

Selecting the Most Significant Determinants that contributes to Islamic banks Profitability using Principal Component Analysis

Rashidah Ismail, Nur Huda Nabihan Md Shahri, Nur Liyana Basir

^{1, 2,3} Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Malaysia

Email: shidah@tmsk.uitm.edu.my, huda_nabihan@tmsk.uitm.edu.my, liyana.basir1357@gmail.com

To Link this Article: <http://dx.doi.org/10.6007/IJARBS/v8-i12/5085>

DOI:10.6007/IJARBS/v8-i12/5085

Published Date: 30 December 2018

Abstract

The existence of Islamic banks that grow rapidly in every corner of the world has caused strong competition among them and other conventional banks in Malaysia. This requires the banks management to choose wisely on the determinants that make them to remain strong and relevant in the Islamic banking sector. The main purpose of this study is to determine internal and external significant determinants for Islamic banks in Malaysia using Principal Component Analysis (PCA). The Kaiser-Meyer-Olkin (KMO) was applied in this study to measure the data sampling is acceptable and adequate. While, and Bartlett's Test is used to validate the internal and external component. The significant banking determinants selected for the study include deposit ratio, operating efficiency and market concentration. The result shows that the internal and external significant determinants are Bank Size and Gross Domestic Product. The study utilized secondary data from year 2010 until year 2016 for 13 Islamic banks in Malaysia.

Keywords: Principal Component Analysis, Significant determinant, Bank Size, Gross Domestic Product.

Introduction

There are many determinants that contribute to banks' profit. The bank management needs to choose which determinants are the most suitable for them to focus on, to deliver the best profit for the Islamic banks. The objectives of this study is to identify the most significant determinants internal (bank size, bank capital, credit risk, liquidity, deposits ratio and operating efficiency) and external (Gross Domestic Product, annual inflation rate and market concentration) that contribute to bank profitability by using Principal Component Analysis.

Literature Review

Several independent variables can be predicted as the significant determinants that influence bank profitability which are divided into microeconomics as an internal factor (bank size, bank capital, credit risk, liquidity, deposit ratio and operating efficiency and macroeconomics as an external factor (Gross Domestic Product, annual inflation rate and market concentration).

Based on the previous study by Djalilov and Piesse (2016), the researchers highlighted on numerous effects of bank size on bank performance such as larger banks have more diversification in their products and services. Menicucci and Paolucci (2016) indicated that well-capitalized banks are less risky and have a positive relationship with profitability and banks can gain more profit in additional deposits, while low deposits may impact negatively on their profitability. Tafri, Hamid, Mydin Meera, & Omar (2009) suggested that an increase exposure to credit risk normally lead to a decrease in the firm profitability. Masood and Ashraf (2012) mentioned that a higher percentage of ratio denotes that banks are more liquid; therefore the liquidity risk become lower. Garcia and Guerreiro (2016) stated that the examples of operational cost are staff costs, general and administrative expenses, depreciation and amortization, provisions net of reversals and impairment on assets net of reversals.

Curak et al., (2012) stated that GDP controls the cyclical output effects to affect numerous factors regarding supply and demand for loan and deposits. Besides, another inflation rate is used to measure the percentage increase in the Consumer Price Index (CPI) for all good and services. Annual inflation is predicted to have positive impact on profitability when income of bank increases rapidly than its costs.

Lastly, market concentration is a ratio of total assets of one bank to total assets of all banks in country (Olson & Zoubi, 2011). Dietrich and Wanzenried (2011) mentioned that lower market concentration might be the result of a positive relationship to the bank profitability.

Principal Component Analysis (PCA) was introduced by Pearson in 1901. It has been considered as the oldest technique in multivariate analysis and modified until it was generalized by Loeve in 1963 (Ionita & Şchiopu, 2010). Jolliffe (2002) mentioned that the objective of the Principal Component Analysis (PCA) is to minimize the dimensionality of a data set that is produced by many variables which are interrelated while withholding the variability available in the data set. In addition, Wuensch (2012) explained that Principal Component Analysis is a method to extract from a set of p variables to a set of m components or factors. Besides, Principal Component Analysis is a tool to reduce multidimensional data to lower dimensions while retaining most of the information that covers standard deviation, covariance, and eigenvectors (Karamizadeh et al., 2013). Robu and Istrate (2015) mentioned that the main advantage of using PCA is that it reduces the initial causal space to an equivalent space of less considerable dimensions.

