CHAPTER-1

INTRODUCTION

Banking is a Service Industry. Public money is the major input of the banks. Money has opportunity cost and banks leave no stone unturned to lure people to invest their money with them. To achieve this, banks must assure the public that money deposited with the banks is not only safe and secure but also multiplying. With the onset of globalization, many new financial instruments and institutions have come up. People today have multiple choice of investment. This vast availability of alternative has made the banking industry rethink and they are forced to work efficiently in order to survive and sustain themselves. Hence, the concept of efficiency is of paramount importance for the banks.

1.1. Concept of Efficiency

Efficiency is defined as the choice of alternatives which produces the largest outputs with the application of given resources or which uses the minimum inputs to produce the given outputs (McKevitt and Lawton, 1994). It measures a firm’s performance at a particular point of time in relation to the target firm i.e. the best operating firm in terms of performance (Ram Mohan and Ray, 2004). It is linked with how a bank simultaneously minimizes its cost and maximizes its revenue based on an existing level of production technology (Tandon et al., 2003; Ahmed, 2008; Kumar, 2006 and Chatterjee et al., 2014). Efficiency is supposed to be attained when a bank is not in a position to reduce the quantity of inputs to produce the same level of outputs or when a bank is unable to generate more outputs from the available level of resources (Resti, 1997). It depicts the minimum level of resources utilized to achieve the given outputs or portrays the extent of consumption of available resources to obtain the maximum output (Saha and Ravishankar, 2000). The efficiency of a firm refers to how well firm uses its resources in comparison to the current best practice firm. It is measured by comparing the actually attained or realized value against the best achievable value (Lovell, 1993). It describes how much distance exists between the quantity of inputs and outputs used by the concerned firm and the quantity of inputs and outputs used by the efficient firm. Thus the information related to efficiency is required by every firm to determine whether the set standards by the firm are achieved or not.
The efficiency of banking system is imperative for the welfare of a society as a whole when it offers innovative and quality service to society at minimum cost (Valverde *et al.*, 2003; Bader *et al.*, 2008 and Gulati, 2011b). Moreover, high efficiency in the banking system leads to better financial stability of the economy and promotes economic growth (Rajan and Zingales, 1995; Levin, 1997; Cetorelli and Gambera, 2001; Egesa, 2010; Gulati, 2011b and Pančurová and Lyócsa, 2013). If banks are fully efficient, these can have improved profitability with more funds intermediated at greater prices and thus provide exclusive services to the consumers (Berger *et al.*, 1993). Banks can take the advantage of competitive environment only if these perform efficiently in the market. Higher efficiency can lead a bank to earn higher profitability which provides safety to them to absorb huge risks (Egesa, 2010). The efficient bank can provide more trustworthy services to the consumers at optimum prices which will help to maintain faith, confidence and reliability of the customers in the banking sector (Zeitun and Benjelloun, 2013).

Bank’s efficiency analysis is imperative for various stakeholders (Ahmed, 2008). Efficiency information assists bank management to determine whether the level of efficiency achieved meets the acceptable standards or not. It can also help to evaluate the variation in the relative efficiency of individual banks against the most efficient banks (Karimzadeh, 2012). It enlightens the scope for managerial improvement by recognizing ‘best practice bank’ and ‘worst practice bank’ with the help of efficiency scores gained by them (Berger and Humphrey, 1997). Besides, efficient banks can achieve a higher rate of return with the minimum cost and at the same time participate in the economic development of the country (Zeitun and Benjelloun, 2013). The efficiency analysis can also help government to assess the effects of deregulation, mergers, or market structure on the bank’s performance (Berger and Humphrey, 1997). From the regulators’ perspective, efficiency provides a framework to quantify the health of individual banks and aids them to work out appropriate interventions to prevent systemic failures (Lacasta, 1988). The stability of financial system of an economy is also a matter of concern for regulators which can be judged from the efficiency of the banking sector (Pančurová and Lyócsa, 2013). Efficiency measures of the banking system could be leading indicators for knowing strengths or weaknesses of the banks
and enable the regulators to take pre-emptive steps when necessary (Reserve Bank of India, 2008). This information also assists investors to recognize the efficient banks to ensure reasonable and regular returns. From the point of view of customers, only efficient banks can offer better services at reasonable prices.

Additionally, the efficiency evaluation can direct an inefficient bank to recognize the areas of inefficiency so that bank can formulate suitable policies and strategies to improve its relative position in the market. Efficient banks have lower operational costs. These can steal business of inefficient banks (Kumar and Gulati, 2008). Alternatively, inefficient banks have lesser probability and face difficulty in surviving in the market (Zeitun and Benjelloun 2013). Inefficient banks are riskier and have a higher possibility of failure (Kumar and Gulati, 2008). Lastly, a bank failure can hinder financial activities of other sectors of the economy as banks are directly linked to the entire economy.

