Empirical Study on Efficiency and Productivity of the Banking Industry in Egypt

Malak REDA

Abstract

In 1991, Egypt introduced a series of financial reforms to boost the efficiency and productivity of Egyptian banks by limiting state interventions and enhancing the role of market forces. Enticed by the tremendous transformation of this industry, the current study measures the efficiency and productivity change of Egyptian commercial banks from 1995 to 2003, using a non-parametric technique called Data Envelopment Analysis (DEA) and Malmquist Productivity Index. Results indicate that over the period covered by this study, Egyptian commercial banks’ technical inefficiency was 22 percent. In general, smaller banks were found to be least efficient. Malmquist results for a panel of 24 banks indicated that the productivity of commercial banks deteriorated by four percent per year on average during the study period. Moreover, most Egyptian banks operate at incorrect scale. A large majority experienced increasing returns to scale (IRS) in their operations, implying that substantial gains could be obtained from altering scale via either internal growth or consolidation in the sector. Our main recommendation for the government is to adopt policies that would foster competition in the banking sector. We also recommend that the industry devise incentive schemes to improve managerial efficiency through greater investment in technology and skills enhancement.

Keywords: Egyptian banks, productivity change, DEA, Malmquist index.

1. Introduction

In 1991, Egypt introduced an extensive economic reform and structural adjustment programme (ERSAP) to transform the economy from an inward-centrally planned one dominated by the public sector to an outward looking economy led by the private sector. Liberalisation and privatisation of the
financial sector in general, and the banking system in particular, were crucial to the intended transformation of the economy. Egypt’s move towards this transformation of the economy was in response to the general globalisation process.

Within the context of globalisation, liberalisation of financial markets worldwide has led to deeper integration of financial institutions (Ragunathan 1999). As a result, financial institutions today face a fast-paced, dynamic, and competitive environment on a global scale. Given such competitive environment, financial sector supervisors, as well as financial institutions, are required to examine their performance as their survival is likely to depend on their productive efficiencies.

Some earlier studies (Berger and Humphrey 1991, and Berger et al. 1993) demonstrate that inefficiencies, particularly in the banking sector, are more important than scale and scope issues. Consequently, firms have been working towards improving their productive efficiencies in the changing and highly competitive environment (Harker and Zenios 2000).

Motivated by the catalytic role the Egyptian banking sector plays in the Egyptian economy, the objective of the current paper is to assess the performance of the Egyptian commercial banks in terms of efficiency and productivity growth over the 1995-2003 period. The study uses a non-parametric technique data envelopment approach (DEA) to measure bank efficiency, and the Malmquist productivity index to assess productivity change. As efficient banking systems tend to contribute extensively to higher economic growth in any country, studies of this nature are very important for policy makers, industry leaders and others, who are reliant on the banking sector. The paper should hopefully fill a demanding gap in this regard as it examines the performance of Egyptian commercial banks following deregulation. By doing so, we should be able to draw some indicative conclusions and make recommendations for the improvement of banking efficiency in Egypt.

Besides, the current study is important because, first and foremost, there is no empirical study, to the best of our knowledge, which addresses technical and scale efficiency issues of Egyptian banks per se. Thus, this paper will complement and extend the international banking efficiency literature, which is substantially skewed towards the banks of developed countries (Berger and Humphrey 1997; Berger et al. 1999; Isik and Hassan 2002 a, b; Yildirim and Philippatos 2003). Second, like in many emerging markets, banks in Egypt remain the dominant financial institutions. They control much of the financial flow and still possess a large portion of the financial assets in the economy. Moreover, many non-bank financial institutions in Egypt, like in other emerging economies, are also affiliates of banks. Thus, competition in the banking market is somewhat constrained.

Inefficient use of resources stemming from weak competitive pressures may result in higher operating costs for banks, leading to more expensive services and higher loan prices. This has the impact of raising costs of funding for all economic units. It eventually makes many public and private projects...
unfeasible ventures by forcing their internal rate of return (IRR) to fall below the hurdle rate. Thus, the study of how efficiently the banks are operating is also important for the government and industrial firms. Banks are in effect the only external source of funds for public and private investments in the country. Third, one of the cardinal missions of bank regulators is to ensure the efficient functioning of the banking system, its safety, and soundness.

The rest of the paper is organised as follows: Section 2 discusses the period of study. Section 3 examines the Egyptian commercial banking briefly, while the methodology of the study is discussed in Section 4. Section 5 details out the variables and the data used in the study. The results are discussed in Section 6. Section 7 is about the conclusions drawn from the study.

2. Period of Study

The study covered a period of nine years, beginning 1995 and ending in 2003. Two factors determined the selection of this period. One was the absence of comparable data of adequate quality before 1995. The other was that 2003 became a turning point that signalled a new phase for the banking sector reforms. It was in June that year that a new banking law (Law 88 of 2003) was passed.

The law, whose executive regulations were issued in March 2004, addressed key deficiencies in previous laws governing the banking system. One of the implications of the law is the improvement in capitalisation of operating banks over a maximum period of three years ending 2006. The requirement of increased capitalisation led to consolidation of several banks through mergers and acquisitions, hence the reduction of the number of operating banks. Similarly, in December 2003, a new board of directors was appointed in the CBE. The new management of the Central Bank is committed to a comprehensive banking sector reform programme that comprises banking sector restructuring, consolidation of banks through mergers and acquisitions, and privatisation of joint venture banks and one of the four public banks. Examining the impact of the reforms undertaken as of 2003 is beyond the scope of the current study. It constitutes a potential area for future research. Moreover, we note that unless there is a fundamental innovation or shock to the banking technology, efficiency measures are expected to be stable over time. Indeed, it is unlikely that efficiencies will fluctuate markedly over a short period of time (Berger and Humphrey 1997). Several studies have found that efficiency is reasonably persistent over time (Berger and Humphrey 1991; Berg et al. 1992; Kwan 1997; Eisenbeis et al. 1996). A relatively long period is needed for developments in the macro economy, market place, and regulatory environment to exert their influence on the banking technology. Isik (2004) assumes that a four-year time interval is adequate to capture any structural change in the performance of banks.
3. Egyptian Banking at a Glance

Egypt’s shift to a market-oriented economy starting 1991 was backed by various regulatory and legislative adjustments, as well as economic reform policies. The latter included the following:

- Price liberalisation, and elimination of most subsidies and government pricing in many industries.
- A single unified floating exchange rate instead of multiple exchange rates. In January 2001, the Egyptian government moved from a fixed foreign exchange system to a crawling peg foreign exchange regime with a +/-1.5 percent band that was eventually raised to a +/- 3 percent band. In late January 2003, the Egyptian government adopted a flexible exchange rate regime.
- Liberalisation of foreign exchange and capital movements.
- Lifting of most interest rate ceilings, reduction in state directed and subsidised credits, and a shift to indirect tools of monetary policy.
- Creation of a domestic capital market, including mutual funds, alternative financial institutions, and the reanimation of the stock market.
- Abolition of most quantitative restrictions, lower tariffs, and increased freedom to export.

Following the introduction of these deregulation policies, the Egyptian banking sector experienced vast changes. Deregulation covered a number of areas, such as interest on loans and deposits, banking services fees and charges, exchange rates, credit ceilings, and allowing foreign branches to operate in local currency. The changes also made it possible for foreign investors to enjoy ownership in Egyptian public sector banks, and allowed public business sector to deal with all banks, whether public or private. More adjustments included areas such as the classification of and provisioning for non-performing loans, capital adequacy, financial reporting, and divestiture of public sector bank’s interest in joint venture banks.

In June 1998, the parliament passed Law 155, allowing private sector (foreign or local) participation in the ownership of the public sector banks. However, no action has been taken regarding the privatisation of any of the four public sector banks during the study period.

Similar to the situation in other developing countries, the banking industry in Egypt continues to dominate the financial sector. Banks accounted for more than 62 percent of aggregate financial assets as of June 2003. Contractual savings sector, which includes social insurance funds, commercial insurance, and pension schemes, accounted for 20 percent of financial assets. The capital market accounted for 16 percent of total financial assets. The remaining two percent pertained to leasing and venture capital companies. Majority of banks operating in Egypt perform the traditional banking role of transforming deposits to loans. They are neither active in provision of mortgage banking, nor factoring operations and financial leasing.
According to the classification by the Central Bank of Egypt (CBE), the country had 62 banks by June 2003. This number included 28 commercial banks, of which four were state-owned; 31 business and investment banks, out of which 11 were private and joint venture banks; and 20 branches of foreign banks. In addition, the country had three public specialised banks that provided mid and long-term financing of agriculture, industry and real estate sectors. In 2003, the total number of licensed bank branches working in Egypt reached 2,582, in addition to 39 branches licensed to operate overseas. Figure 1 portrays the structure of the Egyptian banking sector in 2003.

