Measuring scale economies in New Zealand banking – traps for the unwary

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

Phone: +64 (06) 350-5799 ext 2337
Fax: +64 (06) 350-5651
E-mail: D.W.Tripe@massey.ac.nz

Abstract:

This paper uses Data Envelopment Analysis (DEA) to investigate the existence of scale economies for New Zealand's major (retail) banks, and in particular, whether larger scale enhances operating efficiencies.

It is found that the appearance of operating efficiencies with increased scale may be a consequence of the technique used to test for them, in that the inclusion of an additional bank causes the appearance of benefit from increased scale to be no longer observed.

Some suggestions are offered for why the commonly assumed benefits from increased scale may not be observed.

1. Introduction

One of the key reasons for investigating the efficiency of financial institutions has been to look for economies of scale. In accordance with standard microeconomic theory it is expected that there should be something like a U-shaped average cost curve. Smaller financial institutions might be operating at less than efficient scale, and if they were able to grow larger, perhaps as a result of a merger, their reduced costs (assuming a competitive market) should allow them to reduce their charges to the public for undertaking financial intermediation. Conversely, larger firms might be operating at greater than minimum cost, and higher prices for financial intermediation might be sustained by their exercise of market power.

This relates to arguments as to the relative merits of the structure-conduct-performance (SCP) and efficient structure (ES) hypotheses for the study of competition between banks. The SCP hypothesis suggests that, in a more concentrated market, banks will be able to exercise some element of monopoly power and increase their profits accordingly. By contrast, the ES hypothesis argues that banks will be more profitable because they are more efficient, and that because they are more efficient, they will achieve a larger market share. Clark (1988) noted that, if the ES hypothesis applied, a market should be made up of firms of relatively similar size, each of them operating at scale efficient levels of output. Various efforts have been made to explore the relative merits of the SCP and ES hypotheses empirically – notable examples include Berger (1995) and Goldberg & Rai (1996).[1]

The existence or otherwise of scale economies, and the specification of optimum scale are thus important for the analysis of mergers, in that it is assumed that firms which are operating at efficient scale should be able to minimise the cost of financial intermediation, and provide financial services to society at minimum cost. Avkiran (1999) looked at the extent to which any efficiency gains from mergers in Australia were passed through to bank customers, while Liu & Tripe (2002) undertook a comparable study for New Zealand, with mixed results.

[1] For a recent review of literature on competition in banking, see Shaffer (2004).

Such research as has been undertaken looking for scale efficiencies in banking has found evidence for their existence to be somewhat limited, while Berger & Humphrey (1991) and Berger et al (1993) report that X-inefficiencies account for a much greater proportion of efficiency differences between banks than do scale inefficiencies. Our object in this paper is to investigate the existence of scale economies in the New Zealand banking market, at least in respect of the major banks that dominate the retail banking business.

Because of the relatively small number of participants in the New Zealand banking market, the scope to undertake this study using standard econometric methods is somewhat limited. Our approach is therefore to use Data Envelopment Analysis (DEA), which is a major non-parametric approach to specification of an efficient frontier.

The rest of the paper proceeds as follows. The next section briefly reviews previous efforts to investigate the efficiency of financial institutions, with a particular emphasis on the investigation of scale effects. The following section looks at the DEA approach in more detail, describes how it will be used in this study, and discusses some of the issues relative to that approach. Section 4 looks at the New Zealand banking market and describes the data used in this study. Section 5 presents the results and discusses these, while Section 6 concludes.

2. Scale economies in banking

The concept of efficiency may be regarded as one of the fundamental precepts of economics. It may be defined as the ratio of the weighted sum of outputs to the weighted sum of inputs (Boussofiane et al, 1991). In general terms, a firm may be said to be operating efficiently if it cannot produce more output without a corresponding relative increase in inputs, or if it cannot reduce its inputs without a corresponding relative decrease in output.[2]

[2] This assumes that there is no change or difference in the quality of inputs or outputs.

Efficiency can be discussed in a variety of different forms, not all of them necessarily totally consistent with the previous definition. Traditional microeconomic theory has long talked of economies of scale, where increased volumes of output are supposed to be able to be produced with less than proportionate increases in quantities of inputs (increasing returns to scale). In due course, however, economies of scale will be exhausted, and increased output will require a more than proportionate increase in inputs, a situation described as diseconomies of scale (decreasing returns to scale).

Other forms of efficiency are economies of scope, which look at the potential benefits of producing multiple outputs from one set of inputs, at lower cost than by firms producing only one output, and X-efficiency, which looks at whether firms are maximising output from the inputs being used, and whether the right combinations of inputs are being used to produce the right combination of outputs.

Common approaches to measurement of efficiency in banks are inclined to focus on ratios of non-interest costs to average total assets (cost to assets), or non-interest costs to gross income (the cost to income ratio). As has been discussed elsewhere (DeYoung, 1997; Tripe, 1998), these ratios have a number of deficiencies, most particularly in that they don't take account of differences in the business that banks undertake, which will in turn be reflected in different combinations of inputs and outputs. DeYoung (1998) suggests that blind pursuit of accounting-based benchmarks might reduce a bank's cost efficiency by cutting back on those expenditures necessary to run the bank properly. There is also the issue that ratios may be contradictory: if a bank performs very well in terms of one ratio but poorly in terms of another, how does one determine an overall performance ranking (Golany & Storbeck, 1999)?

To take account of these issues, one therefore needs to look at banks on a multiple input and multi-product basis. Banks use a mixture of inputs to produce a mixture of outputs, and their reported average cost figures will depend just as much on the mix of inputs and outputs as on the rate at which those inputs are used to produce outputs (Mester, 1987). In looking at bank costs one needs to consider more than just operating costs, which account for only part of overall bank costs. Berger & Humphrey (1992) highlight the effect of tradeoffs between price and service, and note that a bank with a less extensive branch network may pay higher interest costs to

attract deposits, although it will have lower operating costs. As a general rule, larger banks will have lower (non-interest) operating costs but higher interest costs, reflecting their dependence on borrowed funds. The extent of banks' branch networks is important for similar reasons (Humphrey, 1990): where a bank does not have branches, it can be argued as reducing its operating costs by transferring these to its customers.