1.2. Dimensions of Efficiency

There are three main efficiency concepts for analyzing the financial performance of a bank, i.e. Revenue Efficiency, Cost Efficiency and Profit Efficiency. These three efficiency concepts are based on an economic viewpoint of the bank as these focus on financial viability of a bank in response to competition, market prices, use of technology and business conditions in which the banks are operating (Adongo et al., 2005). Fried, Lovell and Schmidt, (2008) mentioned that Cost Efficiency and Revenue Efficiency are the production performance indicators but Profit Efficiency is better measure as it incorporates both cost and revenue objective simultaneously. All these efficiency measures assist to focus on both input and output aspects of a bank along with their prices. The efficiency concepts are explained in detail as follows:

A. Revenue Efficiency (RE)

Revenue Efficiency refers to how competently a bank offers services to the customers by using its available resources. It informs about the link between the revenue generated by a particular bundle of inputs and maximum possible revenue produced from the same bundle of inputs (Berger and Mester, 1997). In other words, the actual revenue generated by a particular firm is compared with the most efficient firm
producing maximum possible revenue from similar input quantities and output prices. Revenue Efficiency measures the comparative performance of bank against the best practice bank i.e. the bank which is producing the maximum output from the available inputs. It is calculated as the ratio of revenues of a given firm to the revenues of a fully efficient firm with the given inputs. Thus Revenue Efficiency measures the ratio between current revenues and optimal revenues, given the output prices and input and output quantities (Lovell, 1993 and Sanchez et al., 2013). The distance of firm from the maximum output indicates the potential increase in outputs that the firm could achieve using the same inputs. It helps to achieve the maximum rate of increase in outputs that would be feasible for all inputs (Das et al., 2005). Revenue Efficiency is said to be achieved when a firm chooses that output combination which gives them maximum revenue at the applicable output prices.

The fully efficient firm has a Revenue Efficiency score equal to 1 while the inefficient firms have efficiency score less than 1. In order to ascertain the Revenue Efficiency, firms should focus on two components of it, i.e. Technical Efficiency (output oriented) and Allocative Efficiency (output oriented) (English et al., 1993). Technical Efficiency (output oriented) reflects the ability of a firm to maximize output from a given set of inputs whereas Allocative Efficiency reflects the ability of the firm to use the inputs to maximize the proportion of output given their respective output prices based on certain regulations. These measures are then combined to provide a measure of efficiency called Revenue Efficiency. Thus Revenue Efficiency (RE) is a multiplicative combination of Technical (TE) and Allocative Efficiency (AE), such as RE = TE * AE (Farrell, 1957). Any dissimilarity between the actual and probable revenue comes from either output-oriented technical inefficiency i.e. a firm is producing a small amount of outputs given the input quantities or output-oriented allocative inefficiency i.e. a firm is producing non-optimal combination of outputs given the output prices. Clearly the decomposition of Revenue Efficiency enables to understand what needs to be done to enhance the performance of the firm. Using output distance function, Färe, Grosskopf, and Lovell (1994) decomposed revenue inefficiency into technical inefficiency and allocative inefficiency. Firms can be revenue inefficient for the reason that they produce less output than a fully efficient firm from the same
quantity of inputs as produced by the efficient firm (output technical inefficiency) or due to the reason they choose to produce an inefficient combination of outputs given their prices (output allocative inefficiency).

- **Technical Efficiency (TE) (output oriented)** is the ability of the firm to maximize output from the given set of inputs. It evaluates the firm’s ability to obtain the maximum possible output from a given set of inputs. It refers to the ability to avoid waste by producing as much output as input usage permits (Lovell, 1993). In other words, Technical Efficiency (output oriented) involves maximizing output while keeping inputs constant. Technical Efficiency is measured as the ratio between observed output and maximum output that can be produced at efficient frontier, given the input quantities (Porcelli, 2009). Technical Efficiency (output-oriented) depicts by how much a firm can proportionally expand the output quantities without altering the input quantities. The Technical Efficiency (output-oriented) is designed to determine a firm’s potential outputs given its inputs if it is working efficiently as a firm operating on the best practice frontier. In order to survive in the competitive environment, firms strive to offer the best possible products to their customers. Therefore, they are more likely to maintain their competitive advantage by increasing their output rather than reducing the input usage.

