The Impact of the Global Financial Crisis on Efficiency and Productivity of the Banking System in South Africa

Andrew Maredza and Sylvanus Ikhide

ERSA working paper 328

January 2013

Economic Research Southern Africa (ERSA) is a research programme funded by the National Treasury of South Africa.

The Impact of the Global Financial Crisis on Efficiency and Productivity of the Banking System in South Africa

Andrew Maredza*and Sylvanus Ikhide†

January 22, 2013

Abstract

South Africa's financial sector is believed to have weathered the contagion and catastrophic effects of the 2008 world wide financial crisis partly on account of a sound regulatory framework and solid macroeconomic policies. In this paper, we seek to measure efficiency and productivity changes during the period of the crisis through an analysis of bank performance over the period 2000 – 2010 using a two stage methodology framework. The recently developed Hicks-Moorsteen total factor productivity (TFP) index approach developed by O'Donnell (2010a) as opposed to the popular Malmquist TFP was utilised. Our first stage results showed that during the crisis period there was a noticeable but mild deviation of total factor productivity and efficiency measures. Second stage analysis using the censored Tobit model showed that the financial crisis was the main determinant of bank efficiency, indicating that total factor productivity efficiency was 16.96% lower during the crisis period compared to the pre-crisis period.

Key Words: Banking Efficiency, Data Envelopment Analysis, Global Financial Crisis, Hicks-Moorsteen, Malmquist, South African Banking, Total Factor Productivity Efficiency, Censored Tobit Model.

JEL Classification: G01, G21, C14, C24

1 Introduction

The primary objective of this paper is to determine if there has been a change in efficiency and productivity of South African banks during the period of the global financial crisis. The failure of the financial system of the United States in 2007 generated larger concern on the financial sector of which the banking

*School of Economics & Decision Science, North West University, Private Bag X2046, Mmabatho, 2735, South Africa. Corresponding author's e-mail: Andrew.Maredza@nwu.ac.za

†Business School, University of Stellenbosch, P.O Box 610, Bellville 7535, South Africa, e-mail: Sylvanus.Ikhide@usb.ac.za

system forms the major component. The financial global crisis showed how catastrophic problems in the financial sector can be for the entire economy. The banking sector is crucial in that it shelters the economy against instability and boost consumer confidence. Mercan et al (2003) maintains that the fragility of the fibre of banks and the fact that they are perceived to be institutions of confidence, threatens not just banking institutions but the entire economy.

The banking industry is considered the backbone of most economies and plays a vital role in attaining economic growth and development. A well-functioning banking sector contributes to economic growth via more efficient allocation of resources and risk diversification. This nexus between banking performance and economic growth demands that an enquiry of this nature be conducted. The efficiency and productivity of such instrumental elements of the economy is critical to the well being of the economy. The financial system and the banking system in particular facilitate the efficient movement of funds from sources of plenty (savers) to units of deficit (borrowers). Hence the growth and development of any economy hinges strongly on the proper functioning of the financial system.

According to South African Reserve Bank (SARB) (2011), the devastating effect of the sub-prime financial crisis in South Africa was fairly cushioned by the solid macroeconomic policies, and prudent regulation in the domestic banking sector. However, as highlighted by Mnyande[1] (2012) in his key note address, South Africa is not immune to the crisis in the medium to long term. He stated that South Africa's inter-linkages with the global economy mean that its growth outcomes are highly dependant on those of the affected European economies. Hence, there is a need to investigate the magnitude of efficiency and productivity changes that were realised during the period of the crisis. This is important for two reasons. Firstly, knowledge of the size of the impact of the crisis on the banking sector is crucial to inform policy makers to provide precise and targeted policy measures as well as whether to tighten bank supervision and regulation. Secondly, pessimists have argued that it is too early to celebrate and that the worst case global challenges may have been merely delayed rather than avoided.

2 The structure of the South African banking industry

The banking system in South Africa is recognized for being well developed, sophisticated and well regulated and is ranked among those of first world economies. According to SARB (2012), the banking sector has a current composition of 17 domestic controlled banks, 11 branches of international banks in South Africa, 2 mutual banks, 1 co-operative bank and 43 representative offices. However, the South African banking sector is dominated by four major banking institutions the so-called "Big Four" namely the Amalgamated Bank of South Africa (ABSA), FirstRand Bank, Nedbank, and Standard Bank. These four largest

[1] Chief Economist and Adviser to the Governor of Reserve Bank.

banks, contributed 84.6 % to the balance-sheet size of the total banking sector at the end of December 2010 (SARB Annual Report, 2010). The rest of the banks hold a very small portion of the market share. Hence, the South African banking industry exhibits a high level of concentration. Concentration is defined as the extent to which most of the market's output is produced by a few firms in the industry.

The significance of the banking sector can be examined in terms of the size of the sector's assets. Figure 1, shows that the size of the banking industry's assets as a percentage of GDP has been increasing steadily over the years. By 2001, the value of banks' assets exceeded GDP for the first time. In 2008, the ratio of total banking sector assets to GDP reached its highest of 139 % from 102 % since 2001 (see figure 1 and 2). Maheshwari (2009) states that for the period 2003-2006, asset growth recorded in the South African banking industry was higher than that witnessed in the global banking industry. Hence, the size of these assets relative to GDP highlights the importance of the banking sector to the South African economy. Banks' assets are predominantly loans and advances to the private non-bank and government sectors. According to Banking Association South Africa (2010), the economic recession in 2009 which was a consequence of the global financial crisis affected consumer affordability and therefore spending patterns, resulting in consumers being reluctant to take on more debt. This resulted in the increase in non-performing loans which had a huge impact on the banks' loan books and so total assets and liabilities declined in 2009.

Total banking-sector assets amounted to R3406 billion at the end of December 2011 (December 2010: R3126 billion), representing a moderate year-on-year increase of 8.9 %. As at December 2010, the banking sector assets represented an increase of 5.3 % from a value of R2967 billion in December 2009. Gross loans and advances, which represented, on average, 74 % of banking-sector assets during 2010, increased marginally by 2.5 % to R2314 billion at the end of December 2010, mainly due to modest growth in home loans and higher overnight and interbank call loan balances (SARB Supervision Report, 2010).

3 Theoretical framework

3.1 Productivity

Productivity is a simple indicator that describes the relationship between output and the inputs that generate that output and as such regarded as an important measure of economic performance. Productivity of a bank is measured by the quantity of output produced per unit of input factor employed. In the singe-output and single-input setting, productivity is simply the ratio of that bank‘s output and input quantities. Suppose that, in period i,bank i produces output $y_t$using inputx_t, its productivity becomes:

$$TFP_{it} = \frac{Y_{it}}{X_{it}} \tag{1}$$

Thus, TFP of bank i in period $t$relative to its TFP in period 0 is:

$$TFP_{is,it} = \frac{TFP_{it}}{TFP_{i0}} = \frac{{}^{Y_{it}}/_{X_{it}}}{{}^{Y_{i0}}/_{X_{i0}}} = \frac{Y_{i0,it}}{X_{i0,it}} \tag{2}$$

Formulation (2) shows that TFP growth is the ratio of the output growth and input growth. Similarly, the productivity of bank $h$in period s for producing output $y_t$with input$x_t$is

$$TFP_{hs} = \frac{Y_{hs}}{X_{hs}} \tag{3}$$

Hence, the productivity index of bank i in period t relative to bank h in period s is given by:

$$TFP_{hs,it} = \frac{TFP_{it}}{TFP_{hs}} = \frac{{}^{Y_{it}}/_{X_{it}}}{{}^{Y_{hs}}/_{X_{hs}}} = \frac{{}^{Y_{it}}/_{Y_{hs}}}{{}^{X_{it}}/_{X_{hs}}} = \frac{Y_{hs,it}}{X_{hs,it}} \tag{4}$$

This particular productivity index shows the change in productivity of bank i in period t when compared to the reference bank h in period s.