Similar issues can arise on the output side. Noulas et al (1990) show that attempts to aggregate outputs into a single index are invalid, and that one is therefore required to take account of the multi-product nature of large banks in seeking to measure their efficiency. Resti (2000) notes that techniques aimed at summarising multiple products had been proven to be too good to be true, since they required separability conditions not usually supported by empirical data.

More generally, Berger et al. (1993) note that financial ratios may be misleading because they do not control for product mix or input prices.[3] Moreover, use of a simple ratio cannot distinguish between X-efficiencies and scale and scope efficiencies. Thanassoulis et al (1996) suggest that use of a multivariate approach gives a more balanced approach to performance measurement than ratio-based performance indicators.

Initial studies of bank efficiency were inclined to be focused on looking for economies of scale, although some attention was also given to economies of scope (Mester, 1987; Clark, 1988). Despite what Humphrey (1985) refers to as the conventional wisdom, earlier studies were generally unable to find evidence for economies of scale beyond a relatively small size for a financial institution (assets greater than $100 million).

A number of methodological difficulties were identified with this earlier research, which cast doubt on the reliability of some of the results obtained. Humphrey (1985) noted the problem that could arise from looking at unit banks and banks with branches together: scale economies should be expected to be observed quite differently

[3] Thus Chu & Lim (1998) argue that it is more appropriate to consider interest expense as an input in efficiency studies, rather than deposits, because not all deposits carry the same interest expense (p 158).

between the different classes of banks. Clark (1988) noted the difficulties in defining bank costs and outputs, and problems with data and statistical methodology. Thus, as Berger et al (1993) noted, studies which looked at larger banks found the minimum average cost point to be associated with rather larger banks, with assets between $2 and $10 billion. In more recent work, however, using data from the 1990s as opposed to the 1980s, Berger & Mester (1997) have suggested that the most efficient scale size might be rather larger.[4]

McAllister & McManus (1993) explained what they perceived to be some of the methodological reasons for the earlier results, with a particular focus on the use of the translog cost function. Using a different approach, they suggested that banks could operate at minimum constant average cost at asset levels between $500 million and $10 billion. Also using a different approach, Siems & Clark (1997) found banks' scale efficiencies to be essentially invariant above a relatively small balance sheet size.

Another problem identified by Berger et al. (1993) was the importance of using a method which only measured scale efficiency for firms that were on the efficient frontier: this was not necessarily common, particularly prior to the work of Berger et al (1987). If one was not focusing on firms on the frontier for assessing economies of scale, one was in danger of confusing scale efficiencies and X-inefficiency. Berger & Humphrey (1991) found that X-inefficiency was a much more significant component of overall inefficiency than was scale inefficiency. Scale inefficiencies accounted for only 5% of costs, whereas X-inefficiencies were around 20%.

Humphrey (1990) highlighted the importance of the definition of costs to be used when trying to measure economies of scale. Larger banks are generally less able to fully fund themselves through retail deposits, and will thus have higher interest costs but lower non-interest costs than smaller institutions. If a researcher tries to measure scale economies using only operating costs, minimum costs will be found at a larger asset size than if all costs are considered.

[4] This would be consistent with the suggestion of Noulas et al (1990) to the effect that optimum scale might change for year to year, and that this might have been a particular issue in the 1980s when deregulation was still working its way through the system in the United States.

More recent work has found stronger evidence for the existence of economies of scale among larger banks, which Stiroh (2000) suggests may reflect the size at which banks achieve minimum cost efficiency levels having increased. Berger & Mester (2003) note that different approaches may lead to different conclusions, but once again stress the importance of treating equity capital as an input, because it affects other costs and provides an alternative source of funding. Equity is also suggested as being important because of its role in absorbing risk (and this may be the source of the scale economies).

Much of this previous research has been based on data from the United States, whereas one of the objectives of this paper is to look at the existence of scale effects in the New Zealand market. Before we can report on that analysis, however, we need to review the method to be employed.

3. Method

Reflecting the small number of banks in the New Zealand market (as is discussed in the next section), this research is undertaken using Data Envelopment Analysis (DEA). Even then, however, we are obliged to use panel data covering a period of 20 quarters to obtain a sufficient number of observations to provide adequate discriminatory power.

One of the problems of DEA is that, as a non-parametric technique, it does not take account of random error in the data. Exploratory tests undertaken using the super-efficiency model suggest that this is not a problem, however: the highest super-efficiency score for the basic model used was 1.248, well within the guideline of 2 suggested by Hartman et al (2001).

The standard approach to measuring scale effects using DEA is to run models on both a constant (CRS) and variable returns to scale (VRS) basis. Scale efficiency is then found by dividing the efficiency score from the CRS model by the efficiency score from the VRS model. Because the data points are enveloped more tightly under the

VRS model, the VRS efficiency scores will be higher, and the scale efficiency measures will therefore be in the range 0 to 1.

A further feature of VRS models (in most DEA software) is that they report whether a Decision-Making Unit (DMU) is operating at increasing, constant or decreasing returns to scale. Constant returns to scale will apply when the CRS and VRS efficiency frontiers are tangential with each other: in other words, when the (local) slope of the efficiency frontier is equal to the ratio of input(s) to output(s) (Cooper et al, 2000). Increasing returns to scale must apply below that level, as the slope of the efficient frontier (which reflects the marginal rate of transformation of inputs to outputs) will be greater than the average rate of conversion (which is equivalent to the average cost). Likewise, decreasing returns to scale must apply must apply above the zone in which constant returns to scale apply. DMUs not on the efficient frontier must first be projected onto the efficient frontier before their returns to scale status can be assessed.

