No.7-2021

Financial Sector Development Studies

Bank Failure Prediction in the Arab Region Using Logistic Regression Model

Rami Obeid

صندوق النقد العربي
ARAB MONETARY FUND

صندوق النقد العربي

ARAB MONETARY FUND

Bank Failure Prediction in the Arab Region Using Logistic Regression Model

Rami OBEID

Arab Monetary Fund
May 2021

Bank Failure Prediction in the Arab Region Using Logistic Regression Model

These economic studies are the product of the staff of the Economic Department at the AMF. The AMF publishes these studies which survey issues pertinent to monetary, fiscal, banking, trade and capital market policies and their impact on Arab economies.

All correspondence should be addressed to:
Economic Department
Arab Monetary Fund
P.O. Box 2818
United Arab Emirates
Telephone No.: +9712-6171552
Fax No: +9712-6326454
E-Mail: economic@amfad.org.ae
Website: www.amf.org.ae

Contents

Abstract 5
Introduction 6
1. Literature review 7
2. Model and data 9
3. Econometric methodology 10
4. Analysis of the results 12
4.1. Descriptive analysis 12
4.3. Bank Failure Prediction 14
5. Conclusion and policy recommendations 16
References 18

List of Tables

Table 1. Categorical Descriptive Statistics for Explanatory Variables 21
Table 2 - Summary statistics of the variables 22
Table 3 - Expectation-Prediction Evaluation for Binary Specification 23
Table 4 - Goodness-of-Fit Evaluation for Binary Specification 24
Table 5 - Logistic Regression Results 25

Bank Failure Prediction in the Arab Region Using Logistic Regression Model

Abstract

Interest increased after the last global financial crisis in 2007, as the crisis revealed the challenges facing central banks in dealing with the financial crises arising from the banking sector, and thus exposing the financial positions of the banking sector to high risks that caused bankruptcy of many commercial banks. This prompted the international financial institutions to develop regulatory requirements and provide more instruments that reduce the costs of resolving the crisis. As the interest has emerged in the necessity of using the predictive aspect of the bank's performance, and not waiting for the problem to occur to be dealt with, this approach will lead to reducing the high cost of the solution in terms of early intervention of the central bank.

In this study, we tried to build a logistic regression model to predict the bankruptcy of the bank, based on the data of 40 commercial banks in the Arab region for the period 2005-2015, where the financial soundness indicators were used to predict the bankruptcy of the bank, in addition to the GDP variable in order to capture the impact of economic risks. The study confirmed the findings of previous studies in terms of the ability of the financial soundness indicators used in the (CAMEL) system to predict bank bankruptcy. While the results did not show a significant statistical impact of the variable GDP.

Bank Failure Prediction in the Arab Region Using Logistic Regression Model

Introduction

The global financial crisis of 2007 revealed an acute shortage of tools available to effectively deal with bank failures. This caused several shocks in all components of the financial system, whose effects appeared on the real economy. The competent supervisory authorities had limited options at their disposal to prevent the difficulties of some institutions from threatening financial stability and exposing it to risks. These challenges prompted countries to use taxpayers' money to save these institutions.

Given this essential role and the absence of effective resolution systems, the financial authorities found it necessary at that time to use taxpayer money to restore confidence in the banking system and avoid further systemic damage during the crisis. Governments have intervened on a large scale to restore financial stability, implementing additional measures to reduce the severity of the problems arising from the failure of systemically important global financial institutions.

Therefore, the financial authorities concluded the need to create a crisis resolution system that provides them with a set of reliable tools that enable them to intervene early and quickly in a weak or failed institution, to ensure the continuity of its necessary financial and economic functions, while minimizing the impact of its failure on the economy and the financial system. Consequently, the central banks concluded that dealing with financial crises should not start when they occur, but rather the failure of the financial institution must be predicted according to early warning systems methodologies. This study will provide a methodology for predicting bank failure in the Arab region based on variables related to the (CAMEL) rating for the banking sector. The study sample included 40 commercial banks from ten Arab countries, based on published disclosures of the annual banks' balance sheets during the period (2005-2015).

