The cost of inefficiency: an empirical study into the relative efficiency of a financial institution's branches

David Murray
David Tripe*
Centre for Banking Studies
Massey University
Private Bag 11-222
Palmerston North
New Zealand

Abstract

This paper goes beyond standard ratio analysis to apply data envelopment analysis (DEA) to evaluate the relative efficiency of the branches within the network of a New Zealand non-bank financial institution, to discover what makes some branches more efficient than others, and identify ways of improving the identified X-inefficiencies. To the authors' knowledge, this is the first reported research to apply DEA in an evaluation of a network of decision making units (branches) of a New Zealand financial institution.

Under output maximisation optimisations, material although not extensive improvements are identified in the inputs and outputs employed by branches in the production and intermediation processes. With the exception of a limited number of branches, scale effects do not appear to play a major role in identified inefficiencies.

Keywords

Financial institution efficiency, Data Envelopment Analysis, New Zealand

* Corresponding author: E-mail D.W.Tripe@massey.ac.nz; phone +64 6 350-5799 ext 2337

1. Introduction

Ideally, financial institutions that are more efficient should benefit their shareholders, customers, and society in general. Shareholders benefit from higher profitability translating into greater returns on their investment. Customers benefit by receiving high-quality services at reasonable prices. Society benefits from the improved soundness of the financial system and a reduced cost of financial intermediation.

The subject of financial institution (FI) efficiency has been widely discussed in the literature. In the most recent major survey, Berger and Humphrey (1997) identified 130 studies, covering 21 countries, from multiple time periods, and of various types of depository FI. The primary focus has been on the corporate performance of individual banks, with fewer bank branch-specific studies being undertaken. This is surprising, given that for most banks the branch is still the primary distribution channel for their retail products and services.

This paper seeks to add to the literature on FI efficiency in New Zealand. This is the first reported research using data envelopment analysis (DEA) to evaluate the efficiency of the branches of a FI operating in New Zealand. It aims to study the relative efficiency of the branches within the network of a co-operative non-bank financial institution; discover what makes some branches more efficient than others; and identify ways of improving efficiency and thereby profitability. The name of the financial institution and the identity of its branches are not disclosed.

Whilst standard ratio analysis has been the tool traditionally employed to measure efficiency, ratios tend to ignore the multidimensional aspects of bank performance (Mukherjee et al, 2002). As DEA measures of efficiency are based on simultaneous consideration of multiple inputs and multiple outputs (Thanassoulis et al, 1996), it seems an appropriate tool for the purposes of this study.

2. Review of the Literature

2.1 Defining Efficiency

Efficiency and productivity are core concepts of economics. Farrell (1957) identified economic efficiency as comprising both technical and allocative efficiency. Technical efficiency refers either to a firm's ability to maximise its output for a given level of input, or minimise inputs for a given level of output, whereas allocative efficiency refers to a firm's ability to use inputs in optimal proportions given their prices. The efficiency measure of Farrell was formulated into a mathematical programming framework (commonly called DEA) by Charnes, Cooper, and Rhodes (hereafter referred to as CCR) (1978), with subsequent modification by Banker, Charnes, and Cooper (BCC) (1984), among others.

Charnes & Cooper (1985) defined efficiency in the following terms "100% efficiency is attained for a unit only when: a) none of its outputs can be increased without (i) increasing one or more of its inputs, or (ii) decreasing some of its other outputs; b) none of its inputs can be decreased without either (i) decreasing some of its outputs, or (ii) increasing some of its other inputs" (p. 72). An efficiency score of less than 100% (one) indicates relative inefficiency.

Efficiency can generally be considered in terms of scale efficiency, scope efficiency, and X-efficiency "Scale efficiency measures whether banks are operating with an efficient level of outputs; scope efficiency measures whether banks are operating with an efficient mix of outputs; and X-efficiency focuses on whether banks are operating with an efficient mix of inputs" (Liu and Tripe, 2002, p. 63).

Berger (1993, p. 264) defines X-efficiency, consistent with that originally given by Leibenstein (1966), as "…. the ratio of the minimum costs that could have been expended to produce a given output bundle to the actual costs expended, and varies between 0 and 100 percent. X-efficiency includes both technical inefficiency, or errors that result in general overuse of inputs, and allocative inefficiency, or errors in choosing an input mix that is consistent with relative prices."

This research focuses on X-efficiency as it represents managements' ability to control costs and employ resources to generate outputs. Although X-efficiency incorporates both allocative and technical efficiency, separation of price and quantity data would not always have been meaningful, and this research thus looks only at technical efficiency.

X-inefficiencies have been shown to be more important in determining overall firm and market performance than scale or scope inefficiencies (Berger & Humphrey, 1991).

2.2 Efficiency in Financial Institutions

Despite the large amount of previous research on FI, relatively little has been undertaken in a New Zealand or Australian context. This may be attributed to the small number of banks and the general lack of readily accessible bank-specific data, which makes it difficult to undertake econometric analysis. There is however a body of research on the efficiency of non-bank financial institutions in Australia, although the focus is on the relative corporate performance of the FIs rather than branch-specific efficiency studies.

Studies undertaken within Australasia have focused on such issues as Avkiran's (1999c) measurement of operating efficiencies, employee productivity, profit performance, and the average relative efficiency of Australian trading banks between 1986 and 1995. Liu and Tripe (2002) explored the efficiency impacts of bank mergers in New Zealand between 1989 and 1998. Sathye (2001) investigated the X-efficiency in Australian banks using data from 1996. Tripe (2003) explored the efficiency of New Zealand Building Societies within a window analysis framework.

While there is no consensus in the literature on the best way to measure efficiency in banking, the majority of studies employ some form of frontier analysis. Berger and Humphrey (1997) suggest that the essence of frontier analysis is its sophisticated method of benchmarking the relative performance of decision making units (DMUs[1]), such that managers can employ it to improve their performance through identifying "best" and "worst practices". It is equally effective when assessing DMUs within an industry and DMUs within a firm.

At least five different frontier approaches to evaluating efficiency have been employed in the literature, three of which are parametric and two non-parametric. The three parametric frontier approaches are the stochastic frontier approach (SFA), the distribution-free approach (DFA), and the thick frontier approach (TFA). DEA and free disposal hull (FDH) are non-parametric approaches.[2] The main econometric problem is in separating X-efficiency differences from random error. Each of the parametric approaches specifies a functional form for their frontiers, but differs in the assumptions made in disentangling random error from X-efficiency.

While the non-parametric approaches suffer from the assumption that there is no random error, they also place little structure on specifying the best-practice frontier. Not assuming a particular production technology or correspondence is an important advantage of DEA because it means that a DMU's efficiency assessment is based on actual observed performances rather than attempting to relate performance to statistical averages that may or may not apply to that DMU (Avkiran, 1999a).

The DEA frontier consists of the piece-wise linear set of best-practice DMUs, with the best-practice DMUs being those observations "for which no other decision making unit or linear combination of units has as much or more of every output (given inputs)" (Berger and Humphrey, 1997, p. 177). As DEA assumes that there are no random fluctuations, all deviations from the best-practice frontier represent inefficiency (Ferrier and Lovell, 1990).

