number of accounts. Accordingly, on cost side, this approach includes only non-interest cost while deposits along with loans and other earning assets are classified as outputs.

On the other hand, the intermediation approach, pioneered by Sealey and Lindley (1967), regards the financial firms as engaged in a multi stage production process involving intermediate outputs (deposits). To produce loans and other earning assets financial firms uses loanable funds (financial inputs), which it borrows from depositors, as well as non-financial inputs such as labour and capital.

There are further three variants of intermediation approach that have been used in literature: the asset approach, the user cost approach and the value-added approach. The asset approach is a reduced form modelling of the banking activity, focusing exclusively on the role of banks as financial intermediaries between depositors and final users of bank assets (Mohan, 2005).

In the asset approach, the researcher determines a priori the activates that can be considered as output of the bank. Since bank liabilities have some characteristic of inputs, as these provide 'raw material' for investible funds, these are considered as inputs. On the other hands, since bank assets are the ultimate users of funds and generate bulk of revenue, these are treated as outputs. (RBI, 1997). Total cost under this approach includes both operating as well as interest cost.

The user cost approach considers a financial product as an input or output on the basis of its net contribution to bank revenue (Hancock, 1985). Fixler and Zieschang, (1992) used weighted contributions of bank assets and liabilities to define bank outputs. In this way, to the extent deposits, loans, as well as other assets and liabilities contribute towards the net revenue of a bank, these may be classified as outputs. In other words, if the financial returns on an asset exceed the opportunity cost of the funds, or alternately,
if the financial costs of a liability are less than the opportunity cost, these are considered as outputs; otherwise, these are considered as inputs (Mohan, 2005).

The value-added approach provides middle ground between production and intermediation approach. It identifies major categories of produced deposits and loans as outputs because they form a significant proportion of value added. Berger and Humphrey (1991; 1992) argue that outputs should be classified according to the value addition provided by them and therefore recommend including deposits as outputs. On the cost side, both the operating costs, as well as financial costs in terms of interest paid on the liabilities are included. Hence, the value added approach can be viewed as a kind of compromise between production and intermediation approaches, wherein the dual nature of deposits is accommodated by including them as both output and input (in terms of cost associated with raising them).

In context of Indian studies, both approaches, intermediation as well as production, have been used, though intermediation is the more frequently used one. Recommending the intermediation approach in Indian scenario, Mohan and Ray (2004) observe that using deposits and loans as outputs would have been appropriate in the nationalized era when maximizing these was indeed the objective of bank but they are, perhaps, less appropriate in the reforms era. Banks are not simply maximizing deposits and loans, they are in the business of maximizing profit.

However, there are advocates for the production approach as well. In a classic study relating to economies of scale in Indian banking, Srivastva (1999) opines that if deposits are viewed as only financial inputs with no output content, it would be difficult to explain why people open bank accounts, store money in the banking system, write checks, deposit money, withdraw cash, etc. This is a lot of activity to undertake without compensation. It may perhaps be more appealing conceptually to view deposits as both outputs and inputs simultaneously.

Berger (1986) suggests that production approach may be more suitable to compare efficiency across branches of a large bank as funding costs are likely to be
same for them and exclusion of interest expense may not affect their relative efficiency. The intermediation approach on the other hand may help to compare independent financial institutions, which may face different financial and non-financial cost that may affect their efficiency. Mester (1987) also observed that majority of banking efficiency studies have followed intermediation approach.

Intermediation approach also seems more reasonable as interest expenses are a major component of total cost. Interest cost typically constitutes one half to two third of total expense. Therefore ignoring interest cost as in the case of production approach may bias the results. Owing to these arguments, this study employs intermediation approach. The intermediation approach has been used earlier in international studies by Berger and Humphery (1991), Grabowski (1993), Miller and Noulas (1996) and in Indian studies by Me.han and Ray (2003), Shanmugam and Das (2004) etc.

