Cost efficiency and non-performing loans: An application of the Granger causality test

This study evaluated the nexus between cost efficiency and non-performing loans (NPLs) in the Zimbabwean banking sector for the period 2009–2014. The study was motivated by the increase in NPLs in the banking sector while banks have been accused of profiteering through excessive bank charges and interest rates. The study contributes to the literature on the relationship between efficiency and NPLs, which is a controversial area. The study established that the average cost efficiency was 81%. It increased from 70% to 88% between 2009 and 2014. It declined in 2012 and 2013 because of slowdown in economic activity. The study established that cost efficiency negatively Granger-causes NPLs, supporting the bad management hypothesis implying that the low level of efficiency was a result of poor credit management which led to a deterioration in the quality of banks’ loan books. Although poor credit policies might look lucrative in the short run, they have detrimental effects on the quality of the loan books of banks in the long run. The policy recommendation drawn from the results is that credit managers should adhere to the international best practice in managing credit.

Introduction

The study of non-performing loans (NPLs) has gained increasing attention in the literature of the banking sector (Batra 2003; Clementina & Isu 2014; Diawan & Rodrik 1992; Sethi & Bhatia 2007; Vatansever & Hepsen 2013). The Basel Committee defined NPLs as any loan that is overdue for more than 90 days (Alton & Hazen 2001; Guy 2011). Non-performing loans are bad debts whose recovery is doubtful because the borrowers are not servicing them. Fofack (2005) defined NPLs as those loans which have ceased earning income for the bank for a long time, that is, the principal and the interest have not been paid for more than 90 days. Vatansever and Hepsen (2013) argued that NPLs are a good measure to ascertain the performance of financial institutions, the economy and the stability of the financial sector. Understanding the evolution of NPLs is important because they provide a good measure of the success of the banks whose measurements depend on the profitability and quality of assets.

Non-performing loans have wider implications on the performance of the banks. They increase uncertainty with regard to the capital of the banks, which reduces the capacity of banks to access financing (Diawan & Rodrik 1992). They lead to the deterioration in the quality of the assets of a bank, its capital, as well as its profitability (Clementina & Isu 2014). Non-performing loans exert operational costs which reduce the capital and liquidity of the banks, distorting the process of credit growth and ultimately the performance of the banks. As long as NPLs are kept on bank books and continuously rolled over, the resources will remain locked up in unprofitable assets, thus hindering the financial intermediation process and impairing economic growth.

Vatansever and Hepsen (2013) argued that loans are income-generating assets of banks which determine the performance and stability of the banking sector. Poorly performing assets that cease to generate income for the banks lead to bank failures. An increase in NPLs affects the profitability, liquidity and solvency of the banks. Batra (2003) held the view that banks become preoccupied with NPL recovery procedures, spending more time and effort in this exercise when they should be growing the business. Law suits preoccupy the banks when NPLs are rising. Bank performance and profits are then affected. Sethi and Bhatia (2007) noted that if the bank had already debited interest and other charges which have not been realised, they are supposed to increase their provisions on the gross income amount. Non-performing loans reduce the return on investment, reduce the capital adequacy ratio, increase the cost of capital, widen the assets and liability mismatch and upset the economic value additions by banks while limiting the recycling of funds.

There are a number of factors that have been identified as determinants of NPLs. Goldstein and Turner (1996) recognised economic downturns, macro-economic volatility, high-interest rates and
excessive reliance on highly overpriced interbank borrowings, insider borrowing and moral hazard as determinants of NPLs. Kassim (2002) identified poor management, lack of sound credit policy, inadequate credit analysis, error in documentation, undue emphasis on profit at the expense of loan quality, fraudulent practices, political instability, abnormal competition, policy and regulatory inconsistency, weak real sector and political and social influence on bank operations. Other factors that have also been identified include moral hazard and adverse selection problems, bank performance and stability (Vatansever & Hapsen 2013). Interest rate spreads also affect the performance of loans. As the interest rate spreads increase, people already indebted to banks might face challenges in repaying them as the cost of servicing increases. This means the greater the spreads, the lower the borrower’s capacity to repay the loan, leading to non-performance (Nkusu 2011).