---

[1] The term DMU is used in frontier efficiency studies to describe the branch, bank, firm, or department within a firm being analysed. "Generically a DMU is regarded as the entity responsible for converting inputs into outputs and whose performances are to be evaluated" (Cooper et al, 2000, p. 22).

[2] FDH is arguably a special case of DEA.

DEA was selected as the most appropriate measure of FI branch efficiency for this study on the basis that:

- its use for this purpose is well established in the literature. As reported by Berger and Humphrey (1997), 11 of the 13 reported bank branch studies applied DEA;
- proprietary data is available, leveraging off DEA's major strength of working with non-standardised variables of differing units (Avkiran, 1999c);
- its ability to consider explicitly the use of multiple resources to provide multiple services while comparing branches (Sherman and Ladino, 1995); and
- its ability to be used with small sample sizes (Evanoff and Israilevich, 1991).

DEA has been applied extensively in the literature to measure the efficiency of bank branch networks. Some examples include Sherman and Gold (1985); Oral and Yolalan (1990); Al-Faraj et al (1993); Drake and Howcroft (1994); Sherman and Ladino (1995); Golany and Storbeck (1999); Zenios et al (1999); and Athanassopoulos and Giokas (2000).

These studies all employ data of one year or less, with the majority of the branches in these studies found to be either 100% X-efficient, or close to it. Berger et al (1997) suggest that this result could reflect great management, or more likely, a problem with the DEA models employed. Athanassopoulos (1998) also identified the modelling problem by suggesting that the limited number of observations prevents effective discrimination between efficient and inefficient branches.

This paper has tried to mitigate the problems of previous studies by developing an appropriate DEA data set, identifying the correct mix of inputs and outputs, and applying the most appropriate DEA models in the analysis.

3. Method

DEA is a mathematical programming technique which can be applied to assess the relative efficiencies of homogeneous DMUs in converting multiple inputs into multiple outputs. It compares the observed inputs and outputs for all DMUs, identifies the best-practice units to define an efficient frontier, and then measures the efficiency of the other DMUs relative to this frontier using a mathematical algorithm. DEA also measures the "slacks" in each of the inputs and outputs of inefficient DMUs to identify where improvements can be made.

The basic (multiplier form of the) DEA problem, in the constant returns to scale version, can be expressed as a requirement to maximise efficiency, for output weights u and input weights v, for i inputs x and j outputs y (with bold to indicate vectors). If we set the weighted sum of inputs as 1, in mathematical notation this gives us a requirement to

$$\max_{uv} (\mathbf{u}y_j)$$

$$\text{st} \quad \mathbf{v}x_i = 1$$

$$\mathbf{u}y_j - \mathbf{v}x_i \leq 0$$

$$u, v > 0$$

As inputs and outputs in DEA can be continuous, ordinal, or categorical variables, measurable in different units, DEA provides considerable flexibility in data selection (Nyhan and Martin, 1999). A DEA data set is the group of DMUs and the values of their inputs and outputs included in the analysis.

3.1 Sample Selection

While there is no clear outline in previous research explaining how a sample of DMUs should be assembled for a DEA study, Dyson et al (2001) argue that DEA requires a data set of relatively homogeneous DMUs, with similar operational goals and characteristics. Moreover, to effectively discriminate between efficient and inefficient DMUs, the sample size should be larger than the product of the number of inputs and outputs (Dyson et al, 1998), or at least three times larger than the sum of the number of inputs and outputs (Stern et al, 1994).

When detailed proprietary data is available, it is common practice to group this data into smaller sets of similar categories which can then be made operational (Berger and Humphrey, 1997). Sherman and Gold (1985) created a data set with 17 of the most common services offered at 14 bank branches under their study. Recognising that their sample size was too small for DEA to effectively discriminate between efficient and inefficient DMUs, they reduced the number of service categories to four on the basis of complexity and resources required to complete each transaction.

Oral and Yolalan (1990) selected the first 20 bank branches of the 583 operating at a bank in Turkey using a scoring system they devised to account for the homogeneity of the branch operations. Berger and Humphrey (1997) suggest that some form of statistical test would be a more appropriate way to do this.

For this study, all 28 branches of the financial institution are included in the data set. This provides a larger sample than in some other studies reported (Sherman and Gold, 1985; Oral and Yolalan, 1990; Vassiloglou and Giokas, 1990). While the branches are of varying sizes, each location offers the same services and uses the same resources, and the branches are therefore considered homogeneous.

3.2 The Organisation under Study

Although the financial institution is not registered as a bank, its primary function is to provide banking services. As a co-operative, its members (who are its customers) both supply the capital and consume the output. The benefits of ownership take the form of lower fees and more favourable interest rates for customers. Berger and Humphrey (1997) noted that customers (owners) of mutual-type organisations may be happy with a lower level of measured efficiency if the increased costs were in the form of higher interest paid or additional services provided.

While it is a profit-seeking organisation, accumulating reserves to support its growth, its co-operative ownership structure makes it a suitable candidate for an efficiency study using DEA. DEA was originally intended for application to public sector and not-for-profit organisations (where the common economic objectives of cost minimisation and profit maximisation may not necessarily fit the business model).

3.3 Model Selection

The selection of appropriate inputs and outputs is a major step in examining the relative efficiency of any DMU using DEA. The outputs should reflect services that are considered primary to the organisation's purpose, and inputs should reflect resources required to produce the outputs, "such that an increase (decrease) in output levels is expected to result in an increase (decrease) in the amount of inputs used" (Sherman and Gold, 1985, p. 302).

The choice between which inputs and outputs to include is influenced by the literature on DEA applications in the banking industry (Mukherjee et al., 2002), the researcher's *a priori* view of financial institution behaviour (Mester, 1987), the objectives and strategies of the DMUs under study (Avkiran, 1999b), and the availability of data.

The two dominant approaches to measuring bank branch efficiency in the literature are the production approach (PA) and the intermediation approach (IA). The PA views banks as producers of services, using resources such as capital, labour, equipment, and total operating expenses to service accounts and develop products such as deposits, loans, and other financial services. The greater a branch's production efficiency, the lower the staff and other operating costs incurred in delivering its revenue-generating financial products (Thanassoulis, 1999).

Interest costs are commonly excluded. Berger and Humphrey (1997) advise that only physical inputs should be included as these alone are needed to perform financial transactions and process documentation. Output is measured by the number and/or type of transactions undertaken, or the number of accounts serviced at the branch (Mukherjee et al., 2002), rather than dollars.

Berger and Humphrey (1997) suggest that the PA is well suited to studies of bank branches as branch managers have little control over decision making on the bank's funding and investment. Also, as the number of transactions is commonly measured as an output, the PA is less sensitive to branch location; therefore, branches in more affluent neighbourhoods (which tend to have lower transaction levels relative to account balances) are less likely to be incorrectly assessed as efficient (Berger et al, 1997). Sherman and Gold (1985), Oral and Yolalan (1990), and Vassiloglou and Giokas (1990) all used the PA to measure operational and service efficiency of bank branches.

