

Decomposition of the technical efficiency of the banking system

**Authors:**

Sanderson Abel¹ 
Alex Bara¹ 
Pierre Le Roux¹

Affiliations:

¹Department of Economics,
Nelson Mandela
Metropolitan University,
South Africa

Corresponding author:

Alex Bara,
alexabara15@gmail.com

Dates:

Received: 11 May 2017
Accepted: 09 Oct. 2017
Published: 24 Apr. 2018

How to cite this article:

Abel, S., Bara, A. & Le Roux, P., 2018, 'Decomposition of the technical efficiency of the banking system', *Journal of Economic and Financial Sciences* 11(1), a160.
<http://dx.doi.org/10.4102/jef.v11i1.160>

Copyright:

© 2018. The Authors.
Licensee: AOSIS. This work
is licensed under the
Creative Commons
Attribution License.

The study investigated the technical efficiency of the commercial banks in Zimbabwe during the period 2009–2015. The study entailed the decomposition of the technical efficiency into pure technical and scale efficiency to understand the sources of the technical inefficiency in the commercial banks in Zimbabwe. To accomplish the task, the study sampled 11 commercial banks of which 6 are domestic and the other 5 are foreign banks. The study used the data envelopment analysis method. The results of the study revealed that commercial banks in Zimbabwe are technically inefficient with an efficiency score of 82.9%. The average pure technical and scale efficiency scores were 96.6% and 85.6%, respectively. The results imply that technical inefficiency of the Zimbabwean commercial banks is mainly a result of scale inefficiency emanating from decreasing returns to scale. The deduction is that commercial banks in Zimbabwe are operating at below their optimum capacity and hence have scope to increase their operations in order to improve on technical efficiency.

Introduction

Banks are vital institutions in any society as they significantly contribute to the development of an economy through facilitation of business. Banks facilitate the development of saving plans and are instruments of the government's monetary strategy. The global financial crisis in 2008 and 2009 impacted the banking system leading to declining banking sector profitability, levels of credit growth and quality of assets (Ganić 2012). This then forced governments to intervene through a raft of policy measures to restore financial sector stability. The measures included central bank liquidity support, state guarantees to financial institutions, targeted facilities and increased capital requirements (Edey 2009). The experience of the crisis called for a closer analysis of the performance of the banking sector.

Analysing bank performance is necessary to distinguish efficiency in resource utilisation among banks (Berger & Humphrey 1997). Analytical information is supposed to guide regulators in designing proactive processes for ensuring the stability of the system. The analysis of the banking sector performance also assists in assessing the effect of government policies such as deregulations, mergers and interest rate restrictions among others, and how they affect the economy while also helping banks to reduce wastage in resources, enhance competition and reduce market prices of financial products (Berger & Humphrey 1997). In competitive environments, performance measures are used to compare the performance of peers and evaluate the trend over time (Berger & Humphrey 1997).

One of the important performance measures is the efficiency with which resources are deployed by the banks. An analysis of the bank efficiency is used to evaluate the sources of banking profitability. An efficient bank is supposed to generate its profits through effective utilisation of resources rather than through exploitation of market power (Berger 1995). Banks that are efficient reduce wastage of resources and enhance competition (Berger & Humphrey 1997). A strong legal framework should be able to ensure managers are efficient in resource allocation as it encourages competition. Bank managers can improve cost efficiency by adopting better technologies or, alternatively, enhance capital through improving profit efficiency by adopting new marketing and pricing methods (Buchs & Mathisen 2005).

The current study seeks to measure the technical efficiency of the commercial banks in Zimbabwe during the period 2009–2015 using the non-parametric approach of data envelopment analysis (DEA). The study decomposes the technical efficiency of commercial banks into pure technical efficiency (PT) and scale efficiency so as to identify the sources of inefficiency in the banking system with emphasis on whether this is a result of managerial underperformance or the wrong

Read online:

Scan this QR
code with your
smart phone or
mobile device
to read online.

choice of scale. Besides, there is literature on bank efficiency for other countries and studies carried out produced diverging results (Kumar & Singh 2015; Li 2014; Marwa & Aziakpono 2015; Roy 2014; Singh & Fida 2015), but studies in the Zimbabwean context remain limited. This motivated the current study to contribute to the literature, drawing lessons from the Zimbabwean banking sector.

A notable study on the banking sector efficiency in Zimbabwe by Abel and Le Roux (2016) established that banks were both revenue and cost inefficient, but it did not discuss the issues of managerial and scale efficiency in the sector. It is important to further decompose the inefficiency into component factors technical or scale because inefficiency in the financial scale has been identified as one of the major determining factors of financial crisis in emerging markets. Disaggregating technical efficiency helps in understanding how banks are earning their profits, which consequently has an implication on financial sector stability.

The Zimbabwean banking case has experienced a number of developments in the banking and financial sector, beginning 2009, which affects the source and level of efficiency for banks, hence the study. The net interest rate margin, which approximates banking sector efficiency, has been increasing over time. Resultantly, players in the banking sector became more reliant on interest income as their source of revenue. This prompted intervention by the monetary authorities to control the lending rates in the banking sector. On the other hand, there have been nine bank failures since 2009, signifying an element of inefficiency in the system. The banking system has been characterised by deteriorating asset quality during the period 2009–2016, which could signal managerial inefficiency in the process of asset creation. Further inefficiency is brought about as banks try to recover the bad loans as they incur higher cost in trying to recover from the menacing non-performing loans (NPLs).

The rest of the study is organised as follows; the 'Background to the study' section discusses the background of the study followed by a review of theoretical and empirical literature in the 'Literature review' section. The study methodology is discussed in the 'Methodology' section, while the 'Presentation of results and analysis' section presents the study results and their discussion. The 'Conclusions' section concludes the study and proffers some recommendations.