1. Literature review

Many previous studies attempted to predict the failure of companies using different methodologies, and one of the most famous researchers in this field was Altman (1968) who conducted a Multiple Discriminant Analysis on the data of 33 bankrupt and similarly bankrupt industrial companies. The study showed that the following five ratios among 22 ratios are the most significant variables to predict the bankruptcy: (1) Working capital / Total assets, (2) Retained earnings / Total assets, (3) Earnings before interest and taxes / Total assets, (4) Book value of equity / Book value of liabilities, (5) Sales / Total assets , the results showed that the model used in the study is effective in predicting bankruptcy even two years before the crisis, but its accuracy decreases dramatically as the period increases. This model was then subsequently improved by Altman and Narayanan (1997) by proposing the zeta model that includes seven variables and classified correctly 96% of companies one year before bankruptcy.

Most studies achieved after 1980 used the logit models to overcome the defects of the Multiple Discriminant Analysis method (See Zavgren (1983); Tennyson et al. (1990)...). The logit analysis fits linear logistic regression model by the method of maximum likelihood.

Zaghdoudi (2013) has attempted to build an early warning model to predict the bank failures in Tunisia with the contribution of the logistic regression method. The results showed that a bank's ability to repay its debt, bank profitability per employee, the coefficient of banking operations and leverage ratio has a negative impact on the probability of failure. Maricica and Georgeta (2012) argue that financial ratios can be used as an early warning signals to discriminate between survival and failed firms.

Baker (2018) has identified the determinants of bank failures in Jamaica by analyzing the financial ratios based on CAMELS rating. The resulting scoring model has been used as a tool to assess potential risks to financial stability. The study has used the logistic regression methodology based on quarterly data over the period (3/2008-12/2017). The main results showed that the non-performing loans ratio and the capital adequacy ratio are significantly and positively related to the probability of default of Jamaican banking sector.

2. Model and data

The study will examine the factors affecting the likelihood of bank failure and provide an analytical framework as an early warning system, based on a (CAMEL) rating of the banking sector, furthermore, we will use GDP variable to capture the impact of the economic risks. The data has obtained from published disclosures about the results of banks' annual balance sheets during the period (2005-2015).

Analytically, the logit analysis fits linear logistic regression model by the method of maximum likelihood will be used in this study. The model takes the following form:

$$ln\left(\frac{P_{it}}{1-P_{it}}\right) = Z_{it} = \beta_0 + \beta_1 CAR_{it} + \beta_2 CIR_{it} + \beta_3 NPL_{it} + \beta_4 ROA_{it} + \beta_5 LIQ_{it} + \beta_6 GDP_{it} + \varepsilon it \qquad (1)$$

$i = 1, 2, \dots, N,\ \ t = 1, 2, \dots, T$

$$P_{it}(Y|X) = \frac{1}{1+e^{-Z_{it}}}, P\epsilon[0,1] \qquad (2)$$

Where i refers to bank and t refers to time, Z_{it} Linear regression extracted from the variables that will be used to predict bank failure. P_{it} is the probability of (Y_i=0) for a bank with low risk of failure or bankruptcy, and the probability of (Y_i=1) in case the bank is at risk of failure or bankruptcy. CAR_{it} is the capital adequacy ratio, CIR_{it} is the cost income ratio, NPL_{it} is the nonperforming loans, ROA_{it} is the return on assets, LIQ_{it} is the liquid assets ratio, and GDP_{it} is the real GDP growth rate, and ε_{it} is the disturbance term.