3.1.1 The Hicks-Moorsteen Index

Productivity in the past has been analysed using the more popular Malmquist TFP index approach. However, recent developments by O'Donnell (2010a) have shown that the Malmquist is a biased and inconsistent approach when applied to technologies exhibiting varying returns to scale (VRS). Coelli (1996) thus emphasises the importance of imposing constant returns to scale (CRS) technology assumption when employing the Malmquist approach. In this particular study, the non-parametric and multiplicatively-complete Hicks-Moorsteen technique which was developed by O'Donnell (2010a) to respond to the inadequacy of the popular Malmquist technique will be utilised. This is an attempt to redress the shortcomings entrenched in the banking literature on researches that were conducted in the past. The Hicks-Moorsteen TFP Index can be formulated as:

$$TFP_{hs,it} = \left(\frac{D_O\left(x_{hs}, y_{it}, s\right)}{D_O\left(x_{hs}, y_{hs}, s\right)} \frac{D_I\left(x_{hs}, y_{hs}, s\right)}{D_I\left(x_{it}, y_{hs}, s\right)} \frac{D_O\left(x_{it}, y_{it}, t\right)}{D_O\left(x_{it}, y_{hs}, t\right)} \frac{D_I\left(x_{hs}, y_{it}, t\right)}{D_I\left(x_{it}, y_{it}, t\right)} \right)^{\frac{1}{2}} \tag{5}$$

Where $\boldsymbol{y}$ and $\boldsymbol{x}$are vectors of quantities and D_O (.) and D_I(.) are output and input distance functions. Formulation (5) above was first proposed by Bjurek (1996) but is commonly known as the Hicks-Moorsteen index. O'Donnell (2011) states that it is called as such because it is the geometric average of two indexes that Diewert (1992) attributed to Hicks (1961) and Moorsteen (1961).

O'Donnell (2011) describes TFP indexes that can be expressed in terms of aggregate quantities as in equation (4) as being *multiplicatively-complete.* However, unlike the Hicks-Moorsten TFP Index, the popular Malmquist TFP index is not included among the class of multiplicatively complete TFP indexes. Thus O'Donnell (2010a, pp.1) argues that for this reason it cannot be regarded

as a valid measure of productivity change except under constant returns to scale technology - which he describes as "except in restrictive special cases".

Briec and Kerstens (2011) proved that the Hicks–Moorsteen productivity index indeed satisfies the determinateness axiom under weak conditions on technology. However, on the contrary the popular Malmquist productivity index failed to satisfy the determinateness condition. In their conclusion the authors state: "We expect the Hicks–Moorsteen productivity index to gain in popularity in future empirical work, especially when infeasible solutions are simply unacceptable" Briec and Kerstens (2011, pp.10).

3.1.2 Productivity Change and its Components

O'Donnell (2010c) shows that any multiplicatively-complete TFP index can be exhaustively decomposed into measures of technical change and several meaningful measures of efficiency change. These measures are presented below from an input orientation perspective.

Depending on the perspective of analysis of the researcher one can choose to analyse efficiency from an input orientation or output orientation. An input orientation measures input reductions that are necessary for a production unit to become efficient without a reduction in output. An output orientation measures the expansion of output that is necessary for efficiency improvement holding inputs constant. Thus output inefficiencies represent the needed increase in output for the inefficient bank to become efficient.

Input-oriented Technical Efficiency (ITE)

It measures the difference between observed TFP and the maximum TFP that is possible while holding the input mix, output mix and output level fixed. ITE is illustrated in Figure 4 as the measure of the horizontal distance from point A to point B. It is equivalent to the measure of the difference in TFP at points A and B: $ITE_{io} = {}^{slope\ OA}/_{slope\ OB}$

Input-oriented Scale Efficiency (ISE)

It measures the difference between TFP at a technically efficient point and the maximum TFP that is possible while holding the input and output mixes fixed (but allowing the levels to vary). Figure 4 represents this measure of efficiency as a movement from point B to point D: $ISE_{io} = {}^{slope\ OB}/_{slope\ OD}$. O'Donnell (2010a) refers to point D as the point of *mix-invariant optimal scale* (MIOS).

Residual Mix Efficiency (RME)

It measures the difference between the maximum TFP possible on a mix-restricted frontier and the maximum TFP possible when input and output mixes (and levels) can vary. In figure 4, it is represented as a movement from point D to point E: $RME_{io} = {}^{slope\ OD}/_{slope\ OE}$

Input-oriented Mix Efficiency (IME)

It measures the difference between TFP at a technically efficient point on the mix-restricted frontier and the maximum TFP that is possible while holding the output level fixed. This measure of efficiency is represented in figure 5 as a

movement from point B to point U: $IME_{io} = {}^{slope\ OB}/_{slope\ OU}$

Residual input-oriented Scale Efficiency (RISE)

It measures the difference between TFP at a technically and mix-efficient point and TFP at the point of maximum productivity. RISE is represented in figure 4 as a movement from point U to point E: $RISE_{io} = {}^{slope\ OU}/_{slope\ OE}$

TFP Efficiency (TFPE)

It measures the difference between observed TFP and the maximum TFP* possible using the available technology. Figure 4 illustrates this measure as a movement from point A to point E: $TFPE_{io} = {}^{TFP_{I0}}/_{TFP_0^*} = {}^{slope\ OA}/_{slope\ OE}$

Where TFP_0^* denotes the maximum TFP possible using the technology available in period 0. Hence, TFP efficiency (TFPE) is a measure of overall productive performance.

4 Brief review of empirical literature

Efficiency studies on the banking system particularly in developing economies are very few. In this section of the paper, a brief empirical literature review relating to the economic performance of the banking sector is presented. This review is intended to show previous work in the area of banking efficiency and productivity in South Africa.

Okeahalam (2006) employed the Bayesian stochastic frontier approach to assess the production efficiency of 61 South African bank branches in the 9 provinces for the year 1999. The author found productive efficiency of banks to be 83.1% suggesting that on average banks could reduce their costs by 16.9% without altering their current output levels. Okeahalam also found that all branches were operating at increasing returns to scale and recommended levels of output to be increased either through regulatory reforms or competitive incentives.

Ncube (2009) conducted a study to analyse the cost and profit efficiency of banks in South Africa. Their study employed the parametric stochastic frontier approach to determine both cost and profit efficiency of four large and four small South African banks over the period of 2000-2005 classified according to the number of employees. The large banks whose number of employees exceeded 10 000 included ABSA, FirstRand Bank, Nedbank and Standard Bank. The small banks with employees less than 10 000 included African Bank, Capitec Bank, Investec Bank and Teba Bank. The average cost and profit efficiencies over the six periods were 92% and 55% respectively. Their study concluded that South African banks were relatively better at controlling cost than generating profit as indicated by the lower profit efficiency score and a higher cost efficiency score.

Kumbirai and Webb (2010) investigated the performance of the five largest South Africa‘s commercial banks for the period 2005 – 2009. Three aspects of bank performance namely profitability, liquidity and credit quality were analysed using financial ratio analysis. The findings showed that overall performance improved considerably in 2005 and 2006. However, the impact of the global financial crisis was evident when the overall performance deteriorated in

2007 until 2009. Using the student t test to test if there was any significant difference in profitability performance for the period 2005-2006 and the period 2008 – 2009, their results indicated that profitability deteriorated during the later period. However, they concluded that the South African banking system remained stable as there were adequately capitalised and profitable.

Oberholzer et al (2010) applied the DEA methodology to five of the largest banks in South Africa for the ten year period 1998 - 2007. The total assets of these five banks as at the end of 2007 represented 85.28% of the total banking assets in South Africa. Two DEA models were employed based on the definition of outputs included. Model 1 used only income statement data as outputs namely the value of interest income and non-interest income. Under model 2, only balance sheet data outputs were considered, that is, the value of deposits, loans and equity. The empirical findings from DEA analysis revealed that the average technical efficiency of all the banks was 89.5% and 79% for Model 1 and Model 2 respectively. Therefore banks had the capacity to increase output by 10.5% and 21% without increasing their inputs, respectively. In addition, they found that the average allocative efficiency of all the banks were 98.5% and 89.3% respectively.

As previously stated, the banking sector in South Africa is highly concentrated with only four large banks controlling over 80% of total deposits and assets. As such many studies conducted in South Africa [Okeahalam (2001), Greenberg & Simbanegavi (2009), Mlambo and Ncube (2011),] have taken a keen interest on investigating the nexus between competition and bank performance.