Dyson et al (2001) have identified problems with this approach to measurement of scale efficiency, however. Because of the way a VRS model envelops points more tightly small and large units will tend to be over-rated in the efficiency assessment. This means that scale inefficiencies identified for such institutions may be spurious, with the actual cause of inefficiency being X-inefficiency.

There are a number of other issues identified by Dyson et al (2001), Avkiran (2002) and others as needing to be resolved prior to applying DEA to efficiency measurement. These issues include model orientation, selection of the input/output set, and selection of the DMUs whose efficiency is to be compared.

When looking at banks as a whole it is common to use an input-minimisation approach (input orientation), reflecting banks' tendency to focus on costs, which reflects in turn what banks' management can actually change. Outputs are generally constrained by demand in the market as a whole, whereas management should have some ability to vary the quantity and costs of inputs, such as staff and funds, to achieve that output.

In selecting inputs and outputs, prior research has often tended to stress a distinction between the production and intermediation approaches, with the intermediation approach existing in three different forms. The approach followed in this research is broadly consistent with the intermediation approach, where deposits and other funds are transformed into loans (Sealey & Lindley, 1977), subject to two further points. Inputs and outputs must be measurable, consistently, across all DMUs in the study, while the input/output set must also cover the full range of resources used and outputs created. Inputs used in this study are interest expense, non-interest expense and equity capital, with outputs net interest income and non-interest income.

In selecting our set of banks to include in the study, we need to ensure that there is sufficient homogeneity between them for their performance to be validly compared. Thus, according to Dyson et al (2001), the DMUs should be undertaking similar activities and producing comparable products or services so that a common set of outputs could be defined. There should also be a similar range of resources available to all DMUs, which should be operating is similar environments.

The focus in this study is thus on banks which undertake a full range of business, including providing services across counters through a bank branch network. The composition of this data set is discussed in greater depth in the next section.

4. New Zealand Banking

Despite the opening up of the banking market following the deregulation of the 1980s, the number of full service banks operating in New Zealand has remained relatively small. As of 2005, following the acquisition of the National Bank of New Zealand (NBNZ) by the ANZ Banking Group (ANZ), there are six banks which provide full service through branch networks. These are the Australian-owned ANZ-National Bank, ASB Bank (ASB), Bank of New Zealand (BNZ) and the Westpac Banking Corporation (Westpac), and two New Zealand-owned institutions, Kiwibank and TSB Bank Limited (TSB).

These two New Zealand-owned institutions are much smaller, and for that reason alone, there would be some question as to whether it was fair to include them for comparative analysis of bank efficiency. Kiwibank only commenced business in early 2002, and although it is now reporting small profits, its income statements are quite different from those of the other banks. It is therefore omitted from the rest of this research.[5]

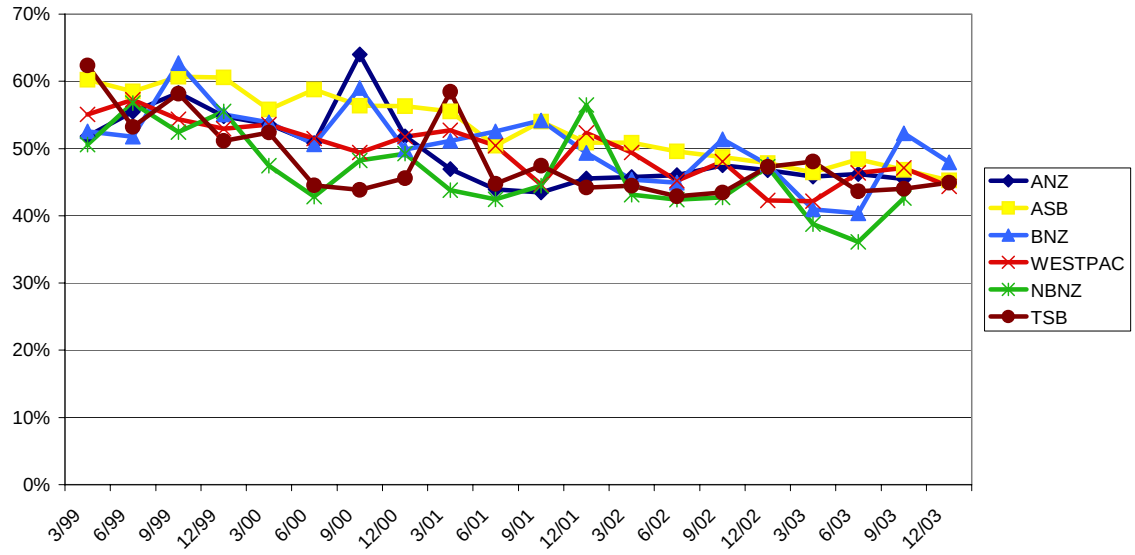
If we were to look at the period subsequent to the acquisition of the NBNZ by the ANZ, we would only have four or five banks for comparative analysis and a relatively short time period. The merger has also cause significant changes in performance figures (and size) for the ANZ. It is therefore considered more appropriate to look at the banks' efficiency up to the time of that merger. The study period is thus up to December 2003 (or September 2003 in the case of ANZ and NBNZ).

The study period commences with the March quarter 1999. This means that the figures are not impacted by the previous retail bank merger in New Zealand, which was the NBNZ's acquisition of Countrywide Bank in 1998. It also reflects a period of relative stability in interest rates: where (gross) interest expense is used as an input, major changes in interest rates can impact on efficiency scores (Tripe, 2003). The study thus covers a period of 19 or 20 quarters for five banks (six where TSB is included), giving a minimum of 98 observations, which is sufficient to provide adequate discriminatory power relative to a total of five inputs and outputs.[6]

Data for the study is obtained from the banks' quarterly disclosure statements (which are required to be published under New Zealand's banking supervision regime). Throughout the period of the study, the banks have been making assiduous efforts to reduce their non-interest expenses, with this being reflected in their cost to income ratios, as shown in Figure 1 below. The trend shown in this graph is the underpinning for the popular view that continued asset growth has allowed the banks to achieve economies of scale. The trend in the ratio of operating costs to average total assets is even more strikingly downwards.

