A STATISTICAL MODEL OF ORGANIZATIONAL PERFORMANCE USING FACTOR ANALYSIS - A CASE OF A BANK IN GHANA

Emmanuel M. Baah[1*], Joseph K. A. Johnson[2], Frank B. K. Twenefour[3]
[1, 2, 3]Department of Mathematics and Statistics, Takoradi Polytechnic,
P. O. Box 256. Takoradi, Western Region, Ghana

ABSTRACT: *The growth and progress of any organization depends on its ability to perform at the highest level on a continuous basis. Knowing what determines organizational performance is important especially in the context of the current Ghanaian banking business environment because factors identified to be driving performance could be given priority attention in order to improve or maintain the organizational performance. This paper reports on an investigation into what is influencing the performance of a given bank in Ghana. Two hundred and ten (210) respondents were used for the study. Among other things, the study results reveal that there are four dimensions informing the performance of the bank, which accounted for 63.5% of the variance in the original variables. In conclusion, the dimensions adduced to be influencing the performance of the bank were: Repeat purchase, Customer experience, Customer satisfaction and Intelligent responsiveness. Management of the bank are encouraged to focus on these dimensions to ensure that the performance of the bank is at the prime.*

KEYWORDS: Organizational Performance, Factor Analysis, Principal Component Factoring

INTRODUCTION

The recent global economic and financial crises had adverse effect on the banking industry of many a country. Fortunately the Ghanaian banking sector remains buoyant. The performance of banking companies in Ghana had until recently been dictated by market forces and how shares and funds fared. For instance, banks have hitherto not bothered so much about what their customers felt about them. However the financial and banking scene is fast changing as a result of the proliferation of financial and banking institutions, especially with the advent of competition from foreign banks. The dynamics of the current Ghanaian banking scene is not the same as it used to be fifteen years ago. Currently banks chase customers with their products and services and do all they can to entice customers into signing up to these products and services. This was not the case fifteen years ago when customers were even pleading with banks to allow them to open accounts with them. Loans could take months before getting approval then, but now one could get a loan from a bank within forty-eight hours or less. The use of Automatic Teller Machines (ATMs) and Debit/Credit cards are among other products and services that are fast catching up with a considerable proportion of the populace now.

It is therefore imperative for banks to determine what factors impinge on their performance through empirical studies so that they can effectively plan for sustained growth. They could build performance indicators around these factors to provide information regarding how they are faring in respect of the quality of services rendered to their customers.

ISSN 2053-2229 (Print), ISSN 2053-2210 (Online)

This study uses factor analysis to model the organizational performance of a bank in Ghana. The bank under discussion was used as a case study because it is a wholly-owned Ghanaian company which was established as a Savings & Loans Company. It grew to become the largest Savings & Loans Company in Ghana before it transitioned to the status of a universal bank.

In all, two hundred and ten (210) respondents conversant with the operations of the four branches of the bank were involved in the study. A questionnaire including 17 items that employs a five-point differential scale ranging from 'strongly disagree' to 'strongly agree' was administered to the respondent. The resulting data was analysed using Statistical Product and Service Solutions (SPSS 16). The major statistical tool used for the study was factor analysis which has been reviewed below.

REVIEW OF METHODS

Factor analysis (including common factor analysis and principal component analysis) is used to examine the interdependence among variables and to explain the underlying common dimensions (factors) that are responsible for the correlations among the variables. The procedure allows one to condense the information in a large set of variables into a smaller set of variables by identifying variables that are influenced by the same underlying dimensions. We can therefore look upon the underlying dimensions or factors, which are of primary interest but directly unobservable, as the new set of variables. Factor analysis facilitates the transformation from the original observable variables to the new variables (factors) with as little loss of information as possible (Everitt and Dunn, 2001; Johnson and Wichern, 1992; Sharma, 1996).

Given the latent factors $F' = (F_1, F_2, \ . \ . \ . \ , F_m)$ and the observable (indicator) variables $X' = (X_1, X_2, \ . \ . \ . \ X_P)$, where $m \ll p$, the data model, in matrix notation, is given by

$$(X-\mu)_{p\times 1} \ = \ L_{p\times m}\, F_{m\times 1} \ + \ \varepsilon_{p\times 1}$$