3. Econometric methodology

In this study, we will use the logit model which is considered as one of the most applied parametric failure prediction models, especially in bank failure prediction (See Dumirguc-Kunt and Detragiache, 1998). This study will attempt to analyses the factors which explain the banks weaknesses in the Arab region based on the significant fragility factors. In line with the literature. The study will be using the financial ratios which can predict potential failures (See Beaver (1966), Altman (1968) and Jagtiani et al. (2003)). Logit model is based on a binomial regression, and it also based on the estimation of the probability of failure P(Z). This probability is defined as a linear function of a vector of covariates Zi and a vector of regression coefficients βi:

$$Z_{it} = \beta_0 + \beta_1 X_{i,t-k} \qquad (3)$$

The following logistic model can be used to express probability of survival (Y = 0) or bankruptcy (Y = 1):

$$P_{i,t}(y) = \begin{cases} P(Z) = \frac{1}{1+e^{-(\beta_0+\beta_1 X_{i,t-k})}}, & y = 1 \\ 1 - P(Z) = 1 - \frac{1}{1+e^{-(\beta_0+\beta_1 X_{i,t-k})}}, & y = 0 \end{cases} \qquad (4)$$

Where:

$P_{i,t}$: Probability that bank (i) will fail at the time (t), P∈[0,1].

X_i: The variables that will be used to predict bank failure.

Z_{it}: Linear regression extracted from the variables that will be used to predict bank failure, Z∈(-∞,∞).

k: The year before the bank failure period.

e: Natural logarithm.

β: Coefficients of the regression.

To calculate the scores, we can use the generalized linear method which used by Nelder and Wedderburn (1972). Then, the estimation of the parameters $\beta_0,\ldots,\beta_d$ to be obtained by maximum likelihood function of the Lagrangian as following:

$$Max\ L(Y|X,b) \qquad (5)$$

We can solve the intercept and β `s by the following equation:

$$L(Y|X,b) = P(Y|X), where: P(Y|X) = \prod_{j=1}^{n}[P_i^{Yi}(1-P_i)^{1-Yi} \qquad (6)$$

So:

$$Log\ LP(Y|X,b) =$$

$$\prod_{j=1}^{n}[Y_j log(\frac{1}{1+e^{-(\beta_0+\beta_1 X_{i,t-k})}}) + (1-Y_j) log(\frac{1}{1+e^{-(\beta_0+\beta_1 X_{i,t-k})}})] \quad (7)$$

Where:

P: Probability that bank (i) will fail at the time (t), P∈[0,1].

X_i: The variables that will be used to predict bank failure.

k: The year before the bank failure period.

b: the intercept and coefficients.

β: Coefficients of the regression.

Finally, we have to calculate the Akaike Information Criterion (AIC) in order to choose the most important performing explicative variables to predict bank failure. The AIC formula as following:

$$AIC= -2LOGL+2P \qquad (8)$$

Where:

L: The maximum likelihood of the fitted model.

P: the number of estimated parameters.

4. Analysis of the results

4.1. Descriptive analysis

Table No. (2) shows the descriptive data of the study variables, where it is noticed that the minimum value of some variables reached zero, due to the bankruptcy of some of the study sample banks in a specific period of time, as well as the minimum value of the (ROA) variable about -5.1%, and this value indicates mainly that some banks were exposed to losses that resulted either from a weakness in the operational efficiency and / or the presence of weak risk management and a high desire for risk not based on accurate evaluation of customers. As for the non-performing loans ratio, it reached the lowest value of about 1.1%, which indicates the existence of high efficiency in managing credit risk. On the other hand, the highest value of the capital adequacy ratio reached about 45%, which is very high if compared to the requirements of Basel, which indicates the bank's ability to absorb potential shocks, but at the same time high ratios of levels of capital adequacy may indicate weakness in the effectiveness of financial intermediation, and the bank's exaggeration in not granting loans, taking into account that newly licensed banks usually have high capital adequacy, as they are still in the stage of preparing to enter the market and attract customers. As for the ratio of liquid assets to total assets, it reached the highest value of 96.1%, and it is known that the higher this ratio, this indicates the bank's ability to fulfill its short-term obligations.