Although the number of public banks is small relative to the total number of operating banks in the country, they account for quite a large share of the banking sector activity. The four public sector commercial banks are the largest operating banking institutions in Egypt in terms of balance-sheet size. They account for nearly 52 percent of total bank assets. They also have a significant market share in retail and corporate banking services through large branch networks and close relationship with state-owned companies. Their share in total deposits amounted to 58 percent in 2003. They were also major participants in the equity capital of most joint-venture banks. The private and joint ventures remain relatively small with modest branching, when compared to the public sector banks. This is despite their rapid growth.

Banks are supervised by CBE, which has made considerable efforts to improve its supervisory system. According to FSAP (Financial Sector Assessment Program) report of 2002 by the IMF, CBE applied most of the Basel Core Principles for Effective Banking Supervision. In order to further strengthen the role the CBE, the Egyptian government, together with the CBE, drafted Law 88 of 2003, titled Central Bank and Banking Sector Law, to increase the degree of independence of CBE in maintaining price stability. The new banking law confirms the supervisory role of the CBE.

As of March 2003, banks had been requested to comply with the new capital adequacy ratio of 10 percent of weighted risk assets. According to new banking law, banks were to raise their paid-up capital to LE 500 million (US$91 million) for Egyptian banks and to US$ 50 million for branches of foreign banks. Other measures that were undertaken by the CBE to improve the supervisory role included the introduction of stress testing techniques, early warning systems, and enhancement of the role of offsite supervision.

Lender of last resort has been a feature of the Egyptian banking system since the issuance of law 163 of 1957, which made the CBE responsible for providing troubled banks with emergency funds. Although a deposit insurance fund was allowed to be established by law in 1992, it has not been implemented.

The CBE is apparently reluctant to license new domestic banks, as it regards the number of existing ones large enough for establishing a competitive market. It introduced an economic needs test in 1993 to limit entry of
Figure 1. Structure of the Egyptian Banking System as of June 30, 2003

Central Bank of Egypt

- Commercial Banks
  - 4 Public Banks (617 Branches)
  - 24 Private & Joint Ventures (383 Branches)
  - 11 Private & Joint Ventures (148 Branches)
  - 20 Foreign Banks (59 Branches)

- Business & Investment Banks
  - Industrial Banks
    - The Egyptian Industrial Development Bank (14 Branches)
  - Real Estate Banks
    - The Arab Egyptian Real Estate Bank (27 Branches)
  - Agricultural Banks
    - Principal Bank for Development & Agricultural Credit (1034 Branches)

- Specialised Banks

Notes: Egyptian banks abroad are not included. Also not included are two banks established under private laws and are not registered with CBE. They are the Arab International Bank and Nasser Social Bank.

- The Egyptian Real Estate Bank had merged with the Arab Real Estate Bank on June 21, 1999.
- The total number of banks as of June 30, 2003 was 62 banks, with a total of 2,582 branches.

new banks. Maintaining an adequate exit mechanism is also crucial for achieving an efficient market. Not allowing banks to fail comes at a high fiscal cost.

Table 1 presents a summary of key indicators of banks broadly broken into public, and private and joint venture banks, for year 2003. The table shows that in terms of equity, the four public banks aggregate equity is almost the same as that of 23 joint ventures. Yet, the assets of public banks are almost twice as much of those of the private sector. In terms of profitability, the mean return on average assets is much higher in the case of private and joint ventures at 0.92, versus 0.16 for the public sector banks. Return on Average equity stood at 8.8 for the private and joint ventures, against 3.9 for the public sector banks.

### 4. Methodology

The term, “productive efficiency”, is commonly used to describe the level of performance of a production unit in terms of its utilisation of resources in generating outputs. Koopmans (1951) defines technical efficiency as a feasible input/output vector, where it is technologically impossible to increase any output without simultaneously reducing another output. This analogy holds for a reduction in any input, or both a reduction in any input and an increase in any output. Farrell (1957) demonstrated that overall efficiency of a production unit is composed of two separate efficiency measures called technical efficiency and allocative efficiency. Farrell measured technical inefficiency as the maximum equi-proportional reduction in all inputs consistent with equivalent production of observed output. A Farrell efficient unit, however, may not be Koopmans efficient, since even after Farrell efficiency is achieved, there may exist additional slack in individual inputs. Allocative efficiency is based on cost considerations, namely input prices. The type of efficiency measured depends on the data availability and appropriate behavioural assumptions (Yin 1999). When only quantities are available, technical efficiency can be
calculated. When both, quantities and prices are available, economic efficiency can be calculated and decomposed into technical and allocative components. Our research focuses on technical efficiency only. Allocative efficiency was not computed due to unavailability of all input and output prices.

Technical inefficiency arises in cases where more of each input is used than what should be required to produce a given level of output. Technical inefficiency is typically attributed to lack of strong competitive pressures, which allow bank managers to continue with less than optimal performance. Because it relies solely on the amounts of inputs and outputs in its calculation and does not involve factor prices, which are mostly market or regulation driven, technical inefficiency is entirely under the control of bank management. It therefore results directly from management laxity and errors (Leibenstein 1966, 1978). Following the same terminology adopted by Isik (2003), we coin the term technical inefficiency to describe managerial inefficiency in banking. In fact, some earlier researchers used technical X-inefficiencies as proxies for management quality. For instance, Barr et al. (1993) and Wheelock and Wilson (1995) make the explicit assumption that technical X-inefficiency is a good measure of management quality in bank failure studies.

Managerial inefficiency or Technical inefficiency consists of two mutually exclusive and exhaustive components. They are pure technical inefficiency (PTE) and scale inefficiency (SE). Pure technical inefficiency is defined as managerial inefficiency devoid of scale effects. When the scale issues are dismantled, technical efficiency (TE) and pure technical efficiency (PTE) scores are the same, as the difference between them refers to scale inefficiency. Thus, PTE refers to proportional reduction in input usage that can be obtained if the bank operates on the efficient frontier. As it results directly from management errors, it is considered one form of managerial inefficiency. Scale inefficiency refers to non-optimal choice of production scale in terms of cost control. A scale efficient firm will produce where there are constant returns to scale (CRS). Thus, when there are increasing returns to scale (IRS), efficiency gains could be obtained by expanding production levels. If decreasing returns to scale (DRS) exist, efficiency gains could be achieved by reducing production levels. As it involves the choice of an inefficient level by management, scale inefficiency is also considered a form of managerial inefficiency. Thus, total managerial inefficiency includes both pure technical inefficiency and scale inefficiency, that is, inefficient level of both inputs and outputs.

4.1. Measuring Frontier Analysis (X-Efficiency) of Banks

At its heart, frontier analysis is essentially a sophisticated way to “benchmark” the relative performance of production units. Frontier analysis provides an overall, objectively determined, numerical efficiency value and ranking of firms (X-efficiency) that is not otherwise available. This attribute makes
frontier analysis particularly valuable in assessing and informing government policy regarding financial institutions.

Frontier efficiency refers to how close financial institutions are to a “best-practice” frontier. Studies of frontier efficiency rely on accounting measures of costs, outputs, inputs, revenues, and profits etc. to impute efficiency relative to the best practice within the available sample. Berger and Humphrey (1997), state that there is no consensus on the preferred method for determining the best-practice frontier against which relative efficiencies are measured. They report that there are at least five types of approaches that have been employed in evaluating the efficiency of financial institutions and branches. These methods differ primarily in the assumptions imposed on the data in terms of:

1. The functional form of the best-practice frontier (a more restrictive parametric functional form versus a less restrictive nonparametric form);
2. Whether or not account is taken of random error that may temporarily give some production units high or low outputs, inputs, costs, or profits;
3. If there is random error, the probability distribution assumed for the inefficiencies (e.g., half normal, truncated normal) is used to disentangle the inefficiencies from the random error.

Thus, the established approaches to efficiency measurement differ primarily in how much shape is imposed on the frontier and the distributional assumptions imposed on the random error and inefficiency. Table 2 summarises the main assumption, benefits and drawbacks of the various approaches to frontier efficiency measurement.

Is there a “Best” Frontier Method?

As shown in Table 2, there are five possible approaches that have been employed in evaluating the efficiency of financial institutions. These differ in their assumptions imposed on the data. The nonparametric approaches do not require explicit specification of the form underlying production relationship. They permit efficiency to vary over time and make no prior assumption regarding the form of the distribution of inefficiencies across observations, except that undominated observations are 100 percent efficient. The key drawback of the nonparametric approaches is their assumption that there is no random error. The parametric approaches specify a functional form for the frontier. The parametric approaches commit “the sin” of imposing a functional form (and associated behavioural assumptions) that presupposes the shape of the frontier.