Method

The first step was collecting data where the data was collected from tradingeconomic.com, Datastream and financial statement of each bank from year 2010 to year 2016. The second step was a stage of identifying the significant determinants using Principal Component Analysis.

Data Collection

Secondary data for return on equity, bank size, bank capital, credit risk, liquidity risk, deposit ratio, operating efficiency, and market concentration were taken from the datastream and the financial statement of each bank from year 2010 to 2016. Meanwhile, data for GDP and the annual inflation rate were taken from the tradingeconomic.com from year 2010 to 2016. In this study, return on equity is a dependent variable while the independent variables are determinants that influence bank profitability such as bank size, bank capital, credit risk, liquidity risk, deposit ratio, operating efficiency, market concentration, GDP and annual inflation rate.

As mentioned before, there are 16 Islamic banks operating in Malaysia. However, due to unavailability of data, there are only 13 Islamic banks in Malaysia involved in this study which are Bank Islam Malaysia Bhd., CIMB Islamic Bank Bhd., Maybank Islamic Berhad, RHB Islamic Bank Berhad., Affin Islamic Bank Bhd., Bank Muamalat Malaysia Bhd., Asian Finance Bank Bhd., Hong Leong Islamic Bank Bhd., Public Islamic Bank Bhd., Standard Chartered Saadiq Berhad, OCBC Al-Amin Bank Bhd., HSBC Amanah Malaysia Berhad, and Alliance Islamic Bank Berhad.

The data were exported into SPSS software to run the Principal Component Analysis.

Principal component analysis is linear combination X_1, X_2, \dots, X_p which are not correlated, and whose variances $Var(X_i) = b'_i \sum b_i, i = 1, 2, \dots, p$ are as large as possible. Given $Cov(X_i, X_k) = b'_i \sum b_k, i, k = 1, 2, \dots, p$, where $b'_i = b_{i1}, b_{i2}, \dots, b_{ip}$ are weights, then:

$$\begin{aligned} b'_i \cdot b_i &= 1 \\ b'_i \cdot b_j &= 0 \text{ for all } i \neq j \end{aligned} \quad (1)$$

The first principal component is the linear combination $b'_1 X$ that maximizes $Var(b'_1 X)$ subject to $b'_1 b_1 = 1$. However, this accounts for the greatest variance in the data. The second principal component is also a linear combination $b'_2 X$ that maximizes $Var(b'_2 X)$ subject to which $b'_1 b_1 = 1$ and $Cov(b'_1 X, b'_2 X) = 0$. However, this accounts for the greatest of the remaining variance in the data and so on. The number of variables in the data set is trimmed down to the rendering or analysis of the correlation among the variables when the first principal component caters for a larger share of the variances in the original.

Suppose each of the bank's financial statements is an observed variable Y_1, Y_2, \dots, Y_p with weights $b_{ij}, i = 1, 2, \dots, p, j = 1, 2, \dots, p$, then the principal component X_1, X_2, \dots, X_p are given by:

$$\begin{aligned} X_1 &= b_{11}y_1 + b_{12}y_2 + \dots + b_{1p}y_p \\ X_2 &= b_{21}y_1 + b_{22}y_2 + \dots + b_{2p}y_p \\ &\vdots \\ X_p &= b_{k1}y_1 + b_{k2}y_2 + \dots + b_{kp}y_p \end{aligned} \quad (2)$$

Let the random variable $Y_i (i = 1, 2, \dots, p)$ have mean μ_i and the standard deviation σ_{ii} . Thus, the transformed standard variables $Z_i (i = 1, 2, \dots, p)$ is given as:

$$Z_i = \frac{Y_i - \mu_i}{\sigma_{ii}} \quad (3)$$

Furthermore, the vector of the standardized variables could be written in matrix notation as:

$$Z = (V^{\frac{1}{2}})^{-1}(Y - \mu) \quad (4)$$

Where $\mu' = (\mu_1, \mu_2, \dots, \mu_k)$ and $V^{\frac{1}{2}}$ is the diagonal standard deviation matrix given by:

$$V^{\frac{1}{2}} = \begin{bmatrix} \sigma_{11} & 0 & \dots & 0 \\ 0 & \sigma_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{kk} \end{bmatrix} \quad (5)$$