Regarding the cost to income ratio, it reached the highest value of about 200%, and this indicates a clear weakness in operational efficiency, which leads to achieve losses, but this value was achieved from banks that suffered from weakness in their financial positions, or banks that went bankrupt and were liquidated. As the cost of liquidation is usually expensive, especially in the event of a weak crisis management system. As for the rate of return on assets, the highest value was 4.7%, and this indicates the bank's ability to generate profits from the assets.

Finally, the highest value of the non-performing loans ratio reached 21%, and this percentage indicates a clear weakness in credit risk management, which exposes the bank to high bankruptcy risks, in the absence of a comprehensive corrective plan with the bank under the supervision of the central bank.

4.2. Bank Failure Prediction

The results of the logistic regression model analysis showed that the elements of (CAMEL) classification play a prominent role in assessing the financial strength of the bank, and have a clear ability to predict bank bankruptcy, as it is noted from Table (5) that the capital adequacy variable has a positive and statistically significant effect in reducing the possibility of bank default. As the capital adequacy coefficient reached about 1.29, and this is consistent with economic theory, since capital adequacy enhances the bank's ability to absorb potential shocks and losses and maintains the integrity of the bank's financial positions. That is, the results indicated that increasing the percentage of CAR by one unit reduces the probability of the bank failure by 1.3 unit This result explains why the Basel Committee on Banking Supervision was interested in enhancing the quality of capital requirements after the global financial crisis in 2008. These requirements have been raised in quantity and quality, so that the requirements of the tier one capital to be met by common equity tier one.

Regarding the percentage of high-quality assets, the results showed that there is a positive significant relationship between this variable and the probability that the bank will not go bankrupt and remain viable. Whereas the presence of high levels of liquidity (liquid assets) keeps the bank continuously able to fulfill its obligations, the results indicated that increasing the percentage of high-quality assets by one unit reduces the probability of the bank failure by 0.17 units. It should be noted that the Basel Committee on Banking Supervision has strengthened its interest in liquidity requirements after the global financial crisis, as it has set, for the first time, quantitative and qualitative requirements for short-term and long-term liquidity, by launching the ratios of liquidity coverage (LCR) and net stable funding (NSFR) in accordance with the requirements of Basel 3.

The results also showed that there is a negative significant relationship at 5% between the cost-to-income ratio and the bank's survival probability, as an increase in the CIR ratio by one unit leads to a decrease in the bank's survival probability by 0.06 units. This result shows the important role of the bank's management in controlling the bank's expenses in a way that generates more profits. That is, the bank's operational efficiency and effectiveness in financial intermediation enhance the robustness and health of its financial positions.

As for the asset quality factor, the results showed a negative significant relationship between the percentage of non-performing loans and the probability of the bank surviving, as a rise in the non-performing debt ratio by one unit without adequate provisions leads to an increase in the probability of default of the bank by 0.6 units. The supervisory authorities have been keen to monitor this ratio continuously after the global financial crisis in 2007, and to urge commercial banks to build adequate provisions to face any potential credit risks. In addition, stress tests were developed to measure the potential impact of the non-performing facilities ratio on the bank's financial position. It is also expected that IFRS 9 will contribute to enhancing the volume of provisions at the bank, as provisions are being built after taking the predictive aspect of clients ’default, not just the realized losses.

Finally, regarding the rate of return on assets, the results showed that there is a positive significant relationship between this variable and the bank’s likelihood of survival, and this is a logical result, as achieving profits indicates the efficiency of banks in employing its assets, and making profits enhances the bank’s ability to build its capital. As well as provisions for any potential or sudden risks.

It is worth noting, however, that the variable of GDP growth has any significant statistical function on the bank’s default or survival.