- **Allocative Efficiency (AE) (output oriented)** evaluates the ability to maximize its profits by comparing the marginal revenue product with marginal costs of inputs. It refers to the ability to combine inputs and outputs in optimal proportion in the light of prevailing prices (Lovell, 1993). When a firm chooses the revenue maximizing mix of output, then Allocative Efficiency occurs (Chen, 2001). Given the prices of outputs, they produce output in the right proportions to maximize revenue (Allocative Efficiency). In other words, a company is efficient from the allocative point of view, when its choice of outputs is the optimum for maximizing revenues. Allocative Efficiency measures the extent to which a firm produces its outputs in a proportion that maximizes revenues, assuming that the firm is already fully technically efficient (Brázdík, 2006). Allocative Efficiency (Output Oriented) is measured as leftover component of Revenue Efficiency and is measured as ratio of Revenue Efficiency of firm to Technical Efficiency of the firm.
Technical and Allocative Efficiency (Output Oriented) concepts can be illustrated with the help of Figure: 1.1.

**Figure: 1.1 Technical and Allocative Efficiency (Output Oriented)**

In Figure: 1.1, the firm is producing a given level of outputs (Y1 and Y2) using the input (x) defined by point P assuming CRS. SS’ is the Production Possibility Curve (PPC) for the firm. A firm which lies on the curve such as Q and Q’ are fully efficient firms. P firm is inefficient as it lies below the PPC. If the inputs employed by the firm are used efficiently, the output of the firm producing at point P can be expanded radically to point Q. This shows the distance PQ represents related to technical inefficiency i.e. the amount by which output could be increased without decreasing the input. Hence, Technical Efficiency (output oriented) can be given as OP/OQ. If the output price information is given, then the iso-revenue line AA’ can be drawn. The point Q is technically efficient as it lies on the production possibility frontier and higher revenue could be achieved by producing at point Q’ because Firm Q’ is operating at the tangency point between the production possibility curve and the price line. Firm Q’ is the point of maximum revenue efficient. To achieve the same level of revenues as at point Q’ with the same inputs and output combination, the output of the firm would
need to be expanded to point \( R \). The Allocative Efficiency is measured as \( 0Q/0R \). Further overall Revenue Efficiency is the product of Technical Efficiency (output oriented) and Allocative Efficiency (output oriented) which forms Revenue Efficiency.

Revenue Efficiency = Technical Efficiency (Output) \( \times \) Allocative Efficiency (Output)

\( (0P/0R) = (0P/0Q) \times (0Q/0R) \)

Two firms with identical input and output quantities and output prices may have different Revenue Efficiency scores because one of the firms has superior product quality or more effective marketing strategies. Revenue Efficiency can be attained if the firm uses best practice technology and chooses the optimal mix of outputs. In other words, RE is achieved by means of the best practice technology choosing the most favorable mix of outputs. A bank can improve its Revenue Efficiency by offering quality services in addition by avoiding improper input-output quantity mix and mispricing of output prices (Rogers, 1998).

**B. Cost Efficiency**

Cost Efficiency depicts by how much a firm can reduce its cost while producing the same amount of goods and services. It is defined as the effective choice of inputs together with input prices which aims to minimize production cost. Cost Efficiency measures the relative performance of the firm as against the best practice firm which is managing its operating cost at the lowest for producing the same output under the similar technological conditions as faced by the concerned firm. It helps to know how close a firm’s cost is to what best practice firm’s cost would be for producing the same level of outputs (Weill, 2004). Cost Efficiency of firm suggests possibility to trim down the cost further so that the firm can operate efficiently at the minimum cost. It is the ratio of the cost of a fully efficient firm with the same output quantities and input prices (i.e. firm operating on the efficient cost frontier) to the given firm’s actual costs, i.e. the firm whose efficiency is calculated (Cummins et al., 2010). It can be written as the ratio of minimum cost to observed cost. Under the cost frontier, i.e. Cost Efficiency, the actual cost expended in producing a particular bundle of outputs is compared to the minimum cost necessary for producing that same bundle.
The Cost Efficiency score of the given firms can vary between 0 and 1. Firm achieving efficiency value of 1 is said to be the most efficient firm while firm having a score of less than 1 is supposed to be inefficient firm. A firm is considered to be inefficient in cost when with the given input prices and inputs-outputs quantities, it does not reach its minimum level of costs. Farrell (1957) propound that the Cost Efficiency (CE) of a firm consists of two components, i.e. Technical Efficiency (TE) and Allocative Efficiency (AE). The Cost Efficiency is a multiplicative combination of Technical Efficiency (TE) and Allocative Efficiency (AE) such as \( CE = TE \times AE \) (Farrell, 1957). Technical Efficiency (input oriented) relates to quantities of inputs while Allocative Efficiency (input oriented) relates to prices of inputs (Barros and Mascarenhas, 2005). This decomposition helps to know the firm’s cost inefficiency using input distance function. To be fully cost efficient, a firm must be both technically and allocatively efficient (Coelli, 1996). The segregation of Cost Efficiency into technical and allocative components helps to know the sources of cost inefficiency. Technical inefficiencies arise due to inadequate technologies or deficiency in the adoption of technology or due to use of less productive input factors while the allocative inefficiencies are due to a suboptimal allocation of input factors.

