Commercial Banks Efficiency in Tanzania

By Jehovaness Aikaeli
Bank of Tanzania
Monetary and Financial Affairs Department
Email: jaikaeli@yahoo.co.uk

A Paper Presented in a CSAE Conference on "Economic Development in Africa"
Held at St. Catherine's College, Oxford, 16th – 18th March 2008

Abstract

Efficient banking system reflects a sound intermediation process and hence the banks' due contribution to economic growth. If commercial banks are functioning efficiently, monetary policies are likely to be effective. This study is motivated by the fact that, though banking sector is the largest part of the financial system in Tanzania, little is known about its efficiency status. Secondary time series data are used in empirical analysis of banks' efficiency. Non-parametric Data Envelopment Analysis (DEA) model is utilised in estimation of technical and scale efficiency, while x-inefficiency is estimated using multi-product translog cost function. Though banks were not full efficient in all respects, they performed fairly well during the 1998-2004 period. Nevertheless, the major conclusions show that banks in Tanzania still have reasons to improve their performance.

1. Introduction

Banks contribute to economic growth through their financial intermediation role. Banking sector in Tanzania has experienced fundamental changes over the last decade following banks and other financial institutions reforms starting from the early 1990s. However, what is still concealed is the extent to which banks are efficient in Tanzania. Banking industry in Tanzania is open to entry and therefore it is highly contestable[1]. Commercial banks in the country can be subdivided into three major categories: large domestic banks; subsidiaries of the major international banks; and small banks. About 50 percent of the total banks' assets are held in the large domestic banks while subsidiaries of the major international banks account for 40 percent and the small banks hold the remaining 10 percent. Risk vulnerability of the Tanzanian banks which was examined through stress test found that banks are generally resilient to shocks (IMF-World Bank, 2003). This IMF-World Bank study also singles out the high degree of liquidity as one of the outstanding features of banks in Tanzania.

Efficiency in banking has been defined and studied in different dimensions including: (i) *scale efficiency,* which refers to relationship between the level of output and the average cost; (ii) *Scope efficiency,* which refers to relationship between average cost and production of diversified output varieties; and (iii) *Operational efficiency,* a wide concept sometimes referred to as x-efficiency, which measures deviation from the cost efficient frontier that represents the maximum attainable output for the given level of inputs. With reference to various definitions, inefficiency is therefore a multifaceted concept with several meanings depending on the perspective in which it is used (Leibenstein, 1966). Scale and scope economies for example, are achieved from the firms' output expansion resulting in an increase in the industry's output. And that reduces costs of production thus leading to the strong technological external economy. Hirshleifer and Glazer (1993) argue that scope economies occur where it is cheaper to produce varieties in a plant than in separate plants, and this is the concept from which banking consolidation stems.

---

[1] Foreign banks find greater entry where the expected rate of economic growth is higher and the banking system is on average less efficient. Although additional research on the activities of foreign banks in developing countries is clearly warranted, initial indications are that, in developing countries, foreign entrants face relatively less effective domestic competition. Developing host countries, therefore, might offer substantial profit opportunities in the provision of financial services. In that sense, foreign penetration in banking might precede, and perhaps help bring about, entry of non-financial sector firms, Clarke et. al. (2001).

The central objective of this study is to determine efficiency status of commercial banks in Tanzania and to investigate the reasons why x-inefficiency might exist. The next section 2 summarises some of important theoretical and empirical issues arising from efficiency studies. Section 3 describes methods of efficiency analysis applied in this study, while section 4 details empirical analysis. The last section 5 outlines conclusions and some policy implications.

2. Theoretical and Empirical Backgrounds

The positive underlying principle that derives efficiency analysis in economics stems from the urge to create and enhance tangible value, while the normative raison d'être for efficiency analysis is founded on the challenge to obtain useful policy information. Because value itself may be subjective, Vilfredo Pareto set a condition for efficiency that: if there is a change which makes at least one individual better off without making any one else worse off, that change is efficient (Debreu, 1959; Varian, 1992; Schenk, 2004). The relationship between cost function and production function, which underlines efficiency assessment, was first established by Shephard (1953, 1970) with assumption of theoretically known efficiency. Quantitative methods for measuring total economic efficiency (with assumption of unknown theoretical efficiency) were pioneered by Farrell (1957). In economic theory there are algebraic and geometric characterisations of production plans that can unambiguously be regarded as non-wasteful (efficient). According to Mas-Colell et al. (1995), a production vector $y \in Y$ is efficient if there is no $y' \in Y$ such that $y' \geq y$ and $y' \neq y$. This concept means a production vector y is efficient if there is no other feasible production vector y' that generates as much output as y using no additional inputs. This philosophy is the basis of the illustrative Production Possibility Frontier (PPF), from which the methods of analysis used in this study originate.

Looking specifically at banks in the face of financial liberalisation in transition economies, the sector experiences an unprecedented consolidation through mergers and acquisitions. Maggi and Rossie (2003) suggest that the only way to lower costs in the banking industry is to improve x-efficiency instead of paying close attention to cross border mergers and acquisitions. According to Jemrić and VujČić (2002), there are questions which continue to dominate financial sector discussions in developing countries: Will small banks manage to exist in the era of globalisation and banking market consolidation? How useful it is to allow new banks? They measured relative efficiency of banks in Croatia for the 1995–2000 period using DEA. They found that foreign-owned banks were the most efficient; new banks were more efficient than old banks; small banks were globally more efficient, but large banks were found to be locally more efficient.

A big number of U.S. empirical studies of banks efficiency have used panel data analysis. These studies overall (Berger et al., 1993; Berger and Humphrey, 1997; Mitchell and Onvural, 1996) conclude that the U.S. banks average cost curve is relatively flat when compared to European banks. Most of empirical work on European banks however, has focused on cost functions using data from single bank or country. They find a U-shaped average cost curve, and to some extent, scope economies exist (Parisio, 1992; Berger et al., 1993; Drake and Simper, 2002). It is noticeable from these results that the choice of specific approach to efficiency study as well as the definitions of inputs and outputs in multi-product financial firms model, will most likely affect the estimates.