5. Conclusion and policy recommendations

The study investigated the possibility of building a logistic regression model that predicts the stumbling of banks in the Arab region, during the period (2005-2015), as data for 40 banks in the Arab region were used, and it was mainly dependent on data from (CAMEL) indicators which already used in the financial soundness indicators. The results confirmed the findings of previous studies in terms of the high predictive power of financial soundness indicators in predicting bank bankruptcy. Accordingly, the study recommends the following:

- The necessity for central banks to continue to adopt an effective monitoring and inspection plans for banks in accordance with international best practices.
- Analyzing the financial soundness indicators of the banking sector on an ongoing basis, and submitting periodic reports to the management of the Central Bank, including the appropriate recommendations.
- Strengthening the crisis management system at the Central Bank, and the recovery plans of commercial banks.
- Building early warning systems at the central bank so that the appropriate indicators are chosen to predict weak banks.
- Development of micro and macro and stress tests, taking into account the predictive aspect.

- Application of international standards related to Basel 3 requirements and IFRS 9.

References

Annual Reports for Commercial Banks.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. The journal of finance, 23(4), 589-609.

Altman, E. I. and Narayanan, P. (1997) An International Survey of Business Failure Classification Models – Financial Markets, Institutions and Instruments, 6(2), May, New York University Salomon Center.

Baker, C. (2018), Predicting Bank Failures in Jamaica: A Logistic Regression Approach, Working Paper, Bank of Jamaica.

Beaver, W. (1966), "Financial ratios as predictors of Failure", Journal of Accounting Research, Vol. 4, Empirical Research in Accounting: Selected Studies, pp. 71-111.

Demirguc-Kunt, A., Detragiache, E. (1998), The Determinants of Banking Crises in Developing and Developed Countries. *IFM staff papers*, 45(1), 81-109.

Jagtiani, J., J. Kolari, C. Lemieux, and H. Shin,2002, "Looking for trouble: Early detection of inadequate capitalization of U.S. commercial banks," International Company and Commercial Law Review, Vol. 13, No. 7, July, pp. 269–279.

Tennyson, B.M., Ingram, R.W. and Dugan, M.T. (1990), ``Assessing the information content of narrative disclosures in explaining bankruptcy", Journal of Business Finance and Accounting, Vol. 17 No. 3, Summer, pp. 390-410.

Maricica, M., & Georgeta, V. (2012). Business failure risk analysis using financial ratios. Procedia-Social and Behavioral Sciences, 62, 728-732.

Nelder, J.A. and Wedderburn, R.W.M. (1972) Generalized Linear Models. Journal of the Royal Statistical Society, Series A, 135, 370-384.

Zaghdoudi, T. (2013). Bank failure prediction with logistic regression. International Journal of Economics and Financial Issues, 3(2), 537-543.

Zavgren, C. (1983). The prediction of corporate failure: the state of the art. Journal of Accounting Literature, 2(1), 1-38.

Table 1. Categorical Descriptive Statistics for Explanatory Variables

Variable	Dep=0	Mean Dep=1	All
C	1.000000	1.000000	1.000000
CAR	6.254875	16.00462	15.11828
LIQ	18.41345	61.20295	57.31300
CIR	116.5298	63.52171	68.34062
ROA	-2.001850	1.236182	0.941816
Variable	Dep=0	Standard Deviation Dep=1	All
C	0.000000	0.000000	0.000000
CAR	3.771647	3.425974	4.450335
LIQ	17.80131	16.60494	20.74644
CIR	31.18340	98.83526	95.90351
ROA	1.460821	0.652530	1.202110
Observations	40	400	440