This brief review has demonstrated the developments that have taken place within the banking industry of South Africa. The significance of this paper hinges on the fact that the study uses the Hicks-Moorsteen TFP DEA-based methodology framework that was developed by O'Donnell (2010a) to address the shortcomings entrenched in the popular Malmquist TFP index approach. Therefore, unlike previous studies in South Africa, this paper employs this methodology, covering the period of the sub-prime financial crisis which highlights significant changes that may have occurred in the banking industry. This is crucial in order to inform formulation of suitable policy for enhancing bank performance and economic growth at large.

5 Methodology

Financial ratios have traditionally been used in the banking sector to measure efficiency and productivity. However, the exploitation of financial ratios as an exclusive measure of performance has been criticised by many researchers as being inadequate (Ncube, 2009; Yeh, 1996; Berger & Humphrey, 1997). (Ncube, 2009) argues that on one hand the ratio approach is straightforward and unsophisticated but on the other hand it does not provide enough insight on the actual efficiencies. However, this study adopts a two stage methodology framework. In the first stage, the Hicks-Moorsteen total factor productivity (TFP)

indices are computed and further decomposed into a measure of technical change and several measures of efficiency change (O'Donnell, 2011). In the second stage, various explanatory factors that are known to affect bank performance are regressed on total factor productivity efficiency. The main purpose of the second stage Tobit regression is to determine which factors influence overall productive efficiency significantly and to estimate the size of factor marginal effects. Sufian (2009) also adopted the same methodology framework when he investigated the efficiency of Malaysian and Thailand banks during the 1997 Asian financial crisis. The author first employed DEA and then a Tobit model for the period 1995-1999.

5.1 The Tobit Model

The censored Tobit model will be applied on the obtained first-stage total factor productivity efficiency (TFPE) scores. As stated earlier, TFPE is a measure of overall productive efficiency which captures the difference between observed TFP and the maximum TFP possible using the given technology. The Tobit model is used to estimate equations whose dependant variable values are constrained within some range. Given that efficiency scores are restricted between 0 and 1, no values should be observed below the lower limit and above the upper limit. Irsova and Havranek (2010), state that unlike OLS, censored regression models generate consistent estimates of coefficients in cases of limited dependant variables. The standard Tobit regression will be censored between 0 and 1 as follows:

$$Eff_i = \begin{cases} 0 & if\ z_i\beta_i + \varepsilon_i \leq 0 \\ z_i\beta_i + \varepsilon_i if\ 0 < z_i\beta_i + \varepsilon_i < 1 \\ 1 & if\ z_i\beta_i + \varepsilon_i \geq 1 \end{cases}$$

The vector $\boldsymbol{z}_i$ includes the variables that affect DMU efficiency and the vector β is a vector of coefficients to be estimated using the software application EViews. In order to avoid the dummy variable trap, only one dummy variable was specified to reflect the period of the crisis. The omitted pre-crisis dummy becomes the reference category. The empirical model to be estimated will therefore take the following form;

$$\begin{aligned} TFPE_t &= \beta_1 + \beta_2 CRISIS_t + \beta_3 NPL_t + \beta_4 TA_t + \beta_5 CIR_t + \beta_6 NETP_t \\ &\quad + \beta_7 SPREAD_t + \beta_8 NON\_INT + \varepsilon_t \end{aligned}$$

Where**:**

TFPE_t - Total Factor Productivity Efficiency Score in period t
CRISIS_t- Crisis Period: 1 if 2008 - 2010, 0 otherwise
NPL_t- Non-Performing Loans to total gross loans
TA_t- Total Banking Assets
CIR_t - Cost to Income Ratio
NETP_t - Net Profit after tax
SPREAD_t - Interest Spread (Lending – Deposit Rates)
NON_INT - Non-Interest Income
ε_t- Error term to capture other possible factors not specified.

5.2 Discussion of Second stage variables and the expected priori

Dummy for Financial crisis (POST)

This variable constitutes the main focus of this paper. While most economies particularly emerging and developed economies were not immune to the contagion, South Africa's banking sector performed well partly on account of prudent bank supervision and solid regulation. The adverse effect of the 2007/2008 subprime financial crisis is expected to have a significant and negative coefficient but marginal effect.

Interest Rate Spread (%)

A reduction in spreads between the lending and deposit rates is an indication of increased intermediation efficiency. Ikhide (2008), states that wide spreads affect intermediation and distort prices weakening the role of the financial system to the economy. An increase in the interest rate spread is expected to decrease intermediation efficiency and therefore overall productive efficiency. The variable is expected to exhibit a negative relationship with TFPE.

Non-performing loans (NPLs) to total gross loans (%)

NPL is an important variable that captures the quality of loans. Bad loans indicate inefficiency in lending and a lack by the management to manage risk. Therefore an increase in NPL is expected to decrease efficiency and productivity. As such the sign of NPL coefficient is expected to be negative.

Bank Size (TA)

The total banking assets will be used to capture the effect of bank size on bank performance. An entity can exhibit inefficiencies due to inappropriate size. For instance a bank can be either too small or too large relative to the optimum size. The coefficient sign can either be negative or positive.

Cost to Income Ratio (CIR)

The ratio is a measure of the proportion of the bank's income that is consumed by its operating costs. This variable was included to capture cost efficiency of the banking sector. A negative relationship is expected between this ratio and bank performance.

Non-Interest Income (NIC)

Many studies on bank efficiency have advocated for the inclusion of non-interest income as a proxy for diversification into non-traditional activities. Smith et al (2003) states that diversification is justified for reducing the bank's exposure to risk. However, Demsetz and Strahan (1995) argue that diversification may not necessarily translate into risk reduction because these banks may shift into riskier activities, resulting in less equity. Hence, the variable is expected to exhibit either a positive or negative sign.

Net Profit (NETP)

It is generally acknowledged that highly profitable banks are more efficient in their operations and hence a positive sign is expected. However, a study by Ncube (2009) on the cost and profit efficiency of the South African banking sector indicated that South African banks were good at controlling costs than generating profits.

6 Data analysis and empirical findings

6.1 Sources of Data

The input and output data for the first stage analysis were collected from published annual reports of banks included in the sample for the period 2000 to 2010. Data that was used in the second stage analysis was obtained from two sources within the Quantec data base. These sources include SARB and Bureau for Economic Research (BER), collected for the period 2002 to 2010. The sample consists of the four largest South African commercial banks namely ABSA bank, FirstRand bank, Nedbank and Standard bank. The computer program *DPIN 3.*[2]developed by O'Donnell (2011) which uses the aggregate-quantity framework for computing and decomposing productivity index numbers was used to analyse the performance of the big four banks in South Africa. This computer program does not require price data which in most cases are difficult to obtain for bank inputs and outputs. It requires only data on inputs and outputs. Moreover it does not require any assumptions concerning the optimizing behaviour of firms. The program*DPIN version 3* enables the analysis of the drivers of productivity by decomposing changes in total factor productivity into measures of technical change, technical efficiency change, scale efficiency change and mix efficiency change (O'Donnell, 2011)

An essential consideration when evaluating efficiency and productivity within the banking sector is the choice of inputs and outputs. When defining the inputs and outputs to be adopted two basic approaches are followed: the production approach and the intermediation approach. The production approach regards banks as firms that use inputs such as labour and capital to produce outputs such as deposits, loans and advances. On the other hand, the intermediation approach recognises the intermediary role of banks as accepting deposits to produce the outputs which are loans and advances. In this paper, the intermediation approach is adopted. Labour, fixed assets, deposits and current accounts, are considered inputs while loans and advances are considered as output variables. These variables and their descriptive statistics are presented in Table 1 and 2 respectively.

6.2 Empirical Results:

6.2.1 First Stage DEA Results

While the number of banks in the sample is limited to only four, it is important to appreciate the fact that these big four banks are a fair[3] representation of the entire banking sector. As mentioned earlier, together the account for over 80 % of the total banking assets in South Africa. Hence the results of this paper can

[2]DPIN 3.0 is a computer program that uses data envelopment analysis (DEA) linear programs (LPs) to estimate the production technology and *levels* of productivity and efficiency (O'Donnell, 2011).