Background to the study

After experiencing serious economic crisis during the period 2000–2008, which resulted in the domestic currency losing value, Zimbabwe officially abandoned its currency in favour of foreign currencies in February 2009 (GoZ 2009). The country adopted a basket of currencies that managed to arrest the hyperinflation and sustained economic activity in the economy. Following the adoption of the multicurrency system, the banking system encountered a number of positive developments, including increase in deposits, loans, assets,

profitability and capitalisation. Prior to the introduction of a multicurrency system, the banking sector experienced flight-to-quality situations where the banking public shifted their deposits to established international banks. The flight-to-quality experience is reflected by the high market shares of the top four banks, which reached 97% at the height of the hyperinflation in 2008. With the introduction of the multicurrency system, the market share of the top four banks significantly declined during the period 2009–2012. As a result of competition under a dollarised environment, in the first 2 years of dollarisation, banks embarked on aggressive lending in order to increase their market share. Banks started issuing medium- to long-term loans financed by short-term deposits and lines of credit as an incentive to attract business. With these incentives banks were able to attract increased numbers of clients, which led to a decline in the market share of the top four banks.

On the downside, the excessive lending by banks created NPLs, one of the other biggest challenges faced in the banking sector, in the dollarised environment. The NPLs increased from 1.8% in February 2009 to 20.1% in September 2014 as shown in Figure 1. The NPLs were attributed to the high cost of borrowing, weak credit risk management, absence of robust credit reference systems, insider loans, over-indebtedness and inappropriate loan structuring (Reserve Bank of Zimbabwe 2013). The weak credit risk management was attributed to the tough operating environment in the country and the absence of a credit-referencing system.

The growth in NPLs limited the capacity of banks to expand financial intermediation. High and rising levels of NPLs in Zimbabwe exerted strong pressure on bank balance sheets, with adverse effect on banks' lending operations, thereby limiting the capacity of banks to expand financial intermediation. The increasing amount of the NPLs led to the problem of disintermediation where the banks had to cut down on their lending and requested borrowers to pledge collateral even for small loans. NPLs affected bank performance, reducing the profitability of the sector through increased provisioning.

The banking sector was also faced by a number of other challenges, among them liquidity challenges, insolvency and bank failures. These challenges led to the collapse of some banking institutions as well as the consolidations and mergers in the sector that took place in an effort to meet the stringent regulatory capital requirements, during the period 2009–2016. Consequently, the number of operating banking institutions declined from 28 in December 2008 to below 20 by mid-2014. The country's central bank, however, argues that failure of the banks was mostly caused by poor corporate governance, insolvency and imprudent lending activities (Reserve Bank of Zimbabwe 2014).

Overall, the growth in the amount of NPLs had an impact on the performance of the banking sector, including its competitiveness. The NPLs also resulted in the collapse of

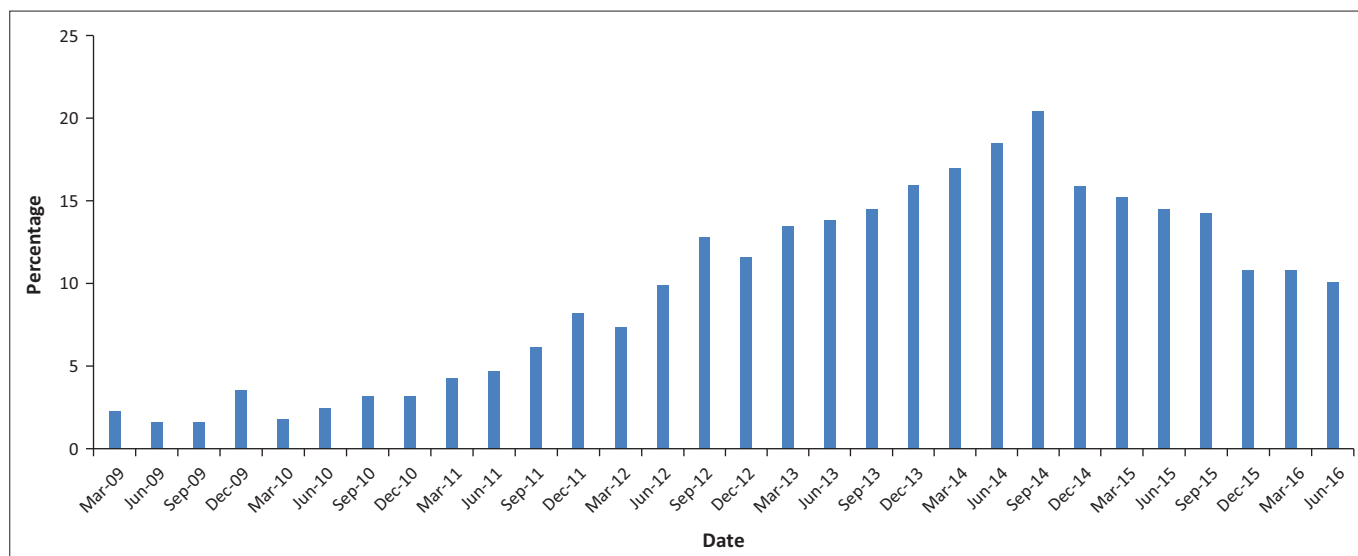


FIGURE 1: Non-performing loans.

some of the banking institutions as it led to insolvency and liquidity challenges.

The dollarised monetary system also created distortions in the interest rates in the banking sector. The developments have largely been the marked disparity between lending and deposit rates that has deterred both borrowers and savers. Lending rates that were being offered by the banks were very high, ranging between 6% and 25% with most banks quoting average lending rates of around 20%. The high lending rates have been a result of high premiums sustained by persistent liquidity shortages that characterised the multicurrency regime. During the same time, the deposit rates quoted by most banks for demand deposits ranged from 0.5% to 5%, while savings deposits ranged from 0.3% to 17% as of 31 December 2015. Given the centrality of the interest rate margin in the economy, this was assumed to be punitive to the productive sectors. The margin shows that the additional costs related to banks intermediary role of linking borrowers with the ultimate lenders. The high interest rate margins that prevailed under the multicurrency system were counterproductive as they acted as a disincentive to both savers and borrowers, culminating in financial disintermediation. This has undesirable repercussions on efforts geared at fostering savings culture as low returns deter depositors (Reserve Bank of Zimbabwe 2015). From an economic perspective, huge interest rate margins signify inefficiency in the banking system.