[5] This reflects the approach followed in previous research – see DeYoung & Hasan (1998).
[6] There are various rules of thumb for specifying sample size relative to inputs and outputs: a common one is reported by Cooper et al (2000) to the effect that the number of observations in the data set should be at least three times the sum of the number of input and output variables.

Figure 1: trends in New Zealand's major banks' cost to income ratios - 1999 to 2003



Descriptive statistics for the five banks (not including TSB) are shown in Table 1, with correlations between the inputs, the outputs and between the inputs and outputs shown in Table 2. The correlation coefficients reported in Table 2 do not appear to be inappropriate for the proposed study.

Table 1: Summary descriptive statistics for the input and output variables used, across five banks.

$M	Minimum	Maximum	Mean	Standard deviation
Interest expense	131.4	479	330.3	78.2
Non-interest expense	71.9	193	138.2	29.8
Equity	603.5	2953.5	1611.2	599.0
Net interest income	81.3	267	183.5	45.1
Non-interest income	38.6	149	97.5	29.2

Table 2: Correlations between inputs variables, output variables, and between input and output variables

	Non-interest expense	Equity	Net interest income	Non-interest income
Interest expense	0.617	0.679	0.672	0.607
Non-interest expense		0.625	0.775	0.854
Equity			0.840	0.604
Net interest income				0.640

We can now look to the results relative to the individual studies undertaken.

5. Results and discussion

Our study begins with an analysis of just the five larger banks (ANZ, ASB, BNZ, NBNZ and Westpac), with their figures analysed in a panel.[7] We run both CRS and VRS models (with the results reported in Tables 3 and 4 respectively), and use the results from the two to generate measures of scale efficiency (with the results reported in Table 5).

The analysis was undertaken using the software DEA-Solver, as described by Cooper et al (2000).

Table 3: Efficiency scores using CRS model, TSB not included.

Quarter ending	ANZ	ASB	BNZ	NBNZ	Westpac
March 1999	0.885	0.798	0.842	0.889	1.000
June 1999	1.000	0.813	0.905	0.818	0.976
September 1999	1.000	0.790	0.728	0.844	1.000
December 1999	0.922	0.761	0.793	0.817	0.939
March 2000	0.864	0.795	0.819	0.871	0.890
June 2000	0.842	0.767	0.840	0.913	0.886
September 2000	0.761	0.842	0.850	0.832	0.852
December 2000	0.843	0.822	0.847	0.828	0.848
March 2001	0.914	0.831	0.835	0.892	0.821
June 2001	0.955	0.931	0.815	0.895	0.841
September 2001	1.000	0.847	0.787	0.875	0.975
December 2001	0.976	0.879	0.847	0.790	0.896
March 2002	0.987	0.856	0.879	0.950	0.840
June 2002	0.981	0.894	0.916	0.976	0.949
September 2002	0.964	0.904	0.845	0.934	1.000
December 2002	0.934	0.915	0.887	0.888	0.993
March 2003	0.955	0.921	1.000	0.981	1.000
June 2003	0.952	0.936	1.000	1.000	0.998
September 2003	0.996	0.937	0.840	0.921	0.909
December 2003	-	0.948	0.883	-	0.996
Average	**0.933**	**0.859**	**0.858**	**0.890**	**0.930**

Under the constant returns to scale model, ANZ and Westpac show as more efficient, and ASB, BNZ and NBNZ as less efficient.[8]

[7] A theoretical justification for use of a panel to increase the number of DMUs is provided by Tulkens &Vanden Eeckaut (1995).

[8] The efficiency differences between the banks are significant at the 5% level, except for the difference between the NBNZ and Westpac, which is significant at the 10% level only. Differences were tested for using the (non-parametric) Mann-Whitney test.

All the banks show some improvement in their efficiency scores when the variable returns to scale model is used (which is to be expected, as the data points will now be enveloped more tightly), but the dramatic improvement is evident for ASB, which now shows as being most efficient.[9] Other things being equal, one would expect this to reflect economies of scale, as is reflected in the average scale efficiency score for ASB, shown in Table 5, of 0.88.

Table 4: Efficiency scores using VRS model, TSB not included.

Quarter ending	ANZ	ASB	BNZ	NBNZ	Westpac	ASB Returns to scale status
March 1999	0.935	1.000	0.872	0.914	1.000	Increasing
June 1999	1.000	1.000	0.942	0.859	0.983	Increasing
September 1999	1.000	0.956	0.775	0.888	1.000	Increasing
December 1999	0.937	0.965	0.831	0.850	0.952	Increasing
March 2000	0.885	0.983	0.851	0.917	0.893	Increasing
June 2000	0.881	0.934	0.847	0.965	0.887	Increasing
September 2000	0.776	1.000	1.000	0.870	0.853	Increasing
December 2000	0.873	0.983	0.852	0.857	0.854	Increasing
March 2001	0.943	0.967	0.841	0.931	0.824	Increasing
June 2001	0.978	1.000	0.817	0.949	0.851	Increasing
September 2001	1.000	0.948	0.794	0.897	1.000	Increasing
December 2001	0.984	0.978	0.849	0.807	0.929	Increasing
March 2002	1.000	0.978	0.901	0.968	0.848	Increasing
June 2002	0.990	0.970	0.918	0.984	0.953	Increasing
September 2002	0.967	0.966	0.848	0.937	1.000	Increasing
December 2002	0.942	0.970	0.922	0.898	1.000	Increasing
March 2003	0.961	0.985	1.000	0.990	1.000	Increasing
June 2003	0.958	0.978	1.000	1.000	1.000	Increasing
September 2003	0.997	0.979	0.841	0.921	0.912	Increasing
December 2003	-	0.986	0.885	-	1.000	Increasing
Average	**0.948**	**0.976**	**0.879**	**0.916**	**0.937**	

Because of this, a further column is added to Table 4, to show the returns to scale status of ASB, as estimated by the variable returns to scale model. ASB is shown to be consistently exhibiting increasing returns to scale, whereas, if we look across all 98 observations, we find that there are 72 instances of increasing returns to scale, 11 of constant returns to scale and 15 of decreasing returns to scale (which are split between BNZ, NBNZ and Westpac).