[3]These four largest South African commercial banks contributed 84.6 % to the balance-sheet size of the total banking sector in 2010 (Bank Supervision Department, 2010).

be interpreted as being representative of the total banking sector. Table 3 and 4 presents the results of the estimation. It is important to recall that efficiency scores ranges from zero to one. A score below unity represents an inefficient bank which is located below the production frontier while a score of one implies that the bank is fully efficient and lies on the frontier of the production technology. Regarding productivity, a value greater than one indicates positive TFP growth from period t to period $t+1$.

TFP efficiency (TFPE) is a measure of overall productive performance. Table 3 indicates that this measure was an average 81.8 % which means that for the period 2000-2010 banks fell short of 18.2 % to realise the maximum productivity that was possible using the available technology. This overall productive efficiency score of this study is a bit lower than that obtained in previous studies of bank efficiency in South Africa. As stated earlier, Okeahalam (2006) found an efficiency score of 83.1% for 61 bank branches in the nine provinces of South Africa. Ncube (2009) found an efficiency score of 85% for 8 South African banks classified as four large and four small.

Table 3 presents the estimates of *levels* of productivity and the various input and output oriented efficiency scores for all the banks for the entire period. An estimated TFP average score of 1.0361 for all banks indicated a positive growth of 3.61 % over the 11 year period of evaluation. An output-oriented technical efficiency (OTE) score of 98.20 implied that with the endowment of inputs that was available to the banking sector, each bank on average had the potential to expand its output by 1.8 %. Similarly the obtained input-oriented technical efficiency (ITE) estimate of 98.78 meant that the banks had the potential to decrease their inputs by 1.22 % without altering their output. In other words, this represented input wastage of 1.22 %. Table 3 results indicate a noticeable change in the scores for the period 2008 – 2009. This period coincides with the worst performance noted during the 2008-2009 when economies were at the height of the financial global crisis. The worst performance is evident across all the different efficiency and productivity scores especially for the period 2009. For example, TFP deteriorated from 106.88 in 2008 to 104.57 in 2009 before declining further to 96.90 in 2010.

Inefficiencies can also occur as a result of resources being employed in the wrong mix. For instance, input-oriented mix efficiency (IME) scores of 1 means that the combination (mix) of inputs being used is in the most efficient state for producing the current outputs level. Hence, the mean score obtained of IME of 98.84 represented an inefficiency of 1.16 % that was due to input resources being employed in the wrong mix.

The study also identified evidence of combined scale and mix inefficiencies arising from banks not operating on their most productive scale and input/output mix. For example the average input-oriented scale-mix efficiency (ISME) and output-oriented scale-mix efficiency (OSME) for the banking sector was 82.77 % and 83.24 % respectively. This means that banks exhibited inefficiency of 17.23 % and 16.76 % due to input utilisation and output production that was associated with both economies of scale and economies of scope respectively. Scale efficiency is achieved when a bank operates on the minimum

point of its average cost curve while economies of scope are achieved when the cost of jointly producing a range of outputs is less than the cost of producing them independently.

Table 4 reports indexes measuring *changes* in total factor productivity (ΔTFP), the technology (ΔTech = ΔTFP*) and various other types of efficiency for the entire banking sector for each period. There are two periods when the industry experienced a significant improvement in TFP growth (ΔTFP) of above 11 %: 2002/2003 (111.07) and 2005/2006 (111.86). This results is also consistent with Maheshwari's, (2009) finding that asset growth recorded in the South African banking industry for the period 2003-2006, was higher than that witnessed in the global banking industry. However, there are two periods when the banking industry experienced a significant deterioration of ΔTFP of below 100 %: the period 2008-2009 (88.75 %) when it reached its lowest and 2009-2010 (95.73 %) during slow recovery. It is important to bear in mind that even though ΔTFP improved from 88.75 % to 95.73 % for the period 2009-2010, the improvement still reflected the aftermath of the world-wide financial crisis.

A change in the maximum TFP (ΔTFP*) which captures a contraction or expansion of the production possibilities frontier improved significantly for the whole period. A technological change (ΔTFP*) estimate of 1.0622 for the whole period of evaluation indicated technological progress of 6.22% in the South African banking industry. This is most likely due to the technological advancements and innovations in the banking industry which followed after gaining independence in 1994. However these positive gains were slowed in 2008/2009 (0.9323) and 2009/2010 (0.9245) as can be seen in table 4 when the maximum TFP* deteriorated. Table 4 shows that the changes in ΔTFP* since 2002/2003 were above unity except for the period 2008/2009 and 2009/2010. Again, this coincided with the period of intensifying global financial crisis 2008/2009 and a slower recovery during 2009/2010. A graphical analysis of the above discussed performance measures are presented in figures 5 and 6. The rest of the findings are in table 3 and table 4.

6.2.2 Second Stage Censored Tobit Model Results

In the second stage, the relationship between TFP efficiency (TFPE) and bank specific factors was assessed using the censored Tobit regression model. A time period was included to reflect the global financial crisis period. In table 5 we presented the descriptive statistics of the determinant factors including the JB statistic for normality of the included series. This is in safe keeping with the requirements of the censored regression. In order to test for model adequacy, the residual and coefficient test were performed before interpreting the Tobit results. In each case, the null hypothesis of the Jarque-Bera (JB) test of normality in residuals and the Ljung-Box Q statistic of no autocorrelation was not rejected.

Of paramount value to the present study was to ascertain the significance and magnitude of the impact of the sub-prime financial crisis on the South African banking system. The results are presented in table 7. A significant and negative coefficient found implies that the financial crisis negatively influenced

performance within the banking sector. A coefficient of 0.1697 means that relative to the pre-crisis period, the average TFP efficiency of banks during the crisis period was 16.97% lower. It is worth mentioning that of the included determinant factors, the crisis dummy had the greatest impact on bank performance. Contrary to the conclusion of a mild deviation of efficiency measures in the first stage analysis, the second stage results appear to suggest that the crisis caused considerable deterioration in total factor productivity efficiency of banks. This is in line with the SARB supervision report (2009, pp.100) that the banking sector balance sheet size declined from R3177 billion at the end of December 2008 (135.4% of GDP) to R2967 billion (118.5% of GDP) at the end of 2009. This represented a negative year-on-year growth of 6.6% or an equivalent deterioration in the banking sector balance sheet size of 16.9% of GDP.

The rest of the second stage regressions are contained in table 7 as follows. All other bank specific factors were significant except the interest spread variable. Non-performing loans negatively and significantly affected bank's productive efficiency. This is acceptable given that an increase in non-performing loans retards the growth of bank assets. Diversification as represented by non-interest income variable had a negative impact indicating that diversification from traditional bank activities lowers TFP efficiency of banks in South Africa. This is also consistent with the finding of Sharma et al (2012) who found that bank diversification practices were negatively affecting the efficiency and performance of Indian banks. Profitability showed a positive and significant coefficient meaning that highly profitable banks are more efficient in their operations. However, the size of the incremental effect was marginal. The same thing applied for the bank size variable which though highly significant had a very trivial contribution effect. This indicates that larger banks tend to achieve higher efficiency. Ismail et al (2012) maintains that larger banks have more capital that can be used to adopt new technology that can enhance their profitability and minimizing their management cost. Cost to income ratio, a traditional proxy for cost efficiency is negatively associated with efficiency. This is expected given that deterioration in cost efficiency as indicated by an increase in cost relative to income decreases the overall productive efficiency.

7 Conclusion

In this paper we have evaluated the total factor productivity of the four largest South African banks for the period 2000 - 2010. The measurement and the decomposition of total factor productivity into several measures of efficiency was performed using the software program *DPIN 3* developed by O'Donnell (2011). To the best knowledge of the author, this is one of the pioneering works to analyse productivity of the South African banking system to adopt the Hicks-Moorsteen TFP index approach. The main motivation driving this research has been to determine the magnitude of the impact of the global financial crisis on the performance of the banking sector in South Africa. The major findings of this paper are as follows:

In the first stage analysis, most of the efficiency and productivity measures clearly deteriorated during the period 2008-2010, the period that coincided with the global financial crisis. Nevertheless, this paper found the deterioration in the efficiency and productivity measures during the crisis to be generally mild. This is consistent with the views of the Bank Supervision Annual Report (2009) that broadly described the financial sector as "remaining vigilant" despite the difficult circumstances that came with the crisis. This vigilance was attributed to bank's low leverage ratios, sound profitability, limited exposure to foreign assets and foreign funding and the fact that capital levels were adequate (SARB, 2009). Mabwe & Webb (2010) also states that for an economy that is integrated into the global financial system, South Africa weathered the global financial crisis well compared to other countries in the region.