Given the continuous changes being experienced in the banking sector in Zimbabwe, there is need to understand how these changes have impacted on the efficiency of the sector. This is more so given the regulatory interventions that have seen the central bank dictating pricing in the sector, a development that affects the efficiency of individual banks. As such, it becomes apparent to try and decompose technical efficiency into PT and scale efficiency to ascertain sources of inefficiencies in the Zimbabwean banking sector.

Literature review

This section discusses both the theoretical and empirical literature on banking sector efficiency. The theoretical literature dwells on the discussion on efficiency and how it impacts the banking system. The empirical literature discusses a number of studies that have looked at the technical, pure technical and scale efficiency in the banking sector and the methodologies used. The section concludes by summarising what have been the main findings and lessons from these studies.

Theoretical literature

Efficiency measures how close a decision-making unit (DMU) gets to its production possibility frontier, composed of sets of points that optimally combine inputs in order to produce one unit of output (Kablan 2010). Alternatively, efficiency is defined as the ability of a firm to derive maximum output given a set of input levels under certain conditions (Coelli, Prassada Rao & Battese 1998). An efficient banking sector reflects a sound intermediation process and makes monetary policies effective. Understanding the primary transmission channel allows policymakers to obtain feedback on how changes in the regulatory environment affect bank efficiency and how efficiency translates into profitability of banks (Kablan 2010). The interest among policymakers and scholars in banking sector efficiency emanates from the role played by banks in ensuring unhindered flows in financial and real resources to where they can earn higher returns (Karimzadeh 2012). Efficiency of the banking system results in the reduction of spreads between lending and deposit rates, which stimulates greater demand for loans and the mobilisation of savings.

The concept of efficiency in banking is multifaceted and has been studied in different dimensions. Allocative efficiency is the extent to which resources are being allocated to the use with the highest expected value. A firm is technically efficient

if it produces a given set of outputs using the smallest possible amount of inputs. Alternatively, technical efficiency is the ability of the firm to maximise outputs from a given set of inputs and is associated with managerial decisions (Lovell 1993). The technical efficiency scores can be decomposed into pure technical and scale efficiency to determine the main source of the technical efficiency. Scale efficiency refers to the relationship between the level of output and the average cost; hence, it relates to the size of operation in the organisation.

Organisations can operate under three possibilities: constant returns to scale (CRS), increasing returns to scale and decreasing returns to scale. CRS attains when the relationship between input and output is constant. In this case output changes proportionately with an increase or decrease in inputs; hence, the organisation is scale efficient. An organisation is said to be experiencing increasing returns to scale if the output increases more than the increase in the inputs. This means that the organisation is suffering the problem of undersize and, hence, has scope to increase its size. Decreasing returns to scale attains when the increase in output is far less than the increase in the inputs. This means that the organisation is overly large above the optimal size. Both the phenomena of increasing returns to scale and decreasing returns to scale show that the organisation is operating outside the optimal scale, thus exhibiting scale inefficiency. The various forms in which efficiency has been studied show that the concept of efficiency is a multifaceted concept with several meanings depending from which perspective it is regarded (Leibenstein 1978).

Empirical literature

There are a number of studies that have been undertaken to estimate the technical efficiency of the banking sector (Bahrini 2017; Eric Fosu Oteng-Abayie 2017; Hacini & Dahou 2016; Jreisat & Hassan 2016; Kumar & Singh 2015; Marwa & Aziakpono 2015; Roy 2014; Singh & Fida 2015; Tandon, Tandon & Malhotra 2014; Tuo 2016; Yadav 2015).

The majority of these studies have been undertaken using the stochastic frontier approach of the DEA. The DEA has been used to decompose the technical efficiency into PT and scale efficiency. The studies have shown that the main source of technical inefficiency could be scale inefficiency or pure technical inefficiency. There is no uniformity among the studies.

Bahrini (2017) evaluated the technical efficiency of Islamic banks in the Middle East and North Africa region during the period 2007–2012 using the bootstrap DEA. The study managed to decompose technical efficiency into its component parts of PT and scale efficiency. The study established that technical inefficiency was mostly a result of pure technical inefficiency. The study further found that the Islamic banks enjoyed stable efficiency during the global financial crisis of 2007–2008 and post-economic crisis (2009–2010). Eric Fosu Oteng-Abayie (2017) estimated the level of

technical efficiency and total factor productivity in Ghana using Stochastic Frontier Analysis. The study established that there was scope for the banks to improve their technical efficiency and total factor productivity. The changes in technical efficiency were the main source of total factor productivity during the study period. Qamruzzaman and Jianguo (2016) studied the efficiency level of financial institutions in Bangladesh over the period 2011–2015 employing the DEA method. The study revealed that under the assumption of CRS, 62% of the banks were efficient while 75% were efficient under variable returns to scale (VRS). Overall, the whole industry showed that the industry was not technically efficient.

Tuo (2016) decomposed the technical efficiency of commercial banks from 2003 to 2014. The study revealed that from early 2003, technical efficiency and PT significantly improved. The scale efficiency also improved during the period, exhibiting a growing tendency. The study concluded that there was a need for the banks to reform and avoid blind expansion as this would not produce higher efficiency. Jreisat and Hassan (2016) studied the efficiency of the Egyptian banking system using the DEA. To ensure comprehensiveness, the study decomposed technical efficiency into PT and scale efficiency. The study established that medium banks were the most efficient while the least efficient banks were the large-sized banks. The study also showed that foreign banks were more efficient compared to domestic banks operating in Egypt. Yadav (2015) examined the technical efficiency of commercial banks in Russia using the data envelopment approach. The study further decomposed technical efficiency into PT and scale efficiency. The study found that technical inefficiency in the Russian commercial banks is as a result of poor input utilisation (i.e. managerial inefficiency) rather than scale efficiency.