The IA views banks as financial intermediaries whose primary business is to gather funds from depositors and lend them to borrowers. Inputs are measured in both physical (e.g. labour, capital, and deposits) and financial (e.g. interest paid and total operating expenses) terms (Berger et al., 1997). Outputs may be considered in terms of both the value of loans intermediated and other investments made.

The IA has proven to be both a durable tool for financial institution analysis and the dominant approach in empirical research (Worthington, 1999). It also has the advantage of including all the costs in banking (Mester, 1987). Berger and Humphrey (1997) argue that the IA is better suited to evaluating the performance of an entire bank, rather than its branch network.

As neither of these approaches fully encapsulates all aspects of bank branch performance, Berger and Humphrey (1997) advocate applying both methods to assess whether results are affected by the choice of output metric.

3.4 Variable Selection

As the input and output variables chosen materially affect the results of efficiency studies, care was taken to identify the most appropriate input-output combinations (models) to measure the branches' relative efficiency. As suggested by Berger and Humphrey (1997), both the production and intermediation approaches were used. The literature supports use of both approaches, with Thanassoulis (1999) suggesting that they are complementary rather than mutually exclusive, and Berger and Humphrey (1997) noting that neither approach in isolation will provide a holistic view of the roles of FIs providing both transactional services and intermediating funds gathered as deposits into money lent to borrowers.

Denizer et al (2000) applied both the production and intermediation approaches in their analysis of the efficiency of Turkish commercial banks, under the key assumption "that banking is a simultaneously occurring two-stage process. During the production stage banks collect deposits using their resources, labour and physical capital. Banks use their managerial and marketing skills in the intermediation stage to transform these deposits into loans and investments" (p. 15).

This research also uses a two-stage DEA analysis. In the first stage, the production processes were assessed, while the second stage saw the intermediation processes measured. This is because some branches may employ fewer (more) resources than others to generate deposits, while at the same time using more (less) resources than others in the intermediation process.

As with other studies that have employed DEA to measure the relative efficiency of a FI's branch network (Sherman and Gold, 1985; Oral and Yolalan, 1990; and Drake and Howcroft, 1994), this research employs proprietary data (not publicly available).

Inputs and outputs and accounting information covering the financial year 1 April 2002 through 31 March 2003 were collected for each branch. The inputs and outputs for which data were available for both approaches are reported in Table 1.

Table 1: Available Inputs and Outputs

Inputs	Outputs
Average number of full time equivalent staff (FTE)	Number and type of transactions
Personnel expense	Number of customer automated service relationships (i.e. ATM/EFTPOS cards, and telephone and Internet banking facilities)
Hours worked in year	Loan originations
Number of deposit and loan accounts by account type	Number and value of deposit and loan accounts by account type
Branch size (M^2)	Number of new loan accounts established
Non-interest expense	Number of customer insurance policies
Number of computers at each branch	Average loan balances
Premises costs	Total loan balances
Value of deposit accounts by account type	Net interest income
Average total deposit balance	Non-interest income
Interest expense	New, total, and active customer numbers
Total costs	Total deposit volume growth
Corporate allocations	Total loan volume growth

A cross-section of 28 does not allow inclusion of all these variables. To economise on variables, we therefore identified highly correlated input (and output) variables, where one of the two inputs (outputs) could be eliminated (Golany and Storbeck, 1999).

To ensure the analysis is relevant, inputs should be selected on the basis that their use is evident in the generation of the chosen outputs (Avkiran, 1999b). Golany and Storbeck (1999) suggest use of regression, but we have followed Avkiran (1999c), reviewing correlations between inputs and outputs: the higher the correlations, the closer their relationship. Avkiran (1999c) recommends a correlation coefficient greater than 0.80 to indicate an adequate strength of relationship between variables for a DEA analysis.

3.4.1 The Production Approach

Using the input and output set in Table 1 a number of DEA models could have been employed. It was, however, considered more appropriate to develop a model that focused on the key aspects of branch production, where services expected to be provided to customers were listed as outputs, and the main resources employed in providing those services listed as inputs.

The following issues were considered in deciding which variables to eliminate:

- The average number of FTEs was strongly correlated with both personnel expenses (0.99) and hours worked (0.97), so little additional information would be generated by including all three. Average FTE was selected as it had the strongest overall correlation with each of the outputs.
- Branch size (M^2) was excluded on the basis of its comparatively low correlation with the available outputs.
- As total non-interest expense was strongly correlated with FTE (0.99), salary costs were removed from non-interest expense and the term operating expenses was adopted.
- The number of new loan accounts established was omitted as it had a comparatively low correlation with the available inputs.

Inputs and outputs used under the PA are reported in Table 2, while the strength of the association between them is reported in Table 3.

Table 2: PA – Inputs and Outputs

Inputs	Outputs
Average number of full time equivalent staff (FTE)	Number of deposit accounts (TOTDEP)
Number of computers at each branch (PC)	Number of credit facilities (TOTLOA)
Branch operating expenses (OPEX)	Number of insurance policies (INSUR)
	Number of branch transactions (TRANS)
	Number of electronic banking facilities (ELEC)

The positive correlation between the observed inputs and outputs are in all but one case (OPEX and TRANS) greater than 0.80. However, due to the strong correlations between OPEX and all other outputs, the 0.6976 correlation between OPEX and TRANS was not considered sufficiently weak to exclude them from the analysis.

Table 3: PA – Input and Output Correlations

	FTE	PC	OPEX	TOTDEP	TOTLOA	INSUR	TRANS	ELEC
FTE	1.0000	0.9809	0.9153	0.9600	0.9245	0.9634	0.8386	0.9597
PC		1.0000	0.8594	0.9226	0.9118	0.9407	0.8684	0.9324
OPEX			1.0000	0.9533	0.8357	0.8934	0.6976	0.9213
TOTDEP				1.0000	0.8886	0.9625	0.7732	0.9880
TOTLOA					1.0000	0.9403	0.9229	0.9120
INSUR						1.0000	0.8387	0.9725
TRANS							1.0000	0.7956
ELEC								1.0000

The inputs and outputs listed in Table 2 have been employed in previous studies using the PA. Many proxies for labour have been used – e.g. FTE by Sherman and Gold (1985), Sherman and Ladino (1995); personnel expense or salaries paid by Athanassopoulos et al. (1997), Denizer et al. (2000); and hours worked by Golany and Storbeck (1999). The number of personal computers (PC) or online terminals has been used by Athanassopoulos et al. (1997), Zenios et al. (1999), and Athanassopoulos and Giokas (2000) as a proxy for capital invested in the branch. Branch operating expenses, comprising non-interest expense less salary costs plus branch rent, have been used by Sherman and Ladino (1995) and Athanassopoulos and Giokas (2000).

For measuring output, the PA commonly uses either the number of accounts or the number of transactions made on them. In this study, both the number of branch transactions completed and the number of open deposit accounts and credit facilities were included. We also recognised the opportunity for customers to complete their own transactions through electronic facilities, and the sales of other financial products to these customers.

Outputs include:

- Total deposits (TOTDEP) – comprising all deposit accounts at a branch, including all current, savings, and term deposit accounts.