where $L_{p\times m}$ is a $p \times m$ matrix of coefficients $l_{ij},\ i = 1, 2, \ . \ . \ . \ , p;\ j = \ 1, 2, \ . \ . \ . \ , m$. The l_{ij}s are referred to as the factor loadings. The entities $\varepsilon' = \left(\varepsilon_1, \varepsilon_2, \ . \ . \ . \ , \ \varepsilon_p\right)$ are thought of as specific error terms or factors associated with $X_1, X_2, \ . \ . \ . \ X_P$ respectively. The mean corrected vector $(X-\mu)_{p\times 1}$, where $\mu = E[X] \ = \ \left(\mu_1, \mu_2, \ . \ . \ . \ , \mu_p\right)'$, is taken to be the response variable (Johnson and Wichern 1992). For an orthogonal factor model, the analysis of the data is done under the assumption that

$$E[F] \ = \ O_{(m\times 1)} \qquad E[\varepsilon] \ = \ O_{(p\times 1)} \qquad Cov(F) = I_{(m\times m)}$$
$$Cov\left(\varepsilon_i, \varepsilon_j\right) = 0 \qquad Cov\left(F_i, \varepsilon_j\right) = 0$$

where O is a matrix of zeros. It follows from the above assumptions that:

i. the factors are independent,
ii. the specific error terms are independent and
iii. the factors and the specific error terms are independent.

However, as mentioned above, the observable variables $X_1,\ X_2,\ . \ . \ . \ ,\ X_P$ are correlated because they are influenced by some common underlying dimensions (factors). The correlation among the indicator variables facilitates the identification of the common latent factors as the indicator variables that are influenced by the same factor tend to 'load' highly on (have a high correlation coefficient with) that common factor and also amongst themselves (Everitt and Dunn 2001; Johnson and Wichern, 1992; Sharma, 1996).

ISSN 2053-2229 (Print), ISSN 2053-2210 (Online)

In the orthogonal model, the coefficients (pattern loadings) l_{ij}, $i = 1, 2, \ldots, p$; $j = 1, 2, \ldots, m$ are the same as the simple correlations (structure loadings) between the indicator variables X_i and the factors F_j, and the variance (communality) that X_i shares with F_j is given by l_{ij}^2 (Sharma, 1996). Thus the total communality of an indicator variable X_i with all the m common factors is given by

$$l_{i1}^2 + l_{i2}^2 + , \ldots , + l_{im}^2 .$$

Principal Component Factoring

Factor analysis can be done using one of several techniques. One of the popular techniques is Principal Component Factoring (PCF), which uses Principal Component Analysis (PCA) to extract the dimensions (factors) influencing the observed variables, by analysing the correlation amongst them. In PCA, new uncorrelated variables (called Principal Components (PC)) are formed which are a linear combination of the original (observable) variables, and the number of new variables is equal to the number of old variables. However, the new variables are so formed that the first principal component accounts for the highest variance in the data, the second principal component accounts for the highest of the remaining variance in the data, the third principal component accounts for highest of the remaining variance not accounted for by the first and second components, and so on. Ideally, one would want a situation where the first few principal components account for much of the variance in the original data, thereby achieving data reduction by replacing the original variables by the first few principal components, for further analysis or interpretation of the correlation amongst the indicator variables (Everitt and Dunn, 2001; Johnson and Wichern, 1992; Sharma, 1996).

Given the observed variables $X_1, X_2, \ldots, X_p$ and the coefficients (weights) w_{ij}, $i = 1, \ldots, p$, $j = 1, \ldots, p$, the principal component $C_1, C_2, \ldots, C_p$ are given by

$$\begin{aligned}
C_1 &= w_{11}X_1 + w_{12}X_2 + \ldots + w_{1p}X_p \\
C_2 &= w_{21}X_1 + w_{22}X_2 + \ldots + w_{2p}X_p \\
&\vdots \\
C_p &= w_{p1}X_1 + w_{p2}X_2 + \ldots + w_{pp}X_p
\end{aligned}$$

To place a limit on the variance of the C_is, $i = 1, \ldots, p$ and to guarantee that the new axes representing the C_is are uncorrelated, the weights w_{ij}, $i = 1, \ldots, p$, $j = 1, \ldots, p$ are estimated subject to the conditions given by Equations 1 and 2 respectively (Everitt and Dunn 2001; Johnson and Wichern, 1992; Sharma, 1996).