- **Technical Efficiency (TE) (input oriented)** scrutinizes how well a firm converts inputs into outputs. Technical Efficiency (TE) refers to the capability of a firm to produce existing level of output with the minimum quantity of inputs (input-oriented). It measures the proportional reduction in input usage required if the firm operates on the efficient frontier i.e. if a firm become fully efficient. Technical Efficiency for a given firm is defined as the ratio of the input usage of a fully efficient firm (firm operating on the efficient frontier) producing the same output to the input usage of the firm under consideration. To be fully efficient in Technical Efficiency firm needs to use its resources to the fullest extent without wasting any of these resources. A fully efficient firm has the Technical Efficiency score of 1 while the inefficient firm has scores between 0 and 1. The technical inefficiency is caused due to the managerial decisions and only management can correct it (Hassan, 2005). Further to know the causes of technical inefficiencies, it can be further decomposed into Pure Technical Efficiency (PTE) and Scale Efficiency (SE). Pure Technical Efficiency (input oriented) is the proportional reduction in input usage to produce the
given output, i.e. if the inputs are not wasted and Scale Efficiency explicitly is the proportional reduction if the firm achieves Constant Return to Scale (CRS). Scale Efficiency is based on an optimal operation scale, namely the output level that minimizes the average costs.

- **Allocative Efficiency (AE) (input oriented)** reveals the ability of a firm to choose the inputs in optimal proportions when their respective prices are given. In other words, it describes whether the firm is using the right mix of inputs in the light of relative price of each input. It measures the firm’s success in choosing the cost minimizing combination of inputs. A firm is said to be allocative efficient if it is using its inputs in the proportion which minimizes its cost. Allocative Efficiency is related to the regulatory environment or macroeconomic conditions (Lovell, 1993). This highlights that Allocative inefficiency is caused by regulations, rules and policies and it might not be corrected and controlled by the management (Hassan, 2005). The price or Allocative Efficiency is calculated as a residual component of the Cost Efficiency of the firm and is obtained from the ratio of Cost Efficiency score to Technical Efficiency scores i.e. \( AE = CE / TE \).

These concepts can be illustrated with the help of the Figure: 1.2.

**Figure: 1.2 Technical and Allocative Efficiency (Input Oriented)**

![Graph showing Technical and Allocative Efficiency](image-url)
Figure: 1.2 permits the measurement of technical and allocative (input oriented) efficiency. This concept is explained using two inputs (x1 and x2) which are used to produce the single output (y). SS’ is the technically fully efficient iso-quant representing fully efficient firms. It shows all possible combinations of inputs to produce an output if the firm is perfectly efficient. AA’ is the budget line or input price line. If a firm uses quantities of inputs to produce a unit of output as represented by point P, the technical inefficiency of the firm is represented by QP. QP represents the amount of input that could be reduced without reducing the output. When it is represented in the ratio it becomes QP/0P which shows by how much inputs could be reduced. Technical Efficiency (TE) of a firm is measured by the ratio of inputs needed at best practice to produce given outputs relative to observed input quantities as 0Q/0P. Q lies on iso-quant curve, so it is technically efficient. If the input price line, i.e. AA’ is known then Allocative Efficiency can be calculated. Allocative Efficiency is the cost of producing observed output at observed factor prices, assuming Technical Efficiency relative to minimize costs at the frontier. Firm R and Q’ are allocatively efficient as they are on price line. Allocative Efficiency of the firm P is defined as ratio of 0R/0Q. If the firm wants to be allocatively and technically efficient, i.e. at point Q’ where the input price line is tangent to the iso-quant SS’, it has to reduce production cost i.e. RQ. The total Cost Efficiency is defined as 0R/0P where the distance RP can be interpreted in terms of cost reduction.

Cost Efficiency = Technical Efficiency (IO) × Allocative Efficiency (IO)

\[(0R/0P) = (0Q/0P) \times (0R/0Q)\]

Thus the product of Technical Efficiency (input oriented) and Allocative Efficiency (input oriented) provides Cost Efficiency. A firm is said to be fully efficient in terms of Cost Efficiency when it uses inputs that are necessary to produce the given outputs (Technical Efficiency) and with the given input prices, they are unable to minimize the cost proportions further (Allocative Efficiency). Cost Efficiency measures the overall possible reduction in costs that can be achieved if a firm is technically and allocatively efficient. A firm can achieve full Cost Efficiency by becoming technically efficient i.e. by using its inputs in the most advantageous way and by adopting the best
cost minimization mix of inputs through considering input prices (Cummins and Rubio Misas, 2001).