As stated by Berger and Humphrey (1997), there is no consensus on the preferred method for determining the best-practice frontier against which relative efficiencies are measured. Berger and Humphrey analysed the results obtained from several comparative studies that examined the variation of results due to the adoption of different parametric and non-parametric techniques. The main conclusion they drew can be summarised in the following statement:
<table>
<thead>
<tr>
<th>Approach</th>
<th>Nonparametric frontiers</th>
<th>Parametric frontiers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Data Envelopment Analysis</td>
<td>Free Disposal Hull (FDH)</td>
</tr>
<tr>
<td></td>
<td>Put relatively little structure on the specification of the best practice frontier</td>
<td>Put relatively little structure on the specification of the best practice frontier</td>
</tr>
<tr>
<td>Characteristics</td>
<td>DEA is a linear programming technique where the best-practice or frontier observations are those for which no other decision making unit or linear combination of units has as much or more of every output (given inputs) or as little or less of every input (given outputs)</td>
<td>FDH typically generate larger estimates of average efficiency than DEA (Berger &amp; Humphrey, 1997)</td>
</tr>
<tr>
<td></td>
<td>DEA frontier is formed as a piecewise linear combinations that connect the set of best-practice observations, yielding a convex production possibilities set.</td>
<td>FDH typically generate larger estimates of average efficiency than DEA (Berger &amp; Humphrey, 1997)</td>
</tr>
<tr>
<td></td>
<td>DEA does not require the explicit specification of the form underlying production relationship.</td>
<td>Allows for random error</td>
</tr>
<tr>
<td></td>
<td>Both DEA &amp; FDH approaches permit efficiency to vary over time and make no prior assumption regarding the form of the distribution of efficiencies across observations except that undominated observations are 100% efficient.</td>
<td>Allows for random error</td>
</tr>
<tr>
<td></td>
<td>No random error</td>
<td>No random error</td>
</tr>
</tbody>
</table>

Source: Author’s Summary from Berger and Humphrey (1997)
“...It seems clear that the estimates of mean or median efficiency for an industry may be more consistently reliable guide for policy and research purposes than are rankings of firms by their efficiency value, especially between nonparametric and parametric approaches” – Berger and Humphrey (1997).

The lack of agreement among researchers regarding a preferred frontier model boils down to a difference of opinion on the lesser of two evils. The parametric approaches “commit the sin” of imposing a particular functional form that presupposes the shape of the frontier. If the functional form is misspecified, measured efficiency may be confounded with the specification errors. Usually, a local approximation, such as the translog, is specified. However, this has been shown to provide poor approximations for banking data that are not near the mean scale and product mix. The translog also forces the frontier average cost curve to have a symmetric U-shape in logs. The nonparametric studies impose less structure on the frontier, but commit “the sin” of not allowing for random error owing to luck, data problems, or other measurement errors. If random error exists, measured efficiency may be confounded with these random deviations from the true efficiency frontier.

We intend to use the Data Envelopment (DEA) Analysis approach, as it has several features that make it preferable to other performance measurement techniques, such as traditional ratio analysis and Stochastic Frontier Analysis (SFA). First, being nonparametric, DEA does not require the specification of an a priori, well-defined functional form for the particular production process being analysed. This “flexibility” makes it particularly useful when it is impossible to determine the mode in which the set of resources (inputs) are employed in combination to realise a multiplicity of products (outputs). Second, DEA permits the simultaneous management of more than one input and output because of its capacity to maximise the relationship between a “virtual” output and a “virtual” input (appropriately weighted sums of the vectors of inputs and outputs) typical of banking activity. Third, depending on the particular model selected, DEA can distinguish technical inefficiency from scale and scope inefficiency, since each bank is compared to a peer group homogenous in terms of size and product mix.

DEA has proven to be a popular technique for performance analysis in general, and in the financial services industry in particular. In this regard, the banking sector has a series of characteristics that make it particularly suitable for study through DEA. These are its multi-input and multi-output nature, the non-linearity of its input-output relationships, the non-physical nature of some fundamental resources and products, and the impossibility of drawing on market prices for some of them. As we intend to use DEA approach, it is useful to define the approach and its applications.
4.2. Data Envelopment Analysis

Broadly speaking, the DEA technique defines an efficiency measure of a production unit by its position relative to the frontier of the best performance established mathematically by the ratio of weighted sum of outputs to weighted sum of inputs (see, for example, Norman and Stoker (1991) for a detailed description of the DEA technique). The estimated frontier of the best performance is also referred to as efficient frontier or envelopment surface. The frontier of the best performance characterises the efficiency of production units and identifies inefficiencies based on known levels of attainment. Thus, a production unit attains 100 percent efficiency only when it is not found to be inefficient in using the inputs to generate the output when compared with other relevant production units.

Basic DEA Models

DEA begins with a relatively simple fractional programming formulation. Assume that there are $n$ DMUs (Decision Making Units) to be evaluated. Each consumes different amounts of $i$ inputs and produces $r$ different outputs, i.e., DMU$j$ consumes $x_{ij}$ amounts of input to produce $y_{rj}$ amounts of output. It is assumed that these inputs, $x_{ij}$ and outputs, $y_{rj}$, are non-negative, and each DMU has at least one positive input and output value. The productivity of a DMU can be written as:

$$ h_j = \frac{\sum_{r=1}^{r} U_r y_{rj}}{\sum_{i=1}^{n} v_i x_{ij}} $$

(1)

Where $h$ refers to the efficiency, $j$ the DMU under study, and $x_{ij}$ the amounts of input consumed by DMU$j$ to produce $y_{rj}$ amounts of output.

In this formulation, $u$ and $v$ are the weights assigned to each input and output. By using mathematical programming techniques, DEA optimally assigns the weights subject to respectively two constraints, namely:

1. The weights for each DMU are assigned subject to the constraint that no other DMU has an efficiency greater than 1 if it uses the same weights, implying that efficient DMUs will have a ratio value of 1.
2. The derived weights, $u$ and $v$ are not negative.

The objective function of DMU$k$ is the ratio of the total weighted output divided by the total weighted input:

$$ \text{Maximise } h_k = \frac{\sum_{r=1}^{r} U_r y_{rk}}{\sum_{i=1}^{n} v_i x_{ik}} $$

(2)
This is a simple presentation of basic DEA model.

According to Denizer et al. (2000), Charnes et al. (1978) employed the optimisation method of mathematical programming to generalise the Farrel (1957) single-output/input technical-efficiency measure to multiple-output/multiple-input case. The characteristic of the Charnes, Cooper and Rhodes (CCR) ratio model is the reduction of the multiple-output/multiple-input situation for each DMU to a single virtual output and a single virtual input ratio. This ratio provides a measure of efficiency for a given DMU, which is a function of multipliers. The objective is to find the largest sum of weighted outputs of DMUk, while keeping the sum of its weighted inputs at the unit value, thereby forcing the ratio of the weighted output to the weighted input for any DMU to be less than one. The CCR model is also known as the constant return to scale model. It identifies inefficient units regardless of their scale size. In the CCR models, both technical and scale inefficiency are present.

Banker et al. (1984) take into account the effect of returns to scale within the group of DMUs to be analysed. The purpose here is to point out the most efficient scale size for each DMU and to identify its technical efficiency at the same time. To do so, the Banker, Charnes and Cooper (BCC) model introduces another restriction (convexity) to the envelopment requirements. This model requires that the reference point on the production function for DMUk will be a convex combination of the observed efficient DMUs. The BCC model, known as variable returns to scale model, gives the technical efficiency of DMUs under investigation without any scale effect.

It is possible to use models that provide input-oriented or output-oriented projections for both CCR (constant returns to scale) and BCC (variable returns to scale) envelopment. An input-oriented model attempts to maximise the proportional decrease in input variables while remaining within the envelopment space. On the other hand, an output-oriented model maximises the proportional increase in the output variables, while remaining within the envelopment space.

Scale Issue

Although commercial banks are homogeneous with respect to their organisational structure and objectives, they vary significantly in size and production level. Even after normalising the data, this suggests that the scale of banks
plays an important role in their relative efficiency or inefficiency. As previously stated, the CCR model comprehends both technical and scale efficiency. The BCC model, introduced by Banker, Charnes and Cooper (1984), separates technical efficiency and scale efficiency. BCC also modified the original CCR linear programming formulation by adding a convexity constraint for the production possibility set to estimate not only technical efficiency, but also returns to scale.

Banker, Charnes and Cooper (1984) showed that the CCR measure captures not only the productive inefficiency of a DMU at its actual scale size, but also any inefficiency resulting from its actual scale size being different from the most productive scale size. A most productive scale size maximises average productivity in the long run. In order to maximise average productivity, a DMU would have to increase its scale size if increasing returns to scale were prevailing, and decrease the scale size if decreasing returns to scale were prevailing. It follows that a technically efficient and scale efficient DMU will be in the most productive scale size.