Banking sector efficiency is used to evaluate the sources of banking profitability, that is, whether profitability is determined by market power or effectiveness in production (Berger 1995; Molyneux, Altnubas & Gardener 1996). Efficient banking systems create social benefits through the reduction of wastage and enhanced competition (Berger & Humphrey 1997). Banking efficiency is essential for improving the accuracy of banks’ ranking in terms of performance (Berger & Humphrey 1997). Measuring bank efficiency allows managers to take action on improving performance. Managers can improve cost efficiency by adopting better technologies; alternatively, they can enhance capital through improving profit efficiency by adopting new marketing and pricing methods. An efficient intermediation process is important for fostering economic growth (Buchs & Mathisen 2005).

The main objective of this article is to investigate the relationship between NPLs and cost efficiency. The study specifically investigates the causality relationship between NPLs and cost efficiency for the Zimbabwean banking sector. The Zimbabwean banking sector is a good laboratory case for such an empirical investigation. The banking sector experienced high level of NPLs during 2009–2014. The level of NPLs increased from 1.8% in February 2009 to 20.1% by September 2014 (Reserve Bank of Zimbabwe 2015). The growth in NPLs limited the capacity of banks to expand financial intermediation. The increase in 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.

The banking sector was perceived to be inefficient because of the number of bank failures and the wide spreads between lending and borrowing rates in the country (Reserve Bank of Zimbabwe 2013). The high lending rates were presumed to be a cover for the potential defaults from the borrowing clients. The closure of a number of banks was a result of poor credit management as loans turned bad, leading to insolvency and liquidity challenges while at the same time insider loans were very high (Reserve Bank of Zimbabwe 2015). Perceiving that the banking sector was profiteering from excessive interest rates and bank charges while paying little on deposits, the government instituted a Memorandum of Understanding (MoU) between the Central Bank and other banks (Reserve Bank of Zimbabwe 2013). The MoU prescribed the lending rates which were not supposed to exceed 12.5% points above the weighted cost of funds (Reserve Bank of Zimbabwe 2013). The high interest rate spreads between lending and deposit rates motivate the need to understand the efficiency of the banking sector.

The rest of the article is organised as follows: the next section discusses the background to NPLs in Zimbabwe. The ‘Relationships between efficiency and non-performing loans’ section discusses relationships between efficiency and NPLs. The ‘Methodology’ section outlines the study methodology and the ‘Result presentation and analysis’ section presents the study findings. Finally, conclusions and recommendations are discussed in the ‘Conclusions and policy recommendations’ section.

Background to non-performing loans in the Zimbabwean banking sector

The unique developments that typified the Zimbabwean banking sector during 2009–2014 require a comprehensive analysis to understand the key performance drivers. The Zimbabwean banking sector has undergone significant transformation over the said period. The period coincided with economic stability and growth after a decade of economic decline. There have been both negative and positive developments in the banking sector during this period; there was growth in deposits, loans and assets over the period as the banking public’s confidence increased. The market share of the top four banks declined significantly during the period, reversing the flight to quality problem experienced before 2009. The big banks then were CBZ, Stanbic, BancABC and FBC bank (MMC capital 2010). There have also been a number of bank failures in the sector as well as mergers and acquisitions. Some of the banks that failed during the period were Kingdom, Interfin Bank, Allied Bank, Capital Bank and Renaissance, among others. The bank failures were a result of insolvency, liquidity challenges, undercapitalisation and corporate governance challenges (Reserve Bank of Zimbabwe 2013). The performance of the banking institutions was also affected by the rising NPLs, absence of a credit reference bureau, human resource flight, liquidity challenges, limited lender of last resort and reduced activity on the interbank market, among other factors (IMF 2013).