[9] The efficiency improvement for ASB is the only one that shows as statistically significant at the 5% level (according to the Mann-Whitney test)

It is further noted that estimated scale efficiency for ASB Bank increased throughout the period of the study, corresponding with a steady increase in the bank's size (from total assets of $14.1 billion as at 31 March 1999 to $30.4 billion as at 31 December 2003). As at 31 March 1999, ASB was only two thirds the size of the next largest bank (ANZ), whereas by 30 September 2003 it had overtaken it in size. The trend in banks' relative sizes is shown in Figure 2.[10]

Table 5: Scale efficiency scores, TSB not included.

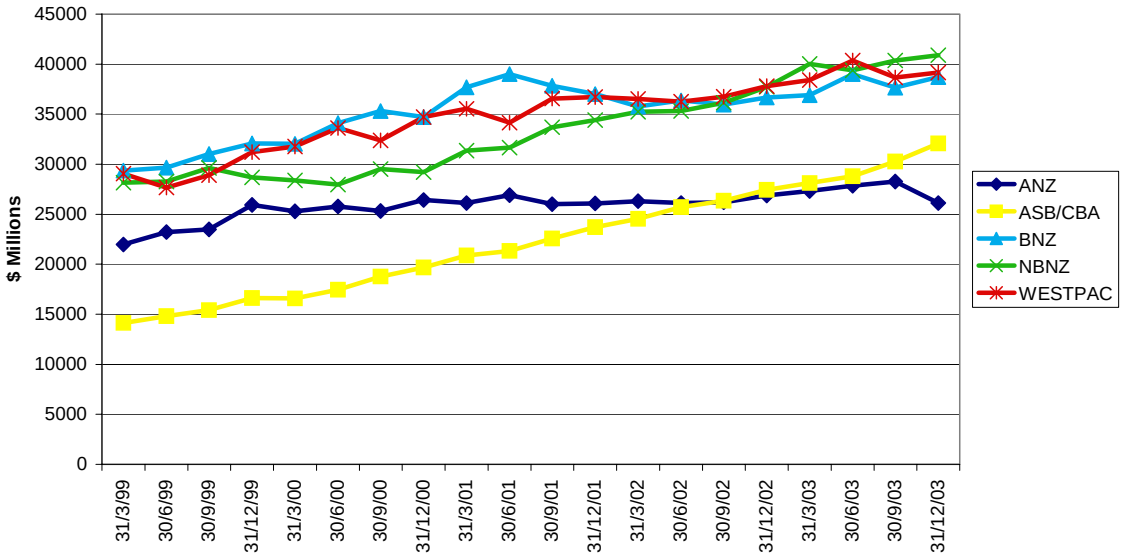
Quarter ending	ANZ	ASB	BNZ	NBNZ	Westpac
March 1999	0.947	0.798	0.965	0.972	1.000
June 1999	1.000	0.813	0.961	0.952	0.992
September 1999	1.000	0.827	0.939	0.951	1.000
December 1999	0.984	0.788	0.953	0.961	0.986
March 2000	0.976	0.809	0.963	0.950	0.997
June 2000	0.956	0.821	0.992	0.946	1.000
September 2000	0.982	0.842	0.850	0.956	0.999
December 2000	0.965	0.836	0.994	0.966	0.992
March 2001	0.969	0.859	0.993	0.958	0.996
June 2001	0.977	0.931	0.998	0.943	0.989
September 2001	1.000	0.893	0.991	0.976	0.975
December 2001	0.992	0.899	0.998	0.978	0.964
March 2002	0.987	0.875	0.975	0.981	0.991
June 2002	0.991	0.922	0.998	0.992	0.996
September 2002	0.997	0.936	0.996	0.997	1.000
December 2002	0.992	0.943	0.962	0.989	0.993
March 2003	0.994	0.936	1.000	0.991	1.000
June 2003	0.994	0.958	1.000	1.000	0.998
September 2003	0.999	0.957	1.000	1.000	0.997
December 2003	-	0.961	0.997	-	0.996
Average	**0.984**	**0.880**	**0.976**	**0.972**	**0.993**

It thus appears reasonable that the scale inefficiency for ASB, particularly during the earlier part of the period of the study, might be a reflection of a genuine scale efficiency effect.

The extent of the scale inefficiency estimated is perhaps rather larger than might be expected, however, having regard to the extent of scale inefficiency found by Berger & Humphrey (1991), particularly as the extent of the difference in size is not especially great.

[10] Figures are shown for the ASB and the (New Zealand branch of the) CBA combined, although the overwhelming majority of the assets (more than 90%) are in fact attributable to ASB. The CBA branch did not obtain separate registration as a New Zealand banks until June 2000.

Figure 2 - the relative sizes of New Zealand's major banks - 1999 to 2003



It was therefore considered appropriate to run a further set of models which included TSB, a bank which is a great deal smaller (with total assets of only $2.03 billion by December 2003). Once again, we ran both CRS and VRS models (with the results reported in Tables 6 and 7 respectively), and use the results from the two to generate measures of scale efficiency (with the results reported in Table 8).

We find that the introduction of the 20 new cases to our data set has almost no impact on the efficiency scores for the banks other than TSB in the CRS model. One would normally expect a reduction in the other firms' efficiency scores, but the minimal change observed (despite some TSB cases now being on the efficient frontier) suggests that the efficient frontier is in fact relatively stable.