Given that the CCR efficiency score is a product of technical and scale efficiency, and BCC measures pure technical efficiency, the ratio of the efficiency scores is represented by:

\[
S_k = \frac{q_{k, CCR}}{q_{k, BCC}}
\]

It yields a measure of the relative scale efficiency of bank k. If S=1, it is said that bank k is operating at the most efficient scale size. If it is less than unity, there is scale inefficiency for bank k. Thus, (1-S) represents the relative scale inefficiency of a bank (Banker et al. 1984). The units that are CCR efficient will also be scale efficient, since scale was already factored in the CCR model. Thus, the two are equal. The units that are BCC efficient, but inefficient based on the CCR model, have a scale inefficiency. Since they were technically efficient, all of the inefficiencies picked by CCR are due to scale. Those units that are CCR efficient are considered most productive scale sizes, as the average productivity of each of those units is maximised.

This can serve as a useful diagnostic tool for decision makers and bank directors. Once technical and scale efficiencies are isolated, the next step is to determine the share of the overall inefficiency that is attributable to technical inefficiency and scale inefficiency.

We adopted input orientated DEA technique to measure the efficiency and productivity scores of Egyptian banks on a yearly basis over the 1995 to 2003 period because of the expressed interest in the Egyptian banking industry to control costs in recent years following liberalisation. Through input-oriented DEA, we can dwell on the sources of input waste in Egyptian banking and draw some policy conclusions.
4.3. **Malmquist Productivity Index**

We used Malmquist total factor productivity (TFP) index to examine productivity change in the banking industry. Malmquist firm-specific productivity indexes were introduced by Caves *et al.* (1982). The approach has been employed frequently to study bank productivity change. Studies that adopted the Malmquist Productivity Index to measure productivity change in banking include Berg *et al.* (1992), Berg *et al.* (1991), and Mlima (1999) for banks of Nordic countries, Rebelo and Mendes (2000) for Portuguese banks, and Isik and Hassan (2003) for Turkish banks.

Malmquist index uses exclusively quantity information, and thus demands neither problematic price information nor a restrictive behavioural assumption in its calculation. Malmquist index allows us to distinguish between shifts in the production frontier (technology change, TECCH) and movements of firms towards the frontier (efficiency change, EFFCH). Thus, Malmquist total factor productivity index change (TFPCH) is simply the product of efficiency change (EFFCH), which is how much closer a bank gets to the efficient frontier (catching or falling behind), and technological change (TECCH), which refers to how much the benchmark production frontier shifts at each bank’s observed input mix (innovation shock).

We obtain the TECCH and EFFCH indexes under the assumption of constant returns to scale (CRS). This assumes that banks operate at an optimum scale for input minimisation. However, in reality, banks could face scale inefficiencies due to decreasing returns to scale (DRS) or increasing returns to scale (IRS) in their operations resulting from market and regulatory constraints. When we relax the CRS assumption and adopt the more realistic variable returns to scale assumption (VRS), we are able to decompose EFFCH index into pure efficiency change (PEFFCH) and scale efficiency change (SECH) components. (PEFFCH) index measures the changes in proximity of firms to the frontier, devoid of scale effects. SECH shows whether the movements inside the frontier are in the right direction to attain the CRS point, where changes in output result in proportional changes in costs. Briefly, TFPCH= TECCH x EFFCH and EFFCH= PEFFCH x SECH. Thus, TFPCH= TECCH x PEFFCH x SECH. Any value greater (lower) than 1 indicates a growth (fall), while any value equal to 1 indicates stagnation in the relevant index.

EFFCH, PEFFCH, and SECH obviously show the changes in technical efficiency (EFF), pure technical efficiency (PEFF) and scale efficiency (SE), which are calculated relative to both contemporaneous and previous year frontiers. Pure technical inefficiency (PEFF), which is technical inefficiency (EFF) devoid of scale effects, is entirely under control and results directly due to, management errors. Thus, it is also called managerial inefficiency in the literature. It occurs when more of each input is used than should be required to produce a given level of output. It is typically attributed to insufficient competitive pressure that allows management to “get away” with slackened productivity. A scale efficient (SE) firm will produce where there are constant returns to scale. In
case there are increasing returns to scale, then efficiency gains could be achieved by expanding production. If decreasing returns to scale is prevailing, efficiency gains could be attained by reducing production levels. Because it involves the choice of an inefficient level, scale inefficiency is also considered a form of technical inefficiency. Thus, total technical efficiency includes both pure technical and scale inefficiency; that is, inefficient level of both inputs and outputs.

There are output-oriented and input-oriented measures of change in productivity (Coelli (1998)). In this study, we used input-oriented Malmquist productivity index, where the input-oriented productivity measures focus on the minimum level of inputs that could be used to produce a given output vector and a given production technology relative to the observed level of inputs. DEA was used in measuring the distance function that makes up the Malmquist TFP index.

**The Model**

Malmquist productivity index is defined using distance functions. Suppose the function that describes the technology of production is given as: \( F(X, Y) = 0 \) where \( X = (x_1, x_2, ..., x_M) \) is the input vector and \( Y = (y_1, y_2, ..., y_s) \) is the output vector.

Caves et al. (1982) provided an alternative interpretation of production technology using the concept of ‘distance function’. They defined the output distance function as \( D_0(X, Y) = \min_{\mu} \{ \mu : F(X, Y/\mu) = 0 \} \) where \( \mu_y \) is the minimum equi-proportional change in the output vector. The output distance function measures the maximum proportional change in output required to place \((X, Y)\) on the efficiency frontier. If the evaluated production unit is efficient, \( D_0(X, Y) = 1 \). Otherwise, \( D_0(X, Y) < 1 \). Distance function may also be computed with input orientation reference technology in a certain time period, and CRS or VRS specification. Let \( D^t_0 \) (CRS) and \( D^{t+1}_0 \) (VRS) denote the output distance function computed with period \( t \) technology and with CRS and VRS specification respectively.

Caves et al. (1982) define the output based Malmquist productivity index to compare performance of a production unit in time period \( t \) and \( t + 1 \) with reference to period \( t \) technology as

\[
M^t_0(X_{t+1}, Y_{t+1}, X_t, Y_t) = \frac{D^t_0(X_{t+1}, Y_{t+1})}{D^t_0(X_t, Y_t)}
\]

Alternatively, we may define output based Malmquist productivity index with reference to period \( t + 1 \) technology as

\[
M^{t+1}_0(X_{t+1}, Y_{t+1}, X_t, Y_t) = \frac{D^{t+1}_0(X_{t+1}, Y_{t+1})}{D^{t+1}_0(X_t, Y_t)}
\]
$M_0 > 1$ indicates higher productivity in period $t + 1$ than in period $t$. Whereas $M_0 < 1$ implies decline in productivity in period $t + 1$ as compared to $t$. $M_0 = 1$ indicates stagnation in productivity between the period $t$ and $t + 1$.

Fare et al. (1994) define an index that incorporates Malmquist indices in both periods. They suggest this to avoid choice of the time period arbitrarily. They specify the output based Malmquist productivity change index as:

$$M_0(X_{t+1}, Y_{t+1}, X_t, Y_t) = \left[ \frac{D_0(X_{t+1}, Y_{t+1})}{D_0(X_t, Y_t)} \right] \left[ \frac{D_0(X_{t+1}, Y_{t+1})}{D_0(X_t, Y_t)} \right]^{1/2}$$

(6)

$$= \left[ \frac{D_0(X_{t+1}, Y_{t+1})}{D_0(X_t, Y_t)} \right] \left[ \frac{D_0'(X_{t+1}, Y_{t+1})}{D_0'(X_t, Y_t)} \right]^{1/2}$$

(7)

$\frac{D_0'(X_{t+1}, Y_{t+1})}{D_0'(X_t, Y_t)}$ is the change in relative technical efficiency between periods $t$ and $t + 1$.

and $\left[ \frac{D_0'(X_{t+1}, Y_{t+1})}{D_0'(X_t, Y_t)} \right]^{1/2}$ captures the shift in technology (technological change) between the two time periods evaluated at $(X, Y)$ and $(X_{t+1}, Y_{t+1})$. Now, for each production unit, we define five Malmquist indices for period $t + 1$ relative to period $t$:

**Total factor productivity change index (TFPCI)**

$$= \left[ \frac{D_0'(CRS)(X_{t+1}, Y_{t+1})}{D_0'(CRS)(X_t, Y_t)} \right] \left[ \frac{D_0'(CRS)(X_{t+1}, Y_{t+1})}{D_0'(CRS)(X_t, Y_t)} \right]^{1/2}$$

(8)

= (Technical efficiency change index) (Technological change index)

**Technological change index (TCI)**

$$= \left[ \frac{D_0'(CRS)(X_{t+1}, Y_{t+1})}{D_0'(CRS)(X_t, Y_t)} \right] \left[ \frac{D_0'(CRS)(X_{t+1}, Y_{t+1})}{D_0'(CRS)(X_t, Y_t)} \right]^{1/2}$$