A case study of commercial banks efficiency in Namibia by Ikhide (2000) used operating ratios and parametric approach to measure efficiency for the 1993–1998 period. Ikhide found substantial existence of economies of scale in Namibian banks but they were not operating at the minimum point of average cost curve. Some other studies of banks efficiency have included measures of non-performing loans in the cost or production function. Berger and De Young (1997), studied problem loans and cost efficiency in U.S. commercial banks and found that low cost efficiency occurred before soaring non-performing loans.

In the literature there is an evidence of relationship between a sound intermediation process and efficiency in the banking system[2] (Horward and Haynes, 2001; Vittas, 1991; Kenny and Moss, 1998). Lindley et al. (2000), explain the circumstance of low banks' capacity in the context of huge deposits inflow (excess liquidity) that overwhelms the ability of banks to produce income-earning assets. Under certain conditions as explains Baltensperger (1972), reserves adjustment is closely related with the optimal bank production position (efficiency). With reference to scale efficiency, and the presumption of negligible low reserve adjustment cost, he argues that a large bank will be more often profitable to adjust its reserves towards

[2] Because from intermediation process banks generate revenue, and that revenue is made from earning assets, it makes sense to take earning assets as a yard stick of intermediation efficiency (Agu (2004).

the optimal level, and it will, on average, stay relatively closer to its optimal position than a small bank[3].

3. Methodology

Models of bank efficiency applied for empirical analysis are presented in this section. The study looks into different contexts of bank inefficiency; hence both parametric and non-parametric methods are employed in the analysis depending upon usefulness of the respective techniques on specific types of inefficiency being analysed.

3.1 Data Envelopment Analysis Model of Bank Efficiency

To measure scale and technical efficiency, we employ a non-parametric DEA technique. DEA appeared formally for the first time in the literature through the work of Charnes, Cooper and Rhodes (1978), and hence the acronym CCR version of DEA. Banker, Charnes, and Cooper (1984) extended the earlier version of CCR, which was taking care of analysis of only constant returns to scale case, to a variable returns to scale analysis, hence the acronym BCC version of DEA. Intermediation approach is chosen as a method of analysis applied in this study. The applied DEA methodology views banks as institutions that collect and allocate funds into loans and other assets. Because bank managers have more control over outputs than inputs, output-oriented DEA is adopted. The CCR mathematical ratio form of DEA, which is afterwards converted into linear programme for a useful dual formulation, is written as

maximise $$h_o = \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{i0}}$$

subject to $$\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i xij} \leq 1;\ j = 1, 2, \ldots\ldots, n \tag{1}$$

$$\frac{u_r}{\sum_{i=1}^{m} v_i x_{i0}} > \varepsilon;\ r = 1, \ldots\ldots\ldots, s \quad \text{and} \quad \frac{v_i}{\sum_{i=1}^{m} v_i x_{i0}} > \varepsilon;\ i = 1, \ldots\ldots\ldots, m, \quad \varepsilon > 0.$$

The model measures relative performance of decision making units (DMUs). These can be individual units (banks in this case) or a group of units (eg. foreign owned banks, large banks, etc). There are n DMUs which are j=1,2,….,n. In the model, $y_{rj} > 0$ and $x_{ij} > 0$ represent

[3] See Myers and Rajan (1995), for the paradox of excess liquidity.

observed amounts of r^{th} output and i^{th} input of the j^{th} DMU. DMU_0 efficiency score is $0 \leq h_0 \leq 1$ as regards the constraints. $\varepsilon > 0$ is a non-archimedean constant that is smaller than any positive valued real number[4] for h_o. u_r, and v_i represent visual multipliers obtained by solving the maximisation problem. In equation (1) the numerator represents a set of desired outputs and the denominator represents a set of inputs. Visual output, $Y_0 = \sum u_r^* y_{ro}$, is summed over $r = 1, 2, \ldots, S$. Visual input, $X_0 = \sum v_i^* y_{io}$, is summed over $i = 1, 2, \ldots, m$. The obtained $h_o^* = Y_0 / X_0$, is efficiency score which satisfies $0 \leq h_o^* \leq 1$. When $h_o^* = 1$ represents full (or 100 percent) efficiency and $h_o^* < 1$ represents some relative inefficiency[5].

For computational tractability, CCR DEA model is presented as a dual problem of maximisation linear programming,

$$\text{minimise} \qquad \Phi_0 - \varepsilon\left[\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+\right] \tag{2}$$

$$\text{subject to} \qquad 0 = \Phi_0 x_{io} - \sum_{j=1}^{n} x_{ij}\lambda_j - s_i^-, \text{ and } y_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+ ,$$

where Φ_0 is efficiency score of a particular decision making unit (DMU_0) in terms of dual problem solution. λ_j is a vector of constants (weighted multipliers of dual variables) while s_i^- is a vector of non-negative slack associated with input inequalities and s_r^+ is a vector of non-negative slack associated with output inequalities. Because equation (2) has finite optimal solution, from duality theory of the linear programming we have

$$h_0^* = \Phi_0^* - \varepsilon\left(\sum_{i=1}^{m} s_i^{-*} + \sum_{r=1}^{s} s_r^{+*}\right) = \sum_{r=1}^{s} u_r^* y_{ro} , \tag{3}$$

which is used to estimate efficiency scores.

The most important difference between CCR and BCC model is the ability of BCC model to analyse efficiency in variable returns to scale cases. The BCC equivalent version to (1) is

[4] In practice ε is handled by the software.

[5] Note that the asterisk denotes an optimal h_o and the vectors of u_r and v_i which maximise it.

$$\text{minimise} \quad \Phi_0 - \varepsilon\left[\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+\right]$$

$$\text{subject to} \quad 0 = \Phi_0 x_{io} - \sum_{j=1}^{n} x_{ij}\lambda_j - s_i^- , \qquad (4)$$

$$y_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+ , \quad \text{and} \quad 1 = \sum_{j=1}^{n} \lambda_j .$$

$0 \leq \lambda_j, s_i^-, s_r^+$ for *i* = 1, 2,……., *m*; *r* = 1, 2,………*S*; *j* = 1, 2,……, *n*. We notice now that λ_js are restricted to sum up to one which eliminates the constraint in CCR that DMU must be scale efficient for it to become technical efficient. Models (2) and (4) are estimated for DEA efficiency scores which range between 0 and 1. DEA inputs in this case are labour, capital, and deposits. Outputs include loans, advances and overdrafts, as well as investment in securities.