TSB shows as having the highest average efficiency score, although the difference between its efficiency and that of ANZ and Westpac (with the next highest average scores, and the higher scores for which were found to be statistically significant in our review of the model whose results were reported in Table 3) are not significant at the 5% level (according to the Mann-Whitney test).

When the VRS model is used, the differences in efficiency scores between the models with and without TSB are very much greater, particularly for ASB, where the difference is significant at the 5% level (according to the Mann-Whitney test). This

suggests that the efficient frontier from the VRS models in not stable, and we can see from the results reported in Table 8 that there is no longer such obvious support for the conclusion that ASB might be scale inefficient. This is confirmed by the lack of any significance in the differences in ASB's efficiency scores, at the 5% level, between the figures reported in Tables 6 and 7.

Table 6: Efficiency scores using CRS model, TSB included.

Quarter ending	ANZ	ASB	BNZ	NBNZ	TSB	Westpac
March 1999	0.885	0.798	0.842	0.887	0.831	1.000
June 1999	1.000	0.812	0.905	0.816	0.934	0.976
September 1999	1.000	0.789	0.728	0.840	0.938	1.000
December 1999	0.922	0.760	0.793	0.815	1.000	0.939
March 2000	0.864	0.795	0.819	0.871	0.964	0.890
June 2000	0.842	0.767	0.840	0.913	0.989	0.886
September 2000	0.761	0.842	0.850	0.832	0.991	0.852
December 2000	0.843	0.822	0.847	0.828	0.954	0.848
March 2001	0.914	0.831	0.835	0.892	0.817	0.821
June 2001	0.955	0.931	0.815	0.895	0.954	0.841
September 2001	1.000	0.847	0.787	0.875	0.906	0.975
December 2001	0.976	0.879	0.847	0.779	1.000	0.896
March 2002	0.987	0.856	0.879	0.948	0.986	0.840
June 2002	0.981	0.894	0.915	0.953	1.000	0.949
September 2002	0.964	0.904	0.843	0.924	1.000	1.000
December 2002	0.934	0.915	0.887	0.871	0.953	0.993
March 2003	0.954	0.921	1.000	0.979	0.919	1.000
June 2003	0.950	0.936	1.000	1.000	0.984	0.998
September 2003	0.995	0.937	0.839	0.899	0.979	0.907
December 2003	-	0.948	0.883	-	0.967	0.993
Average	**0.933**	**0.859**	**0.858**	**0.885**	**0.953**	**0.930**

We do, however, find TSB to be significantly scale inefficient at the 5% level. A review of the figures reported in Table 8 suggests that this scale inefficiency was most evident at the beginning of the period, when the bank was at its smallest (with total assets of only $1.1 billion as at 31 March 1999). As we can see from the final column of Table 7, TSB also shows a predominance of increasing returns to scale during the earlier part of the period studied (although constant returns to scale are found more often during the latter part of the period). If we remove the first three observations, there is no longer support for the existence for scale inefficiency.

Table 7: Efficiency scores using VRS model, TSB included.

Quarter ending	ANZ	ASB	BNZ	NBNZ	TSB	Westpac	TSB's returns to scale status
March 1999	0.886	0.799	0.842	0.889	1.000	1.000	Increasing
June 1999	1.000	0.812	0.908	0.818	1.000	0.976	Increasing
September 1999	1.000	0.790	0.733	0.844	1.000	1.000	Increasing
December 1999	0.923	0.760	0.796	0.817	1.000	0.939	Constant
March 2000	0.865	0.797	0.822	0.871	0.990	0.890	Increasing
June 2000	0.843	0.770	0.840	0.914	1.000	0.887	Increasing
September 2000	0.763	0.845	1.000	0.833	1.000	0.852	Increasing
December 2000	0.845	0.825	0.847	0.829	0.986	0.854	Increasing
March 2001	0.915	0.834	0.836	0.893	0.897	0.821	Increasing
June 2001	0.956	0.933	0.815	0.895	0.989	0.842	Increasing
September 2001	1.000	0.849	0.794	0.875	0.933	1.000	Increasing
December 2001	0.976	0.881	0.847	0.789	1.000	0.929	Constant
March 2002	0.987	0.858	0.879	0.950	0.994	0.840	Increasing
June 2002	0.982	0.896	0.916	0.976	1.000	0.949	Constant
September 2002	0.964	0.906	0.848	0.934	1.000	1.000	Constant
December 2002	0.934	0.917	0.922	0.898	0.955	1.000	Constant
March 2003	0.955	0.923	1.000	0.990	0.920	1.000	Increasing
June 2003	0.952	0.938	1.000	1.000	0.988	1.000	Constant
September 2003	0.996	0.938	0.841	0.921	0.983	0.912	Constant
December 2003		0.949	0.885		0.977	1.000	Constant
Average	**0.934**	**0.861**	**0.869**	**0.891**	**0.981**	**0.935**	

This analysis highlights the dangers of using the VRS model in DEA studies, both in terms of reduced frontier stability and for drawing conclusions in respect of frontier stability. The problem is the one highlighted by Dyson et al (2001), and commented on above in Section 3 of this paper: because of the way a VRS model envelops points more tightly, small and large units may tend to be over-rated in the efficiency assessment. This means that scale inefficiencies identified for such institutions may be spurious, with the actual cause of inefficiency being X-inefficiency.

Once we introduced a further set of DMUs, the smaller ASB units were no longer on the VRS frontier, and the scale effects were no longer evident. With TSB, it is the observations corresponding to the bank being at its smallest size that suggest scale inefficiency, but we cannot be sure that this genuinely reflects scale inefficiency, rather than X-inefficiency.

Table 8: Scale efficiency scores using VRS model, TSB included.