(9)

2. This is the geometric mean of output based Malmquist productivity indices with reference to period $t$ and period $t + 1$ technology.
Technical efficiency change index (TECI)

\[
\frac{D_{0}^{t+1}(CRS)(X_{t+1},Y_{t+1})}{D_{0}^{t}(CRS)(X_{t},Y_{t})}
\]  

(10)

Pure technical efficiency change index (PTECI)

\[
\frac{D_{0}^{t+1}(VRS)(X_{t+1},Y_{t+1})}{D_{0}^{t}(VRS)(X_{t},Y_{t})}
\]  

(11)

Scale efficiency change index (SECI) = Technical efficiency change index/ Pure technical efficiency change index

\[
= \frac{D_{0}^{t+1}(CRS)(X_{t+1},Y_{t+1})}{D_{0}^{t}(CRS)(X_{t},Y_{t})} \div \frac{D_{0}^{t+1}(VRS)(X_{t+1},Y_{t+1})}{D_{0}^{t}(VRS)(X_{t},Y_{t})}
\]  

(12)

A value of less than 1 in the index indicates a decline in efficiency. A value equal to 1 indicates stagnation. If it is greater than 1, it indicates a growth between period \(t\) and \(t+1\) from the perspective of period \(t\) technology.

5. Variables and Data

Variables Selection

The number of variables used in DEA is critical. Inclusion of many variables is not a viable option in DEA. As the number of variables in the model increases, more and more production units become efficient. On the other hand, when relevant variables are omitted, DEA underestimates efficiency and the effect of this is more severe than when irrelevant variables are included in the DEA model. Lack of a standard structured approach to variable selection in DEA makes the task of variable selection even more difficult.

Berger and Humphrey (1997) commented on the difficulty of variable selection in performance appraisal of banks using the DEA technique by stating that there was no “perfect approach” on the explicit definition and measurement of banks’ inputs and outputs.

The choice of variables in efficiency studies affects the expected results significantly. A number of studies have presented results that differ due to variable selection (Favore and Pappi 1995; Hunter and Timme 1995). There are, however, certain limitations on variable selection due to reliability of the data.
Definition of commercial bank’s function:

The definition of a bank’s function is one of the complications in efficiency studies. It affects variable selection and associated results. In order to provide guidelines for variable selection and application, it is useful to define the banking process.

Two approaches in the banking literature discuss the activities of banks: The production approach and the intermediation approach (Denizer, Dinc, and Tarimcilar 2000). Both approaches apply the traditional microeconomic theory of the firm to banking and differ only in the specification of banking activities.

In the production approach, banking activities are described in terms of production of services to depositors and borrowers. Traditional production factors such as labour and capital are used as inputs to produce desired outputs. Although this approach recognises the multi-product nature of banking activities, earlier studies ignored this aspect, mainly because the techniques to deal with scope and scale issues were not well developed (Denizer, Dinc, and Tarimcilar 2000). As stated earlier, this approach suffers from a basic problem of measurement of outputs. Is it the number of accounts, the number of operations on these accounts, or the value amounts? The generally accepted approach is to use value amounts because of availability of such data. Yet the number of accounts and the number of operations could also be used.

The intermediation approach describes banking activities as transforming the money borrowed from depositors into the money lent to borrowers. This transformation activity originates from the different characteristics of deposits and loans. Deposits are typically liquid and risk less, while loans, on the other hand, are regarded as illiquid and risky. In this approach, the deposits collected and funds borrowed from financial markets constitute inputs, while outputs are measured by the volume of outstanding loans and investments.

Similar to many studies on banking efficiency (e.g., Aly et al. 1990; Zaim 1995; DeYoung and Nolle 1998; Berger and Mester 1997, Resti 1997; and DeYoung and Hasan 1998), we adopted the intermediation approach. The use of intermediation approach in bank productivity presents fewer data problems than with the production approach. The literature suggests that it is the most appropriate approach for evaluating the entire banking industry as it is inclusive of interest expenses, which account for 50-66 percent of total costs of banks as confirmed by Rao (2002). Accordingly, we model commercial banks as multi-product firms, producing three outputs and using three inputs. All variables except for the input factor labour were measured in Egyptian pounds.

The input vector included: (1) labour [LABOUR], the number of full-time employees on the payroll; (2) capital [CAPITAL], the book value of premises and fixed assets; and (3) loanable funds [FUNDS], the sum of...
deposit (demand and time) and non-deposit funds. Thus, the total costs include both interest expense and operating costs and are proxied by the sum of labour, capital, and loanable funds expenditures.

The output vector included: (1) Loans and Overdrafts [LOANS]; (2) Off-balance sheet items [OFF—B/S], namely guarantees and warranties (letters of guarantee, bank acceptance, letters of credit, guaranteed pre-financing, endorsements and others), commitments, foreign exchange as well as other off-balance sheet activities; and (3) other earning assets [OTHER—EA], such as loans to special sectors, inter-bank funds sold, and investment securities (treasury bills, government bonds and others).

We followed Isik and Kabir (2002a) in our choice of inputs and outputs. Similar to their study, we included off-balance sheet activities as one of our outputs. As Berger and Mester (1997) reported, no earlier frontier study accounted for off-balance sheet items, although they were comparable to loans in terms of risk and revenue. Thus, in a sense, our study is the second non-parametric efficiency study that takes the off-balance sheet items fully and directly into account and the second frontier efficiency study that treats these items as variable output. It is critical to include such activities in the efficiency studies because they are often four or five times greater than the on-balance sheet items when measured by notional values (Saunders 1993). These items are often effective substitutes for loans that demand similar information gathering, origination, monitoring, and control costs. Thus, they might require similar revenues if they are competitive substitutes for direct loans. Moreover, the risk weights suggested by the Basle Accord imply that off-balance sheet items have approximately the same perceived credit risk as loans (Berger and Mester 1997).

In order to account for heterogeneous business operations among banks, such as additional overhead costs that may result from operating large branch networks, we normalised all production variables by the number of branch offices. This treatment is in accordance with Berger and Mester (1997) and Denizer et al. (2000). We also deflated all variables on a yearly basis according to the corresponding inflation rate. The aim of deflating the figures of inputs and outputs is to eliminate the adverse impact of inflation on real magnitudes. We used CPI as our basis for deflation. The used rates are presented in Table 3 as follows:

---

3. Non-deposit funds include borrowed funds from Interbank, Central bank, domestic and foreign-owned banks, among others, as well as funds raised by issuing securities.
4. Other off-balance sheet activities contain various types of guarantees, mostly regarding foreign trade, foreign currency and interest rate transactions in addition to loan commitments.
6. Only a prior non-frontier study, Jagtiani et al. (1995) took a number of off-balance sheet activities into account and found that cost scale and product mix economies were affected insignificantly by these activities (Berger and Mester 1997).
Data Collection and Processing

The accounting data and banks financials have been obtained from Bankscope Database. As for the number of employees and number of branches per bank, data are drawn from the financial information provided by the Kompass Egypt Financial Yearbooks (Fiani and Partners, various issues). The Bankscope database provided useable data for a maximum of 35 banks over the period of study. Other sources of data on banking system included annual reports of the Central Bank of Egypt, and Economic Reviews.

DEAP software Version 2.1 developed by Time Coelli (1996) was used for data processing. We solved input oriented models for both constant returns to scale (CCR) and variable returns to scale (BCC) for every year, from 1995 to 2003, for commercial banks in Egypt.

6. Results and Discussion

6.2. DEA Efficiency

We used data envelopment analysis to measure the technical efficiency for up to 35 commercial banks for the years 1995 to 2003. The sample under study comprised all locally incorporated banks (excluding foreign bank branches, Islamic banks, and specialised banks). We define commercial banks as all those labelled commercial and investment banks, based on the CBE classification. This is because although investment banks are registered to function as “investment” banks, in reality, they tend to perform the same tasks as those registered as “commercial” banks.

The commercial banks under study represented around 75 percent of total banking assets, deposits and loans. The 35 banks included in our study comprised four public banks, seven joint venture banks with public ownership constituting the majority of ownership (JVMP), 11 private banks (JVMPR), and 13 joint venture banks with majority of foreign ownership. We define foreign ownership to be that of non Egyptian individuals or companies/banks. Private Banks are those with majority of local private ownerships. The total number of observations is 2,637, which span the time

<table>
<thead>
<tr>
<th>Year</th>
<th>Inflation (annualised)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995</td>
<td>9.4</td>
</tr>
<tr>
<td>1996</td>
<td>7.1</td>
</tr>
<tr>
<td>1997</td>
<td>6.2</td>
</tr>
<tr>
<td>1998</td>
<td>4.7</td>
</tr>
<tr>
<td>1999</td>
<td>3.8</td>
</tr>
<tr>
<td>2000</td>
<td>2.8</td>
</tr>
<tr>
<td>2001</td>
<td>2.4</td>
</tr>
<tr>
<td>2002</td>
<td>2.4</td>
</tr>
<tr>
<td>2003</td>
<td>3.2</td>
</tr>
</tbody>
</table>


Foreign banks branches were not included as they have their own regulations and reporting measures. Similarly, specialised banks have not been included as they have defined functions that relate to specific activity such as agricultural credit, industrial development, and real estate development. Islamic and specialised banks were excluded as their functions differ from the universe of commercial banks.
horizon of 1995 to 2003 (DMUs per year x 9 yrs). The use of a large number of observations would lead to better results.