- Total loans (TOTLOA) – comprising all credit facilities at a branch, including home loans, personal loans, and overdraft and creditline facilities.
- Insurance policies (INSUR) held by customers – comprising health, general (e.g. home, contents, and vehicle), and travel insurance, and loan protection-type products available to borrowing customers. These reflect non-traditional banking activities which are increasingly important for FIs and this institution in particular, as it has a relatively high ratio of non-interest income to total income.
- Transactions completed in branch (TRANS) – includes all new loans advanced and credit facilities established, new term deposit accounts established, transfers between customer accounts, deposits and withdrawals made at tellers, and cheques prepared for customers by staff.
- Number of customer automated service relationships (ELEC) – includes all debit cards, and telephone and Internet banking facilities actively being used by a branch's customers.

The product of the number of inputs and outputs in this data set (3 x 5) is 15, well below the guideline proposed by Dyson et al. (1998), whilst also meeting the suggestion of Stern et al. (1994), that the sample size be at least three times the sum of the number of inputs and outputs (5 + 3 = 8; 8 x 3 = 24).

3.4.2 The Intermediation Approach

The available input and output set allows more than one DEA model to be employed. However, as with the PA, it was considered more appropriate to develop a model that focused on the key aspects of branch intermediation.

FTEs and operating expenses were not included as they were used in the PA analysis, while salary costs were eliminated to avoid double counting. Corporate allocations were excluded as generally outside the direct control of branch management, being applied to branches on the basis of the size of their asset and liability books.

The correlation between the available inputs is reported in Table 4.

Table 4: IA – Input Correlations

	Branch Rent	Total Deposit Volume	Interest Expense	Non-Interest Expense	Total Expenses	Total Cost
Branch Rent	1.0000	0.8119	0.8123	0.6602	0.8019	0.7993
Total Deposit Volume		1.0000	0.9981	0.8820	0.9965	0.9899
Interest Expense			1.0000	0.8753	0.9971	0.9893
Non-Interest Expenses				1.0000	0.9097	0.9339
Total Expenses					1.0000	0.9971
Total Cost						1.0000

On the basis of their high correlation, there would be little value in including total deposit volume, interest expense, total expenses, and total cost all as inputs in the same model.

According to Athanassopoulos et al. (1997), under the IA, interest and non-interest expense are the main inputs for studies of branch efficiency. It was therefore necessary to quantify the strength of their relationship with the available outputs, from the correlations between interest and non-interest expense and the available outputs. (Interest expense includes all interest paid to depositors, whilst non-interest expense includes all branch-controllable expenses and any costs associated with the provision of credit.) The results of this analysis are shown in Table 5.

Table 5: IA - Strength of Relationship

	Interest Expense	Non-Interest Expense
Interest Expense	1.00	0.88
Non-Interest Expense	0.88	1.00
Total Customer Numbers	0.93	0.96
Active Customer Numbers	0.89	0.97
New Customer Numbers	0.63	0.82
Interest Income	0.99	0.93
Non-Interest Income	0.71	0.91
Total Income	0.98	0.95
Total Loan Value	0.85	0.97

Interest expense is most strongly correlated with interest income, total income, and total customer numbers. Excluding total customer numbers, interest income, and total income (as already highly correlated with interest expense), non-interest expense is most strongly correlated with active customers, total loan value, and non-interest income. New customers had the lowest correlation with the inputs and were therefore dropped from the analysis. Table 6 reports the correlations between the remaining outputs.

Table 6: IA – Output Correlations

	Total Customers	Active Customers	Interest Income	Non-Interest Income	Total Income	Total Loans
Total Customers	1.00	0.99	0.96	0.87	0.97	0.93
Active Customers		1.00	0.93	0.90	0.96	0.94
Interest Income			1.00	0.77	0.99	0.90
Non-Interest Income				1.00	0.83	0.90
Total Income					1.00	0.93
Total Loans						1.00

Inputs and outputs used under the IA are shown in Table 7. The data set of five inputs and outputs comfortably meets the requirements relative to cross-section size, discussed in Section 3.4.1.

Table 7: IA - Inputs and Outputs

Inputs	Outputs
Interest expense (INTEXP)	Active customer numbers (ACTCUS)
Non-interest expense (NINTEXP)	Interest income (INTINC)
	Non-interest income (NINTINC)

Active customers is a subset of total customers, and excludes those with low balances not transacting on their accounts and/or under 18 years of age. As active customers are more likely to be revenue generating, they have been preferred to total customers. While use of active customers as an output may be questioned, it is not unprecedented. Hassan and Tufte (2001) considered members to be an appropriate output, justifying it by saying that “membership has been a primary target of all cooperatives and credit unions in many countries” (p. 1074).

Interest income and non-interest income (or variants thereof) are commonly used outputs in previous research – e.g. Oral and Yolalan (1990), Avkiran (1999c). Interest income represents interest received from loans and treasury operations (investment of un-lent branch deposits). Non-interest income includes loan fee income, transactions and service fee income, and commissions earned from sales of insurance products.

3.5 Use of DEA

DEA compares the observed inputs and outputs for all DMUs within the data set, identifies the best-practice units to comprise the efficient frontier, and measures the degree of efficiency of the other DMUs relative to this frontier. It thus provides a measure of relative efficiency in that efficiency is determined relative to other DMUs within the data set only.

A major advantage of DEA is that it identifies peers, or peer groups (also known as reference sets) for inefficient DMUs. A peer is an efficient DMU with a similar combination of weights as an inefficient DMU. Where two or more efficient units act as peers for an inefficient unit, they provide a peer group. These reference sets represent the branches most similar to the relatively inefficient branches in their mix of inputs and outputs (Sherman and Ladino, 1995).

Reference sets are very important in DEA as they can be used to identify the amount of resources that could be saved, the additional output that could be generated if an inefficient DMU became efficient, and identify best-practice DMUs that can be used as role models for improving efficiency.

Efficiency measures in DEA focus on whether a DMU can generate the same output with fewer inputs (input minimisation) or increase output using the same level of inputs (output maximisation). Both approaches offer useful insights into improving efficiency.

Under an input minimisation approach, DMUs found to be relatively inefficient can focus on producing current levels of output with fewer inputs. This is most appropriate where the DMUs are focused on controlling costs or potentially downsizing their operations. Under the output maximisation approach, DMUs found to be relatively inefficient can focus on increasing output with the same quantities of inputs. This optimisation is most likely to be appropriate where the DMUs are in a growth phase or where the focus is on raising productivity without reducing inputs.