If inefficiency exists, the performance of any given bank over time will depend on both its position relative to the corresponding frontier (technical efficiency) and the position of the frontier itself (technical change). These enable us to distinguish between improvements emanating from the bank's catch-up to the frontier and that resulting from the frontier shifting up over time. For this purpose, the output oriented Mulmquist index is used to assess sources of the factor productivity change in banks. The index (5) decomposes total factor productivity change into efficiency change and technological change.

$$M_O^{t+1}(y_t, x_t, y_{t+1}, x_{t+1}) = \left[\frac{D_O^t(y_{t+1}, x_{t+1})}{D_O^t(y_t, x_t)} \times \frac{D_O^{t+1}(y_{t+1}, x_{t+1})}{D_O^{1+t}(y_t, x_t)}\right]^{1/2} . \qquad (5)$$

Where *M* denotes Mulmquist productivity index of the most recent production point (x_{t+1}, y_{t+1}), using period (*t*+1) technology relative to the earlier production point (x_t , y_t), using period *t* technology. Subscript *o* indicates output orientation, *D* is output distance function, while *y* and *x* are outputs and inputs, respectively.

3.2 Stochastic Frontier Model

Because DEA model of bank efficiency adopted in this study is output oriented, it does not suffice analysis of x-inefficiency and so stochastic frontier model of Aiger, Lovell, and Schmidt (1977) is employed to measure x-inefficiency. X-inefficiency is a concept that was first introduced by Leibenstein (1966), which implies deviations from the efficient production frontier attributed to managerial inability to control costs and to maximise output. X-inefficiency is an operational concern and it outweighs scale and scope inefficiencies in terms of proportion of cost it accounts to banks, Berger et al. (1993).

A standard multi-product translog cost function is specified and deviations from the cost frontier are estimated based on it. Let total cost for the *n*th banking firm be TC_n, with the measure of its outputs Q_i and input prices P_j. The two-component error term cost function for the firm is

$$\ln TC_n = f(\ln Q_i, \ln P_j) + \varepsilon_n, \tag{6}$$

where $\varepsilon_n = \mu_n + \delta_n$, of which μ_n is uncontrollable component of error term (ε_n), while δ_n is its controllable component, which accounts for inefficiency. Because of respective continuous management styles in Tanzanian banks, the probability of a bank to be inefficient at one period of time is almost the same as the probability for it to be inefficient at any other subsequent period, therefore exponential distribution assumption fits best the analysis at hand. Exponential conditional distribution of δ_n given ε_n, is written as $N(-\sigma_\mu A, \sigma_\mu^2)$, where $A = \varepsilon_n / \sigma_\mu + \sigma_\mu / \sigma_\delta$. X-inefficiency of bank *n* denoted C_n can be specified as the expected value of δ_n conditional on ε_n.

$$C_n = E(\delta_n / \varepsilon_n) = \sigma_\mu \left[\frac{\phi(A)}{1 - \Phi(A)} - A \right], \tag{7}$$

ϕ represents a standard density function while Φ is a cumulative density function in this formulation. In the stochastic frontier model, the ratio of standard deviations of μ_n and δ_n is defined as $\lambda = \sigma_\delta / \sigma_\mu$, while $\sigma^2 = \sigma_\delta^2 + \sigma_\mu^2$. Estimates of x-inefficiency (C_n) are obtained by evaluating equation (7) at estimates of σ_δ^2 and σ_μ^2 .

To estimate parameters for the prediction of x-inefficiency in (7), multi-product translog cost function (6) is now expressed as

$$\begin{aligned} \ln TC = {} & \alpha_0 + \sum_i \alpha_i \ln Q_i + \sum_j \beta_j \ln P_j + \tfrac{1}{2} \sum_i \sum_k \gamma_{ik} \ln Q_i \ln Q_k + \tfrac{1}{2} \sum_j \sum_h \psi_{jh} \ln P_j \ln P_h \\ & + \sum_i \sum_j \varpi_{ij} \ln Q_i \ln P_j + \varepsilon. \end{aligned} \tag{8}$$

In this study the variables included in equation (8) comprise total cost (all interest and non-interest expenses, as well as tax provisions). The two major outputs are loans and investments. Input prices include: a unit cost of capital, which is measured by the ratio of rental and equipment expenses to total fixed plant and equipment; a unit cost of labour, measured as the ratio of wages and salaries to the number of employees; and a unit cost of funds, which is defined as the ratio of interest expenses to total deposits including borrowed funds.

3.3 Tobit Model for Estimation of Factors which Derive X-Inefficiency

To ascertain the factors which derive x-inefficiency in banks, the Tobit model is applied. This model has strength of estimating equations whose dependent variable values are restricted within some range. To get inefficiency index of an observation, we undertake exponential transformation of the difference between 1 (full or 100 percent efficiency) and the estimated efficiency score, such that, efficient observations are assigned 0 inefficiency index value. In this regard, some observations have positive inefficiency indices while others, the efficient ones, have zero index values. This phenomenon can be well handled by the two-limit Tobit model because estimates of the simple linear models such as ordinary least squares would be biased.

The original Tobit regression model which is referred to as censored regression model (with reference to Tobin, 1958, who first proposed the model) is specified generally in terms of the indexed function[6] as

$$
\begin{aligned}
&y_i^* = \boldsymbol{x}'_i\boldsymbol{\delta} + \varepsilon_i \\
&yi = 0, \text{ if } y_i^* \leq 0 \text{ and} \qquad (9)\\
&yi = y_i^*, \text{ if } y_i^* > 0,
\end{aligned}
$$

where y_i is a new random variable transformed from the original one, y^*. $\boldsymbol{x}_i$ is a column vector of independent variables which is a transpose of $1xK$ row of $\boldsymbol{x}$, and $\boldsymbol{\delta}$ is a vector of parameters. A column vector of disturbances is represented by ε_i. Nevertheless, for the data with lower and upper truncations, this model needs some adjustment in line with Maddala

[6] See Green 2003 for further exposition.