Table 2. Summary statistics of the variables

	CAR	LIQ	CIR	ROA	NPL
Mean	15.11828	57.31300	68.34062	0.941816	7.129773
Median	15.66100	59.34202	58.14605	1.195250	5.600000
Maximum	45.12631	96.13044	200.0000	4.700000	21.00000
Minimum	0.000000	0.000000	0.000000	-5.100000	1.100000
Std. Dev.	4.450335	20.74644	25.90351	1.202110	5.327502
Skewness	0.045932	-0.592419	1.92180	-2.186341	1.207960
Kurtosis	9.964909	3.049711	8.17201	10.34991	3.703739
Jarque-Bera	889.5039	25.78237	791.7040	1340.929	116.0851
Probability	0.000000	0.000003	0.000000	0.000000	0.000000
Sum	6652.043	25217.72	28263.1	414.3988	3137.100
Sum Sq. Dev.	8694.606	188952.1	4037695.	634.3855	12459.82
Observations	440	440	440	440	440

Table 3. Expectation-Prediction Evaluation for Binary Specification
Equation: EQ01
Success cutoff: C = 0.5

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)<=C	39	1	40	0	0	0
P(Dep=1)>C	1	399	400	40	400	440
Total	40	400	440	40	400	440
Correct	39	399	438	0	400	400
% Correct	97.50	99.75	99.55	0.00	100.00	90.91
% Incorrect	2.50	0.25	0.45	100.00	0.00	9.09
Total Gain*	97.50	-0.25	8.64			
Percent Gain**	97.50	NA	95.00			

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
E(# of Dep=0)	38.32	1.68	40.00	3.64	36.36	40.00
E(# of Dep=1)	1.68	398.32	400.00	36.36	363.64	400.00
Total	40.00	400.00	440.00	40.00	400.00	440.00
Correct	38.32	398.32	436.64	3.64	363.64	367.27
% Correct	95.80	99.58	99.24	9.09	90.91	83.47
% Incorrect	4.20	0.42	0.76	90.91	9.09	16.53
Total Gain*	86.71	8.67	15.77			
Percent Gain**	95.38	95.38	95.38			

*Change “%Correct” from default (constant probability) specification.

**Percent of incorrect (default) prediction corrected by equation.

Table 4. Goodness-of-Fit Evaluation for Binary Specification

Andrews and Hosmer-Lemeshow

Grouping based upon predicted risk (randomize ties)

	Quantile of Risk		Dep=0		Dep=1		Total	H-L
	Low	High	Actual	Expect	Actual	Expect	Obs	Value
1	0.5050	0.8716	35	34.8874	9	8.9126	44	0.00463
2	0.8745	0.9467	0	0.25075	44	43.7492	44	0.03190
3	0.9470	0.9578	2	2.07178	42	41.9282	44	0.00011
4	0.9579	0.9676	0	0.01327	44	43.9867	44	1.5E-05
5	0.9676	0.9735	3	2.98637	41	40.8136	44	2.9E-06
6	0.9735	0.9802	0	0.00122	44	43.9988	44	6.3E-07
7	0.9805	0.9835	0	0.00048	44	43.9095	44	1.2E-07
8	0.9837	0.9880	0	0.00352	44	43.9665	44	3.4E-08
9	0.9882	0.9919	0	0.00607	44	43.9539	44	5.7E-09
10	0.9920	0.9992	0	0.00416	44	43.9558	44	2.8E-10
		Total	40	40.0000	400	400.000	440	0.03665

H-L Statistic	69.7670	Prob. Chi-Sq(8)	0.0000
Andrews Statistic	272.2089	Prob. Chi-Sq(10)	0.0000

Table 5. Logistic Regression Results

	Estimate	Standard Error
CONSTANT	-11.388**	4.455
NPL	-0.601***	0.229
CAR	1.287***	0.427
LIQ	0.177**	0.069
ROA	0.579*	2.381
CIR	-0.057*	0.088
GDP	0.038	1.209
McFadden R-squared	0.9215	
Akaike info criterion	0.071	
LR statistics	247.039	
	[0.000]	

Notes: The values in brackets are the *p*-values of the tests. *** , ** and * denote significance at the 1%, 5% and 10% levels, respectively.

http://www.amf.org.ae

صندوق النقد العربي
ARAB MONETARY FUND