As stated earlier, there are three variants of the intermediation approach, out of which, the present study employs the asset approach, primarily keeping in mind the argument of Mohan ands Ray (2005). However, the traditional view on bank inputs and outputs do not account for non interest income of banks. Of late, non-interest income has become an important component of banks' total revenue and therefore deserves to be included in the outputs. Accordingly, we consider three outputs, loans and advances, investments and non-interest income, all measured in terms of value. Our input vector includes labour, fixed capital, and total funding, that is deposits and borrowings\(^1\). Price of labour is obtained by dividing total expenses on labour by total number of employees. Price of capital is obtained by dividing other operating expenses (difference between total operating expense and establishment expense) by total fixed assets. Price

\(^1\) Unlike some of the studies in Indian context (e.g. Kumbhakar and Sarkar, 2003 or Zoho, 2004), we do not distinguish between different types of deposits, e.g. demand and time deposits. This is done to keep number of inputs within reasonable limits so as to avoid the problem of dimensionality. When too many inputs or outputs are specified, there are greater number of constraints, which tend to be incomparable. As a result, efficiency may be over estimated. With in the Indian literature on bank efficiency measurement, we find both the extremes with regard to specification of inputs. While Casu (2005) specifies only one input, i.e. total cost, in order to avoid the problem of dimensionality, studies like Das and Ghosh (2005) and Kumbhakar and Sarkar (2003) specify three kinds of deposits, saving, demand and time, as separate inputs. We prefer to take a middle approach in this regard, and specify three inputs, with all kinds of deposits representing one of them.
of funds is calculated as ratio of total funding (deposits plus borrowings) to total interest expense.

SECTION II

4.2: DATA SOURCES AND SAMPLE

Data for the present study comes from various issues of the reports viz.: (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, all published by the Indian Banks' Association; and (v) Statistical Tables Related to Banks in India, published by the Reserve Bank of India.

We use bank specific data for a period of 24 years, from 1984 to 2007-08. The choice of 1984 as the starting point was dictated mainly by the availability of data. While it would have been better to include data from 1981, as the last nationalization was completed in 1980, we were not able to get data for the year 1983, forcing us to choose 1984 as the beginning point.

Further, we include banks of all ownerships on which data on all required variables was available. However, since we use a long panel, it becomes necessary to filter out the potential outliers. This becomes even more important in wake of the frontier efficiency method used in this study which is particularly sensitive to outliers. Therefore, to screen the data, two criteria are used. First, with regard to public sector, old private and foreign banks, we excluded the banks in case of which data was not available on a continuous basis for at least 10 years, while, with regard to new private banks, we excluded those banks for which continuous data was not available for at least five years. The reason behind using this condition was to rule out such banks which were either observed for very few years in our sample, or for which data was not available on a continuous basis, as such banks were more likely to prove outliers, particularly in context of a long panel. For example, some of the old private banks such as Parur Central Bank, Purbanchal Bank etc. closed down in the pre-reform period itself. Such banks often reported unexpected variation in data towards end of their operations, thus acting as potential
outliers. Further, the condition was kept lenient for the new private sector banks as these entered the industry only after 1994-95. Secondly, with regard to foreign banks, we excluded the banks which operated with five or less branches. This is done because such banks primarily operate to service the clients of their parent banks and may choose their input/output mix on considerations totally different from all other banks (Das et al., 2005). A complete list of banks used in this thesis is provided in appendix A1.

One problem in the data set we have is that till the year 1987, data is available on calendar year basis, i.e. January to December, whereas from 1988-89, the data pertains to financial years i.e. April-March. As a result, the data for 1988-89 is for 15 months, January 1988 to March 1989. A priori this does not pose a problem, as all the variables will be proportionately inflated. For example, in efficiency analysis, though the inputs and outputs will both be inflated by 25%, the ratio of inputs and outputs will be comparable to say a ratio of inputs and outputs of 12 months data.

Nevertheless, this may pose problem at some points, say e.g. when we include non-banking variables like growth rate GDP or ratio of GDP to deposits, as the ratio for 1987-88 will be inflated by 25 percent. Therefore, to correct this anomaly, we divided the data for 1987-88 by 1.25 to convert it to a 12 months basis. This has been done earlier by Bhattacharya et al. (1997).

Further, all the financial variables were converted to constant prices using the GDP deflator. Finally, at a number of points in this study, we divide the sample into pre-reform and post reform period. The cut off point in this respect is chosen to be 1991-92. That is, data from 1984 to 1990-91 is considered as pertaining to the pre-reform period while the remaining period (1991-92 to 2007-08) is classified as the post-reform period.