(1983). Below is the two-limit specification of the doubly-truncated Tobit model, and it is the version applied in estimations of the sources of x-inefficiency in this study.

$$y_i = \boldsymbol{x'}_i \boldsymbol{\delta} + \varepsilon_i$$

$$y_i^* = L_{1i} \text{ if } y_i^* \leq L_{1i} \qquad (10)$$

$$= y_i \text{ if } L_{1i} < y_i < L_{2i}$$

$$= L_{2i} \text{ if } y_i \leq L_{2i},$$

where y is a latent variable while y^* is observed dependent variable. L_{1i} is a lower limit and L_{2i} is an upper limit.

The model is therefore specified with x-inefficiency index as a function of regressors hypothesised as determinants of x-inefficiency in Tanzanian banks,

$$Ineff = f(K, L, AQ, S, EL)^{7}, \qquad (11)$$

where *Ineff* denotes x-inefficiency index estimated from the multi-product translog cost function. *K* is capital adequacy measure, of which we use the proportionate spending on the capital goods relative to other non-tax expenses as a proxy. *L* is labour compensation or the incentive to work proxy. This variable is represented by the total labour compensation relative to other non-tax expenses in a bank. *AQ* is a proxy for assets quality, captured in this case by the ratio of non-performing loans to total loans, while *S* is a bank size variable measured by total assets in this case. *EL* is excess liquidity variable constructed by taking the total bank liquid assets less the amount sufficient to finance its statutory required reserves, deposit outflows and short-term maturing obligations. X-inefficiency function is expressed mathematically as a two-limit Tobit estimation model,

$$(Ineff)_i = \beta_0 + \beta_1 K_i + \beta_2 L_i + \beta_3 AQ_i + \beta_4 S_i + \beta_5 EL_i + \varepsilon_i \quad \text{if LHS} > 0$$

$$(Ineff)_i = 0, \text{ otherwise.} \qquad (12)$$

[7] Inclusion of a specific proxy variable for the management quality would be useful, but it is not included due to lack of consistent time series data on such items like education and experience of the respective bank managers, management training and development activities, etcetera.

This means x-inefficiency index is estimated for all inefficient observations; otherwise, observations that are efficient have indices of zero inefficiency.

4. Empirical Analysis of Bank Efficiency

4.1 The Data

This study uses secondary time series data sourced from the Bank of Tanzania annual and quarterly reports. Monthly data for the 1998-2004 period are used. Nevertheless, the data from those reports could not suffice all required statistics for analysis. Therefore unpublished commercial banks' monthly returns information was also requested from the Bank of Tanzania to supplement the available data.

4.2 Technical and Scale Efficiency

4.2.1 Overall Commercial Banks Efficiency

During the study period, banks in Tanzania were generally operating at the decreasing part of their average cost curve. Nevertheless, the findings show that from early 2004 commercial banks in the country started operating on the rising part of their average cost curve. This was consequent to decreasing returns to scale production precipitated mainly by a stiff competition in the banking sector (Table 1).

Table 1: Annual Average DEA Efficiency Scores - All Banks

	1998	1999	2000	2001	2002	2003	2004
CRS	0.930	0.952	0.947	0.962	0.965	0.989	0.986
VRS	0.959	0.963	0.967	0.969	0.966	0.990	0.996
Scale	0.970	0.988	0.979	0.993	0.998	0.999	0.990

CRS = constant returns to scale, VRS = variable returns to scale

In 1998 commercial banks were 97.0 percent scale efficient and they operated under increasing returns to scale environment. That was a minimum overall average efficiency score recorded during the period and there were improvements to reach an exciting average score of 99.9 percent in 2003 showing that there was convergence to the full efficiency line over time.

Under constant returns to scale, a restrictive version of DEA model, the overall average technical efficiency was 96.1 percent while that of variable returns to scale model was 97.3 percent, implying inefficiency divergences of 3.9 percent and 2.7 percent, respectively. These rates tell us the extent to which banks could reduce inputs and yet produce the same amount of outputs in both cases. It is observable from the summary of input slacks that technical inefficiency arose more from inefficient use of labour, fixed assets and equipment rather than underutilisation of deposit input in the intermediation process.

4.2.2 Efficiency by the Type of Banks

Table 2 summarises the annual average DEA efficiency scores of the three bank categories. Starting with scale efficiency for the large domestic banks, and investigating the sample of three banks from this group, scale efficiency of large domestic banks rose significantly from 71.9 percent in 1998 to 97.5 percent in 1999, the performance rate that remained stable for almost all the subsequent years.

Table 2: Annual Average DEA Efficiency Scores - Group Wise

		1998	1999	2000	2001	2002	2003	2004
	CRS	0.646	0.954	0.881	0.918	0.910	0.948	0.990
LB	VRS	0.901	0.979	0.940	0.963	0.940	0.954	0.993
	Scale	0.719	0.975	0.937	0.952	0.964	0.993	0.997
	CRS	0.981	0.943	0.986	0.973	0.965	0.985	0.981
IB	VRS	0.984	0.973	0.992	0.985	0.973	0.989	0.983
	Scale	0.997	0.970	0.994	0.987	0.992	0.995	0.998
	CRS	0.956	0.924	0.951	0.966	0.972	0.968	0.980
SB	VRS	0.973	0.932	0.954	0.969	0.977	0.974	0.985
	Scale	0.982	0.992	0.997	0.997	0.993	0.994	0.995

LB = large domestic banks, IB = subsidiaries of major international banks, SB = Small Banks

It is arguable that from 2000 onwards, scale efficiency of the large domestic banks was persistently rising although its divergence from the full efficient frontier was comparatively higher than that of other bank groups. The rising efficiency of large domestic banks in the study period is a sign of the success of continuing reforms in the financial sector. Banks in this group are those which were formerly Government owned and the ones that had extreme inefficiencies before reforms.