Second stage analysis showed that the financial crisis was the main determinant of bank efficiency for the period reviewed, with total factor productivity efficiency being 16.96% lower during the crisis period compared to the pre-crisis period. Non-performing loans, bank size, cost to income ratio, profitability, non-interest income variables were found to be significant factors that influence bank performance. However, there is need for policy makers to be mindful of the risk of over-regulation because there is normally a danger of excessive regulation after a crisis which retard the ability of banks to operate normally. Mboweni (2009) recommended a fine balance to be achieved but emphasized for those elements of the financial system that were previously not included to be added to the regulatory net.

References

[1] BANKING ASSOCIATION SOUTH AFRICA., (2010). *Annual Review* 2010. Available at: http://www.banking.org.za/

[2] BERGER, A.N. & HUMPHREY, D.B., (1997). Efficiency of Financial Institutions: International Survey and Directions for Future Research, *European Journal of Operational Research*, 98, pp. 175-212.

[3] BJUREK, H., (1996). The Malmquist Total Factor Productivity Index. *Scandinavian Journal of Economics*, 98(2), pp. 303-313

[4] BRIEC, W. & KERSTENS, K., (2011). The Hicks–Moorsteen Productivity Index Satisfies The Determinateness Axiom, *School of Management*, The Manchester. 79(4), pp. 765 -775.

[5] COELLI, T., (1996). A Guide to DEAP version 2.1: A Data Envelopment Analysis (Computer) Program, *CEPA Working Paper* 96/07. Armidale: Department of Econometrics, University of New England.

[6] DEMSETZ, R. & STRAHAN, P., (1995). 'Diversification, size, and risk at bank holding companies', *Federal Reserve Bank of New York Economic Policy Review*, 1 (2), pp. 13-26.

[7] DIEWERT, W. E., (1992). Fisher Ideal Output, Input, and Productivity Indexes Revisited. *Journal of Productivity Analysis,* 3, pp. 211-248

[8] FARRELL, M., & FIELDHOUSE, M., (1962). Estimating Efficient Production Functions under Increasing Returns to Scale. *Journal of Royal Statistical Association*, 125(2), pp. 252-267.

[9] GANLEY, J. & CUBBIN, J., (1992). *Public Sector Efficiency Measurement: Applications of Data Envelopment Analysis*, New York: Elsevier Science Publishers.

[10] GREENBERG, J.B. & W. SIMBANEGAVI., (2009). Testing for Competition in the South African Banking Sector, *Faculty of Commerce University of Cape Town*. 4 November.

[11] HICKS, J. R., (1961). Measurement of Capital in Relation to the Measurement of Other Economic Aggregates. In: F. A. Lutz and D. C. Hague (eds.). *The Theory of Capital*. London, Macmillan.

[12] IKHIDE, S.I., (2008). Measuring the Operational Efficiency of Commercial Banks in Namibia, *South Africa Journal of Economics*, 76(4), pp. 586-595.

[13] IRŠOVÁ, Z. & HAVRÁNEK, T., (2010). Measuring Bank Efficiency: A Meta-Regression Analysis, *Prague Economic Papers*, 4, pp. 307-328.

[14] ISMAIL, F., RAHIM, R.A.B. & MAJID, M.S.A.B.D. (2011). *Determinant of Efficiency in Malaysian Banking Sector*, Universiti Malaysia Sarawak (UNIMAS), Small Grant Scheme (SGS) [03(S83)/819/2011(7)].

[15] MABWE, K. & WEBB, R., (2010). A Financial Ratio Analysis of Commercial Bank Performance in South Africa, *African Review of Economics and Finance*, 2(1), pp. 31-53.

[16] MAHESHWARI, S., (2009). *South African Banking Sector Analysis*. Global Press Release Distribution. 23 April 2009. Available at: http://www.rncos.com/Report/IM080.htm.

[17] MBOWENI, T.T., (2009). The Implications of the Global Crisis for South Africa and Policy Responses Address, Governor of the South African Reserve Bank, Road Freight Association Convention, Vanderbijlpark, 19 May 2009.

[18] MERCAN, M., REISMAN, A., YOALAN, R. AND EMEL, A.B., (2003). The Effect of Scale and Mode of Ownership on the Financial Performance of the Turkish Banking Sector: Results of a DEA-based Analysis. *Socio-Economic Planning Sciences* (37), pp.185-202.

[19] MLAMBO, K., & NCUBE, M., (2011). Competition and Efficiency in the Banking Sector in South Africa, *African Development Review*, 23, pp. 4–15.

[20] MNYANDE, M., (2012). *South Africa's macroeconomic outlook in the wake of the sovereign debt crisis and reforms to financial sector and regulation*, South African Reserve Bank, Keynote address at the 6^{th} Annual Risk and Return South Africa Conference, Cape Town, 1-2 March 2012, South Africa.

[21] MOORSTEEN, R. H., (1961). On Measuring Productive Potential and Relative Efficiency. *The Quarterly Journal of Economics,* 75(3) pp. 151-167

[22] NCUBE, M., (2009). *Efficiency of the Banking Sector in South Africa*, Fourth African Economic Conference on Fostering Development in an Era of Financial and Economic Crises, 11–13 November 2009, Addis Ababa, Ethiopia.

[23] OBERHOLZER, M., VAN DER WESTHUIZEN G. & S. VAN ROOYEN., (2010). The influence of banks' internal performance on market performance: a non-parametric approach, *Southern African Business Review*. 14 (2), pp. 65-88.

[24] O'DONNELL, C. J., (2008). An Aggregate Quantity-Price Framework for Measuring and Decomposing Productivity and Profitability Change. *Centre for Efficiency and Productivity Analysis Working Papers WP07/2008*, University of Queensland.

[25] O'DONNELL, C.J., (2010a). DPIN Version 1.0: A Program for Decomposing Productivity Index Numbers. *Centre for Efficiency and Productivity Analysis Working Papers WP01/2010*, University of Queensland.

[26] O'DONNELL, C. J., (2010b). Measuring and Decomposing Agricultural Productivity and Profitability Change. *Australian Journal of Agricultural and Resource Economics*, 54(4), 527-560.

[27] O'DONNELL, C. J., (2010c). Nonparametric Estimates of the Components of Productivity and Profitability Change in U.S. Agriculture. *Centre for Efficiency and Productivity Analysis Working Papers WP02/2010*, University of Queensland.

[28] O'DONNELL, C.J., (2011). *DPIN Version 3.0: A Program for Decomposing Productivity Index Numbers.* Centre for Efficiency and Productivity Analysis Working Papers, University of Queensland.

[29] OKEAHALAM, C.C., (2001). *Structure and Conduct in the Commercial Banking Sector of South Africa*. Presented at TIPS 2001 Annual Forum.

[30] OKEAHALAM, C.C., (2006). Production Efficiency in the South African Banking Sector: A Stochastic Analysis, *International Review of Applied Economics*, 20, pp. 103–23.

[31] SHARMA, A.K, SHARMA, D. & BARUA, M.K., (2012). *Efficiency and Productivity of Indian Banks: An Application of Data Envelopment Analysis and Tobit Regression*, National Conference on Emerging Challenges for Sustainable Business 2012, pp. 81-90.

[32] SMITH, R., STAIKOURAS, C. & WOOD, G., (2003). *Non-interest income and total income stability*, Working Paper no. 198, Bank of England.

[33] SOUTH AFRICAN RESERVE BANK, (2010). *Annual Economic Report 2010*, 90^{th} Anniversary. Available at: http://www.reservebank.co.za/PublicationsandNotices/Reports. [Accessed 12 July 2012].