C. Profit Efficiency (PE)

Profit Efficiency is related to the economic goal i.e. profit maximization (Berger and Mester, 1997). It is balanced measure of efficiency as it takes into account both banks’ cost and revenue issues simultaneously (Berger and Mester, 1997 and Maudos and Pastor, 2003). Profit Efficiency is calculated as the ratio of actual profits generated by a firm to maximum profits produced by the best firm with given inputs, outputs and their concerned prices. In other words, the actual profit generated by a particular bundle of inputs and outputs is compared to the maximum possible profits that can be produced from the same bundle of inputs and outputs. It is associated with assessing how close a firm is to a firm generating maximum possible profits considering both input and output prices. If the actual profits earned by the firm are closer to the maximum attainable profits then the firm has higher efficiency. Profit Efficiency measures the ratio between current profits and the most advantageous profits. Profit Efficiency is said to be achieved when a firm chooses that output and input combination which gives them maximum profits at the applicable input-output prices. A profit efficient firm maintains a production route at which lowest cost is incurred to produce maximum revenues known the input and output prices (Maudos et al., 2002).

Profit Efficiency depicts the ability of the firm to attain maximum achievable profits. It provides complete vision of the firm to various stake holders in comparison to Cost and Revenue Efficiency which provide a partial view only. Besides, in this profit maximization era, merely reducing cost cannot make the firms efficient in terms of profit; in addition they need to focus on revenue side as well. Furthermore, Profit Efficiency can help to identify whether the firm is facing revenue incompetence or cost inadequacy. Directional distance function is introduced to measure Profit Efficiency by computing the directional distance of actual input and output quantities from the optimal input and output quantities (Fried et al., 1993). Like Cost and Revenue Efficiency scores, Profit Efficiency can be decomposed into Technical Efficiency and Allocative Efficiency
by using this directional distance function (Chambers et al., 1998), but the decomposition is to some extent arbitrary due to the reason that Technical Efficiency could be either input oriented or output oriented (Kumbhakar and Lovell, 2003). Färe et al., (1994) suggested that direction distance function can be used to solve this problem. Under this approach, one need not make illogical choice of input oriented or output oriented approach. Direction Distance Function requires to specify the appropriate direction of the firm. As a result, to measure the Profit Efficiency, non-oriented approach is used which allows both for increase and decrease of inputs and outputs so as to exploit the input-output prices. This entails to optimize the performance of the bank by either fairly lessening the inputs by producing the given outputs or by enhancing the outputs proportionately without reducing the usage of available inputs. Therefore Technical Efficiency reflects the capability of a bank to achieve utmost output from the least amount of possible input use whereas Allocative Efficiency indicates the capacity to use the inputs and outputs in optimal proportions, given their respective prices and the production technology. This concept can be explained with the help of Figure: 1.3:

**Figure: 1.3 Profit Efficiency Concept**

![Profit Efficiency Concept](source: Coelli et al., 2005)
Figure: 1.3 represents a production process in which a single input (x) is used to produce a single output (y) with a curve that represents the production frontier. Firm D is inefficient as it lies inside the frontier and is producing y1 output with x3 input. D must contract the input in order to produce the given output as firm A is producing y1 output with x1 input. On the other hand, D can expand the output with the given available input as Firm C is producing y3 output with x3 input. However Firm A and C are technically efficient, as they both lie on the frontier but are not tangent to the maximum profits and they both are not allocatively efficient. Firm B is fully efficient as it is lying on the frontier as well as it is tangent to the maximum attainable profits.

The Profit Efficiency can be of two types, standard and alternative profit efficiency. The difference in these efficiency scores is due to various market forces taken into consideration. Berger and Mester (1997) describes that Alternative Profit Efficiency is used when firms have certain power to set output prices whereas Standard Profit Efficiency is used when firms behave as price takers in both input and output markets and have no control over the prices.

**Technical Efficiency (Input Oriented and Output Oriented)**

Technical Efficiency enlightens firm’s capability to obtain the maximum possible output from a given set of inputs or firm’s potential to minimize the usage of its input to produce the given set of outputs. It refers to the ability to avoid waste by producing as much output as input usage allows, or by using as little input as output permits (Lovell, 1993). Generally, Technical Efficiency investigates how well a firm implements an effective production plan to convert its inputs to outputs (Avkiran, 2000). Technical Efficiency (output oriented) is measured as the ratio between the observed output and the maximum output, under the assumption of fixed input. Alternatively, Technical Efficiency (input oriented) is the ratio between the minimum input and the observed input under the assumption of fixed output (Porcelli, 2009). A particular method of producing the output is technically efficient if there are no other ways of producing the given output with the least inputs or producing maximum outputs with the given inputs. Technical Efficiency concept can be explained with the help of Figure: 1.4:
Figure: 1.4 Technical Efficiency