$$w_i' \cdot w_i = 1 \qquad \ldots\ldots\ldots 1$$

and

$$w_i' \cdot w_j = 0 \quad \text{for all} \quad i \neq j \qquad \ldots\ldots\ldots 2$$

where

$$w_i' = (w_{i1}, w_{i2}, \ldots, w_{ip})$$

Given the mean μ_i and the standard deviation σ_{ii} of the variable X_i, the transformed variables Z_i, $i = 1, \ldots, p$ given by

$$Z_i = \frac{X_i - \mu_i}{\sigma_{ii}}$$

ISSN 2053-2229 (Print), ISSN 2053-2210 (Online)

could be used to form the principal components (Johnson and Wichern, 1992). Expressed in matrix notation, the vector of standardized variables could be written as

$$Z = \left(V^{1/2}\right)^{-1}(X - \mu)$$

where $\mu' = \left(\mu_1, \mu_2, \ . \ . \ . \ , \mu_p\right)$ and $V^{1/2}$ is the standard deviation matrix given by

$$V^{1/2} = \begin{bmatrix} \sigma_{11} & 0 & \dots & 0 \\ 0 & \sigma_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{pp} \end{bmatrix}$$

$E[Z_i] = 0$, $\mathrm{Var}[Z_i] = 1$, $i = 1, \ . \ . \ . \ , \ p$ and $Cov(Z) = \left(V^{1/2}\right)^{-1}\Sigma\left(V^{1/2}\right)^{-1} = \rho$ where the variance-covariance matrix Σ and the correlation matrix ρ of X are given by

$$\Sigma = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12}^2 & \dots & \sigma_{1p}^2 \\ \sigma_{21}^2 & \sigma_{22} & \dots & \sigma_{2p}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1}^2 & \sigma_{p2}^2 & \dots & \sigma_{pp}^2 \end{bmatrix}$$

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

and

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

is the covariance between variables X_i and X_j, each of which has n observations.

The p principal components $C' = \left[C_1, \ C_2, \ . \ . \ . \ . \ , \ C_p\right]$ are then given by

$$C = A'Z$$

where $A = \left[e_1 \, e_2 \ , \ , \ , \ e_p\right]$ and the e_i s, $i = 1, \ 2, \ . \ . \ . \ , \ p$ are the eigenvectors of ρ. The eigenvalue-eigenvector pairs $(\lambda_1, e_1), (\lambda_2, e_2), \ . \ . \ . \ , \ (\lambda_p, e_p)$ of ρ are such that $\lambda_1 \geq \lambda_2 \geq \ . \ . \ . \geq \lambda_p \geq 0$, $e_i' \cdot e_i = 1$ and $e_i' \cdot e_j = 0$.

$$\mathrm{Var}(\,C_i\,) = e_i'\rho e_i = \lambda_i$$

and

$$\sum_{i=1}^{p} Var(\,C_i\,) = \sum_{i=1}^{p} Var(\,Z_i\,) = p$$

ISSN 2053-2229 (Print), ISSN 2053-2210 (Online)

Thus the proportion of the variance in the data that is accounted for by the C_j is given by λ_j/p.
The correlations between a given PC C_i and a given standardized variable Z_j, referred to as the loading of variable Z_j on C_i, is given by

$$Corr(C_i, Z_j) = e_{ij} \cdot \lambda_j^{(1/2)}$$

The loading reflects the degree to which each Z_j influences each C_i given the effect of the other variables $Z_k, j \neq k$ (Hair et al., 2006; Johnson and Wichern, 1992; Sharma, 1996).

In Principal Component Factoring the initial communalities of the indicator variables are one.

RESULTS

Table 1: KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.719
Bartlett's Test of Sphericity	Approx. Chi-Square	494.786
	Df	55
	Sig.	.000

Source: Results from analysis of data, 2014.

Table 1 shows the KMO measure and the results of the Bartlett's test, which are used to judge the adequacy of the sample size and whether or not the correlation matrix is suitable for factor analysis respectively. Both the KMO value of 0.72 (which is greater than the minimum threshold of 0.50 (Sharma, 1996)) and the Bartlett's test p value of 0.000 suggest the sample size is adequate and that, at least, some of the variables are inter-correlated and therefore the data is suitable for factor analysis.

Table 2: Communalities

Variable	Initial	Extraction	Variable	Initial	Extraction
My bank's products are suited to my needs	1.000	.500	I will continue to operate with the bank because that is where people often operate	1.000	.549
Statements are sent to me on time and are accurate	1.000	.873	I will continue operate with the bank because the bank's staffs are able to handle most of my questions satisfactorily	1.000	.668
I am usually content with the services I get from the bank	1.000	.538	I will continue to operate with the bank because of their customer loyalty schemes	1.000	.680
The bank usually communicates to me proactively	1.000	.637	I will continue to operate with the bank because of special likeness I have for them	1.000	.670

ISSN 2053-2229 (Print), ISSN 2053-2210 (Online)

The bank's staff are able to handle most of my questions satisfactorily	1.000	.688	I will continue to operate with the bank because of their efficient services	1.000	.585
The bank's staff understands my specific needs for financial services	1.000	.603			

Source: Results from analysis of data, 2014.