[34] SOUTH AFRICAN RESERVE BANK, (2010). *Selected South African Banking Sector Trends*, Bank Supervision Department, December 2010. Available at: http://www.reservebank.co.za/ [Accessed 12 July 2012].

[35] SOUTH AFRICAN RESERVE BANK, (2011). *Annual Economic Report 2011*, 90^{th} *Anniversary.* Available at: http://www.reservebank.co.za/PublicationsandNotices/Reports. [Accessed 10 July 2012].

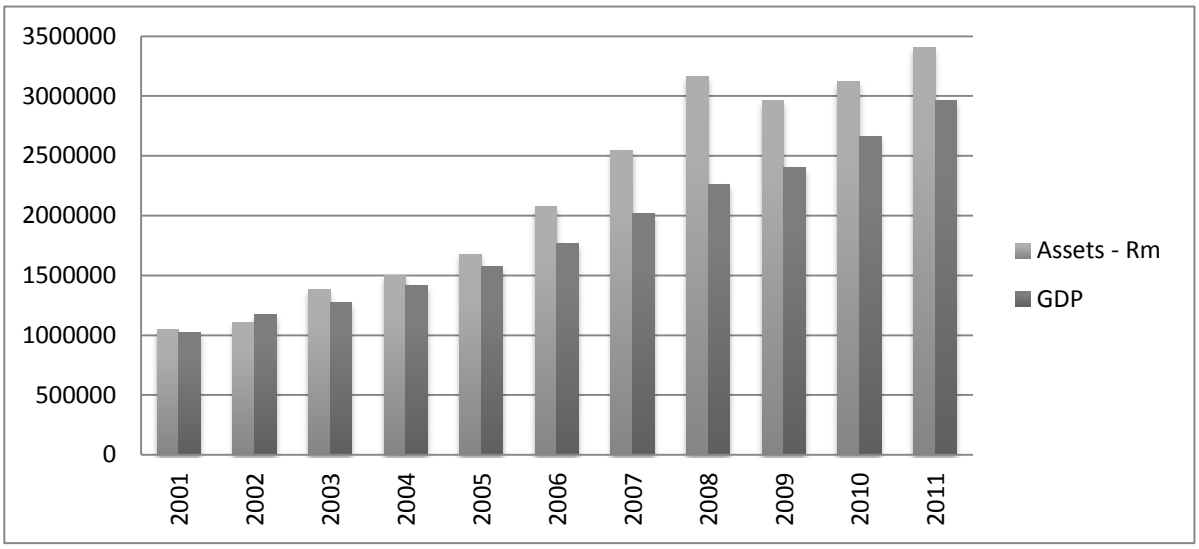
[36] SOUTH AFRICAN RESERVE BANK, (2012). *Selected South African Banking Sector Trends*, Bank Supervision Department. March 2012. Available at: http://www.reservebank.co.za/ [Accessed 10 July 2012].

[37] SUFIAN, F., (2009). The Impact of the Asian Financial Crisis on Bank Efficiency: The 1997 Experience of Malaysia and Thailand, *Journal of International Development,* 22, pp. 866-889.

[38] YEH, Q.J., (1996). The Application of Data Envelopment Analysis in Conjunction with Financial Ratios for Bank Performance Evaluation, *Journal of the Operational Research Society*, 47, pp. 980-988.