Figure: 1.4 depicts a simple production process in which a single input (x) is used to produce a single output (y). The curve 0f represents the production frontier, which is the maximum output attainable from each input level. All points A, B, C between the production frontier and the x-axis form the feasible production set. If a firm’s actual input-output point lies on the frontier, it is perfectly efficient. If it lies below the frontier, then it is inefficient. So, a firm is said to be technically efficient if production occurs on the boundary of the production frontier 0f and technically inefficient if production occurs in the interior of the production frontier 0f. Point A represents an inefficient point, whereas points B and C represent efficient points. A is inefficient because with the same input, it can attain higher output as B or it could reduce the input to the level associated with the point C without reducing any output production. Technical Efficiency (input oriented) is beneficial to recognize if the inputs are costly to the firm. Achieving full Technical Efficiency (input oriented) means producing specified outputs at the least cost. Technical Inefficiency on input basis describes that costs can be reduced by reducing some inputs but not increasing any others. Technical Efficiency (output oriented) scores depict whether a firm can produce
more outputs from available inputs. Accomplishing Technical Efficiency means producing maximum outputs from given set of inputs i.e. outputs could be increased without reducing the inputs. Thus the analysis of Technical Efficiency can have an output augmenting orientation or input conserving orientation.

The above efficiency score evaluation assumes that a firm’s efficiency measure is based on Constant Return to Scale (CRS) assumption, but this assumption does not always hold. A firm might operate at a scale of operation that is not optimal i.e. a firm may operate within the Increasing Return to Scale (IRS) or within the Decreasing Return to Scale (DRS). In other words, the firm might be operating under the assumption of Variable Return to Scale (VRS). This further helps to decompose Technical Efficiency into two parts, Pure Technical Efficiency (PTE) and Scale Efficiency (SE) (Coelli et al., 2005). Pure Technical Efficiency illustrates the capability of firm in using its most advantageous level of inputs to produce maximum outputs or otherwise its competence to produce utmost outputs with the minimum level of resources. On the other hand, Scale Efficiency refers to the ability of the firm to operate at optimal scale. Thus the nature of technical inefficiencies can be due to the inefficient implementation of the production plan in converting inputs to outputs (pure technical inefficiency) or due to the divergence of firm from the most productive scale size (scale inefficiency). Inaccuracy of management in using the optimal level of input leads to Pure Technical inefficiency, while wrong choice of firm scale or size results in scale inefficiency.

Thus decomposing Technical Efficiency permits us to gain insight into the main sources of inefficiencies (García Sánchez, 2009). Both PTE and SE are bound to be either 0 or 1 or lie in between them. The firm having the Pure Technical Efficiency score of 1 confirms that the firm is operating on variable return to scale frontier and the Scale Efficiency score equal to 1 indicates that a firm is operating with Constant Return to Scale (CRS). Figure: 1.5 depicts the different Pure Technical Efficiency and Scale Efficiency.
Figure: 1.5 Technical Efficiency, Pure Technical Efficiency and Scale Efficiency

Figure: 1.5 shows that B and C are efficient firms, but the firm C is fully efficient. This is due to the reason that a ray from the origin is tangent to production frontier and hence defined as the point of maximum possible productivity. C firm is an example of exploiting scale economies. A firm operating at point C is located on the Constant Return to Scale frontier, i.e. it is operating at the most productive scale size or at the technically optimal productive scale and cannot be more productive. Point B is on the production frontier and is efficient in terms of Pure Technical Efficiency as a firm is operating on the variable return to scale of operation. It can still improve its product by exploiting scale economies, i.e. can operate on Constant Return to Scale. Since, B could produce more output per unit of input by adopting the scale employed by the efficient firm C. Scale Efficiency measure can be used to indicate the amount by which productivity can be increased by Firm B by moving to the point of technically optimal productive scale. Therefore, a scale inefficient firm may be efficient, but it can still exploit scale economies to improve its productivity (Coelli et al., 1998).

Return to Scale

The efficiency measurement with Constant Return to Scale (CRS) assumes that all firms have operated with optimal scale and minimize their scale inefficiency to zero.
(Yasmeen, 2011). On the other hand, the Variable Return to Scale (VRS) makes it possible to determine how the scale of operation hinders the firm’s performance. This allows decomposing of the Technical Inefficiencies into Scale Inefficiencies and Pure Technical Inefficiencies. VRS reveals that production technology may depict Increasing, Constant and Decreasing Return to Scale. In CRS model, one can assume that the scale of economies does not change as the size of the organization unit increases, whereas in VRS model, one can assume that scale of economies change as the size of the organization unit increases. CRS efficiency scores will never be higher than that of VRS efficiency scores. The CRS model estimates the gross efficiency of a DMU. This efficiency comprises Pure Technical Efficiency and Scale Efficiency. Pure Technical Efficiency describes the efficiency in converting inputs into outputs, while Scale Efficiency recognizes that economy of scale cannot be attained at all scales of production and there is one most productive scale size. The VRS model takes into account the variation of efficiencies with respect to scale of operation. CRS and VRS concept can be explained with the help of Figure: 1.6:

**Figure: 1.6 Constant Return to Scale (CRS) and Variable Return to Scale (VRS)**
Figure 1.6 illustrates the shape of envelopment surfaces for a single input and single output under CRS and VRS Models. Four points, i.e. A, B, C, and D are used to estimate the efficient frontier under both scale assumptions. The Constant Return to Scale frontier is simply a ray from the origin 0f through point C assuming that all firms are operating at an optimum scale. CRS allows benchmarking of DMUs with similar size DMUs whereas VRS model ignores the assumption of firm operating at an optimum scale and introduces a convexity condition to the basic model (Coelli et al., 1998). Thus, points A, C, D represent the VRS model. Firm C is assigned 100% efficiency in case of the CRS assumption, whereas A, C and D are considered to be 100% efficient in case of VRS assumption. This indicates that inefficiencies assigned to Firm A and D in case of CRS assumption is purely due to their scales of operations. If DMUs lie on the Production frontier, then they are efficient in terms of Pure Technical Efficiency, but none of them is considered overall technically efficient because they operate on an inefficient scale. From the above figure following conclusions can be drawn:

i. C is fully efficient as per VRS and CRS.

ii. A and D are efficient for VRS i.e. Pure Technical Efficiency, but not for CRS. The difference between them is due to scale of operation.

iii. B is fully inefficient.

CRS gives accurate results only as long as firms are operating at an optimal scale (Coelli et al., 2002). The Constant Return to Scale (CRS) is advantageous as it allows for comparison between small and large firms in a situation where the frequency distribution is skewed due to the presence of small and large firms in the sample. In such a situation, the use of Variable Return to Scale (VRS) raises the possibility that large firms would appear as efficient in the sample for the simple reason that there are no truly efficient firms (Berg et al., 1991).

VRS includes both Increasing and Decreasing Return to Scale. Increasing Return to Scale (IRS) exists when the proportional increase in all inputs results in a more than proportional increase in outputs. Decreasing Return to Scale (DRS) exists when an increase in inputs results in a lower increase in outputs. This can be understood taking single input x and single output y with the help of Figure 1.7:
1. Firm A is considered to be inefficient whereas firm B and C is efficient as they are on the production frontier. If a firm moves from point A to point B, it becomes technically efficient by producing more output with the same level of inputs. As the slope of ray OB is higher than that of the ray OA, it indicates higher levels of productivity of the firm. Firm B can still improve its productivity by moving from point B to point C as the slope of OC is higher than OB. This movement of firm from point B to point C is considered as productivity gain from exploiting the return to scale, i.e. reducing its output size to move from Decreasing Return to Scale to Constant Return to Scale.

Similarly, a firm producing at point E can improve its productivity by moving from point E to point D i.e. technically efficient frontier point. Likewise firm D can attain higher productivity by moving from point D to point C as the slope of OC is higher than OD. To achieve that point, Firm D can increase its output size by moving from Increasing Return to Scale to Constant Return to Scale.

1.3. **Need of the Study**

Liberalization, Privatization and Globalization (LPG) reforms started in India in 1991 welcomed Private and Foreign Sector Banks to operate in Indian Banking
Industry. The reforms introduced made sea changes i.e. from fully regulated environment, Indian banks gradually moved into a market driven competitive system. Moreover, increasing competition in banking sector fragmented the share of profitability of banks and forced them to work efficiently. In such a stiff competition, economies of scale and scope had to be exploited in order to sustain in the market.

The series of banking reforms as entry deregulation, branch de-licensing, dismantling of administrated interest rate structure, reduction in CRR and SLR and adoption of prudential norms etc were introduced in the Indian Banking Sector with an intention to improve the efficiency of banks. Efficiency of banking system is one of the most important issues in the financial market as it can affect the stability of banking industry and ultimately the effectiveness of whole monetary system (Yılmaz, 2013). All stakeholders are more concerned about how economically and efficiently banks convert their expensive inputs into valuable financial products and services (Isik and Hassan, 2002a). The prologue of various banking reforms and deregulations in Indian Banking sector in early 1990s encouraged the researchers and various academicians to examine the performance of banks in terms of various efficiency measures.