The level of NPLs as shown in Figure 1 increased from 1.8% in February 2009 to 20.1% in September 2014 (Reserve Bank
of Zimbabwe 2015). 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 banks’ balance sheets, with adverse effect on banks’ lending operations, hence 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. Non-performing loans affected bank performance, reducing the profitability of the sector through increased provisioning. The Central Bank attributed the rising NPLs 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 the amount of NPLs has an impact on the performance of the banking sector including its competitiveness. The increase in the amount of loans led some banks to suffer from insolvency and failure. Eight bank failures were experienced during 2009–2015. The bank closures were mostly a result of poor corporate governance, imprudent lending activities, insolvency and inadequate capitalisation. The episode of bank failures was not new to the Zimbabwean banking sector as it had also occurred during 2003–2004. During that period, the failures were mostly a result of poor corporate governance, insolvency and abuse of Reserve Bank advances (Reserve Bank of Zimbabwe 2004).

Figure 2 shows that Zimbabwe did not compare favourably with regional counterparts. At the time NPLs were increasing in Zimbabwe, regional counterparts were experiencing declining NPLs. This shows that while the quality of bank assets was declining in the Zimbabwean banking sector, regional counterparts were experiencing an improvement in the quality of assets and consolidating their banking sector stability.

Realising the unhealthy growth in NPLs and its impact on the operations of the banking institutions, the Central Bank introduced a number of initiatives to resolve the problem.

Firstly, the Central Bank issued a directive outlawing the issuance, renewal and rollover of insider loans (Reserve Bank of Zimbabwe 2014b). This was after the realisation that insider loans had grown up to US$175.3 million (about 7% of total banking sector loans) and were mostly non-performing. Individuals and companies were encouraged to apply for loans in non-related banks. Secondly, banks were asked to provide adequately for the NPLs so as to reflect their true positions in terms of credit risk in their portfolio and improve their risk management through stress testing (Reserve Bank of Zimbabwe 2014b). Those banks that failed to adhere to the order were to be penalised. The Central Bank promised to invoke relevant provisions of the Banking Act. Thirdly, the Central Bank in collaboration with the Ministry of Finance advised that they were developing the legal framework for credit reference bureaus to strengthen the credit risk management in the banking system (Government of Zimbabwe 2014). This was meant to deal with clients who had developed the habit of borrowing from many sources and failing to repay. The credit reference system was perceived as the best way to deal with errant clients. Fourthly, the Central Bank announced the setting up of a Special Purpose Vehicle (SPV) in the name of Zimbabwe Asset Management Corporation Private Limited (ZAMCO), which would purchase all secured NPLs from the banks (Reserve Bank of Zimbabwe 2014b). The SPV would assist the banks in cleaning their balance sheets by buying all collateralised NPLs. Lastly, the Central Bank proposed the setting up of a Commercial Court to expedite the resolutions of all commercial and bank-related cases before the courts although this has not been operationalised; hence, its effects on the banking sector will be felt later. The measure is meant to reduce the NPLs and strengthen the balance sheets of the banks.

**Relationships between efficiency and non-performing loans**

This section reviews the theoretical and empirical literature on the relationship between banking sector efficiency and NPLs.

**Theoretical literature**

Berger and DeYoung (1997) proposed four hypotheses to explain the relationship between NPLs and the efficiency of
the banking sector, which are bad luck, bad management, skimping behaviour and moral hazard.

The bad luck hypothesis posits that a surge in NPLs can be a result of unanticipated events outside the control of bank management, such as economic slowdown (Berger & DeYoung 1997). Under this scenario, a bank will incur additional expenses in monitoring these NPLs which ultimately negatively impact the efficiency of the bank. The additional costs will be in the form of dealing with defaulting costs such as legal costs; the decline in the value of pledged securities and the cost associated with debt rescheduling and other associated costs including the time banks take in making decisions on these problem loans. Under the bad luck hypothesis, it is then perceived that it is the increase in the problematic loans that causes the decline in the efficiency of the bank. This implies that there is a negative relationship between NPLs and bank efficiency.

The bad management hypothesis postulates that the low level of efficiency of the bank is a result of the poor management decision which is reflected in the poor credit management (Berger & DeYoung 1997). According to this hypothesis, the poor credit monitoring by bank management such as the poor screening of borrowers leads to an increase in the amount of NPLs. The hypothesis implies that the inefficiency of bank management leads to NPLs. Like in the bad luck hypothesis, there is a negative relationship between the two, though inefficiency precedes deterioration in the quality of assets of the bank.