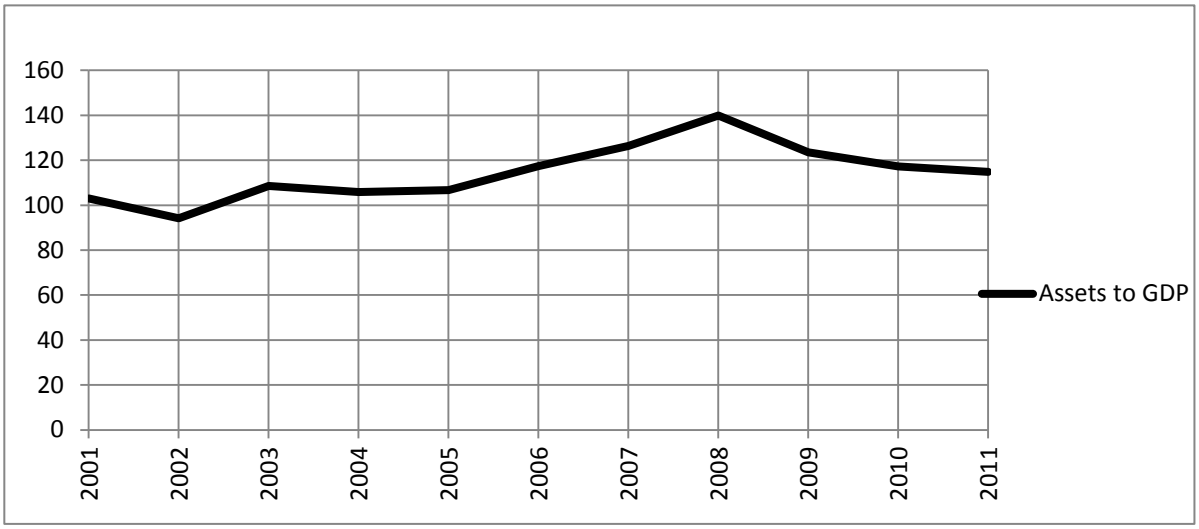
8. APPENDICES

Figure 1: Total Banking-sector Assets to Gross Domestic Product



Source: Computed by author using data from SARB

Figure 2: Total Banking-sector Assets to Gross Domestic Product Ratio



Source: Computed by author using data from SARB

Figure 3: Measuring and Decomposing TFP Change

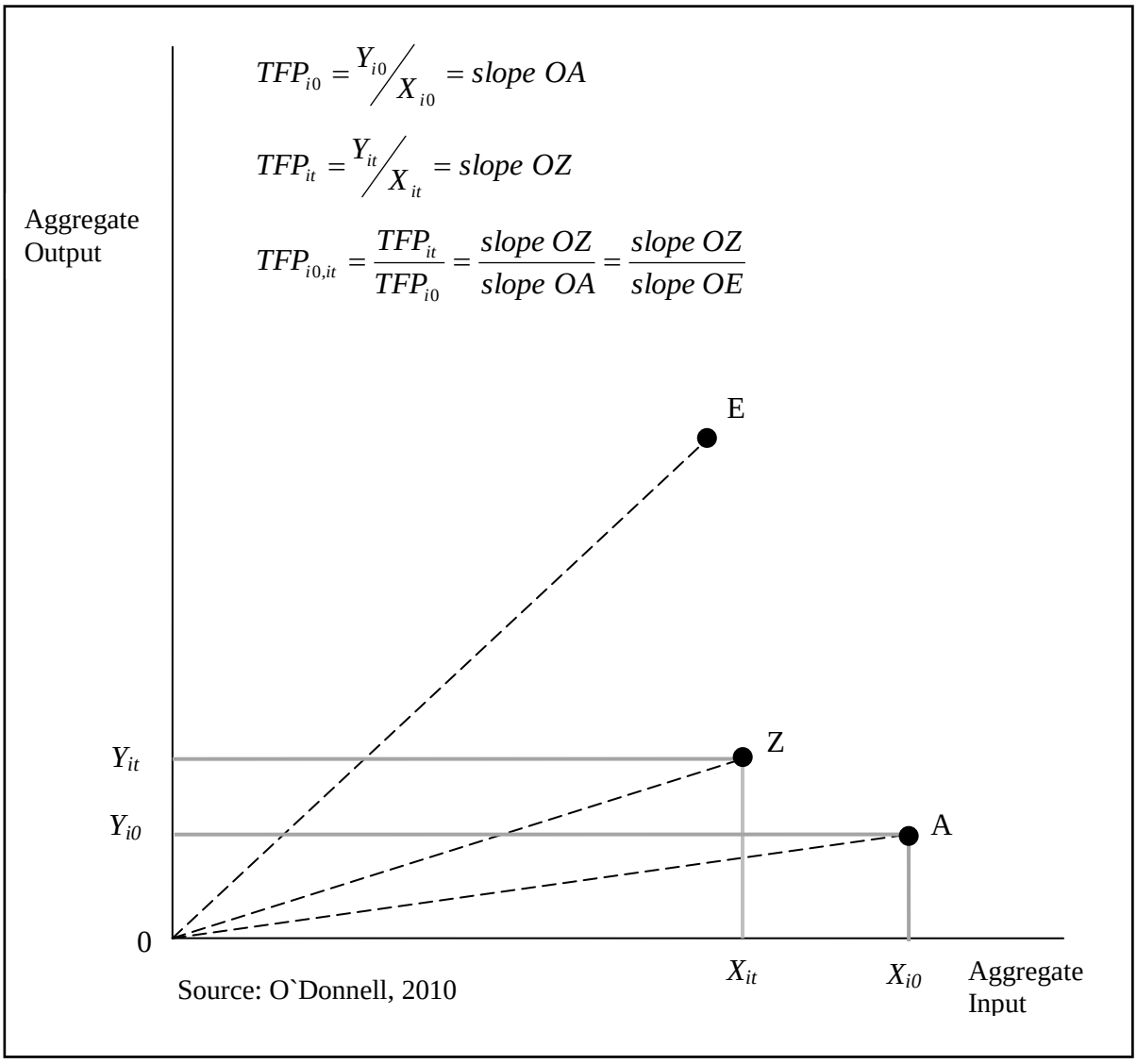


Figure 4: Two Input Oriented Decompositions of TFP Efficiency

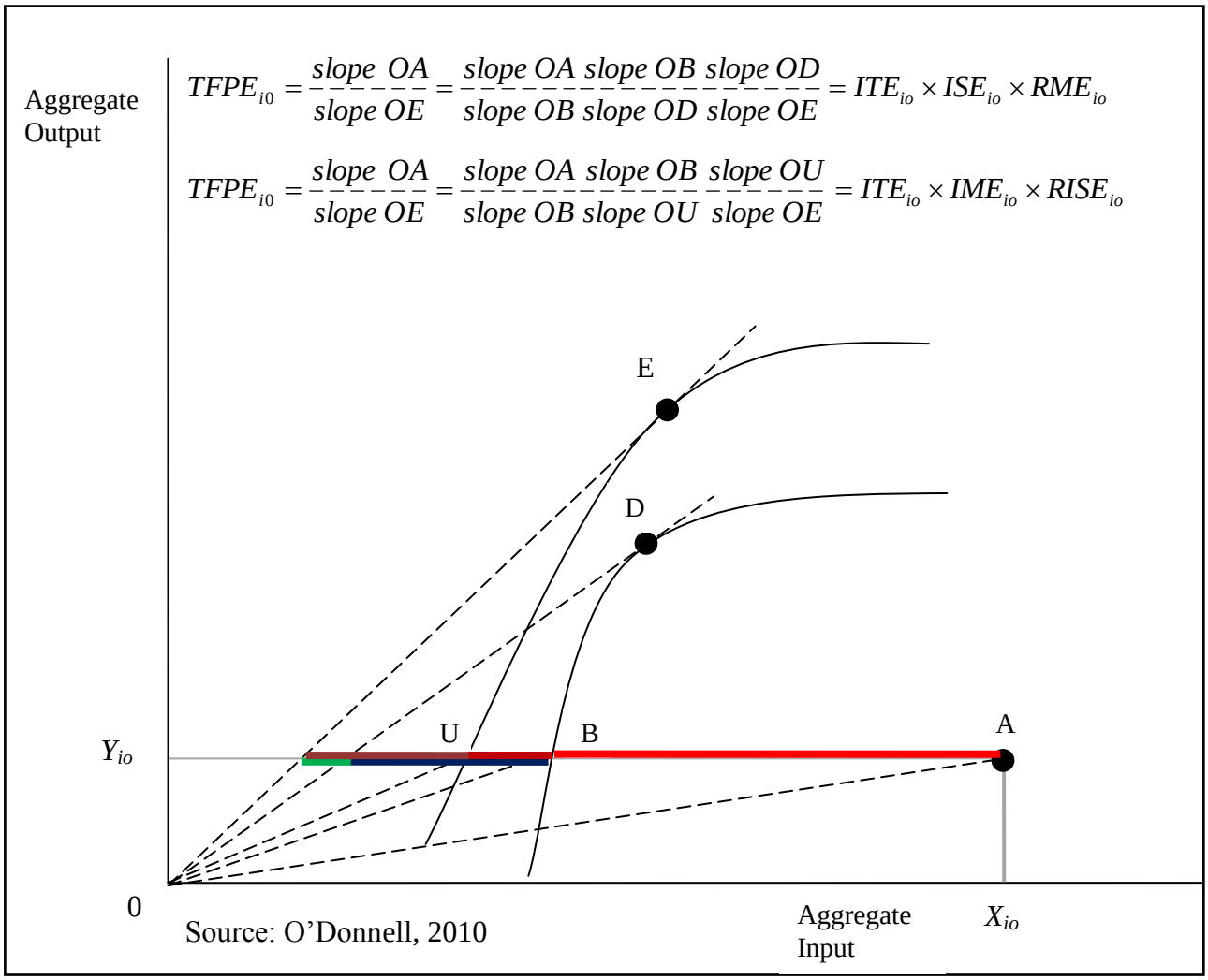


Figure 5: TFP, TFPE, OTE & ITE

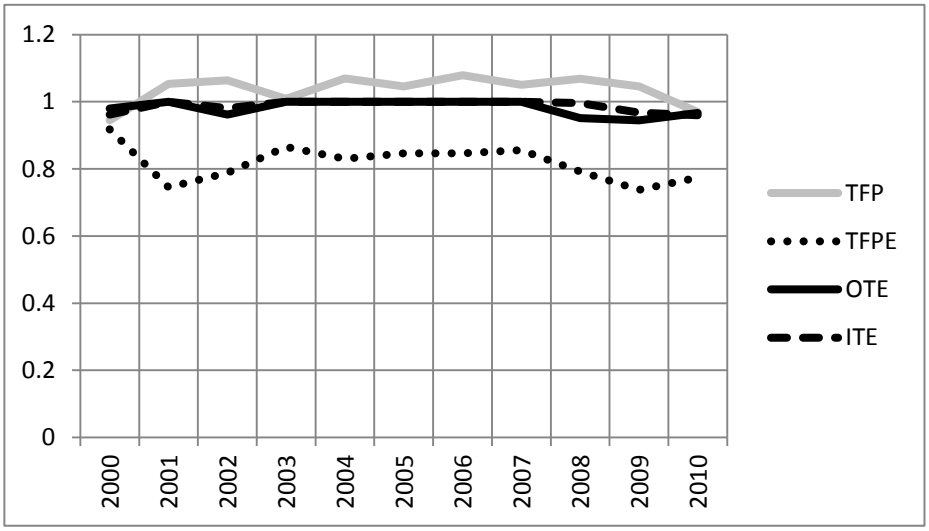


Figure 6: ΔTFP, ΔTFP* & ΔTFPE

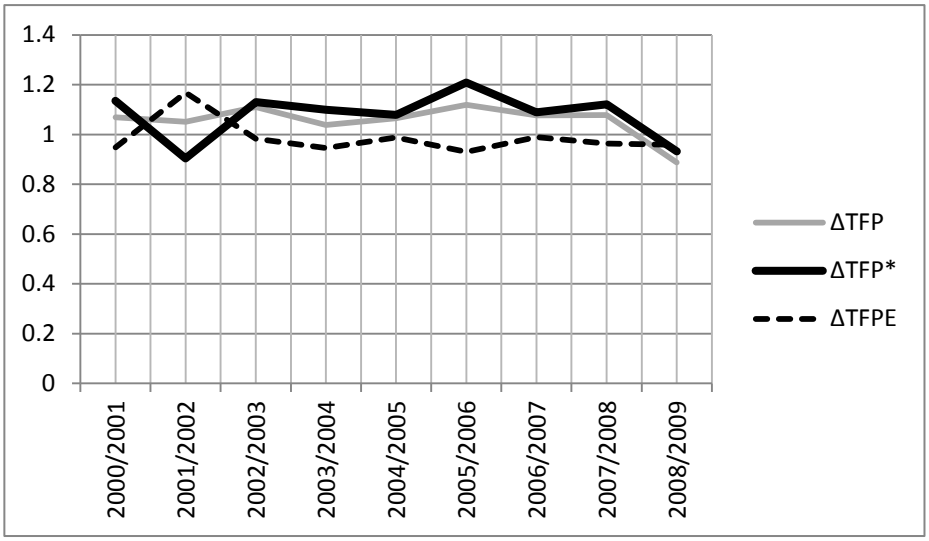


Figure 7: Actual Vs Fitted Values of TFPE: Tobit Model

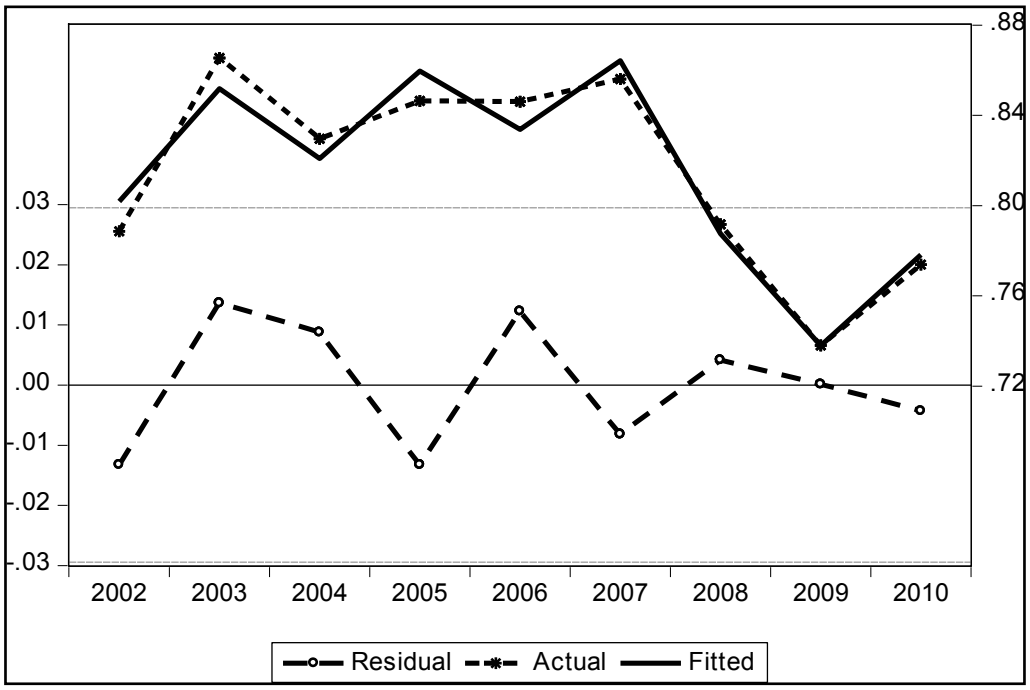


Table 1: Variables Used in Efficiency and Productivity Analysis.

VARIABLE	DEFINITION
OUTPUTS	
Y_1	Advances and Loans[1]
INPUTS	
X_1	Number of Full-Time Employees
X_2	Fixed Assets (Property and Equipment)
X_3	Deposits and Current Accounts

Table 2: Descriptive Statistics: First stage Variables

VARIABLE	MEAN	STD. DEV	MIN	MAX
Loans & Advances	294683	134920	102652	535994
Labour	33347	7177	18664	53351
Capital	3791	1616	1793	10018
Deposits & Current Accounts	312671	128920	117592	599500

Table 3: *Levels* of Productivity and Efficiency

YEAR	TFP	TFPE	OTE	ITE	OME	IME
2000	0.9441	0.9176	0.9794	0.9616	1.0000	1.0000
2001	1.0535	0.7449	1.0000	1.0000	1.0000	1.0000
2002	1.0634	0.7886	0.9614	0.9803	1.0000	0.9832
2003	1.0085	0.8655	1.0000	1.0000	1.0000	0.9939
2004	1.0698	0.8297	1.0000	1.0000	1.0000	0.9387
2005	1.0453	0.8464	1.0000	1.0000	1.0000	1.0000
2006	1.0794	0.8461	1.0000	1.0000	1.0000	1.0000
2007	1.0499	0.8561	1.0000	1.0000	1.0000	1.0000
2008	1.0688	0.7917	0.9511	0.9959	1.0000	0.9766
2009	1.0457	0.7381	0.9443	0.9676	1.0000	0.9962
2010	0.9690	0.7739	0.9662	0.9600	1.0000	0.9841
MEAN	**1.0361**	**0.8180**	**0.9820**	**0.9878**	**1.0000**	**0.9884**

[1] According to Banking Supervision Department SARB (2010), loans and advances represented on average, 74 per cent of banking sector total assets during 2010.

Table 4: *Changes* in Productivity and Efficiency

YEAR	ΔTFP	ΔTFP*	ΔTFPE	ΔOTE	ΔOSE	ΔITE	ΔISE
2000/2001	1.0690	1.1356	0.9482	1.0225	0.9607	1.0454	0.9446
2001/2002	1.0511	0.9037	1.1676	0.9614	1.0895	0.9803	1.0630