Kumar and Singh (2015) evaluated the technical and scale efficiency of commercial banks in India for the period 2006–2010. The study employed the Charnes, Coopers and Rhodes 1978 (CCR model) and BCC model (Banker, Charnes & Coopers 1984). The study established that there was an increase in efficiency as a result of deregulation of the banking sector. The increase in efficiency was a result of an increase in both PT and scale efficiency. The study established that there was a great disparity in the technical efficiency between the banks during the study period. Private sector banks fared better than public sector banks with scale inefficiency being predominant over PT. Marwa and Aziakpono (2015) evaluated the technical and scale efficiency of the Tanzanian Saving and Credit Cooperatives. The study sample was composed of 103 Saving and Credit Cooperatives for the year 2011. The technical, pure technical and scale efficiency scores were found to be 42%, 52% and 76%, respectively. The sources of inefficiencies were both technical and scale; hence, the study recommended that the smaller firms should increase their scale of operations while those firms operating beyond the optimal scale may need to downsize. There was need for a reduction in the wastage in the utilisation of resources by

the banks. Singh and Fida (2015) using the DEA evaluated the degree of technical, pure technical and scale efficiencies in commercial banks of Oman. The study identified that scale inefficiency was higher than pure technical inefficiency in the total technical efficiency of the Oman banking sector. The scale inefficiency was mostly attributed to decreasing return to scale. The profitability of the banking system and liquidity were found to be the main drivers of efficiency. Hacini and Dahou (2016) studied the technical, pure technical and scale efficiency of the Algerian banks for the period 2000–2012. The study found that the technical efficiency of the banks improved during the study period. The average technical efficiency over the period was estimated at 95%. It was established that the main source of technical inefficiency was the scale inefficiency. Most of the banks were operating either under CRS or decreasing returning.

Tandon et al. (2014) also investigated the technical, pure technical and scale efficiencies of the Indian banks with different ownership structure for the period 2009–2012. The study established that of the 44 banks in the sample, only 7 lie of the efficiency frontier. There was no significant difference that was observed among public, private and foreign banks. The technical efficiency of the private and public banks was almost similar. There was scope for increasing the scale efficiency of the foreign banks. The main determinants of efficiency were found to be non-interest income (NNI). Roy (2014) analysed the efficiency of the Indian banking system during the Basel changes using the DEA. The study identified that inefficiency in the banking system was caused mostly by improper size allocation. Li (2014) studied the technical and scale efficiencies of Thai commercial banks using the DEA during the post-financial crisis period from 1997 to 2006. The average technical efficiency was estimated at 90.73%. The average PT was estimated to be higher than the average scale efficiency. The result implies that the technical inefficiency was mostly a result of scale inefficiency instead of PT. The major source of scale inefficiency was found to be decreasing returns to scale. The study also established that banks that were majorly owned by foreigners had lower technical, pure technical and scale efficiencies. Smaller banks were found to have higher technical and scale efficiencies. Similar technical, pure technical and scale efficiencies were identified between private- and government-owned banks.

Sok-Gee (2011) estimated the technical, pure technical and scale efficiency of commercial banks in China for the period 2001–2007 using the DEA. The commercial banks in China were found to be on average technically inefficient. The technical inefficiency of the commercial bank was driven mostly by pure technical inefficiency, implying that banks were facing the challenge of resource allocation to strike a balance between input and output mix. Foreign banks exhibited a significant decline in technical efficiency between 2003 and 2006, which was attributed to increased cost of production during their expansion process with the country's entrance into the World Trade Organization and the gradual liberalisation of the banking system. Gulati (2011), using the

DEA, estimated the extent of technical, pure technical and scale efficiencies of Indian domestic banking during the period 2006–2007. The study found that there were 9 banks in the sample of 51 banks that were operating on the efficiency frontier. Inefficiency in the Indian banking sector was mostly a result of managerial inefficiency rather than scale inefficiency. Most of the efficient banks were the new private sector banks. A significant scale efficiency difference was ascertained between large and medium-sized banks. The study also found that profitability and exposure to off-balance sheet activities were the main determinants of technical efficiency. Mongid and Tahir (2010) studied the rural Indonesian banking system with a view to estimating the technical and scale efficiency using the DEA for the period 2006 and 2007. The study established that the degree of scale efficiency was higher than the technical efficiency. The result meant that the greater proportion of inefficiency was mostly a result of producing below the efficient frontier instead of producing at an inefficient scale. The majority of the banks were found to be operating at suboptimal scale, calling for expansion to reach the optimal scale.

Tahir, Bakar and Sudin Haron (2009) evaluated the technical and scale efficiency of commercial banks in Malaysia during the period 2000–2006 using the DEA. The study identified that the degree of scale efficiency was lower than the technical efficiency. The result implies that the greater part of the technical inefficiency was attributed to producing below the production frontier. Foreign banks were found to be less efficient as compared to domestic banks. Inefficiency of foreign banks was attributed to scale inefficiency while inefficiency of the domestic banks is a result of pure technical inefficiency as a result of producing below the production frontier. Kumar and Gulati (2008) evaluated the technical, pure technical and scale efficiencies of Indian public sector banks. The study established that the overall technical efficiency of the public sector banks was 88.5%, and there were only seven technically efficient banks in the sample. The study also found that public sector banks were more pure and technically inefficient as compared to scale inefficiency; hence, it implies that the public sector banks suffered more from managerial inefficiency as compared to the scale efficiency. The predominant form of scale inefficiency identified in the study is the decreasing returns to scale. The results from the regression show that the banks' exposure to off-balance sheet activities had a positive effect on technical efficiency. Al-Muharrami (2008) investigated the technical, pure technical and scale efficiency for Gulf Cooperation Council banks for the period 1993–2002. Smaller banks were found to have overall technical efficiency than bigger banks. Big banks were more successful in adopting best technology while medium-sized banks had successes in adopting optimal levels of output while Islamic banks were successful in both technology adoption and choosing optimal levels.