The skimping hypothesis posits that those banks that are profitable might appear more efficient in the short run through using less resources in the monitoring and underwriting of loans (Berger & DeYoung 1997). According to skimping hypothesis, NPLs are not apparent in the short run, but in the long run when a higher proportion of borrowers start defaulting. The bank will encounter the problem with these borrowers and incur the costs of dealing with these clients.

The risk-averse management hypothesis (Koutsomanoli-Filippaki & Mamatzakis 2009) suggests that bank managers are risk averse. The risk-averse managers incur huge costs in loan monitoring and screening, which then decreases the efficiency of the banks (Berger & DeYoung 1997). The increase in monitoring and screening cost is meant to compress default in credit portfolio. The negative relationship between efficiency and NPLs is a result of the fear by management of a financial crisis and information asymmetry.

**Empirical literature**

Berger and DeYoung (1997) investigated the relationship between problem loan and banking sector efficiency. Their study employed the Granger causality techniques to test four hypotheses regarding the relationships among loan quality, cost efficiency and bank capital. The results of the study revealed that problematic loans preceded reduction in cost efficiency and also that cost efficiency preceded problem loans. The study supported both the skimping and bad luck hypotheses. They concluded that cost efficiency may be an important signal of potential problematic loans in the future. Williams (2004), applying the Granger causality test, undertook a robustness test for the result of Berger and De Young (1997) using a sample from European savings bank from 1990 to 1998. His results revealed that the banking sector was following the bad management hypothesis. Rossi, Schwaiger and Winkler (2005) also investigated the causality between cost efficiency and loan loss provision applying the Granger causality technique in banking sectors of Central and Eastern European countries. Their study failed to identify any evidence of bad management hypothesis despite identifying that there was significant negative correlation between problem loans and efficiency. They concluded that bank failure and efficiency of banks were a result of external factors outside the control of management, supporting the bad luck hypothesis. Podpiera and Weill (2007) investigated the management behaviour. Their study revealed that bad management was applied in the banking sector and thus rejected the possibility of bad luck hypothesis. The study used loan loss provision because of lack of data on NPLs. Tabak, Carveiro and Caujêiro (2011) undertook a study on causality between bank efficiency and default in Brazil. They also employed a Granger causality test and discovered the evidence of bad management hypothesis. They concluded that in maximising profits banks reduce the cost of loan monitoring which improves short-term efficiency. Consequently, when the default rates increase in the long term, banks become inefficient. The main reason for the increase in NPLs is the covered inefficiency. In a study of the Indian banking sector, Reddy (2011) found support for bad management, moral hazard and bad luck hypotheses. Rossi et al. (2009) studied the nexus between banking efficiency and NPLs in large commercial banks in Austria for the period from 1997 to 2003 and found that the Austrian banking system followed the bad luck and bad management hypotheses. Koutsomanoli-Filippaki and Mamatzakis (2009) studied the relationship between asset quality and the efficiency of the European Union member banks for the period 1998–2006. Their study failed to reject the bad management hypothesis in less financially developed countries. Williams (2004) studied a sample of European banks over the period 1990–1998 and concluded that the banks followed the bad management hypothesis. The result meant that the low cost efficiency led to an increase in the amount of NPLs, increasing the chances of bank failures.

**Methodology**

The estimation of the nexus between cost efficiency and NPLs is done in two stages. Firstly, the study estimates the cost efficiency of the banking sector and, secondly, it undertakes a causality test between cost efficiency and NPLs.

**Measuring cost efficiency of the banking sector**

The measuring of the banking sector efficiency is performed using the non-parametric (mathematical programming)
method (Stavarek & Řepková 2012:360). The non-parametric methods are based on linear programming tools. The efficiency frontier in non-parametric estimations is formed as a piecewise linear combination of best-practice observations. Hassan and Sanchez (2007) argue that the non-parametric approach does not require a production function to estimate the efficiency of the firm against some imposed benchmark through mathematical programming. The study specifically employs the data envelopment analysis (DEA) approach. This approach determines the efficiency frontier based on the input and output variables without knowing a priori the relationship among the variables. The method does not require the specification of the functional form a priori, hence removing the possibility of measurement errors. The method is able to handle multiple inputs and outputs measured in different units (e.g., dollars, time, employees and location). It is also able to identify the efficient units among the sample, hence making comparisons possible among the banks in the sample. This method has the ability to quantify the inefficiency of decision-making units (DMUs) and identify their efficient peers, hence providing information on what changes in inputs and outputs are required for inefficient units to reach the efficiency frontier as defined by efficient DMUs.