2002/2003	1.1107	1.1304	0.9832	1.0457	0.9892	1.0214	1.0106
2003/2004	1.0383	1.0992	0.9456	1.0000	1.0121	1.0000	1.0121
2004/2005	1.0656	1.0785	0.9876	1.0000	0.9928	1.0000	0.9928
2005/2006	1.1186	1.2088	0.9299	1.0000	0.9845	1.0000	0.9845
2006/2007	1.0772	1.0886	0.9898	1.0000	1.0204	1.0000	1.0204
2007/2008	1.0780	1.1209	0.9634	0.9511	0.9982	0.9959	0.9517
2008/2009	0.8875	0.9323	0.9583	0.9916	0.9853	0.9713	1.0062
2009/2010	0.9573	0.9245	1.0365	1.0281	0.9907	0.9912	1.0275
MEAN	**1.0453**	**1.0622**	**0.9910**	**1.0000**	**1.0023**	**1.0005**	**1.0013**

Table 5: Descriptive statistics: Second stage Variables

VARIABLE	TFPE	CRISIS	TA	SPREAD	NPL	NON_INT	NETP	CIR
Mean	0.815122	0.333333	2170424.	4.222222	48.73333	77.99167	175.2222	-104.4889
Maximum	0.865500	1.000000	3166502.	5.000000	280.0000	159.0000	386.0000	173.0000
Minimum	0.738100	0.000000	1102860.	3.000000	-200.0000	-31.00000	-199.0000	-299.0000
Std. Dev.	0.043662	0.500000	800877.6	0.833333	178.5115	64.66699	209.1583	150.9220
Jarque-Bera	0.762585	1.593750	0.938690	0.860249	0.690418	0.445787	1.556127	0.787837
Probability	0.682978	0.450735	0.625412	0.650428	0.708072	0.800200	0.459294	0.674409
(*) and (**) denote the null hypothesis of normality was rejected at 1% and 5% significance levels respectively. The J-B statistic was insignificant implying that all the series were normally distributed.								

Table 6: Diagnostic & Coefficient Tests

	H_O	Test Statistic	P-value	Conclusion
Diagnostic Test				
Jarque-Bera	Residuals are normally distributed	JB = 0.781	0.677	Errors are normally distributed
Ljung-Box Q	No 8^{th} order auto-correlation in residuals	LB_Q (8) =10.079	0.260	No 8^{th} order auto-correlation
Coefficient Test				
Redundant Likelihood Ratio	Independent variables all have zero coefficients	LR = 25.809	0.0005	No redundant variables

Table 7: Dependant Variable, Total Factor Productive Efficiency (TFPE)

Independent Variable	Coefficient	z-statistic	P-value
CRISIS	-0.169665	-3.293372	0.0010**
NPL	-0.000416	-3.366852	0.0008**
TA	2.98E-07	3.693039	0.0002**
CIR	-0.000274	-2.186807	0.0288*
NET_P	0.001723	3.502119	0.0005**
SPREAD	-0.072005	-1.419820	0.1557
NON_INT	-0.004143	-3.541242	0.0004**
CONSTANT	0.541535	2.103336	0.0354*
* and ** denotes significance at 0.05 and 0.01 level respectively			