The initial and final (Extraction) communalities are shown in Table 2. In PCF, all variables are assigned an initial variance (total communality) of one, as indicated in the Section 2. The final (Extraction) communalities of each variable represent the variance accounted for by the chosen factor solution for the variable. Eleven variables remained in the final factor solution out of the 17 variables at the start of the analysis. The other six were removed from the analysis because their communalities were less than 0.50 or they were cross-loading (loading on more than one factor) in the preliminary analysis. From Table 2, all the final communalities are at least 0.50. At least 50% of the initial communality of each variable was accounted for in the final factor solution. The factor solution is thus far considered to be satisfactory as at least half of the variance of each variable is shared with the factors.

Table 3: Total Variance Explained

	Initial Eigenvalues		
Component	Total	% of Variance	Cumulative %
1	3.105	28.232	28.232
2	1.605	14.593	42.825
3	1.250	11.365	54.190
4	**1.029**	**9.357**	**63.547**
5	.868	7.887	71.434
6	.777	7.062	78.496
7	.648	5.889	84.385
8	.560	5.092	89.477
9	.427	3.886	93.363
10	.389	3.534	96.897
11	.341	3.103	100.000

Source: Results from analysis of data, 2014.

ISSN 2053-2229 (Print), ISSN 2053-2210 (Online)

International Journal of Mathematics and Statistics Studies

Vol.3,No.1, pp.38-46, January 2015

Published by European Centre for Research Training and Development UK (www.eajournals.org)

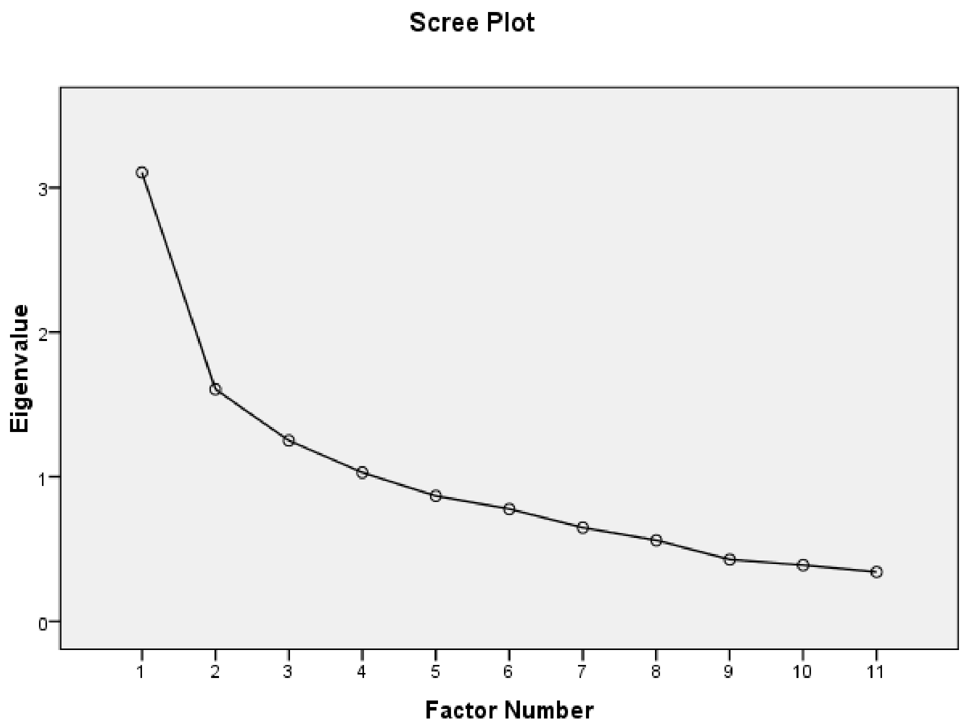

Figure 1: Plot of eigenvalues against factor number.

Number of factors extracted

Three criteria were used to decide on the number of factors to retain for interpretation: eigenvalue-greater-one rule, scree plot and the percentage of variance explained. Four components have eigenvalues greater than one, so going by eigenvalue-greater-one rule, four factors can be retained for interpretation. The scree plot of Figure 1 suggest extracting five factors as the plot begins to straighten out after the fifth component. The first four components explain 63.5% of the variance in the data -- more than the suggested 60% threshold (Hair et al., 2006), while the first five components account for 71.4% of the variance in the data. However the fifth component has eigenvalue less than 1 (0.868). Putting the results from three criteria together, one is inclined to retain four factors for interpretation. The four factors retained for interpretation are shown in Table 4.