With particular reference to India, studies on efficiency performance of Indian Scheduled Commercial Banks reveal that less number of researchers have considered the economic objectives of Cost Efficiency, Revenue Efficiency and Profit Efficiency to evaluate the performance of Indian Banks. Only handful of studies as Chatterjee and Sinha (2006), Sahoo et al. (2007), Kalluru and Bhat (2009), Kaur and Kaur (2010), Gulati (2011b), Kumar (2013) and Raina and Sharma (2013) specifically evaluate the Cost Efficiency of banks. Only one study was found that exclusively evaluated the Revenue Efficiency of Indian Banks (Ram Mohan and Ray, 2004). Although, Das and Ghosh (2009) and Ray and Das (2010) evaluated the Profit Efficiency of Indian Scheduled Commercial Banks along with Cost Efficiency, but none of the studies exclusively evaluated the Profit Efficiency of Indian Scheduled Commercial Banks. In addition the review of literature with specific reference to India, suggests that a small number of studies determined the factors affecting Profit Efficiency (Ghosh, 2009 and Das and Ghosh, 2009) and Cost Efficiency (Gulati, 2011b and Raina and Sharma, 2013)
while none of the studies has evaluated the factors affecting Revenue Efficiency of Indian banks.

Thus, the above information makes it imperative that all the three efficiency concepts and factors affecting efficiency be studied in detail.

1.4. Objectives of the study

Based on the above observations the main objectives of the proposed study are:

1. To analyze and evaluate the Revenue Efficiency scores of Scheduled Commercial Banks (SCBs) operating in India. In addition, Revenue Efficiency is analyzed across bank ownership. The study also determines the nature of Return to Scale (RTS) of banks and identifies the number of banks operating as leaders and laggards in India.

2. To analyze and evaluate the Cost Efficiency scores of Scheduled Commercial Banks (SCBs) operating in India. In addition, Cost Efficiency is analyzed across bank ownership. The study also determines the nature of Return to Scale (RTS) of banks and identifies the number of banks operating as leaders and laggards in India.

3. To study Profit Efficiency of Scheduled Commercial Banks (SCBs) operating in India vis-à-vis Revenue and Cost Efficiency. In addition, Profit Efficiency is analyzed across bank ownership.

4. To identify the factors i.e. Bank Specific Factors, Industry Specific Factors and Economy Specific Factors affecting the Revenue, Cost and Profit Efficiency of Scheduled Commercial Banks (SCBs) operating in India.

1.5. Organization of Chapters

This study is divided into eight chapters organized as follows:

Chapter 1 introduces the topic, highlights the need of the study and defines objectives.

Chapter 2 reviews the literature available on the efficiency performance of the banks. Reviewing of literature helped us to discover areas that need further exploration with respect to efficiency evaluation of banks.
Chapter 3 explains the sample, time period and methodological framework used to measure Revenue, Cost and Profit Efficiency of the banks. It also explains the variables used and techniques adopted for the study.

Chapter 4 evaluates the Revenue Efficiency (RE) scores of all Indian Scheduled Commercial Banks. In addition, Revenue Efficiency across bank ownership has been studied to verify which sector is performing better. The nature of Return to Scale of all Scheduled Commercial Banks is analyzed in this chapter. Further the nature of Return to Scale (RTS) of banks across different ownership is analyzed. An endeavor has also been made to determine and identify the number of Indian Scheduled Commercial Banks operating as Leaders and Laggards in India according to Revenue Efficiency and its components scores.

Chapter 5 estimates the Cost Efficiency (CE) scores of all Indian Scheduled Commercial Banks. Additionally, Cost Efficiency across bank ownership has been studied to verify which sector is performing better. The nature of Return to Scale of all Scheduled Commercial Banks is analyzed in this section. Further the nature of Return to Scale (RTS) of banks across different ownership is analyzed. An Endeavour has also been made to determine and identify the number of Indian Scheduled Commercial Banks operating as Leaders and Laggards in India according to Cost Efficiency and its components scores.

Chapter 6 assesses the Profit Efficiency (PE) of all Indian Scheduled Commercial Banks vis-à-vis Revenue Efficiency (RE) and Cost Efficiency (CE). In addition, Profit Efficiency across bank ownership too has been studied.

Chapter 7 determines the impact of bank specific, industry specific and economy specific factors affecting the Revenue, Cost and Profit Efficiency of Indian Banking Sector.

Chapter 8 summaries the study along with the recommendations of the study and outlines the future scope of research.
1.6. Limitations of the study

No study is free from limitations and the present study is no exception to the phenomenon. Input-Output selection is based on a researcher’s judgment. Thus input-output selection decision becomes major limitation of the study. A Researcher’s judgment may bring bias in selection of inputs and outputs and lead to varied results. Furthermore, the study considers only the quantitative aspect of efficiency of banks while qualitative aspect is ignored. Banking being a service industry is influenced by many qualitative factors such as customer satisfaction, customer retention, customer loyalty, customer relationship management and e-CRM etc. Last but not the least the study suffers from limitations associated with using secondary data.