Analysing extensions of the major international banks in Tanzania (comparatively medium size banks), three out of six banks in this group were examined. The results indicate that they had higher scale efficiency scores than that of large domestic banks. Their annual average scale efficiency stood over 98.0 percent during the period. The viability of expansion of the major international banks in the country is supported by the findings because banks were operating under increasing returns to scale and so they had no reason to stop production expansion by opening new branches, launching new products, and so on.

Small banks' group is formed by 14 banks and 5 of them were included in the analysed sample. In 1998 their average scale efficiency was 98.2 percent and it rose to the rates above 99.0 percent throughout the subsequent years. Production in small banks took place at about 0.7 percent below efficient production frontier. It is approximately true that small banks were full scale efficient in the reviewed period.

With reference to technical efficiency, DEA input slacks show that for months in which large domestic banks were inefficient there were either positive labour input slacks and/or positive fixed assets and equipment slacks. These findings imply that inappropriate use of labour and fixed capital were a main source of technical inefficiency in large domestic banks. On average, technical efficiency divergence from the efficient frontier for the large domestic banks in the study period was about 10.8 percent and 4.7 percent for CRS and VRS models, respectively. Indicating that if they were full efficient, large domestic banks could reduce inputs used in their intermediation process by these rates to yet realise the same level of loans and investments they made.

Although technical efficiency for the large domestic banks was relatively low in the late 1990s, their performance converged to the efficient frontier in the last couple of years reviewed. Table 3 denotes Malmquist indices of commercial banks efficiency change. The table indicates performance enhancement for the large domestic banks resulting from an increase in the average technological innovations by 13.1 percent which translated into the rise in their total factor productivity by the same rate.

Table 3: Malmquist Indices of Efficiency Change

	TeffΔ	TechΔ	PTeffΔ	SeffΔ	TfpΔ
LB	1.000	1.131	1.000	1.000	1.131
IB	1.000	1.056	1.000	1.000	1.056
SB	1.002	0.992	1.000	1.002	0.994
Mean	1.001	1.058	1.000	1.001	1.059

TeffΔ = Technical efficiency change (relative to CRS technology)
TechΔ = Technnological change
PTeffΔ = Pure technical efficiency change (relative to VRS technology)
SeffΔ = Scale efficiency change
TfpΔ = Total factor productivity change

Though by the magnitude of efficiency scores large domestic banks lagged behind the major foreign banks and the small banks, Malmquist indices provide evidence that they had higher rate of technological developments than any other bank categories.

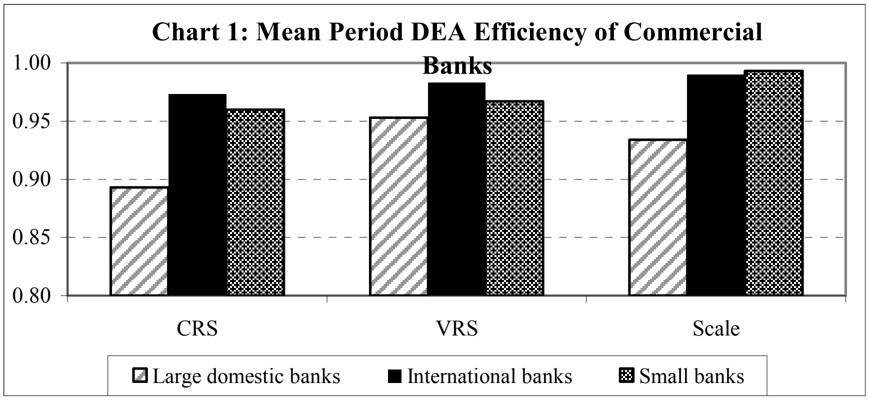
Regarding extensions of the major international banks, they had the highest average technical efficiency scores. Annual average technical efficiency of international banks under both CRS and VRS models stayed above 94.2 percent. On average the major international banks missed their frontier in technical efficiency context by 2.7 percent and 1.7 percent for CRS and VRS models, respectively. These divergence proportions point out the rates by which they should have reduced inputs use to produce the same levels of outputs. Malmquist indices of efficiency change show that extensions of the major international banks in the country made an improvement in technological innovations by about 5.6 percent over the period. This rate was lower than that of the large domestic banks. However, the argument for this is probably that because existing technical inefficiency that needed some improvement was on average very low, the endeavour to catch up could be of less management attention.

Small banks had higher technical efficiency scores than that of large domestic banks but lower than the scores of international banks. During the period, small banks had average technical efficiency rates of 96.0 percent and 96.6 percent in CRS and VRS models, respectively. Inefficiency divergences from the frontier were 4.0 percent and 3.4 percent for the two models. From the Malmquist indices, this group of banks indicates a positive development in both technical and scale efficiency, however, by a diminutive rate of 0.2 percent. Small banks were operating on the rising part of the average cost curve, and as for all other groups, technical inefficiency emanated from underutilisation of labour, fixed assets and equipment.

Table 4 summarises average period DEA efficiency scores of the bank groups estimated from both CRS and VRS models and these are also depicted on Chart 1.

Table 4: Period Average DEA Efficiency Scores

	CRS	VRS	Scale
All Banks	0.961	0.961	0.973
Large domestic banks	0.893	0.953	0.934
International banks	0.973	0.983	0.990
Small banks	0.960	0.967	0.993



These illustrations show that in connection with technical efficiency, the major international banks ranked highest compared to the rest of bank groups. Small banks constitute the second efficient group while large domestic banks form the least efficient one. Regarding scale efficiency, small banks ranked first followed by international banks and then large domestic banks.

4.3 X-inefficiency

Table 5 depicts the estimates of overall commercial banks stochastic frontier model. The significance of σ_{δ}^{2} reveals that banks experienced some x-inefficiency during the period. From the Wald test statistic, the specified stochastic frontier equation explains variations in x-efficiency at 1 percent level. X-inefficiency indices are predicted from the estimates of the stochastic frontier model using the distribution of inefficiency term δ_n conditional on ε_n (equation 7).