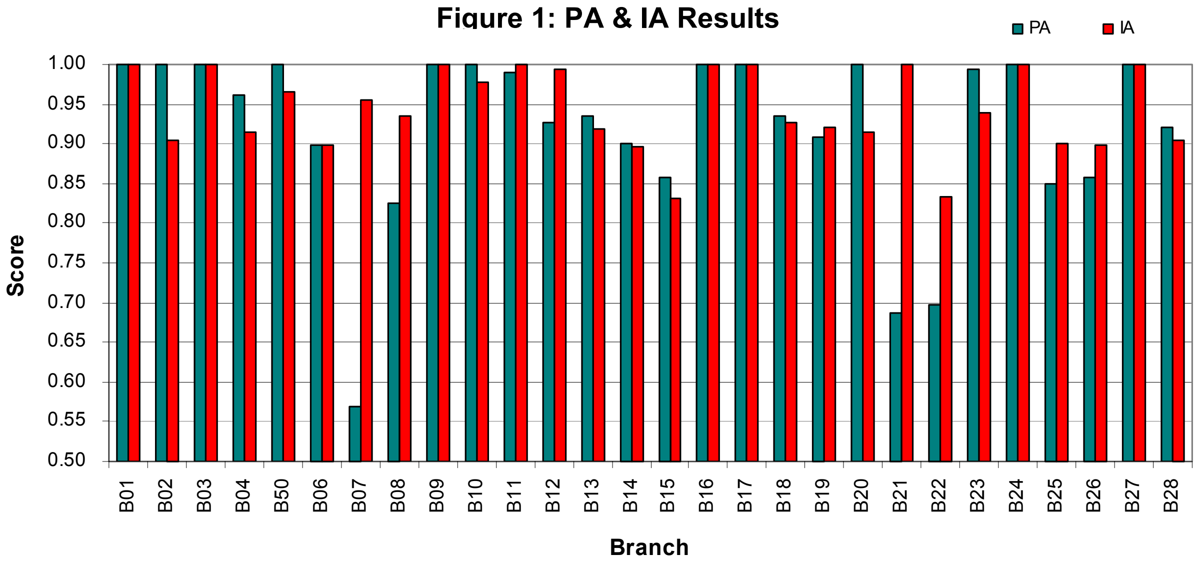
The output maximisation approach can indicate both increasing outputs and decreasing inputs (input slacks), just as it is possible under input minimisation to suggest increasing some or all of the outputs (output slacks), whilst also reducing inputs. Avkiran (1999a) suggests that where this occurs, inputs are over-employed (under output maximisation) or outputs are under-produced (input minimisation). For this research, given that the organisation studied is focused on growth, the output maximisation optimisation was considered appropriate.

Following Tripe (2003), the constant returns to scale approach has been used for this research as the results are thought to be more reliable for small samples. The CCR model generates the same efficiency scores for both the input minimisation and output maximisation approaches.

The following section reports the empirical results of the analysis of the institution's branch network. The software used for the study was "DEA-Solver", developed by Cooper et al (2000).

4. Empirical Results

The results of both the production and intermediation approaches using the CCR model under an output maximisation optimisation are displayed in Figure 1. 25% of branches are efficient under both approaches, 36% are more efficient producers than intermediators, and 32% are more efficient intermediators than producers. Only B06 is equally inefficient under both approaches.



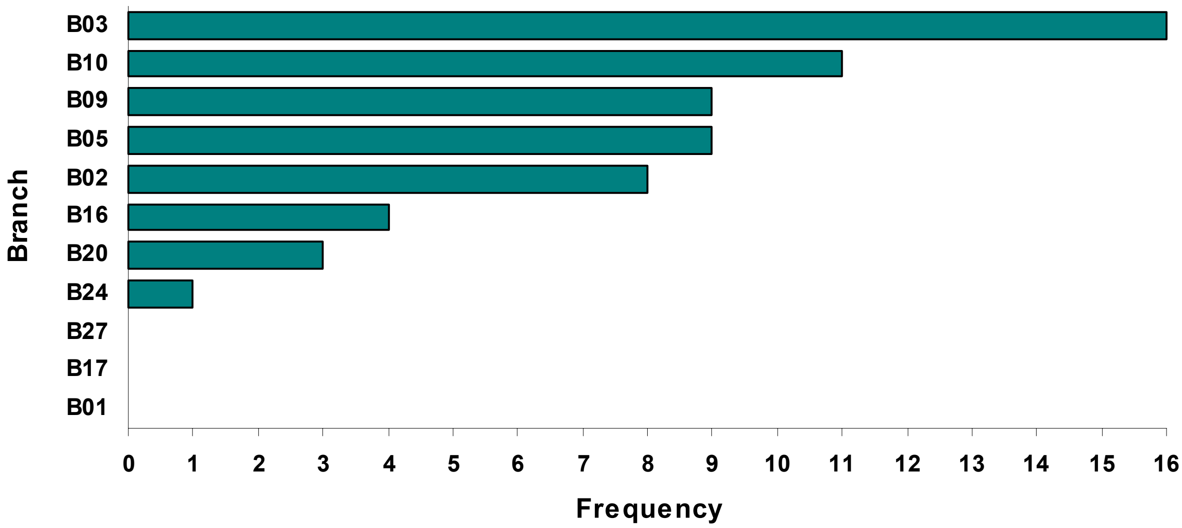
When interpreting results, it is important not to simply list each branch in descending order of efficiency score to identify which are more efficient. This is because DEA assesses a branch's efficiency by comparing it against others in its reference set, rather than all others in the sample. According to Avkiran (1999a), only branches within the same reference set can be strictly rank ordered, while DEA's benefits are realised from identifying potential improvements rather than creating some form of pecking order. It is, however, possible to use DEA to identify the global leader (the branch that appears most commonly in reference sets). But, according to Avkiran (1999a), applying the global leader as the benchmark for all branches will not produce optimal results for all DMUs. Rather, the branches that contribute the most to the calculation of efficiency scores for inefficient branches should be used as benchmarks.

4.1. The Production Approach

With a mean efficiency score of 0.92 and a standard deviation of 0.11, the PA model revealed 11 efficient and 17 relatively inefficient branches. Efficiencies ranged from 0.57 to 1.00. Figure 2 displays the relatively efficient branches, along with the total number of times those branches appear in the reference sets of inefficient branches.

Despite showing as efficient, B27, B17, and B01 do not appear in the reference sets of any inefficient branches. This might be because these branches are potential outliers, but a form of sensitivity analysis can be used to check for this. If we remove the global leader from the DEA data set and rerun the analysis, B27 moves into the reference set of ten branches whilst B17 moves into one. We can therefore conclude that only B01 is a potential outlier.