Summary statistics for the inputs and outputs we used are presented in Table 4, using both monetary variables and non-monetary variables. DEA is able to integrate, unlike multiple inputs and outputs, to make simultaneous comparisons that would otherwise not be possible. In Table 4, we report the average of each input-output variable we used to measure efficiency and the standard deviation for each year (1995-2003). The percentage change in the average compared to average of the previous year is also shown in Table 4, between 1995 and 2003. Average deposits and loans grew by 173 percent and 172 percent, respectively. Yet, as apparent from the entries loans, growth rate started to decline considerably from 1998. In 2000, deposits on the opposite growth rate were much higher than that of loans. This reflects the recessionary economic conditions that the country was undergoing in 1998, and the reluctance of banks to provide fresh loans, given the nonperforming loans at the time, especially in the public sector banks.

Other earning assets were growing at 26 percent in 1995 and 28 percent in 1996. By 1997, growth in other earning assets had started to decline. It became negative in 1998. This is probably due to the Income Tax Law that was introduced in 1998, which led to the cancellation of tax provision that allowed banks to deduct interest on income incurred in order to purchase tax free securities (government treasury bills and bonds). We note, however, that other earning assets started to grow slowly again to reach a growth rate of 84 percent in 2002. A possible interpretation to this immense growth in other earning assets is that banks were moving away from the traditional lending role, given the recessionary environment towards safer and lower risk earning activities, such as management and funding of government securities.

Average labour growth rates were very low over the whole period. In fact, labour growth witnessed a sharp decline and even became negative in 1997 and 1998, possibly mimicking the economic downturn and the privatisation of some joint venture banks. Another plausible reason is the automation in the bank business and increased consciousness of using resources productively in the more competitive business environment of the 1990s.

Table 5 displays the annual means of the efficiency measures of Egyptian commercial banks over the 1995-2003 period, along with other descriptive statistics, namely standard deviation, median, minimum, and maximum DEA result. The averages of technical efficiency (TE), pure technical efficiency (PTE) and scale efficiency (SE) over the study period based on the yearly frontier are 82 percent, 89 percent, and 92 percent, respectively. The above results suggest that average managerial, pure managerial, and scale
### Table 4. Summary of Inputs and Outputs used in DEA

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>% Change</td>
<td>20%</td>
<td>15%</td>
<td>10%</td>
<td>17%</td>
<td>11%</td>
<td>9%</td>
<td>16%</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>Std Dev.</td>
<td>835,333</td>
<td>869,719</td>
<td>948,058</td>
<td>1,000,361</td>
<td>1,203,023</td>
<td>1,438,705</td>
<td>1,604,374</td>
<td>1,928,884</td>
<td>2,168,929</td>
</tr>
<tr>
<td>OFF/B/S</td>
<td>1,819,491</td>
<td>1,958,696</td>
<td>2,118,451</td>
<td>2,438,453</td>
<td>2,502,268</td>
<td>2,608,565</td>
<td>2,883,217</td>
<td>3,345,985</td>
<td>4,117,677</td>
</tr>
<tr>
<td>% Change</td>
<td>8%</td>
<td>8%</td>
<td>15%</td>
<td>3%</td>
<td>4%</td>
<td>11%</td>
<td>16%</td>
<td>11%</td>
<td>23%</td>
</tr>
<tr>
<td>Std Dev.</td>
<td>1,007,296</td>
<td>970,619</td>
<td>1,104,875</td>
<td>1,373,759</td>
<td>1,457,926</td>
<td>1,567,270</td>
<td>1,741,415</td>
<td>2,091,553</td>
<td>2,593,637</td>
</tr>
<tr>
<td>OTHER—EA</td>
<td>603,918</td>
<td>761,514</td>
<td>971,247</td>
<td>1,046,654</td>
<td>751,584</td>
<td>806,012</td>
<td>1,097,698</td>
<td>2,014,330</td>
<td>2,687,819</td>
</tr>
<tr>
<td>% Change</td>
<td>26%</td>
<td>28%</td>
<td>8%</td>
<td>-28%</td>
<td>7%</td>
<td>36%</td>
<td>36%</td>
<td>84%</td>
<td>35%</td>
</tr>
<tr>
<td>Std Dev.</td>
<td>835,333</td>
<td>869,719</td>
<td>948,058</td>
<td>1,000,361</td>
<td>1,203,023</td>
<td>1,438,705</td>
<td>1,604,374</td>
<td>1,928,884</td>
<td>2,168,929</td>
</tr>
<tr>
<td>CAPITAL</td>
<td>28,961</td>
<td>31,404</td>
<td>36,565</td>
<td>41,733</td>
<td>51,014</td>
<td>57,183</td>
<td>63,260</td>
<td>72,981</td>
<td>79,218</td>
</tr>
<tr>
<td>% Change</td>
<td>8%</td>
<td>16%</td>
<td>14%</td>
<td>22%</td>
<td>22%</td>
<td>12%</td>
<td>11%</td>
<td>15%</td>
<td>9%</td>
</tr>
<tr>
<td>Std Dev.</td>
<td>8,100</td>
<td>7,315</td>
<td>7,803</td>
<td>7,981</td>
<td>11,503</td>
<td>12,537</td>
<td>14,016</td>
<td>15,216</td>
<td>16,916</td>
</tr>
<tr>
<td>LABOUR</td>
<td>1,698</td>
<td>1,849</td>
<td>1,807</td>
<td>1,761</td>
<td>1,786</td>
<td>1,823</td>
<td>1,866</td>
<td>2,015</td>
<td>2,138</td>
</tr>
<tr>
<td>% Change</td>
<td>9%</td>
<td>-2%</td>
<td>-3%</td>
<td>1%</td>
<td>2%</td>
<td>2%</td>
<td>2%</td>
<td>8%</td>
<td>6%</td>
</tr>
<tr>
<td>Std Dev.</td>
<td>658</td>
<td>605</td>
<td>588</td>
<td>573</td>
<td>570</td>
<td>570</td>
<td>578</td>
<td>642</td>
<td>671</td>
</tr>
<tr>
<td>FUNDS/Deposits</td>
<td>4,477,805</td>
<td>4,960,140</td>
<td>5,120,360</td>
<td>5,838,120</td>
<td>6,373,587</td>
<td>7,025,457</td>
<td>8,143,069</td>
<td>9,793,196</td>
<td>12,227,315</td>
</tr>
<tr>
<td>% Change</td>
<td>11%</td>
<td>3%</td>
<td>14%</td>
<td>9%</td>
<td>10%</td>
<td>16%</td>
<td>20%</td>
<td>25%</td>
<td>25%</td>
</tr>
<tr>
<td>Std Dev.</td>
<td>4,775,918</td>
<td>1,710,590</td>
<td>1,843,145</td>
<td>1,932,303</td>
<td>2,045,290</td>
<td>2,249,215</td>
<td>2,614,276</td>
<td>3,378,079</td>
<td>4,066,908</td>
</tr>
</tbody>
</table>

Source: Author’s Investigations.

1. All bank variables are in annual terms expressed in LE, except for labour, which is measured by the number of employees by the end of the respective year.
2. Outputs: (1) LOANS, and (2) off-balance sheet items (OFF—B/S) include guarantees and warranties (letters of guarantee, bank acceptance, letters of credit, guaranteed pre-funding, endorsements and others), commitments, foreign exchange, and interest rate transactions as well as other off-balance sheet activities; (3) other earning assets (OTHER—EA) consist of loans to special sectors, interbank funds, and investment securities (treasury bills, government bonds, and other securities). Inputs: (1) labour (LABOUR) is the quantity of labour and is measured by the number of full-time employees on the payroll; (2) capital (CAPITAL) is the book value of premises and fixed assets; (3) loanable funds (FUNDS) is the sum of deposit (demand and time) and non-deposit funds as at the end of the respective year.

* The intermediation model defines banks as financial intermediaries that convert deposits by means of labour and capital into primary and secondary assets.
inefficiencies (average dispersions of banks from the best-practice banks) are
22 percent, 12 percent and nine percent respectively8.