A review of the literature has shown that banking systems are susceptible to inefficiencies as a result of either operating at the wrong scale or because of managerial inefficiencies.

Depending on the peculiar situation in the country of study, there have been some differences in the efficiency of the banking system. Neither scale inefficiency nor pure technical inefficiency is predominant in these studies.

Methodology

The DEA is one of the prominent methods used to measure the efficiency of the commercial banks. The method is derived from the concept of Pareto efficiency and was initially introduced by Charnes, Cooper and Rhodes (1978) in order to measure relative efficiency. The method estimates efficiency using data on inputs and outputs of some DMUs. The method identifies relative efficient DMUs, which are used as reference points. The relative efficient points are then used to define the efficiency frontier and evaluate the inefficiency of other DMUs that lie below that frontier (Casu & Molyneux 2000; Noulas 2001). The efficient DMUs are only efficient in relation to others in the sample. The model is based on linear programming techniques that allow calculating the relative efficiency of DMU (Hassan & Sanchez 2007). Farrell (1957) proposed the use of relative efficiency, which involves multiple inputs and outputs though not necessarily in equal proportion. The major goal of the relative efficiency technique is to develop a frontier of the most efficient decision units (DMUs) and then ascertain the distance of the less efficient units from the frontier (Bader 2008). The current study seeks to decompose the technical efficiency (T) into PT and scale efficiency (S). Following Chan and Karim (2011), the model for measuring the technical efficiency is shown in equation (1).

$$\theta^* = \min \theta$$

Subject to

$$\sum_{j=1}^n \lambda_j x_{ij} \leq \theta x_{io} \quad i = 1, 2, \dots, m \quad [\text{Eqn 1}]$$

$$\sum_{j=1}^n \lambda_j y_{rj} \leq y_{ro} \quad i = 1, 2, \dots, m$$

$$\sum_{j=1}^n \lambda_j = 1$$

$$\lambda_j \geq 0$$

The model has n DMUs being investigated. x and y represent the i th input and the r th output for the DMU_o, respectively. λ are unknown weights and $j = 1, 2, \dots, m$ represents the number of DMUs. The optimal value of θ^* shows the distance of the banks from the efficient frontiers. This means that the most technically efficient banks have $\theta^* = 1$ and the inefficient banks will have $\theta^* < 1$.

Efficiency can be estimated under the assumption of VRS or CRS. The choice between CRS and VRS affects the shape of the envelopment surface and resultantly on the number of efficient DMUs. CRS attains proportional increase if all inputs lead to a proportional increase in output. Models that apply the CRS are called CCR (Charnes, Cooper and Rhodes) models. The CCR model develops the Farrells' efficiency measurement

concept from several inputs and one output to several inputs and several outputs. In this model (Charnes et al. 1978) using a linear combination, different inputs and outputs are changed into one virtual input and output. The ratio of these virtual combinations of outputs to inputs will be the estimations of efficiency boundary for the measurement of relative efficiency, given that the yield is constant (Karimzadeh 2012).

The CRS model is more restrictive and yields fewer numbers of efficient units and lower efficiency scores compared to the alternative VRS (Karimzadeh 2012). VRS can be decreasing or increasing. Increasing returns to scale entail a proportional increase in all factor inputs of production, leading to a more proportional increase in output while the converse is true for decreasing returns to scale where a proportional decrease in factor inputs leads to a less-than-proportional decrease in output (Titko & Jureviciene 2014). Models applying the VRS are called the BCC (Banker et al. 1984) model. In contrast to constant yield in the CRR model, the BCC model assumes a variable output in respect of the scale. Tahir et al. (2009) argue that a firm that is efficient under VRS is considered technologically efficient; the VRS score represents PT, whereas a firm which is efficient under CRS is technologically efficient and also uses the most efficient scale of operation. Scale efficiency (S) is derived from the measures of technical efficiency (T) and PT.

$$S = \frac{PT}{T} \quad [\text{Eqn 2}]$$

Or

$$S = \frac{CRS}{VRS} \quad [\text{Eqn 3}]$$

where $0 \leq S \leq 1$ because $CR \leq VR$. If the value of S equals 1, the firm is scale efficient and all values less than 1 reflect scale inefficiency.

The definition of the inputs and outputs adopted by the study follows the intermediation approach, which views banks as intermediaries that facilitate the transfer of funds from surplus agents to deficit agents rather than producers of loans and deposit account services. Three inputs, capital (K), interest expense (IE) and non-interest expenses (NIE), are considered for the evaluation of the bank's performance. Two outputs considered for the study are total loans (TL) and NNI. These outputs represent bank revenue and the major income-generating business activities (Liu et al. 2010). The empirical model for this study becomes

$$\theta^* = \min \theta$$

Subject to

$$\lambda_j K_j + \lambda_j IE_j + \lambda_j NIE_j \leq \theta x_{io} \quad i = 1, 2, \dots, 3$$

$$TL_j + \lambda_j NNI_j \leq y_{ro} \quad i = 1, 2, \dots, m$$

$$\sum_{j=1}^n \lambda_j = 1$$

$$\lambda_j \geq 0$$

where K is capital, IE is interest expense and NIE is non-interest expense, TL is total loan and NNI is non-interest income.

The study sample involved 11 commercial banks in the Zimbabwean banking system. Of the chosen banks are five foreign banks and six domestic banks. The study relied on published annual financial statements composed of balance sheet and income statements for the period 2009–2015 for the bank-specific data. The data are therefore balanced annual panel data sets.