Quarter ending	ANZ	ASB	BNZ	NBNZ	TSB	Westpac
March 1999	0.999	0.999	1.000	0.998	0.831	1.000
June 1999	1.000	1.000	0.997	0.998	0.934	0.999
September 1999	1.000	1.000	0.993	0.995	0.938	1.000
December 1999	0.999	1.000	0.996	0.997	1.000	1.000
March 2000	0.999	0.997	0.997	1.000	0.974	1.000
June 2000	0.998	0.996	1.000	0.999	0.989	0.999
September 2000	0.998	0.996	0.850	0.999	0.991	1.000
December 2000	0.997	0.996	1.000	0.999	0.967	0.992
March 2001	0.999	0.996	1.000	0.999	0.910	1.000
June 2001	1.000	0.997	1.000	0.999	0.965	1.000
September 2001	1.000	0.997	0.991	1.000	0.971	0.975
December 2001	1.000	0.997	1.000	0.987	1.000	0.964
March 2002	1.000	0.997	0.999	0.998	0.992	1.000
June 2002	0.999	0.998	0.999	0.977	1.000	1.000
September 2002	1.000	0.998	0.994	0.989	1.000	1.000
December 2002	0.999	0.998	0.962	0.971	0.997	0.993
March 2003	0.998	0.998	1.000	0.989	0.999	1.000
June 2003	0.998	0.998	1.000	1.000	0.995	0.998
September 2003	0.999	0.998	0.998	0.976	0.996	0.994
December 2003	-	0.999	0.997	-	0.990	0.993
Average	**0.999**	**0.998**	**0.989**	**0.993**	**0.972**	**0.995**

6. Discussion and conclusion

The results reported above cast doubt on the likelihood that New Zealand banks actually enjoy scale efficiencies from increasing their size, as might be suggested by the trend in cost to income ratios (reviewed in Section 4 above). The multivariate efficiency scores generated by DEA models do not show such a strong improving trend through time. In this respect we may look at a cost to income ratio as a one-input/one output DEA model, but with the output as a denominator (so that a lower number indicates superior performance), and with optimal efficiency not standardised to one. The key differences are that the efficiency models used in this study have two further inputs – interest cost and equity – while the two outputs provide a decomposition of the gross income used as the output in cost to income ratio analysis.

My suspicion is that it is the utilisation of funds, reflected in both equity and interest costs, that provides the major cause of different outcomes in relation to finding benefits from increased scale. This would suggest that, although banks have been

reducing their utilisation of those resources that are aggregated into non-interest expense, they have compensated for this by increasing their utilisation of the resources reflected in interest expense and equity. A key challenge for future research is then in devising a way of proving this, having regard to the further complications that interest expense is also impacted by the general level of interest rates, which also changes through time. Possible approaches include investigation of the weights and weighted data in DEA models, including trying use of models with weight restrictions. Analysis is also complicated, however, by the way in which the margins between banks' cost of deposits and the (benchmark) 90-day bank bill rate moves according to whether the general level of interest rates is rising or falling.

Berger & Mester (2003) note the way in which changes in the cost of funds and their impact on costs can lead to different results for cost productivity than those that might be inferred from a review of cost ratios alone. This is because cost productivity takes into account changes in the business conditions faced by banks, and there is a general perception that market conditions in New Zealand have become more competitive, which has impacted both on net revenues, and on banks' cost of funds.

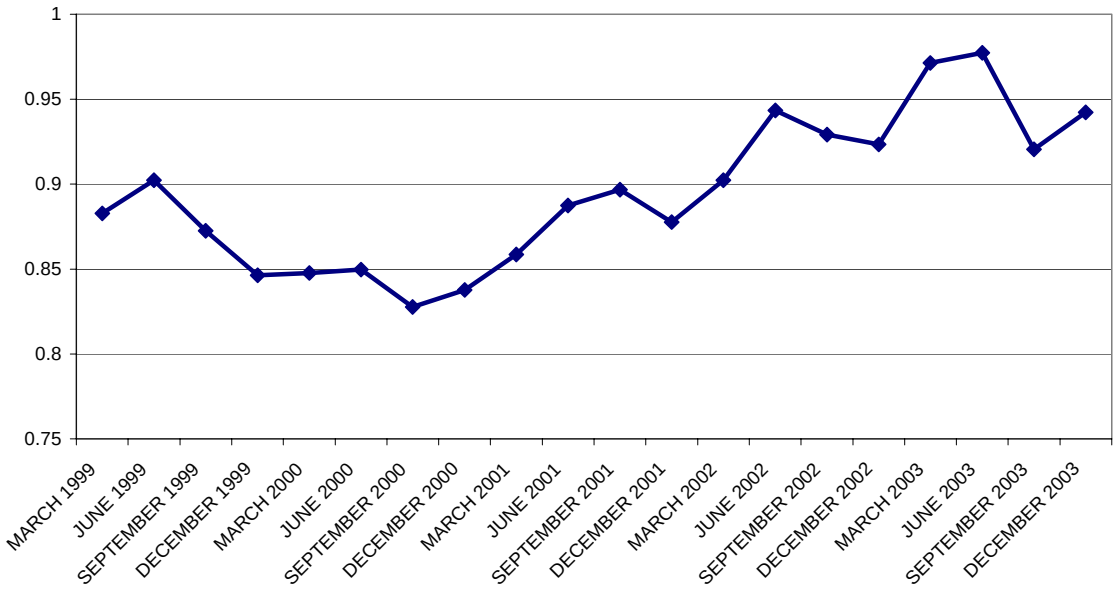
A further factor that explains the rather lesser degree of improvement in efficiency scores is in our selection of output measures. Banks' revenues, in the form of net interest income and non-interest income, have also been reducing over time relative to the resources of funds (reflected in interest costs) and equity that the banks have been using to generate them.

We have nonetheless affirmed that there are dangers in using VRS models to look for scale economies, and provided a practical demonstration of the issues raised by Dyson et al (2001). Figure 3 shows that banks appear to have become slightly more efficient, but this cannot obviously be attributed to the effects of scale.