Farrell (1957) states that efficiency is composed of two components: technical and allocative efficiency. Technical efficiency is the ability of a firm to maximise outputs from a given set of inputs. Allocative efficiency is the ability of the firm to use inputs in optimal proportions given their respective prices. A combination of both provides a measure of economic efficiency. One of the measures of economic efficiency is the cost efficiency, which measures how far apart is a bank’s cost from that of the best practice bank that produces the same output level and under the same environmental conditions (Lovell 1993).

Assuming there are $n$ DMUs and the $j$th DMU, DMU, produces $s$ outputs ($y_1, y_2, ..., y_s$) by using $m$ inputs ($x_1, x_2, ..., x_m$), the cost efficient model is summarised as follows:

$$\text{min } \sum_{i=1}^{m} P_i x_{iw}$$

Subject to

$$\text{CRS } \sum_{j=1}^{n} \lambda_j x_{ij} \leq x_{iw}, \quad i = 1, 2, 3, \ldots, m \quad [\text{Eqn 1}]$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{rw}, \quad r = 1, 2, 3, \ldots, s$$

$$\lambda_j, x_{i,j} \geq 0$$

where $P_i$ is the unit price of the input $i$. These prices may be different among the DMUs. The variable $\lambda$ is the intensity variable that minimises costs.

The definition of the inputs and outputs follows the intermediation approach. This approach 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 are considered for the evaluation of the banks performance: deposits, labour and capital. The outputs are total loans (short-, medium- and long-term loans) and total income (sum of interest income and non-interest income).

Efficiency scores also differ depending on the assumptions underlying their calculations. These can be calculated under variable returns to scale (VRS) or constant returns to scale (CRS) (Banker, Charnes & Cooper 1984; Charnes, Cooper & Rhodes 1978). The choice between CRS and VRS determines the shape of the envelopment surface and as a result the number of efficient DMUs. Constant returns to scale attains proportional increase if all inputs lead to a proportional increase in output. Models that apply the CRS approach are called the CRS models. The CRS 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 efficient scores compared to the alternative VRS model (Karimzadeh 2012). The choice of CRS is based on its restrictiveness.

The estimation of cost efficiency requires the specification of the prices of inputs. The price of a deposit is calculated as interest expense over total deposits. The price of labour is equal to labour costs over total assets (assets are used because the number of employees is not available) and the price of capital is calculated as total expenses minus labour expenses over total assets. The study uses dollar values for measuring inputs and outputs because, firstly, banks compete to increase their market share for dollar amounts rather than the number of accounts. Secondly, the various accounts offered by banks have different costs; for example, time deposit accounts differ from savings accounts. Thirdly, banks offer a number of services, in which case the dollar amount is the only common denominator (Kolari & Zardkoohi 1987).

**Granger causality test between cost efficiency and non-performing loans**

To measure the dynamic links between cost efficiency and NPLs, the study estimates the causal relationship between bank efficiency and NPLs for a panel of banking institutions. The study employs the Granger causality test to examine the causal relationship between efficiency and NPL (Nair-Reichert & Weinhold 2001).

The Granger causality model followed is:

$$NPL_{it} = \alpha + \sum_{j=1}^{k} \gamma_{ij} NPL_{i,t-j} + \sum_{i=1}^{l} \beta_{ij} CEF_{i,t-j} + \epsilon_{it} \quad [\text{Eqn 2}]$$
where $\text{NPL}$ represents the non-performing loans, $\text{CEF}$ represents efficiency score, $i$ ranges from 1 up to $N$ cross units, $t$ represents time periods $(1, 2, ..., T)$, $\alpha$ denotes the intercept, $k$ represents the number of lags and $\varepsilon$ represents the error term including not only the disturbance term but also the cross-unit specific effects.