The results imply that Egyptian banks could have produced the same output using one fifth of the resources used. Decomposition of total technical (in)efficiency into its components signals that the main source of inefficiency in Egyptian banking is managerial inefficiency (pure technical efficiency, PTE). PTE is simply technical efficiency (TE) devoid of scale effects. In other words, the difference between technical efficiency and “pure” technical efficiency represents the cost of operating at an incorrect scale. Finding that the dominant effect on technical efficiency is due to pure technical efficiency rather than by scale efficiency was reported in several studies for other countries. Such studies include Aly et al. (1990) for US banking, Fukuyama (1993) for Japanese banks, and Zaim (1995) for Turkish banking.

It should be noted here that the figures reported for yearly frontiers should not be used as a basis for comparison across years, as DEA measures relative efficiency and not absolute efficiency9. A higher value of average

---

8. The association between efficiency (Eff) and inefficiency (Ineff) is Ineff = (1-Eff)/Eff. Thus, 71.45 percent managerial efficiency corresponds to 39.96 percent inefficiency rather than 28.55 percent (Berger and Mester 1997, and Isik and Hassan 2002).

9. DEA efficiency is measured relative to the observed best-practice frontier, and therefore, the frontier can change from year to year.
efficiency does not imply higher average performance compared to the performance with respect to a lower average efficiency.

Efficiency Results by Size

Investigation of managerial efficiency among different sub-groups of banks is critical for regulators, current owners, and potential investors. Although these groups operate under the same environment, they may exhibit variations in operational performance due to differences in their managerial skills and practices, natures of business, and government patronage.

The average efficiency results of banks in four groups formed according to size of bank as measured by total assets are presented in Table 6. Based on total assets, Egyptian banks were classified as follows: Micro banks with assets below LE 2 billion; Small Banks with total assets ranging between LE 2-10 billion; Medium Banks with total assets ranging between LE 10-50 billion; and Large Banks with Assets exceeding LE 50 billion.

An analysis of efficiency by size reveals that in each year, the medium sized banks appear to be the most technically efficient (Figure 2). Large banks efficiency comes next. In general, it seems that the smaller banks appear to be less efficient than others, with micro banks being the poorest performers. Average technical inefficiency of micro, small, medium, and large banks over the 1995-2003 period was 39 percent, 23 percent, four percent, and 15 percent respectively. Decomposing technical efficiency to its components – pure technical efficiency and scale efficiency – taking size into consideration, reveals that medium banks’ inefficiency is mainly due to pure technical efficiency, whereas inefficiency in large banks is attributed to scale (disecono-
Returns to Scale in Egyptian Banking Sector

Scale inefficiency appears to affect the overall managerial inefficiency of Egyptian banks, particularly the small ones. It is worthwhile to examine their returns to scale. The law of diminishing returns indicates what happens to output when a bank changes only one input, say labour or capital, and holds all other inputs constant, whereas, returns to scale (RTS) tell us what happens to a bank’s output if it changes all inputs. Thus, we define RTS as the increases in output that result from increasing all inputs by the same percentage. Obviously, there are three possible cases:

1) Increasing returns to scale (IRS) occur when one percent increase in inputs produces more than one percent increase in outputs.
2) Constant returns to scale (CRS) occur when one percent increase in inputs results in exactly one percent increase in outputs.
3) Decreasing returns to scale (DRS) occur when one percent increase in inputs produces less than one percent increase in outputs.

Table 7. Returns to Scale Results (%) by Bank Size

<table>
<thead>
<tr>
<th></th>
<th>All banks</th>
<th>Large</th>
<th>Medium</th>
<th>Small</th>
<th>Micro</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRS</td>
<td>81%</td>
<td>31%</td>
<td>84%</td>
<td>75%</td>
<td>89%</td>
</tr>
<tr>
<td>CRS</td>
<td>9%</td>
<td>56%</td>
<td>16%</td>
<td>8%</td>
<td>6%</td>
</tr>
<tr>
<td>DRS</td>
<td>10%</td>
<td>13%</td>
<td>0%</td>
<td>18%</td>
<td>5%</td>
</tr>
<tr>
<td>TOTAL</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Source: Authors’ investigations.

Figure 2. Technical Efficiency Results Classified by Bank Size (1995-2003)

Source: Authors’ investigations.

In both micro and small banks, scale inefficiency dominates pure technical inefficiency, reflecting that considerable gains could be achieved if these banks adjust their size of operation.
3) Decreasing returns to scale (DRS) happen when one percent increase in inputs leads to less than one percent increase in outputs.

Table 7 displays the returns to scale by classified banks sizes during the 1995-2003 period. It shows that most Egyptian banks operate at incorrect scale. The majority experience increasing returns to scale (IRS) in their operations. On average, 81 percent showed IRS, 10 percent DRS, and only nine percent CRS in their production technologies between 1995 and 2003. The scale inefficiency due to IRS might be attributed to small banks that desire to grow to reach the “right” scale, but fail to do so. The scale inefficiency due to DRS might be related to the established large banks, which transgressed the “right” scale perhaps to meet excess market demand for financial services and products induced by a growing Egyptian economy.

Table 7 shows that IRS were prevailing in medium, small, and micro banks. IRS are relatively speaking low in large banks. Large banks experienced CRS. Medium size banks did not report any instance of DRS.

6.3. Productivity Change

The average annual values, Total Factor Productivity Index (TFPCI), Technological Change (TCI), Technical Efficiency Change (TECI), Pure Technical Efficiency Change (PTECI) and Scale Efficiency Change (SECI) for the years 1996 to 2003 are reported in Table 8. The Malmquist Index cannot be constructed without a reference technology, which could be the technology of any year in a multi-period setting. We report the results relative to the technology fixed at the initial year 1995.

The results in the analysis with all the banks reported in Panel C, Table 8, indicate that there was deterioration in productivity, at four percent per year, during the sample period (1995 and 2003). Egyptian commercial banks experienced productivity growth in 1997, then in the three years that followed, productivity declined considerably, suggesting deterioration in bank performance from 1998 to 2000. One possibility for this deterioration could be the economic slow down that prevailed during that time. In 2001, the commercial banking sector defied the negative trend in growth in previous years, but only marginally. In 2002, productivity of banks increased by 16 percent. It was followed by a decline in productivity of nine percent in year 2003.

Over the sample period, the average annual rate of technical efficiency change is -1 percent, whereas the rate of technological change is -3 percent. We therefore infer that the average deterioration in productivity, though modest, is due to the deterioration in technological change rather than technical efficiency.

Pure technical efficiency, which measures performance only due to managerial activity, remained constant over the study period. Looking at the changes through the years, we note that managerial efficiency’s largest decline
of -6 percent was in 2003, and the largest increase of 20 percent in 2001. Scale efficiency of the banks, on the other hand, deteriorated on average by one percent per year. The observed rate of deterioration in technical efficiency may therefore be considered to be due to scale efficiency rather than managerial (PTE) efficiency.

7. Conclusion

Employing Data Envelopment Analysis type Malmquist Index, we examined total factor productivity change in Egyptian commercial banks during