$E[Z_i] = 0, Var [Z_i] = 1, i = 1, 2, \dots, k$ and $Cov (Z) = (V^{1/2})^{-1} \Sigma (V^{1/2})^{-1} = \rho$ where the variance-covariance matrix Σ and the correlation matrix ρ of Y are given by:

$$\Sigma = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12}^2 & \dots & \sigma_{1k}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 & \dots & \sigma_{2k}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{k1}^2 & \sigma_{k2}^2 & \dots & \sigma_{kk}^2 \end{bmatrix} \quad (6)$$

$$\rho = \begin{bmatrix} 1 & \frac{\sigma_{12}^2}{\sigma_{11}\sigma_{22}} & \dots & \frac{\sigma_{1k}^2}{\sigma_{11}\sigma_{kk}} \\ \frac{\sigma_{12}^2}{\sigma_{11}\sigma_{22}} & 1 & \dots & \frac{\sigma_{2k}^2}{\sigma_{22}\sigma_{kk}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sigma_{1k}^2}{\sigma_{11}\sigma_{kk}} & \frac{\sigma_{2k}^2}{\sigma_{22}\sigma_{kk}} & \dots & 1 \end{bmatrix}$$

And

$$\rho_{ij} = \frac{\sum_{k=1}^n (y_{ki} - \mu_i)(y_{kj} - \mu_j)}{n}, i \neq j$$

is the covariance between variables Y_i and Y_j , each of which has n observations. The p principal components $x' = [x_1, x_2, \dots, x_p]$ are derived from the eigenvectors of the correlation matrix of ρ of Y given by:

$$X = B'Z \quad (7)$$

where $B = [e_1, e_2, \dots, e_k]$ and the e_i s, $i = 1, 2, \dots, k$ are eigenvectors of ρ . The eigenvalue-eigenvector pair $(\lambda_1, e_1), (\lambda_2, e_2), \dots, (\lambda_k, e_k)$ of ρ are such that:

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k \geq 0, e_i' \cdot e_i = 1 \text{ and } e_i' \cdot e_j = 0 \quad (8)$$

$$var(X_i) = e_i' \rho e_i = \lambda_i$$

And

$$\sum_{i=1}^k \text{var}(X_i) = \sum_{i=1}^k \text{var}(Z_i) = P$$

Subsequently, the proportion of the total variance in the data which is explained by the X_j is given by proportion $\frac{\lambda_j}{p}$. This is called communality. Where $j = 1, 2, \dots, p$ and the λ_j 's are the eigenvalue of p . The loading of the standardized variable Z_j which is a correlation between each principal component X_i and the corresponding standardized variable Z_j is given by:

$$\text{Corr}(X_i, Z_j) = e_{ij} \cdot \lambda_j^{\left(\frac{1}{2}\right)} \quad (9)$$

Where the loading of the standardized variable Z_i is between -1 and 1 inclusive.

Therefore, this show the degree to which every Z_i affects every X_i which is conditioned on the influence of the others variable $Z_m, j \neq m$. The variable becomes more influential in naming and interpreting the component when there is a higher absolute value of loading of a variable on a component. Similarly, the size of the sample determines the significance of the magnitude of a loading. A loading of an absolute value of at least 0.30 and 0.35 with respective sample sizes of 350 and 250. This shows that a smaller sample size requires a greater absolute loading values. Assare-Kumi (2016) recommended a loading of magnitude which is more that equal to 0.50.

However, the principal component being used as an index involves the establishment of principal component scores and factor loading. The principal component scores which are the i^{th} principal component is derived by putting the standardized observed values of the variables into the equation as shown below:

$$X_i = e_i'Z \quad (10)$$

The number of observed variables being tested is the same as the number of components derived in the principal component analysis. This is because only the first few components account for a bigger amount of the variance in the original data. Thus, these few components are sustained, interpreted, and are involved in subsequent analysis.

Results And Discussion

IDENTIFYING THE SIGNIFICANT DETERMINANTS USING PRINCIPAL COMPONENT ANALYSIS.

Table 4.1

Kaiser-Meyer-Olkin and Bartlett's Test

Test	Results
Kaiser-Meyer-Olkin	0.547
Bartlett's Test of Sphericity	0.000

Table 4.1 shows Kaiser-Mayer-Olkin and Bartlett's Test. Based on Assare-Kumi et al. (2016), the Kaiser-Meyer- Olkin (KMO) statistic used to measure the data sampling is acceptable if

the value is greater than or equal to 0.5. In this study, the KMO value is 0.547 which means that the data is adequate.

Assare-Kumi et al. (2016) also mention that when the probability value of Bartlett's Test of Sphericity is less than the significance level which is significant value 0.05, the Principal Component Analysis can be used for the analysis. Besides, in this study the Bartlett's test is significant with $p = 0.000$. Therefore, Principal Component Analysis can be used for the analysis.