To test the Granger non-causality for $\text{CEF}$ to $\text{NPL}$, the null hypothesis is $H^i_0: \beta = 0$, for all $i = 1, ..., N$. The alternative hypothesis states that there is a causality relationship from $\text{NPL}$ to $\text{CEF}$ for at least one cross unit of the panel: $H^i_0: \beta = 0$, for all $i = 1, ..., N$, $\beta \neq 0$ for $i = N_i + 1, N_i + 2, ..., N; 0 \leq \frac{N_i}{N} \leq 1$.

To test causality from $\text{NPL}$ to $\text{CEF}$, the following model is used:

$$\text{CEF}_i = \alpha + \sum_{i} \phi_i^{(1)} \text{CEF}_{i,t-i} + \sum_{i} \theta_i^{(1)} \text{NPL}_{i,t-i} + \varepsilon_i$$  \hspace{1cm} \text{[Eqn 3]}

To test the Granger non-causality for $\text{NPL}$ to $\text{CEF}$, the null hypothesis is $H^i_0: \beta = 0$, for all $i = 1, ..., N$. The alternative hypothesis states that there is a causality relationship from $\text{CEF}$ to $\text{NPL}$ for at least one cross unit of the panel: $H^i_0: \beta = 0$, for all $i = 1, ..., N$, $\beta \neq 0$ for $i = N_i + 1, N_i + 2, ..., N; 0 \leq \frac{N_i}{N} \leq 1$.

The estimation of the Granger causality equations (2 and 3) will be done using the two-stage generalised method of moments (GMM) panel estimation technique which is designed to handle auto-regressive properties in the dependent variable when lagged values are included as explanatory variables. The study method follows (Arellano & Bond 1991; Arellano & Bover 1995; Blundell & Bond 1998) and deals with the problem of endogeneity and measurement error. The study included 11 commercial banks that operated during 2009–2014. Half-yearly data were used in the study. The data were sourced from published financial statements of the banks, which included the balance sheet and income statements for the period 2009H1–2014H2 for the bank-specific data. The data are therefore balanced quarterly panel data sets.

**Result presentation and analysis**

This section presents the results of the estimation of the cost efficiency of the Zimbabwean banking sector and the Granger causality test.

**Cost efficiency of the banking sector**

Figure 3 reveals that banking sector efficiency increased between 2009 and 2011. The average cost efficiency of the banking sector during 2009–2014 was 81%. This means that average banks in the sample were operating outside their cost efficiency frontier. The results showed that banks were encountering a 19% level of inefficiency, implying that banks could have produced the same amount of output by using 19% less resources if they had performed as the most efficient bank in the sample. This means that banking consumers could have obtained the same amount of services at a lower cost than they paid for.

Banking sector cost efficiency declined during 2012 and 2013 before increasing again in 2014. The decline in the cost efficiency coincided with the slowdown in economic activity. During this period, there was considerable increase in NPLs. This led to a reduction in incomes of the banks as well as the levels of provisioning for bad debts. This could have impacted the levels of efficiency as the banks were forced to reduce loan growth in order to reduce the amount of toxic assets in their books. This situation was compounded by the effects of the MoU which the government put in place in 2013. Effectively, the MoU reduced the amount of price competition in the banking sector as the parameters of pricing were set by the government. Incomes for the banks were therefore reduced while at the same time the amount of loan loss provision was increasingly reducing financial intermediation efficiency.

The rescission of the MoU at the end of 2013 witnessed the cost efficiency scores rebounding as banks were now able to revert to their pricing models. This increased the level of competition among the banks on the pricing front and in attracting customers.

**Granger causality test between cost efficiency and non-performing loans**

The results of the GMM estimation of the dynamic equations represented in equations 2 and 3 are shown in Table 1. The test for autocorrelation in the residuals in differences shows that AR (1) test rejects the null hypothesis of no existence of first-order serial autocorrelation to the bank efficiencies and NPL equations, as expected, indicating a negative first-order serial autocorrelation. On the other hand, AR (2) test does not detect evidences of second-order serial correlation.
in all equations, implying that all instruments are valid. The study adopts specifications with lower numbers of instruments. The table reports the coefficients of the lags of the dependent variable as well as the independent variable. Of primary interest are the coefficients of the lags of the independent variable.