Presentation of results and analysis

The technical efficiency of the Zimbabwean commercial banks was estimated using the DEA and was decomposed into pure technical and scale efficiency. The summary statistics of the DEA is presented in Table 1.

The results in Table 1 show that commercial banks in Zimbabwe were technically inefficient with the average efficiency score of 82.9% during the period 2009–2015. This result implies that the average commercial bank suffered a 17.1% level of technical inefficiency. In other words, there was increased scope for the commercial banks to increase their output if they had operated at the same efficient level as the most efficient bank in the sample. The standard deviation, of 0.2025, shows that there was greater dispersion in terms of technical efficiency among the commercial banks during the study period.

The results also indicate that in Zimbabwe, on average, banks are relatively better off in PT (at 96.59%) compared to scale efficiency (at 85.64%). Table 2 shows the decomposition of the average annual technical efficiency scores into pure technical and scale efficiency scores for the period 2009–2015. The results are different from the studies by Singh and Fida (2015) and Hacini and Dahou (2016), who identified that scale

efficiency was higher than that of PT in Oman. The results are also at variance with Mongid and Tahir (2010), who established that the degree of scale efficiency was higher than the technical efficiency in the rural Indonesian banking system.

Table 2 shows that the average PT score was 96.6% for the period 2009–2015. The PT score was increasing over the period 2009–2015, implying that the managerial efficiency of the banks was improving during that period. The average PT was 1 in 2014, meaning that all the commercial banks attained pure efficiency of 1 during that year. All banks in the sample were experiencing managerial efficiency in 2014. A decline in the PT was experienced in 2015. Borrowing from Table 1 reveals that there was little dispersion in the PT score with standard deviation of 0.07 during the whole study period. As shown in Table 1, the average PT varied between 0.67 and 1 during the whole study period.

The results in Table 2 show that the average scale efficiency of the commercial banks in Zimbabwe for the period 2009–2015 was 0.8564. The scale efficiency scores were not stable, as reflected by the declines experienced between 2009 and 2011, followed by an increase up to 2014 before further declining in 2015. Based on the descriptive statistics in Table 1, the dispersion of the scale efficiency was relatively higher compared to the dispersion in PT. The minimum technical efficiency score was 0.2560 while the maximum was 1. The standard deviation, a measure of dispersion for the scale efficiency, was 0.189. This does not compare favourably with that of PT, which was 0.07 during the period 2009–2015.

Based on the results in Tables 1 and 2, the source of the technical inefficiency of the Zimbabwean commercial banks is the scale inefficiency instead of PT. The results are in conformity with studies by Li (2014) for the Thai banking sector and Fida (2015) for Oman. The results are also in line with Tahir et al. (2009) who identified that the degree of scale efficiency was lower than the technical efficiency in Malaysian commercial banks. The results are, however, in contrast with Sok-Gee (2011), who found that technical inefficiency of the commercial bank in China for the period 2001–2007 was driven mostly by pure technical inefficiency implying that banks were facing the challenge of resource allocation to strike a balance between input and output mix.

The results mean that the banks are mostly suffering from the problem of operating at the wrong scale of operations. The study result could be explained by the developments in the banking sector, wherein most banks were highly constrained by limited capital and liquidity constraints for them to operate and optimal level of scale efficiency. Furthermore, given the banking sector challenges, mostly cash and subsequent growth on demand for electronic banking services in 2014–2015, most banks were failing to meet demand for banking services and this pushed

TABLE 1: Descriptive statistics of data envelopment analysis measures.

| Variable | Technical efficiency | Pure technical efficiency | Scale efficiency |
|--------------------|----------------------|---------------------------|------------------|
| Mean | 0.8296 | 0.9659 | 0.8564 |
| Median | 0.8939 | 1.0000 | 0.9283 |
| Maximum | 1.0000 | 1.0000 | 1.0000 |
| Minimum | 0.2325 | 0.6731 | 0.2560 |
| Standard deviation | 0.2025 | 0.0737 | 0.1889 |
| Observations | 77 | 77 | 77 |

TABLE 2: Average technical, pure technical and scale efficiency.

| Year | Technical efficiency | Pure technical efficiency | Scale efficiency |
|------------------------|----------------------|---------------------------|------------------|
| 2009 | 0.7646 | 0.9090 | 0.8394 |
| 2010 | 0.7420 | 0.9394 | 0.7858 |
| 2011 | 0.7682 | 0.9800 | 0.7844 |
| 2012 | 0.8668 | 0.9779 | 0.8841 |
| 2013 | 0.8661 | 0.9735 | 0.8853 |
| 2014 | 0.9218 | 1.0000 | 0.9218 |
| 2015 | 0.8659 | 0.9830 | 0.8804 |
| Average for the period | 0.8296 | 0.9659 | 0.8564 |

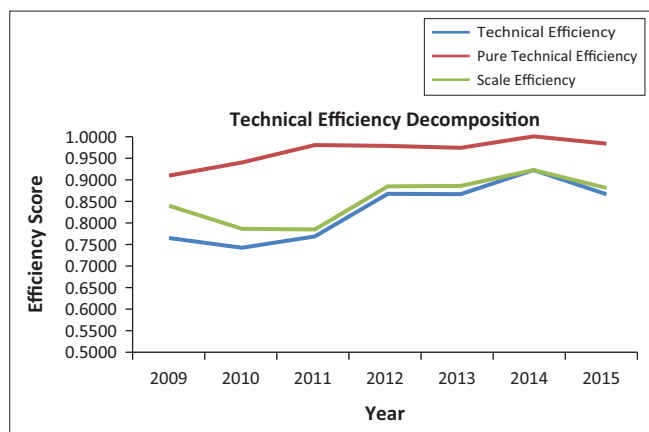


FIGURE 2: Trend of the average technical, pure technical and scale efficiency.