We have also seen that there was, up to the acquisition of the NBNZ by the ANZ, a tendency for the sizes of New Zealand's major banks to converge, which suggests that the size then achieved may have allowed them to operate at an efficient scale, consistent with propositions advanced by Clark (1988) and Stiroh (2000). This would be consistent with the trend in average efficiency across the five larger banks over the

period of the study, as shown in Figure 3.[11] It will be interesting to see, over a longer period of time, whether the (now) much larger ANZ-National Bank will be as efficient as the other banking firms operating in the New Zealand market.

Figure 3: average (all-bank) efficiency through time - 1999 to 2003



It is a pity that we have to wait until after a bank merger has been bedded down to find our whether there were any social benefits of it to justify it. In the meantime, there is also opportunity to and justification for undertaking further work looking at competitive conditions in the New Zealand banking market, and to ascertain their relationship with bank efficiency. This may give us a better appreciation of the realities of scale efficiency for New Zealand banking.

[11] The relevant data have been taken from the (undistorted and apparently stable) CRS model which did not include TSB.

7. References:

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

Avkiran, N. K. (2002). *Productivity Analysis in the Service Sector with Data Envelopment Analysis.* (Second Edition). Camira, Queensland: N K Avkiran.

Berger, A. N. (1995, May). The profit-structure relationship in banking – tests of market-power and efficient-structure hypotheses. *Journal of Money, Credit, and Banking.* 27 (2). 404-431.

Berger, A. N.; Hanweck, G. A. & Humphrey, D. B. (1987). Competitive viability in banking – scale, scope, and product mix economies. *Journal of Monetary Economics.* 20. 501-520.

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. (1992). Megamergers in banking and the use of cost efficiency as an antitrust defence. *The Antitrust Bulletin.* 37. 541-600.

Berger, A. N.; Hunter, W. C. & Timme, S. G. (1993). The efficiency of financial institutions: a review and preview of research past, present and future. *Journal of Banking and Finance.* 17. 221-249.

Berger, A. N. & Mester, L. J. (1997). Inside the black box: what explains differences in the efficiencies of financial institutions? *Journal of Banking and Finance.* 21. 895-947.

Berger, A. N. & Mester, L. J. (2003). Explaining the dramatic changes in performance of US banks: technological change, deregulation, and dynamic changes in competition. *Journal of Financial Intermediation.* 12. 57-95.

Boussofiane, A.; Dyson, R. G. & Thanassoulis, E. (1991). Applied data envelopment analysis. *European Journal of Operational Research.* 52. 1-15.

Chu, S. F. & Lim, G. H. (1998). Share performance and profit efficiency of banks in an oligopolistic market: evidence from Singapore. *Journal of Multinational Financial Management.* 8. 155-168.

Clark, J. A. (1988, September/October). Economies of scale and scope at depository financial institutions: a review of the literature. *Economic Review (Federal Reserve Bank of Kansas City).* 16-33.

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

DeYoung, R. (1997). Measuring bank cost efficiency: don't count on accounting ratios. *Financial Practice and Education.* Spring/Summer. 20-31.

DeYoung, R. (1998). Management quality and X-efficiency in national banks. *Journal of Financial Services Research.* 13 (1). 5-22.

DeYoung, R. & Hasan, I (1998). The performance of de novo commercial banks: a profit efficiency approach. *Journal of Banking and Finance.* 22. 565-587.

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.

Golany, B. & Storbeck, J. E. (1999, May-June). A Data Envelopment Analysis of the Operational Efficiency of Bank Branches. *Interfaces.* 29. 14-26.

Goldberg, L. G. & Rai, A. (1996). The structure-performance relationship for European banking. *Journal of Banking and Finance.* 20. 745-771.

Hartman, T. E.; Storbeck, J. E. & Byrnes, P. (2001). Allocative efficiency in branch banking. *European Journal of Operational Research.* 134. 232-242.

Humphrey, D. B. (1985). Costs and scale economies in bank intermediation. Chapter 24 in Aspinwall, R. C & Eisenbeis, R. A. (eds), *Handbook for Banking Strategy*. (pp 745-783).

Humphrey, D. B. (1990, September/October). Why do estimates of bank scale economies differ? *Economic Review (Federal Reserve Bank of Richmond).* 38-50.

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

McAllister, P. H. & McManus, D. (1993). Resolving the scale efficiency puzzle in banking. *Journal of Banking and Finance.* 17. 389-405.

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

Noulas, A. G.; Ray, S. C. & Miller, S. M. (1990, February). Returns to scale and input substitution for large U.S. banks. *Journal of Money, Credit and Banking.* 22 (1). 94-108.

Resti, A. (2000). Efficiency measurement for multi-product industries: a comparison on classic and recent techniques based on simulated data. *European Journal of Operational Research.* 121. 559-578.

Sealey, C. W. & Lindley, J. T. (1977). Inputs, outputs, and a theory of production and cost at depository financial institutions. *Journal of Finance.* 32 (4). 1251-1266.

Shaffer, S. (2004). Patterns of competition in banking. *Journal of Economics and Business.* 56. 287-313.

Siems, T. F. & Clark, J. A. (1997, December). Rethinking bank efficiency and regulation: how off-balance-sheet activities make a difference. *Financial Industry Studies (Federal Reserve Bank of Dallas).* 1-12.

Stiroh, K. J. (2000). How did bank holding companies prosper in the 1990s? *Journal of Banking and Finance.* 24. 1703-1745.

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

Tripe, D. W. L. (1998). *Cost to Income Ratios in Australasian Banking.* Available at http://centre-banking-studies.massey.ac.nz/research_prog.asp

Tripe, D. W. L. (2003). Trends in New Zealand bank efficiency over time. *Applied Econometrics and International Development.* 3 (1). 55-80.

Tulkens, H. & Vanden Eeckaut, P. (1995). Non-parametric efficiency, progress and regress measures for panel data: methodological aspects. *European Journal of Operational Research.* 80. 474-499.