Figure 2: PA Global Leader

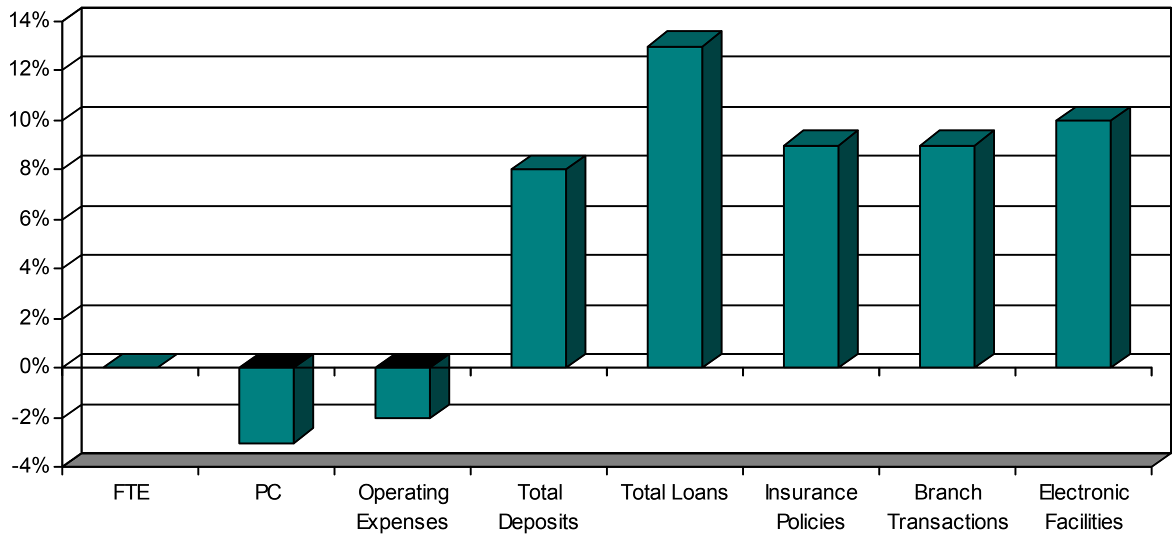


B03 stands out as the global leader, appearing in all but one (B15) of the inefficient branch's reference sets. However, benchmarking at the individual branch level is best achieved through identification of the efficient branch that most contributed to the calculation of the inefficient branch's efficiency score (Avkiran, 1999a). Cooper et al. (2000) argue that it is the existence of the efficient branches within each reference set that forces other branches to be inefficient.

Improvements

Once inefficiencies have been identified, measures can be taken to improve efficiency and performance. Figure 3 shows the potential improvements available for the financial institution if every branch was operating on the efficient frontier, calculated by using each inefficient branch's reference set to project it onto the efficient frontier.

Figure 3: PA – Potential Improvements



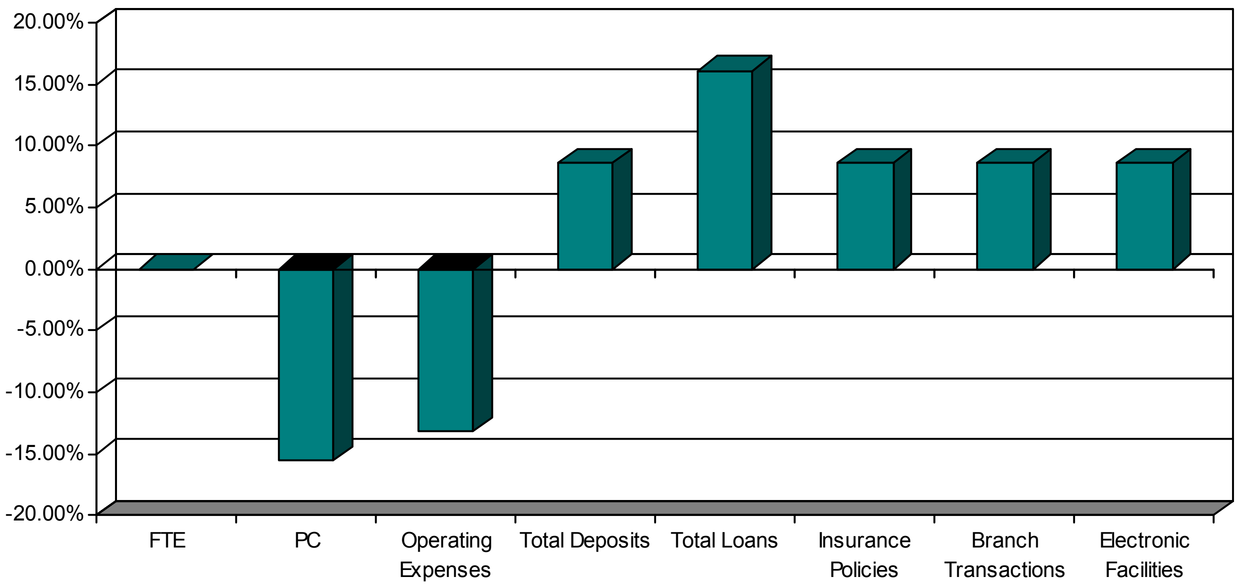
Detailed Branch Analysis

Rather than analysing each inefficient branch individually, a detailed analysis of B28 was undertaken. B28 was chosen on the basis that its efficiency score was closest to the mean. The reference set of B28 comprises B02, B03, B05, and B09, with B09 the most comparable.

Figure 4 shows that B28 could improve performance in all output categories while at the same time reducing two inputs. These input slacks suggests an over-utilisation of inputs in the production process, leaving room for additional outputs to be generated while simultaneously reducing inputs. Projecting B28 onto the efficient frontier would see a 16% reduction in capital invested and a 13% reduction in operating expenses. Based on an average deposit and loan account balances at B28, approximately $7.5M in additional deposit and $2.3M in additional loan volume could potentially be generated.

As B28 is a geographically isolated branch, with many customers having to travel to visit it, it is unreasonable to expect the volume of branch transactions to increase. Moreover, when considering the additional cost of servicing customers in a branch, this might not be a desirable outcome. An increase in the number of electronic banking facilities was also identified as an area for potential improvement, and despite the lower socio-economic demographic (and therefore reduced likelihood of computer and telephone access), it is an area that requires focus. Additional branch revenue could also be generated from the commission earned through a 9% increase in the number of insurance policies sold by the branch.

Figure 4: PA – B28 Improvements



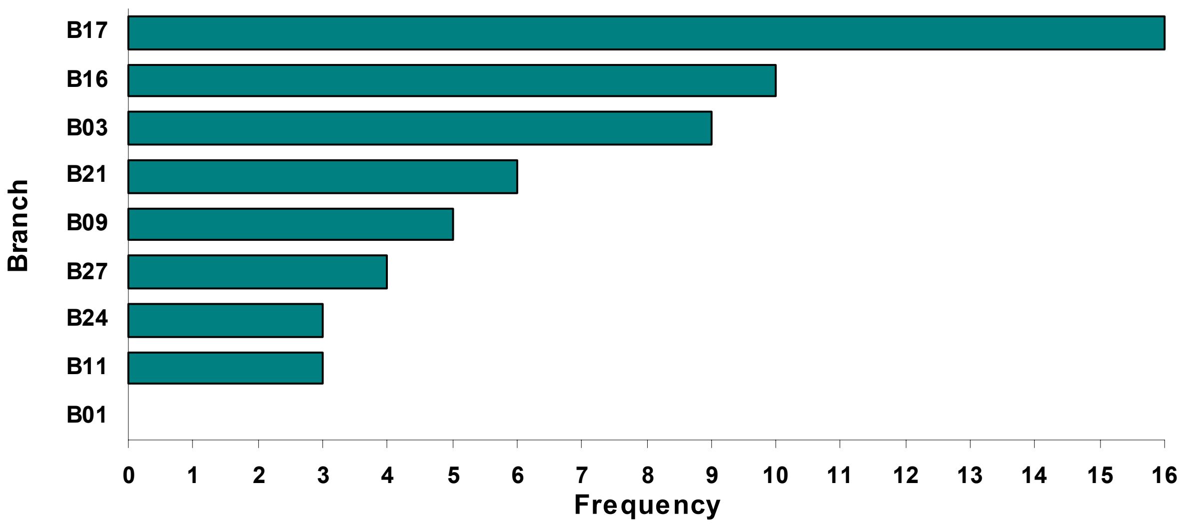
These potential improvements could be employed to set performance targets for B28, albeit with caution until the “organisational impediments to manifestation of inputs as outputs” (Avkiran, 1999b, p. 62) are more fully understood. We noted that the 16% reduction in capital invested equates to a reduction of 0.93 PCs. This would obviously not be a practical performance improvement target.