TABLE 3: Sources of scale inefficiency.

| Year | Number of banks with increasing returns to scale | Number of banks with decreasing returns to scale | Number of banks with constant returns to scale | Total number of banks |
|------|--|--|--|-----------------------|
| 2009 | 8 | 0 | 3 | 11 |
| 2010 | 8 | 0 | 3 | 11 |
| 2011 | 8 | 0 | 3 | 11 |
| 2012 | 7 | 0 | 4 | 11 |
| 2013 | 9 | 0 | 2 | 11 |
| 2014 | 5 | 0 | 6 | 11 |
| 2015 | 6 | 0 | 5 | 11 |

their returns, causing most of them to be operating at below-scale efficiency levels. The cautious trading and aggressive collection of bad loans implemented by most banks in response to growth in NPLs as well as interventions by the Central Bank to manage the NPLs also contributed to the slowdown in scale efficiency by most banks, particularly from 2014 to 2015.

Figure 2 shows the trend in the technical, pure technical and scale efficiency for the period 2009–2015. The figure reveals that the PT was higher than the scale efficiency, giving credence to the fact that the main source of the technical inefficiency during the whole period was scale inefficiency rather than pure technical inefficiency. Because scale efficiency refers to the relationship between the level of output and the average cost; hence, it means banks are facing the challenge of determining the optimal size of operation. Table 3 shows the sources of the scale inefficiency identified above.

Table 3 shows that the majority of the banks were operating under increasing returns to scale. The result shows that eight banks were operating under increasing returns during the period 2009–2011 and the highest operating under the same were registered in 2013 with nine banks. During the whole study period, there was no bank which operated under decreasing returns to scale. This means that the majority of banks were operating below their optimum capacity, implying that they had scope to increase their operations.

Increasing returns to scale output increases more than the increase in the inputs.

Conclusion

The study evaluated the technical efficiency of the commercial banks in Zimbabwe using the method of DEA. The study further decomposed the technical efficiency into pure technical and scale efficiency in order to determine whether the inefficiencies among commercial banks were a result of managerial inefficiencies or because of wrong choice of scale of operation. The study has shown that managerial efficiency scores were higher than technical efficiency scores, implying that commercial banks in Zimbabwe are technically inefficient. The technical inefficiency is a result of scale inefficiency, that is, the majority of banks were operating at the wrong scale of operations. Specifically, the banks were operating under decreasing returns to scale, where there is still opportunity to increase operations to obtain optimum scale. Therefore, the study recommends that banks should review and rescale their scope of operations so that they optimise the scale of operations to levels that guarantee both pure technical and scale efficiency.

Acknowledgements

Competing interests

The authors declare that they have no financial or personal relationships which may have inappropriately influenced them in writing this article.

Authors' contributions

All authors equally contributed to the research and writing of this article.

References

- Abel, S. & Le Roux, P., 2016, 'An evaluation of the nexus between banking competition and efficiency in Zimbabwe', *Journal of Studies in Economics and Econometrics* 40(3), 1–19.
- Al-Muharrami, S., 2008, 'An examination of technical, pure technical and scale efficiencies in GCC banking', *American Journal of Finance and Accounting* 1(2), 152–166.
- Bader, M.K.I., 2008, 'Cost, revenue and profit efficiency of Islamic versus conventional banks: International evidence using data envelopment analysis', *Islamic Economic Studies* 15(2), 23–76.
- Bahrini, R., 2017, 'Efficiency analysis of Islamic Banks in the Middle East and North Africa Region: A Bootstrap DEA approach', *International Journal of Financial studies* 5(7), 1–13. <https://doi.org/10.3390/ijfs5010007>
- Banker, R.D., Charnes, A. & Cooper, W.W., 1984, 'Some models for the estimation of technical and scale inefficiencies in Data Envelopment Analysis', *Management Science* 30, 1078–1092. <https://doi.org/10.1287/mnsc.30.9.1078>
- Berger, A.N., 1995, 'The profit-structure relationship in banking – Tests of market-power and efficient-structure hypotheses', *Journal of Money, Credit and Banking* 27, 404–431. <https://doi.org/10.2307/2077876>
- Berger, A.N. & Humphrey, D.B., 1997, 'Efficiency of financial institutions: International survey and directions for further research', *European Journal of Operational Research* 98, 175–212. [https://doi.org/10.1016/S0377-2217\(96\)00342-6](https://doi.org/10.1016/S0377-2217(96)00342-6)
- Buchs, T. & Mathisen, J., 2005, *Competition and efficiency in banking: Behavioral evidence from Ghana*, IMF Working Paper, No. 05/17, IMF, Washington, DC.
- Casu, B. & Molyneux, P., 2000, *A comparative study of efficiency in European Banking*, Financial Institutions Centre Working Paper Series, (February), pp. 1–17, PA.
- Chan, S. & Karim, M.Z.A., 2011, 'Efficiency of foreign banks: Evidence from selected ASEAN countries', *African Journal of Business Management* 5(14), 5617–5626.
- Charnes, A., Cooper, W.W. & Rhodes, E., 1978, 'Measuring efficiency of decision making units', *European Journal of Operational Research* 2, 429–444. [https://doi.org/10.1016/0377-2217\(78\)90138-8](https://doi.org/10.1016/0377-2217(78)90138-8)