Table 5: Multi-product Translog Cost Function Estimates

All Commercial Banks

Number of obs = 83
Wald chi2(20) = 1924.37
Prob>chi2 = 0.0000

Log likelihood = 88.462922

TC	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
α_1	10.1102500	2.3167110	1.38	0.007	-4.2302	24.450740
α_2	-1.6270500	1.1110420	-1.46	0.143	-3.8047	0.550550
β_1	-25.7538300	5.1963550	4.96	0.000	-35.9385	-15.569160
β_2	35.3804700	9.6518050	3.67	0.000	16.4633	54.297660
β_3	10.6169200	3.4219210	3.10	0.002	3.9101	17.323760
γ_{11}	0.4415200	0.3383580	1.30	0.192	-0.2217	1.104680
γ_{22}	0.0718200	0.0278050	2.58	0.010	0.0173	0.126320
γ_{12}	-0.1272600	0.0597360	-2.13	0.033	-0.2443	-0.010180
ψ_{11}	0.1083600	0.0780560	1.39	0.165	-0.0446	0.261350
ψ_{22}	0.2371700	1.1223170	0.21	0.833	-1.9625	2.436870
ψ_{33}	0.1624800	0.1101280	1.48	0.140	-0.0534	0.378330
ψ_{12}	0.9463300	0.3256440	2.91	0.004	-0.3081	1.584580
ψ_{13}	0.5605100	0.1787030	3.14	0.002	0.2103	0.910760
ψ_{23}	-1.3664000	0.3496970	-3.91	0.000	-2.0519	-0.681070
ω_{11}	0.3353300	0.1793920	1.87	0.062	-0.0163	0.686940
ω_{12}	-1.5944500	0.4817480	-3.31	0.001	-2.5387	-0.650240
ω_{13}	0.2335800	0.1660230	1.41	0.159	-0.0918	0.558980
ω_{21}	0.0485600	0.0532940	0.91	0.362	-0.0559	0.153010
ω_{22}	0.2712600	0.0992640	2.73	0.006	0.0767	0.465820
ω_{23}	-0.0852200	0.5723600	-1.49	0.136	-0.1974	0.026950
α_0	-301.8129000	78.7414700	-3.83	0.000	-456.1433	-147.482400
lnsig2μ	-6.1438250	0.4839534	-12.70	0.000	-7.0924	-5.195292
lnsig2δ	-5.0180840	0.4093681	-12.26	0.000	-5.8204	-4.215737
sigma_μ	0.0463325	0.0112114			0.0288	0.074449
sigma_δ	0.0813461	0.0166503			0.0545	0.121497
sigma2	0.0087639	0.0021756			0.0045	0.013028
mbda	1.7557050	0.0254419			1.7058	1.805570

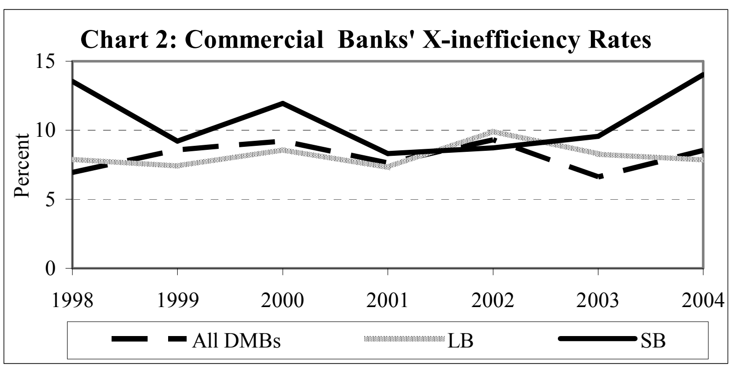
Likelihood-ratio test of sigma_δ = 0: chibar2(01) = 30.04 Prob > = chibar2 = 0.000

The summary of annual average x-inefficiency indexes expressed in percentage is shown on Table 6. In spite of the fact that the overall x-inefficiency was at less than 10 percent during the study period, there were no significant improvements over time. Instead, x-inefficiency was on increase from 7.0 percent in 1998 to the highest rate of 9.3 percent recorded in 2002. The record of 2004 designates an upwards reverse of the trend of x-inefficiency compared with the previous year's situation. In view of x-inefficiency indexes over the period, the change towards improvement seems to be almost negligible. This implies the existence of certain static management approaches which deliver with some sustained inefficiency rates varying little from the period mean of 8.1 percent. Changes were up and down with the standard deviation of 1.1.

Table 6: X-inefficiency Indexes - Annual Average (%)

All Commercial Banks							
	1998	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**
All DMBs	6.96	8.60	9.23	7.65	9.31	6.64	8.56
Large Domestic banks							
	1998	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**
LB 2	7.32	7.82	8.99	8.28	8.73	8.98	9.33
LB 3	8.49	7.03	8.22	6.34	11.13	7.57	6.20
Average	7.90	7.42	8.60	7.31	9.93	8.28	7.76
Small Banks							
	1998	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**
SB 1	24.29	15.93	16.99	15.18	11.94	16.90	22.43
SB 2	9.76	6.91	16.03	6.52	11.18	7.92	11.45
SB 3	6.56	4.83	2.83	3.25	3.12	3.89	3.16
Average	13.54	9.22	11.95	8.32	8.75	9.57	12.35

The trend of annual average of commercial banks x-inefficiency is shown on Chart 2. Group-wise examination of x-inefficiency in commercial banks revealed that out of the nine banks included in the respective bank groups, four were full x-efficient, and among which were all the three sampled major international banks. The analysis of the three banks included in the sample of large domestic banks shows that one of them was full x-efficient while the remaining two were x-inefficient (Table 6).



The period average x-inefficiency in the large domestic banks was 8.2 percent, being only 0.1 percent above the overall banks x-inefficiency rate. X-inefficiency indexes of the large domestic banks reflect closely the indexes of the overall banking sector. These are not shocking results because they are pertinent to financial sector settings of the developing economy with its banking industry dominated by some few large banks. Average x-inefficiency in large domestic banks kept on rising until it reached a highest rate of 9.9 percent in 2002. Nonetheless, x-inefficiency trend in these banks declined in the last couple of years reviewed.