The results for the Granger causality test between cost efficiency and NPLs reveal that causality is unidirectional, running from cost inefficiency to NPLs. The results suggest that there is no causality running from NPLs to cost efficiency.

Table 1 shows that cost efficiency negatively Granger-causes NPLs. As efficiency runs from efficiency to NPLs, suggesting that the sector conforms to the bad management hypothesis, the result agrees with prior studies by Rossi et al. (2005), Podpiera and Weill (2007) and Reddy (2011) who also found support for bad management hypothesis. The bad management hypothesis postulates that the low level of efficiency of the bank is a result of the poor management decision which is reflected in the poor credit management (Berger & DeYoung 1997). This means that the increase in the NPLs was a result of inefficiency of credit managers. During 2009–2010, banks were involved in aggressive competition for clients after their balance sheets were decapitated by the adoption of dollarisation after the country had experienced hyperinflation prior to 2008. As a result banks sought to increase their customer base through offering salary-based loans and loans to corporates without proper screening of customers. Most of the loans were as much as 10 times that of an individual’s salary, whereas before 2009 the maximum loan amount was three times that of an individual’s salary. This then implies that bank management inefficiency was the reason behind the growth in deterioration of asset quality. This adversely affected the performance of the banking sector through increased provisioning of loans and a reduction in profitability. Banks had to cut down on lending as a result of the growing NPLs which stunted financial intermediation. The banks had to incur higher costs in trying to recover from the NPLs as they used various methods including litigation.

The bad management hypothesis gives credence to the action that was adopted by the Central Bank as it sought to reduce the amount of NPLs. These included the outlawing of insider loans, adequately providing for non-performance in their books of accounts and the taking over of all collateralised NPLs by the SPV.

Firstly, the Central Bank issued a directive to banks outlawing the issuance, renewal and rollover of insider loans (Reserve Bank of Zimbabwe 2014a). This was after the realisation that insider loans had grown significantly and were mostly non-performing. Individuals and companies were encouraged to apply for loans at non-related banks.

**Conclusions and policy recommendations**

The study evaluated the nexus between cost efficiency and NPLs employing a two-stage approach. In the initial stage, the study estimated the cost efficiency of the banking sector. In the second stage, the study undertook a causality test between cost efficiency and NPLs. The study established that the average cost efficiency of the banking sector during 2009–2014 was 81%, meaning that average banks in the sample were operating outside the industry cost efficiency frontier. The results showed that banks were encountering a 19% level of inefficiency, implying that banks could have produced the same amount of output by using 19% less resources. On average, the banking sector cost efficiency increased between 2009 and 2014 from 70% to 88%. The cost efficiency declined in 2012 and 2013. The decline in the cost efficiency coincided with the slowdown in economic activity. The results of the study suggest that cost inefficiency negatively Granger-causes NPLs, implying that the bad management hypothesis holds for banks. This means that
the increase in the NPLs was a result of the inefficiency of credit managers. The policy recommendation drawn from the results is that credit managers should adhere to the international best-practice method of managing credit. Though poor credit policies might look lucrative in the short run, there is negative implication in the long run on the bank’s loan book quality. There is a need for the banks to cooperate with the Central Bank on the various initiatives being put in place to clean their balance sheets so as to start on a new slate. Going forward there is a need for further studies to understand the relationship between non-performing assets and profit efficiency, market structure and banking sector productivity.

Acknowledgements

Competing interests

The author declares that he has no financial or personal relationships which may have inappropriately influenced him in writing this article.

References

Alton, R.G. & Hazen, J.H., 2001, ‘As economy flounders, do we see a rise in problem loans?’ Federal Reserve Bank of St. Louis, St. Louis, MO.


Reserve Bank of Zimbabwe, 2014b, Monetary Policy Statements, Reserve Bank of Zimbabwe, Harare, Zimbabwe.