4.2. The Intermediation Approach

With a mean efficiency score of 0.94 and a standard deviation of 0.05, the IA model identified nine relatively efficient and 19 relatively inefficient branches. Efficiencies ranged from 0.83 to 1.00.

Figure 5 displays the relatively efficient branches and the number of times those branches appear in the reference sets of relatively inefficient branches. B17 stands out as the global leader, appearing in all but three inefficient branches' reference sets. Consistent with the PA, B01 remained a potential outlier, and B03, B09, B16, B24, and B27 are again assessed as being relatively efficient branches. To test whether B01 was an outlier, the global leader was removed from the data set and the analysis rerun to see whether B01 appeared in any of the inefficient branches' reference sets, but it did not.

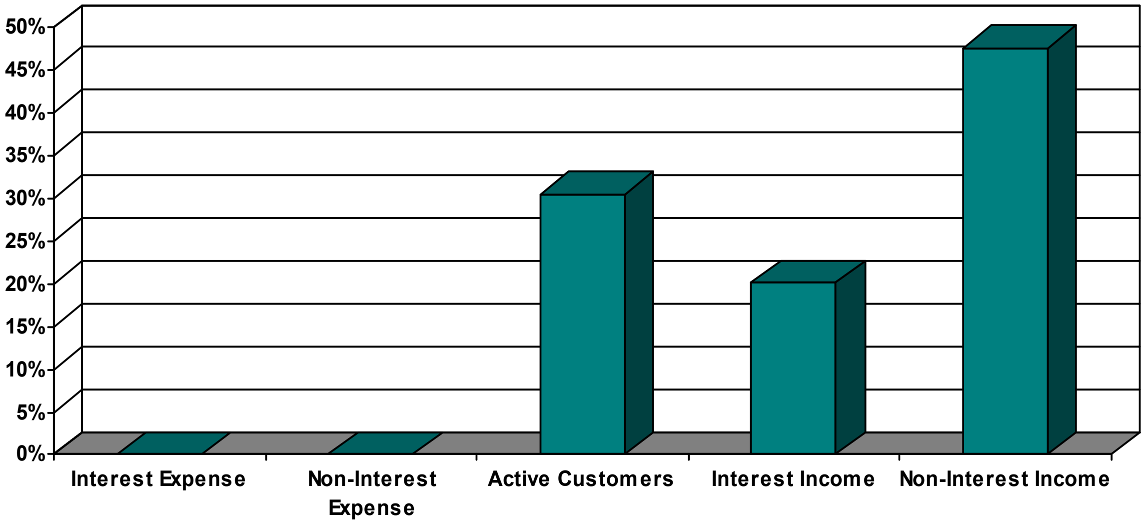
Figure 5: IA – Global Leader



Detailed Branch Analysis

B22 was selected for the detailed branch analysis as it has the greatest potential for improvement of all branches under study. The reference set of B22 comprises B09 and B24, with B24 the most comparable. No other branch shared the same reference set. As shown in Figure 6, while there were no input slacks, B22 could substantially improve performance in all three output categories.

Figure 6: IA – B22 Improvements



B22 serves a relatively small permanent population and is located within a township where the local economy is highly reliant on tourism and primary industry.

The greatest potential for output improvement is non-interest income. With less than 45% of both its peers' insurance policy numbers, B22 can improve this output most effectively through concentrating on increasing sales of insurance products. This would also help B22 improve its last ranking in product penetration per customer within the institution's branch network.

The number of active customers at B22 also presents difficulties, particularly given the relatively small permanent resident population. However, improving its active to total customer ratio to the level of B09 would see B22 almost halfway toward achieving the additional 321 active customers identified by the DEA analysis.

Increasing interest income can be achieved through either growing lending volume or improving the margin funds are lent at. Given B22's loan book already has a stronger weighting of higher margin lending than its peer branches, it is the low loan volume that is contributing most to the branch's inefficiency.

4.3. Scale Effects

DEA can be used to generate measures of scale efficiency to test the nature of the size-efficiency relationship at the branch level. Although constant returns to scale were assumed for this research, we sought, for completeness, to check for the existence of any scale effects.

The standard approach to testing for scale inefficiency is to run the DEA analysis under a variable returns to scale model, and compare the efficiency scores with those achieved under the constant returns to scale model. If the scores generated differ for the majority of the DMUs, it can be assumed there are scale inefficiencies. Given that the efficiency score obtained under the CCR model includes both technical and scale efficiency, and that the BCC model measures pure technical efficiency[3], the degree of relative scale efficiency can be determined by the ratio of the CCR and BCC scores. Specifically, for each DMU, scale efficiency (SE) = (CCR / BCC). Where the ratio results in a score of one the DMU is operating at the most efficient scale. Where the score is less than one, the DMU is scale inefficient. Therefore, subtracting the resultant SE score from one (i.e. 1 - SE) will represent the relative scale inefficiency of a DMU.

Using the BCC model we found that 61% of branches showed some scale inefficiency under the PA, and 68% of branches under the IA. Levels of scale inefficiency were small, with the exception of B07, B15, and B22. Although these were among the smaller branches in the network, they were not the only small branches, and it is not clear that scale efficiency is necessarily a major problem.

5. Conclusion

The purpose of this research was to study the relative efficiency of the branches within the institution's network, investigate what has made some branches more efficient than others, and to identify ways of improving efficiency and therefore profitability. Using both

[3] Technical efficiency measures inefficiencies which can be attributed to both the input/output configuration and the size of the DMU, whilst pure technical efficiency measures efficiency without scale effects (Avkiran, 1999a).

the production and intermediation approaches and an output maximisation optimisation, data envelopment analysis (DEA) was employed in a two-stage process to investigate the relative efficiency of the 28 branches that make up the institution's network.

Eleven branches were found to be efficient using the production approach, compared with nine under the intermediation approach. Seven branches were found to be efficient under both. Potential improvements under both approaches were identified, and improvements quantified for each inefficient branch. By focusing on these identified improvements, management of the financial institution can improve the organisation's overall profitability. The output maximisation optimisation was particularly useful in aiding understanding of the capacity of the institution's branch network to cope with increases in business volumes before additional resources are required.

While the many strengths of DEA have been highlighted, this research also suffers from DEA's limitations, the most important of which is in respect of the possibility of random error distorting results for individual branches.

This study would benefit from a more detailed investigation of scale effects present in the institution's branches, with a view to determining their optimal scale, whether it is appropriate to include weight restrictions on some variables to ensure each variable is adequately considered in the analysis; and whether including environmental and market conditions variables such as branch customer demographics, the number of nearby competitor locations, or the level of customer satisfaction would have impacted on the results.