With regard to small banks, the notable finding is that none of the three banks in their sample was full x-efficient, and the results indicate that they had highest x-inefficiency record over the period. Small banks had a period average of x-inefficiency rate of 10.5 percent, with a standard deviation of 2.0. Ideally, assuming the same management competence across banks, one could argue that small banks would be more x-efficient than the large banks because management task is easier for the small firms than the large firms. But heterogeneity of x-inefficiency indexes across small banks compared to large banks is an important issue to note. And this observation points out the absence of standard management across small banks.

Factor Price and Substitution Elasticities: To see the responsiveness of input demands to their prices and to establish whether factors were substitutes or complements, elasticities are computed from the data of sample banks. Commercial banks own factor price responsiveness was inelastic for each of the three inputs (Table 7).

Table 7: Own Price Elasticities

Factor	Capital	Labour	Funds
Elasticity	-0.04	-0.05	-0.43

Factor price elasticities were very low implying that changes in the prices of these factors did not cause considerable changes in the quantities demanded. One of the reasons for low price elasticities is the absence of close factor substitutes. These low elasticities suggest that there is lack of flexibility in the usage of factors of production in the banking industry. Table 8 shows multi-factor Allen elasticities of substitution among inputs. In Tanzanian commercial banks factors of production labour, capital and funds were found to be complements to each other rather than substitutes.

Table 8: Elasticities of Substitution

Factors	σ_{KL}	σ_{KF}	σ_{LF}
Elasticity	-3.51	-1.02	-0.99

K, L, and F stand for capital, labour and funds.

Causes of less factor substitutability in developing economies are many, ranging from technological barriers, high cost sensitivity to changes, and variety of structural impediments.

4.4 Determinants of X-inefficiency

X-inefficiency indexes were regressed against five explanatory variables comprising capital adequacy, labour compensation or the incentive to work proxy, assets quality, bank size and the excess liquidity (equation 11). Using consolidated data for all sampled x-inefficient commercial banks, the Tobit model is estimated and the results are presented on Table 9. With exception of one variable (assets quality) which is insignificant, the tests suggest plausible robustness of all the estimates.

Table 9: Tobit Estimates of the Sources of X-inefficiency

All Commercial Banks
Two-limit Tobit estimates — Number of obs = 84
LR chi2(5) = 26.35
Prob > chi2 = 0.0001
Log likelihood = 93.780952 — Pseudo R2 = -0.1634

Ineff |	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
β_1 |	-3.80e-10	9.52e-11	-3.99*	0.000	-5.69e-10	-1.90e-10
β_2 |	-852.5821	197.1919	-4.32*	0.000	-1245.083	-460.0814
β_3 |	-6.7301	10.99212	-0.61	0.542	-28.60937	15.14917
β_4 |	1.88e-13	9.38e-14	2.01**	0.048	1.71e-15	3.75e-13
β_5 |	.0507309	.0184045	2.76*	0.007	.0140976	.0873642
β_0 |	.1851182	.0479533	3.86*	0.000	.0896695	.280567
se |	.0746839	.0058615	(Ancillary parameter)			

* significant at 1% level, ** significant at 5% level.

Starting with capital adequacy variable, *K*, we see that it is significant and bears a negative sign. The intuition is, as long as the bank spends more on improvement of its capital base, it increasingly adds to its efficiency gain. Spending on capital goods in commercial bank caters-among other things-for equipment like ATMs, computers, and network development to enhance efficient service provision. Under-investment in fixed capital goods is thus among the factors that derive x-inefficiency in banks.

The incentive to work variable, *L*, is significant and has a negative sign indicating that when incentive to work increases x-inefficiency in banks decreases. This variable is integrally important because it even reflects the quality of labour a bank is able to attract. The competent bank management will most likely be employed in the well paying banks. In this regard, poor labour compensation or the low remuneration is found to be one of the sources of operational inefficiency in banking firms.

Results suggest that the bank size, *S*, is positively related to x-inefficiency. This finding is pertinent to the managerial utility maximisation theory of the firm. The underpinning conceptual framework is, the large firms are more vulnerable to managerial utility maximisation which may rather be motivated more by other external factors than the internal firm performance goals. As the firm grows, the separation of ownership and management increases, agency problem heightens, and the management self interests easily entrench firm objectives.

In connection with banks liquidity, estimated parameters established a significant relationship between bank efficiency and excess liquidity, *EL*. This efficiency-excess liquidity relationship has not been tested explicitly in other previous studies of bank efficiency. In this study excess liquidity is found to be positively related with x-inefficiency index in the Tanzania's banking firms. This finding confirms the hypothesis that accumulation of excess liquidity in banks precipitates inefficiency.

5. Concluding Remarks and Policy Implications

The remarkable comment from the findings of this study is that efficiency status of commercial banks in Tanzania is not disappointing to financial sector reforms because the scores turned out to be fairly high. With reference to DEA estimates during the period, commercial banks operated on the decreasing part of their average cost curves and this gave them a room to expand with increasing returns to scale. However, the results portray upward turning average cost curves especially during 2004. In terms of technical efficiency, foreign banks ranked the highest, followed by small banks and then large domestic banks; while regarding scale efficiency, small banks ranked the highest followed by international banks and then large domestic banks. Evidences from the findings indicate that x-inefficiency in banks was the outcome of inadequate fixed capital, poor labour compensation, less management capacity as banks expanded, and the overwhelming accumulation of excess liquid assets.

Following the major conclusions drawn from the empirical analysis, next are crucial policy implications that practitioners, policymakers, and other stakeholders may take into consideration. Because the average cost curves of commercial banks were eventually turning up, enhanced competition in the existing banks should be of the prime emphasis rather than speeding up to the establishment of many more new banks. It is important to encourage international banking. The critical challenge is on the aspect of capital flights from the economy through foreign banks, which may not have compensating effect if banks in the country do not equally rise up for outreach to other economies.