<table>
<thead>
<tr>
<th>Year</th>
<th>Technical Efficiency Change (1)=(3*4)</th>
<th>Technological Change (2)</th>
<th>Pure Technical Efficiency Change (3)</th>
<th>Scale Efficiency Change (4)</th>
<th>Total Factor Productivity Change = (1*2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Private and JVs Banks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td>0.94</td>
<td>0.96</td>
<td>0.98</td>
<td>0.96</td>
<td>0.90</td>
</tr>
<tr>
<td>1997</td>
<td>0.92</td>
<td>1.13</td>
<td>0.96</td>
<td>0.97</td>
<td>1.04</td>
</tr>
<tr>
<td>1998</td>
<td>1.21</td>
<td>0.74</td>
<td>1.09</td>
<td>1.11</td>
<td>0.89</td>
</tr>
<tr>
<td>1999</td>
<td>0.95</td>
<td>0.90</td>
<td>0.98</td>
<td>0.97</td>
<td>0.86</td>
</tr>
<tr>
<td>2000</td>
<td>1.01</td>
<td>0.96</td>
<td>1.01</td>
<td>1.00</td>
<td>0.97</td>
</tr>
<tr>
<td>2001</td>
<td>1.01</td>
<td>0.97</td>
<td>1.02</td>
<td>0.99</td>
<td>0.98</td>
</tr>
<tr>
<td>2002</td>
<td>1.04</td>
<td>1.11</td>
<td>1.00</td>
<td>1.03</td>
<td>1.15</td>
</tr>
<tr>
<td>2003</td>
<td>0.85</td>
<td>1.04</td>
<td>0.93</td>
<td>0.91</td>
<td>0.88</td>
</tr>
<tr>
<td>Geometric mean</td>
<td>0.99</td>
<td>0.97</td>
<td>0.99</td>
<td>0.99</td>
<td>0.95</td>
</tr>
<tr>
<td>Panel B: Public Banks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td>1.07</td>
<td>0.88</td>
<td>1.00</td>
<td>1.07</td>
<td>0.94</td>
</tr>
<tr>
<td>1997</td>
<td>0.99</td>
<td>1.43</td>
<td>1.00</td>
<td>0.99</td>
<td>1.42</td>
</tr>
<tr>
<td>1998</td>
<td>1.14</td>
<td>0.65</td>
<td>1.00</td>
<td>1.14</td>
<td>0.74</td>
</tr>
<tr>
<td>1999</td>
<td>1.00</td>
<td>0.80</td>
<td>1.00</td>
<td>1.00</td>
<td>0.80</td>
</tr>
<tr>
<td>2000</td>
<td>1.00</td>
<td>1.01</td>
<td>1.00</td>
<td>1.00</td>
<td>1.01</td>
</tr>
<tr>
<td>2001</td>
<td>1.00</td>
<td>1.13</td>
<td>1.00</td>
<td>1.00</td>
<td>1.13</td>
</tr>
<tr>
<td>2002</td>
<td>0.96</td>
<td>1.26</td>
<td>1.00</td>
<td>0.96</td>
<td>1.21</td>
</tr>
<tr>
<td>2003</td>
<td>0.99</td>
<td>1.14</td>
<td>1.00</td>
<td>0.99</td>
<td>1.14</td>
</tr>
<tr>
<td>Geometric mean</td>
<td>1.02</td>
<td>1.01</td>
<td>1.00</td>
<td>1.02</td>
<td>1.03</td>
</tr>
<tr>
<td>Panel C: All Banks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td>0.96</td>
<td>0.95</td>
<td>0.98</td>
<td>0.98</td>
<td>0.91</td>
</tr>
<tr>
<td>1997</td>
<td>0.93</td>
<td>1.16</td>
<td>0.96</td>
<td>0.97</td>
<td>1.08</td>
</tr>
<tr>
<td>1998</td>
<td>1.20</td>
<td>0.73</td>
<td>1.08</td>
<td>1.11</td>
<td>0.87</td>
</tr>
<tr>
<td>1999</td>
<td>0.96</td>
<td>0.89</td>
<td>0.98</td>
<td>0.98</td>
<td>0.85</td>
</tr>
<tr>
<td>2000</td>
<td>1.01</td>
<td>0.96</td>
<td>1.01</td>
<td>1.00</td>
<td>0.97</td>
</tr>
<tr>
<td>2001</td>
<td>1.01</td>
<td>0.99</td>
<td>1.02</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>2002</td>
<td>1.03</td>
<td>1.13</td>
<td>1.00</td>
<td>1.02</td>
<td>1.16</td>
</tr>
<tr>
<td>2003</td>
<td>0.86</td>
<td>1.06</td>
<td>0.94</td>
<td>0.92</td>
<td>0.91</td>
</tr>
<tr>
<td>Geometric mean</td>
<td>0.99</td>
<td>0.97</td>
<td>1.00</td>
<td>0.99</td>
<td>0.96</td>
</tr>
</tbody>
</table>

Source: Authors’ investigations
the 1995-2003 period, along with its mutually exclusive and exhaustive components, these being change in efficiency (catching up or falling behind) and change in technology (innovation or shock). We also elaborated efficiency change in Egyptian banking further by studying its sources, listed as pure technical efficiency change (improvement in management) and scale efficiency change (improvement towards optimal size). Additionally, we studied the returns to scale of Egyptian banks over time by ownership and by size.

Our results indicate that most of the Egyptian banks are facing substantial scale problems, especially due to increasing returns to scale. One implication is that for most Egyptian banks, increases in outputs would result in less than proportional increases in inputs (and costs). Thus, banks with the IRS could achieve significant cost savings and efficiency gains by increasing the scale of their operations. In other words, substantial gains could be obtained from altering scale via internal growth or consolidation in the sector. Indeed, in a perfectly competitive and contestable market, scale inefficient banks should be absorbed by efficient ones to exploit cost advantages. Thus, those banks that experience IRS should either eliminate their scale inefficiency or be ready to become a prime target for acquiring banks that can “create value” from underperforming banks by streamlining their operations and eliminating their redundancies and inefficiencies (Evanoff and Israilevich 1991 and; Cummins et al. 1999).

Further investigation of returns to scale shows that the majority of the banks are experiencing increasing returns to scale. While one way to exploit such economies of scale opportunities is through internal growth, the other is through external growth, i.e. through mergers and acquisition (M&A). Because internal growth opportunities are limited and more time demanding given the sluggish economy, external growth through consolidation in the industry should be encouraged. The Central Bank has already begun to stimulate mergers and acquisition among banks to foster efficiency.

The persistence of scale inefficiency problems in the Egyptian banking sector over a prolonged period of time may imply that the financial markets in Egypt are not efficient enough to eliminate such non-optimal behaviour. There may be some regulatory or market specific characteristics that somehow inhibit the utilisation of scale economies by curbing internal or external growth of the Egyptian banks. In fact, close inspection of the financial system in Egypt reveals some clues for the persistence of scale inefficiencies in this less than perfect market. It appears that external growth of banks is constrained by corporate control and governance issues surrounding the traditional Egyptian financial system.

The Central Bank of Egypt (CBE) has been persistently trying to consolidate the banking industry. The banking Law 88 of 2003 has resulted in a massive ongoing wave of mergers and acquisitions. The CBE aims to decrease the number of commercial banks substantially to about 27. It would be of great importance to assess the success of such efforts in improving the ban-
king sector efficiency, and particularly its impact on returns to scale in Egyptian banking. This is an area that is worth exploring in the near future.

In addition, it is evident that inexistence of effective competition allows inefficient banks to continue with slackened efficiency and still remain in business. The implication of this is that greater policies that would encourage competition are crucial for the improvement of Egyptian banks efficiency.

Further policy suggestions for improved efficiency may include the need for the government to implement financial reform packages that foster competition in the banking market, and that the industry devises incentive schemes to improve managerial efficiency. It is argued that contestable markets and freedom of entry promote efficiency, encourage innovation, and give highly favourable welfare outcomes. For a market to be contestable, there should be no significant entry barriers. Large economies of scale and high sunk costs in addition to other entry costs are examples of such barriers. But in the case of banking, government regulations through permits and licenses are far more important than these other barriers. The Egyptian banking system has suffered from the use of restrictive regulations that prevented new entry and made the incumbent banks far from being contestable.

Furthermore, our study has also some important research implications. Studies from different regulatory environments and market structures may help us conceive the impact of these differences on bank performance. For instance, the Egyptian market is highly concentrated as compared to those of advanced economies. The three-bank concentration ratio is 65 percent for Egypt, whereas it is 19 percent for the US, 22 percent for Japan, 41 percent for France, 45 percent for Germany, and 56 percent for the U.K. The Egyptian concentration ratio is closer to those of other emerging markets. The ratio is 59 percent for Tunisia, 64 percent for Bangladesh, 69 percent for Peru, 74 percent for Pakistan, and 87 percent for Uruguay (Demirguc-Kunt and Levine 1999).

In some cases, the recent finance literature reports that there is a negative association between market concentration and bank performance (Berger and Mester 1997, and Berger et al. 1998). The quiet life theory suggests that the lower the intensity of environmental pressures, the lower is the effort expended by managers to derive the maximum output from a given amount of inputs (Berger and Hannan 1998).

Thus, it is possible that banks of concentrated markets become less motivated to operate efficiently. They do not face strong competition from new banks and non-bank financial institutions. Moreover, the lack of developed money and capital markets also provides comfort for banks of emerging countries, since “disintermediation” from depositors and borrowers has not threatened their businesses to the extent that it did in developed markets. Therefore, other things being held constant, comparisons of banking efficiency in Egypt vis-à-vis in less concentrated markets may let us understand the dynamics between market concentration and efficiency.
In brief, our recommendations for improved efficiency of the Egyptian banking system are: Enhancing and monitoring corporate governance in all banks; allowing for increased competition by allowing new entrants to the market, thus leading to market contestability and increased private ownership and competition in the banking sector; and undertaking of serious steps towards the restructuring of public sector banks.

The restructuring should not be limited to the inclusion of risk management and proper credit risk measures. It should encompass the assessment of functioning of the wide branch networks. In this respect, poor performing branches may be detached from the public bank and offered for private or foreign banks to transform them. We also recommend accelerated action to divest government stakes in joint ventures; increasing safety-net without resorting to the lender of last resort by establishing a deposit insurance fund. Reforms could include improved efficiency of the judicial enforcement of laws, facilitation of commercial judiciary process, greater enforceability of contracts, and creation of special economic courts. Lastly, operational efficiency can be improved through greater investment in technology and skill enhancements.

References


Central Bank of Egypt, Annual Reports. 2003/04.