ISSN 2053-2229 (Print), ISSN 2053-2210 (Online)

Table 4: Rotated Component Matrix

Variable	Component 1	2	3	4
I will continue to operate with the bank because of their customer loyalty schemes	.804			
I will continue operate with the bank because the bank's staffs are able to handle most of my questions satisfactorily	.767			
I will continue to operate with the bank because that is where people often operate	.713			
I will continue to operate with the bank because of special likeness I have for them	.665			
I will continue to operate with the bank because of their efficient services		.692		
My bank's products are suited to my needs		.689		
I am usually content with the services I get from the bank		.667		
The bank's staff understands my specific needs for financial services			.760	
The bank's staff is able to handle most of my questions satisfactorily			.739	
The bank usually communicates to me proactively			.507	
Statements are sent to me on time and are accurate				.933

Source: Results from analysis of data, 2014.

Table 4 shows the extracted factors and the loadings of the various variables on the factors, after the initial factor solution had been rotated via the Varimax method so that each variable load highly on only one factor. The method also ensures that the factors are uncorrelated. As indicated in Section 2, the loadings represent the extent of correlation between a variable and a factor. The higher the absolute value of a loading of variable on a factor, the more influential the variable is on the factor. A loading of 0.40 is considered significant for a sample size of 210 (Hair et al., 2006). However, a higher value of 0.50 was used to ensure that only variables of practical significance are included in the final factor solution. Loadings below 0.50 were omitted with the remaining ones sorted in descending order of magnitude to facilitate easy interpretation of the final factor solution. The factors (Components 1, 2, 3 and 4 (Table 4)) are presumed to be the underlying dimensions informing performance of the bank. The factors were named based on the loadings of the variables shown so that the higher the absolute value of a variable's loading on a factor, the more influential the variable is in naming the factor. The factors were named as follows:

Factor 1: Repeat purchase
Factor 2: Customer experience
Factor 3: Customer satisfaction
Factor 4: Intelligent responsiveness

CONCLUSION

For an organization to be successful, it has to proactively examine all facets of it operations periodically with the view of identifying areas that needs strengthening and to fashion strategies for achieving that. This is especially true of an organization such as a bank, whose core activities involve directly interfacing with customers whose expectations may be dictated by the ever-changing economic environment. Factor analysis is a tool that can be used to uncover the latent factors that influence the success or otherwise of a business entity operating in a competitive business environment such as that of the banking industry of Ghana.

ISSN 2053-2229 (Print), ISSN 2053-2210 (Online)

This paper reports on the results of an investigation aimed at understanding what is influencing the performance of a given bank in Ghana. A questionnaire including 17 items that employs a five-point differential scale was administered to a sample of 210 respondents who are familiar with the operations of the bank and the resulting data was subjected to factor analysis. The KMO value of 0.72 and the p-value of 0.000 for the Bartlett's test of sphericity meant that the data was suitable for factoring. Using a combination of three criteria to decide on the number of factors to retain for interpretation, a four factor solution was arrived at, which accounted for 63.5% of the original variance in the data. A minimum threshold of a loading of 0.50 was used to include only variables that are of practical significance in the final factor solution, which was arrived at by rotating the initial factor solution by the Varimax method, thereby ensuring that the factors in the final factor solution are uncorrelated. The dimensions adduced to be influencing the performance of the bank are: Repeat purchase, Customer experience, Customer satisfaction and Intelligent responsiveness. Management of the bank can focus on these dimensions as they strive to ensure the performance of the bank is optimal.

REFERENCES

Everitt, B. S. and Dunn, G. (2001), *Applied Multivariate Data Analysis*. London, UK: Arnold/Hodder Headline Group.

Hair, J. F. et al. (2006), *Multivariate Data Analysis*. 6th ed. Upper Saddle River, New Jersey, USA: Pearson Prentice Hall.

Johnson, R. A. and Wichern, D. W. (1992), *Applied Multivariate Statistical Analysis*. 3rd ed. New Jersey, USA: Prentice-Hall, Inc.

Sharma, Subhash (1996), *Applied Multivariate Techniques*. USA: John Wiley and Sons, Inc.

ISSN 2053-2229 (Print), ISSN 2053-2210 (Online)