Table 4.2

Eigenvalues Explained

Component / Group	eigenvalue
Internal Component	2.852
External Component	1.476

Table 4.2 shows the Eigenvalues Explain. In this study, there were two fixed components which were extracted, one from internal component and one from external component. The validity of components can be checked via eigenvalue of one more (Pallant, 2011). This step was conducted to ensure the validity of the internal component and external component which were used in this study. So, the internal and external component is valid to use in this study.

Table 4.3

Factor Loading for each Determinants

Independent determinant	Component	
	Internal Component	External Component
Bank Capital	-0.891	
Bank Size	0.889	
Deposit Ratio	0.820	
Market Concentration	0.723	
GDP		0.774
Annual Inflation Rate		-0.716

Table 4.3 shows the Table of Factor Loading for each Determinants. Based on table above, bank capital, bank size, deposit ratio and market concentration was a fit well with each other in internal component while GDP and inflation was fit in external component. The credit risk, liquidity risk and operating efficiency was not fit for both components since the loading value was not appear in Rotated Component Matrix table in SPSS results. In addition, each of negative number show the inverse relationship factor for the component.

So, the significant determinants choose based on the highest factor loading for each component (Pallant, 2011). The highest positive factor loading for internal and external components are bank size and GDP with value 0.889 and 0.774 respectively. Therefore, the most significant determinant that contributes to Islamic banks profitability is bank size (internal factor) and Gross Domestic Product.

Conclusion

Since the early existence of the Islamic banks in the banking sector, Islamic banks have a fierce competition with conventional banks in terms of the products and other services offered. Thus, the bank management of the Islamic Bank must choose wisely which determinants that they need to focus on in order to give the highest profit to the Islamic banks because some of the determinants do not really give an impact towards bank profitability. In this study, Principal Component Analysis method was used to find out the internal and external determinant which were significant. The results showed that the significant internal determinant was bank size while the significant external determinant was GDP.

Acknowledgements

Our sincere appreciation is directed to the Ministry of Education, Malaysia and Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia.

References

- Abidin, S. N. Z., & Jaffar, M. M. (2014). Forecasting share prices of small size companies in bursa Malaysia using geometric Brownian motion. *Applied Mathematics & Information Sciences*, 8(1), 107.
- Assare-Kumi, A. A., Darkwah, K. A., Nortey, E. N. N., & Chapman-Wardy, C. (2016). Variable Reduction and Determinants of Working Capital Management on Profits for Ghanaian Banks. *European Scientific Journal*, 12(7), 316-330.
- Blyth, K. & Kaka, A. (2006). A novel multiple linear regression model for forecasting S-curves. *Engineering, Construction and Architectural Management*, 13(1), 82-95.
- Ichsani, S., & Suhardi, A. R. (2015). The Effect of Return on Equity (ROE) and Return on Investment (ROI) on Trading Volume. *Procedia-Social and Behavioral Sciences*, 211, 896-902.
- Ioniță, I., & Șchiopu, D. (2010). Using Principal Component Analysis in Loan Granting. *The Bulletin of Petroleum-Gas University of Ploiesti*, 62(1/2010), 88–96.
- Jolliffe, I. (2002). *Principal component analysis*. John Wiley & Sons, Ltd.
- Karamizadeh, S., Abdullah, S. M., Manaf, A. A., Zamani, M., & Hooman, A. (2013). An overview of principal component analysis. *Journal of Signal and Information Processing*, 4(03), 173.
- Mohd Alias, L. (2011). *Introductory Business Forecasting: A Practical Approach 3rd Edition* (3rd ed.). Shah Alam, Malaysia: Universiti Teknologi MARA Press.
- Pallant, J. (2011). *SPSS Survival Manual: A step by step guide to data analysis using SPSS* (4th ed.). Crows Nest, New South Wales: Allen & Unwin.
- Robu, I. B., & Istrate, C. (2015). The analysis of the principal components of the financial reporting in the case of Romanian listed companies. *Procedia Economics and Finance*, 20, 553–561.
- Wong, J. M. W., Chan, A. P. C., & Chiang, Y. H. (2011). Construction manpower demand forecasting: A comparative study of univariate time series, multiple regression and econometric modelling techniques. *Engineering, Construction and Architectural Management*, 18(1), 7-29.
- Wuenschk, K. L. (2012). *Principal Components Analysis - SPSS*. Retrieved from <http://core.ecu.edu/psyc/wuenschk/MV/FA/PCA-SPSS.docx>