In some banks human resources are not proportional to the range of activities they have to do. Either more people than required are employed or the employees do not work as they should do. Such banks should accordingly implement policies aiming at enhancing efficiency and reactivation of the work morale. It will be optimal to introduce and corroborate universal banking in Tanzanian commercial banks. Banks should harness their underutilised resources, which can be used in the production of new (non-traditional) variety of products. Because technological change improves banks' productivity, bank supervisors should design practicable regulation for technological standards requirement.

Capital inadequacy is among the critical impediments to x-efficiency. Banks may want to optimise productivity through rearrangement of their capital-labour balance. From the estimates of the sources of x-inefficiency, poor labour compensation is among the sources of uninspired work effort in banks. It implies that banks should offer competitive compensations to their employees to reactivate workers exuberance in order to raise productivity. Moreover, the central bank has to ensure that bank managers have enough professional competence and experience in order to guarantee mandatory skills in management of the banking business. Last but not least, the central bank has to take an action to contain the sub-optimally rising trend of Treasury bills rate. Otherwise, the rising T-bills rate will further crowd out private sector credit. Owing to the competitive nature of returns from alternative assets, commercial banks consider T-bills as a more priority investment avenue than loans and other less liquid assets. Excess holding of Treasury bills in commercial banks add to liquidity which does not only reduce credit allocated to the private sector but also undermine commercial banks' efficiency.

References

Aigner, D., K. Lovell, and P. Schmidt, (1977) "Formulation and Estimation of Stochastic Frontier Production Function Models". *Journal of Econometrics*, Vol. 6, pp. 21-37.

Agu, C. (2004),"Efficiency of Commercial Banking in the Gambia". *University of Nigeria, Nsukka*.

Baker, R., Charnes, A., and W. Cooper (1984), "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis". *Management Science*, Vol. 30, No. 9, pp. 1078-1092.

Baltensperger, E. (1972), "Economies of Scale, Firm Size, and Concentration in Banking", *Journal of Money, Credit and Banking*, Vol. 4, No. 3, pp. 467-488.

Berger, A. N., William, H., and T. Stephen (1993), "The Efficiency of Financial Institutions: A Review and Preview of Research Past, Present, and Future". *Journal of Banking and Finance*, Vol. 17, pp. 221-249.

Berger, A. N., and D. B. Humphrey (1997), "Efficiency of Financial Institutions: International Survey and Directions for Future Research". *Working Papers Wharton School, University of Pennsylvania*.

Berger, S. A. and R. De Young (1997), "Problem Loans and Cost Efficiency in Commercial Banks". *Journal of Banking and Finance*, Vol. 21, pp. 849-870.

Charnes, A., Cooper, W., and E. Rhodes (1978), "Measuring the Efficiency of Decision Making Units", *European Journal of Operational Research*, Vol. 6, pp. 429-444.

Clarke, G. et al. (2001), "Foreign Bank Entry: Experience, Implications for Developing Countries, and Agenda for Further Research", *World Bank Working Paper series 2698*.

Debreu, G. (1959): *Theory of value, an Axiomatic Analysis of Economic Equilibrium*. John Wiley, New York.

Drake, L. and R. Simper (2002), "Economies of Scale in UK Building Societies: A Re-appraisal Using an Entry/Exit Model". *Journal of Banking and Finance,* Vol. 26, pp. 2365-2382.

Farrell, M. J. (1957), "The Measurement of Productivity Efficiency". *Journal of the Royal Statistical Society*, Series A, CXX, Part 3, pp. 253-290.

Green W. (2003), *Econometric Analysis*. Prentice-Hall, New York. Fifth Edition.

Hirshleifer J. and A. Glazer (1993), *Price Theory and Applications*. Prentice-Hall of India, New Delhi, Fifth Edition.

Howard, M. and H. Haynes (2001), "Commercial Banks Efficiency in Barbados". *Savings and Development,* Vol. 25, No. 3.

Ikhide, S. (2000), "Efficiency of Commercial Banks in Namibia". *BON Occasional Paper* No. 4

IMF-World Bank (2003), Tanzania: Financial System Stability Assessment, Including Reports on the Observance of Standards and Codes on Banking Supervision, *IMF Publications*.

Kenny, C. J. and T. J. Moss (1998), "Stock Market in Africa: Emerging Lions or White Elephant?" *World Development,* Vol. 26, No. 5, pp. 829-43.

Leibenstein, H. (1966), "Allocative Efficiency versus 'X-efficiency". *American Economic Review,* Dec., pp. 1252-1258.

Lindley, J. et al. (2000), "Excess Reserves During the 1930s:mpirical Estimates of the Cost of Converting Unintended Cash Inventory into Income Producing Assets". *Macon State College Paper*.

Maddala, G. S. (1983), *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University Press.

Masco-Colell, A. et al. (1995), *Microeconomic Theory*, 1st edition. Oxford University Press.

Mitchell, K. and N. M. Onvural (1996), "Economies of Scale and Scope at Large Commercial Banks: Evidence from the Fourier Flexible Functional Form", *Journal of Money, Credit and Banking*, Vol. 28*,* pp*:*178-197.

Myers, S. and R. Rajan (1995), "The Paradox of Liquidity". *NBER Working Paper* No. 5143.

Parisio, L. (1992), "Economies of Scale and Scope in the Italian Bank Industry: Evidence From Panel Data. *Rivisita Internazionale di Scienze Economiche e Commerciali,* Vol. 39, pp. 959-78.

Schenk, R. (2004), *What is Economic Efficiency? Journal of Economic Education, CyberEconomics,* Vol. 34, Issue 2, pp. 191.

Shephard, R. W. (1953), *Cost and Production Functions*. Princeton University Press, Princeton, NJ.

Shephard, R. W. (1970), *The Theory of Cost and Production Functions*. Princeton University Press, Princeton, NJ.

Tobin, J. (1958), "Liquidity Preference as Behaviour Towards Risk", *Review of Economic Studies.*

Varian, H. R. (1992), *Microeconomic Analysis*, 3rd edition. W.W. Norton & Inc., New York.

Vittas, D. (1991), "Measuring Commercial Bank Efficiency", *Policy Research Working Papers Series, Country Economic Development, The World Bank, Washington D. C.,* November.