Avkiran (1999a) has noted that extension of the data set by inclusion of further years of data could prove advantageous in determining whether the data was in fact free from measurement or other random error (a major assumption of DEA). This is now in train, with a view to utilising DEA to provide a foundation for an ongoing review of the performance of the institution's branches.

References

Al-Faraj, T., Alidi, A.S., & Bu-Bshait, K.A. (1993). Evaluation of bank branches by means of data envelopment analysis. *International Journal of Operations and Production Management, 13(9),* 45-52.

Athanassopoulos, A.D. (1998). Nonparametric frontier models for assessing the market and cost efficiency of large-scale bank branch networks. *Journal of Money Credit and Banking, 30(2*), 172-192.

Athanassopoulos, A.D., & Giokas, D. (2000). The use of data envelopment analysis in banking institutions: evidence from the Commercial Bank of Greece. *Interfaces, 30(2),* 81-95.

Athanassopoulos, A.D., Soteriou, A., & Zenios, S. (1997). Disentangling within- and between- country efficiency differences of bank branches. Working Paper Series, Wharton School of the University of Pennsylvania. Available at http://knowledge.wharton.upenn.edu/PDFs/78.pdf.

Avkiran, N.K. (1999a). An application reference for data envelopment analysis in branch banking: helping the novice researcher. *International Journal of Bank Marketing, 17(5),* 206-220.

Avkiran, N.K. (1999b). *Productivity analysis in the service sector with data envelopment analysis* (1st ed.). Camira: N K Avkiran.

Avkiran, N.K. (1999c). The evidence on efficiency gains: the role of mergers and the benefits to the public. *Journal of Banking and Finance, 23,* 991-1013.

Banker, R. D.; Charnes, A. W. & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis. *Management Science.* 30 (9) 1078-1092.

Berger, A. N. (1993). “Distribution-free” estimates of efficiency in the U.S. banking industry and tests of the standard distributional assumptions. *Journal of Productivity Analysis.* 4. 261-292.

Berger, A. N. & Humphrey, D. B. (1991). The dominance of inefficiencies over scale and product mix economies in banking. *Journal of Monetary Economics.* 28. 117-148.

Berger, A.N., & Humphrey, D.B. (1997). Efficiency of financial institutions: international survey and directions for future research. *European Journal of Operational Research,* 98(2). 175-212.

Berger, A.N., Leusner, J.H., & Mingo, J.J. (1997). The efficiency of bank branches. *Journal of Monetary Economics, 40(1),* 141-162.

Charnes, A. & Cooper, W. W. (1985). Preface to topics in Data Envelopment Analysis. *Annals of Operations Research.* 2. 59-94.

Charnes, A., Cooper, W.W., & Rhodes, E.L. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2(6)*, 429-444.

Cooper, W. W., Seiford, L. M., & Tone, K. (2000). *Data Envelopment Analysis.* Boston: Kluwer Academic Publishers.

Denizer, C.A., Dinc, M., & Tarimcilar, M. (2000). Measuring banking efficiency in the pre- and post-liberalization environment: evidence from the Turkish banking system. *Working Paper, World Bank*. Available at http://econ.worldbank.org/view.php?type=5&id=1286.

Drake, L., & Howcroft, B. (1994). Relative efficiency in the branch network of a United Kingdom bank – an empirical study. *Omega International Journal of Management Science, 22(1),* 83-90.

Dyson, R.G., Thanassoulis, E., & Boussofiane, A. (1998). Data envelopment analysis. *Warwick Business School*. Available at http://www.csv.warwick.ac.uk/~bsrlu/dea/deat1.htm.

Dyson, R. G.; Allen, R.; Camanho, A. S.; Podinovski, V. V.; Sarrico, C. S. & Shale, E. A. (2001). Pitfalls and protocols in DEA. *European Journal of Operational Research.* 132. 245-259.

Evanoff, D. D. & Israilevich, P. R. (1991). Productive efficiency in banking. *Economic Perspectives (Federal Reserve Bank of Chicago).* 11-32.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society.* 120. 253-281.

Golany, B., & Storbeck, J.E. (1999). A data envelopment analysis of the operational efficiency of bank branches. *Interfaces, 29(3),* 14-26.

Hassan, M.K. & Tufte, D.R. (2001). The x-efficiency of a group-based lending institution: the case of the Grameen bank. *World Development, 29(6),* 1071-1082.

Leibenstein, H. (1966). Allocative efficiency vs. “X-efficiency. *American Economic Review.* 56. 392-415.

Liu, B., & Tripe, D.W. (2002). New Zealand bank mergers and efficiency gains. *Journal of Asia-Pacific Business.* 4 (4). 61-81.

Mester, L.J. (1987, January/February). Efficient production of financial services: scale and scope economies. *Federal Reserve Bank of Philadelphia Business Review,* 15-25.

Mukherjee, A., Nath, P. & Pal, M.N. (2002). Performance benchmarking and strategic homogeneity of Indian banks. *International Journal of Bank Marketing, 20(3),* 122-139.

Nyhan, R.C., & Martin, L.L. (1999). Comparative performance measurement. *Public Productivity and Management Review, 22(3),* 348-364.

Oral, M. & Yolalan, R. (1990). An empirical study on measuring operating efficiency and profitability of bank branches. *European Journal of Operational Research.* 46(3). 282-294.

Sathye, M. (2001). X-efficiency in Australian banking: an empirical investigation. *Journal of Banking and Finance, 25,* 613-630.

Sherman, H.D., & Gold, F. (1985). Bank branch operating efficiency – evaluation with data envelopment analysis. *Journal of Banking and Finance, 9(2),* 297-315.

Sherman, H.D. & Ladino, G. (1995). Managing bank productivity using data envelopment analysis. *Interfaces, 25(2),* 60-73.

Stern, Z.S., Mehrez, A., & Barboy, A. (1994). Academic departments efficiency via DEA. *Computer and Operations Research, 21(5),* 543-556.

Thanassoulis, E. (1999). Data envelopment analysis and its use in banking. *Interfaces, 29(3),* 1-13.

Thanassoulis, E., Boussofiane, A. & Dyson, R.G. (1996). A comparison of data envelopment analysis and ratio analysis as tools for performance assessment. *Omega-International Journal of Management Science, 24(3),* 229-244.

Tripe, D.W.L. (2003). *An exploration into the efficiency of financial institutions – the example of New Zealand building societies.* A paper presented at the 16th Annual Australasian Finance and Banking Conference, Sydney.

Vassiloglou, M., & Giokas, D. (1990). A study of the relative efficiency of bank branches: an application of data envelopment analysis. *Journal of Operational Research Society, 41(7),* 591-597.

Worthington, A.C. (1999). Measuring technical efficiency in Australian credit unions. *The Manchester School, 67(2)*, 231-248.

Zenios, C.V., Zenios, S.A., Agathocleous, K., & Soteriou, A.C. (1999). Benchmarks of the efficiency of bank branches. *Interfaces, 29(3)*, 37-51.