- Coelli, T., Prassada Rao, D. & Battese, G., 1998, *An introduction to efficiency and productivity analysis*, Kluwer Academic Publishers, The Netherlands.
- Edey, M., 2009, 'The global financial crisis and its effects', *Economic Papers: A Journal of Applied Economics and Policy* 28(3), 186–195. <https://doi.org/10.1111/j.1759-3441.2009.00032.x>
- Farrell, M.J., 1957, 'The measurement of productive efficiency', *Journal of the Royal Statistical Society (A, general)*, 120, 253–281.
- Ganić, M., 2012, 'The impact of the global financial crisis on the banking sector of Western Balkans: Cross-Country Comparison Analysis', *Journal of Economic and Social Studies* 2(2), 177–196. <https://doi.org/10.14706/JECOSS11228>
- Government of Zimbabwe, 2009, *National budget statement*, Government Publications, Harare, Zimbabwe.
- Gulati, R., 2011, 'Evaluation of technical, pure technical and scale efficiencies of Indian banks: An analysis from cross-sectional perspective', A paper submitted for presentation in The 13th Annual Conference on Money and Finance in the Indian Economy, Indira Gandhi Institute of Development Research, Mumbai, 25–26th February.
- Hacini, I. & Dahou, K., 2016, 'Efficiency of the Algerian banks in the post liberalization period', *Financial Studies* 1, 90–106.
- Hassan, K. & Sanchez, B., 2007, *Efficiency determinants and dynamic efficiency changes in Latin American banking industries*, Networks Financial Institute, Indiana University, Working Paper, No. 32, Networks Financial Institute, US.
- Jreisat, A. & Hassan, H., 2016, 'Banking efficiency in Egypt: An application of data envelopment', *International Journal of Economics and Business Research* 11(2): 101–119. <https://doi.org/10.1504/IJEBR.2016.075302>
- Kablan, S., 2010, *Banking efficiency and financial development in Sub-Saharan Africa*, International Monetary Fund, Washington, DC, IMF Working Paper, No. WP/10/136
- Karimzadeh, M., 2012, 'Efficiency analysis by using data envelop analysis model: Evidence from Indian Banks', *International Journal of Latest Trends in Finance, Economics, Science* 2(3), 228–237.
- Kumar, N. & Singh, A., 2015, 'Measuring technical and scale efficiency of banks in India using DEA', *Journal of Business and Management* 17(1), 66–71.
- Kumar, S. & Gulati, R., 2008, 'An examination of technical, pure technical, and scale efficiencies in Indian public sector banks using data envelopment analysis', *Eurasian Journal of Business and Economics* 1(2), 33–69.
- Leibenstein, H., 1978, *X-Inefficiency Xists: Reply to an Xorcist*, American Economic Association, pp. 203–211.
- Li, L., 2014, 'Technical and scale efficiencies of Thai Commercial Banks after the 1997 Financial Crisis', *NIDA Business Journal* 4, 82–96.
- Liu, J.S., Lu, L.Y., Lu, W. & Lin, B.J.Y., 2012, 'Data envelopment analysis 1978–2010: A citation-based literature survey', *Omega* 41(1), 3–15.
- Lovell, K.C.A., 1993, 'Production frontiers and productive efficiency', in H.O. Fried, C.A. Knox Lovell & S.S. Schmidt (eds.), *The measurement of productive efficiency: Techniques and applications*, Oxford University Press, Oxford, pp. 3–67.
- Mongid, A. & Tahir, I.M., 2010, 'Technical and scale efficiency of Indonesian rural banks', *Banks and Bank Systems* 5(3), 80–86.
- Noulas, A.G., 2001, 'Deregulation and operating efficiency: The case of the Greek banks', *Managerial Finance* 27(8), 35–47. <https://doi.org/10.1108/03074350110767321>
- Nyankomo Marwa, N. & Aziakpono, M., 2015, *Technical and scale efficiency of Tanzania saving and credit cooperatives*, ERSA Working Paper 510, Economic Research Southern Africa, South Africa.
- Oteng-Abayie, E.F., 2017, 'Technical efficiency and total factor productivity of rural banks in Ghana', *Cogent Economics & Finance* 5, 1–20. <https://doi.org/10.1080/23322039.2017.1366088>
- Qamruzzaman, Md. & Jianguo, W., 2016, 'An assessment of financial efficiency using Data Envelopment Analysis (DEA)-Multistage approach: A case study of Banks in Bangladesh', *IOSR Journal of Economics and Finance* 7(6), 96–103.
- Reserve bank of Zimbabwe, 2013, *Monetary Policy Statements*, Harare, Zimbabwe.
- Reserve bank of Zimbabwe, 2014, *Bank Annual Report*, Harare, Zimbabwe.
- Reserve bank of Zimbabwe, 2015, *Bank Annual Report*, Harare, Zimbabwe.
- Roy, D., 2014, 'Analysis of technical efficiency of Indian banking sector: An application of data envelopment analysis', *International Journal of Finance & Banking Studies* 3(1), 150–160.
- Singh, D. & Fida, B.A., 2015, 'Technical efficiency and its determinants: An empirical study on banking sector of Oman', *Problems and Perspectives in Management* 13(1), 168–175.
- Sok-Gee, C., 2011, 'Technical efficiency of commercial banks in china: Decomposition into pure technical and scale efficiency', *International Journal of China Studies* 2(1), 27–38.
- Tahir, I.M., Bakar, N.M.A. & Haron, S., 2009, 'Estimating technical and scale efficiency of Malaysian commercial banks: A non-parametric approach', *International Review of Business Research Papers* 5(1), 113–123.
- Tandon, D., Tandon, K. & Malhotra, I., 2014, 'An evaluation of the technical, pure technical and scale efficiencies in the Indian banking industry using data envelope analysis', *Global Business Review* 15(3), 545–563. <https://doi.org/10.1177/0972150914535141>
- Titko, J. & Jureviciene, D., 2014, 'DEA application at cross-country benchmarking: Latvian vs Lithuanian banking sector', *Procedia – Social and Behavioral Sciences* 110, 1124–1135. <https://doi.org/10.1016/j.sbspro.2013.12.959>
- Tuo, M., 2016, 'An empirical analysis of Chinese Commercial Banks' Efficiency and influencing factors – Under the constraint of non-performing loans', *American Journal of Industrial and Business Management* 6, 455–446. <https://doi.org/10.4236/ajibm.2016.64042>
- Yadav, R., 2015, 'Technical efficiency, pure technical efficiency and scale efficiency of Russian commercial banks: An empirical analysis (2007–2014)', *International Journal of Economic Research* 6(6), 52–59.