ADBI Working Paper Series

**FAIR PREMIUM RATE OF THE DEPOSIT
INSURANCE SYSTEM BASED ON
BANKS' CREDITWORTHINESS**

N. Yoshino, F. Taghizadeh-
Hesary, and F. Nili

Asian Development Bank Institute

Naoyuki Yoshino is the dean of the Asian Development Bank Institute. Farhad Taghizadeh-Hesary is an assistant professor of economics at Keio University, and a visiting professor at the Graduate School of Economics, University of Tokyo, Japan. Farhad Nili is senior advisor to the executive director of the World Bank Group.

The Working Paper series is a continuation of the formerly named Discussion Paper series; the numbering of the papers continued without interruption or change. ADBI's working papers reflect initial ideas on a topic and are posted online for discussion. ADBI encourages readers to post their comments on the main page for each working paper (given in the citation below). Some working papers may develop into other forms of publication.

Abstract

Purpose
Deposit insurance is a key element in modern banking, as it guarantees the financial safety of deposits at depository financial institutions. It is necessary to have at least a dual fair premium rate system based on the creditworthiness of financial institutions, as considering a singular premium system for all banks will have a moral hazard. In this paper, we develop a theoretical as well as an empirical model for calculating dual fair premium rates.

Design/methodology/approach
Our definition of a fair premium rate in this paper is a rate that can cover the operational expenditures of the deposit insuring organization, provides it with sufficient funds to enable it to pay a certain percentage share of deposit amounts to depositors in the case of bank default, and provides it with sufficient funds as precautionary reserves. To identify and classify healthier and more stable banks, we use credit rating methods that employ two major dimensional reduction techniques. For forecasting nonperforming loans (NPLs), we develop a model that can capture both macro shocks and idiosyncratic shocks to financial institutions in a vector error correction model (VECM).

Findings
The response of NPLs/loans to macro shocks and idiosyncratic innovations shows that using a model with macro variables only is insufficient, as it is possible that under favorable economic conditions some banks perform negatively for bank-level reasons such as mismanagement, or vice versa. Final results show that deposit insurance premium rates need to vary in relation to banks' creditworthiness.

Value
The results provide interesting insight for financial authorities to assist them in setting fair deposit insurance premium rates. A high premium rate reduces the capital adequacy of individual financial institutions, which endangers the stability of the financial system; a low premium rate reduces the security of the financial system.

Keywords: deposit insurance premium rate, forecasting nonperforming loans, idiosyncratic shocks

JEL Classification: G28, G21, E44

Contents

1. INTRODUCTION

Since the start of the recent global financial crisis, triggered by the collapse of Lehman Brothers in September 2008, there has been an ongoing international debate about the reform of financial regulation and supervision intended to prevent the recurrence of a similar crisis. Strengthening deposit insurance systems is one of the fundamental steps in this reform. Deposit insurance is a key element in modern banking, as it guarantees the financial safety of deposits at depository financial institutions. If an insured depository institution fails to fulfill its obligations to its depositors, the insuring agency will step in to honor the principal and accrued interests up to a predetermined ceiling. An important issue under this system is how to price deposit insurance (Horvitz 1983; Kane 1986; Yoshino, Taghizadeh-Hesary, and Nili 2013). For determining fair premium rates to be paid by depository financial institutions to the insuring agency, the consensus method tends toward the adoption of a risk-based deposit insurance scheme according to bank defaults. To achieve this goal, several models for assessing bank defaults have been proposed: Buser, Chen, and Kane (1981); Acharya and Dreyfus (1989); Bartholdy, Boyle, and Stover (2003); and, more recently, Yoshino and Hirano (2011).

However, in the literature on banking and finance we have found only a few studies dealing with the deposit insurance system (Horvitz 1983; Hwang, Lee, and Liaw 1997; Inakura and Shimizutani 2010; Yoshino, Taghizadeh-Hesary, and Nili 2013) and hardly any studies on how to estimate and forecast fair premium rates for deposit insurance. In one of the most recent studies, Yoshino, Taghizadeh-Hesary, and Nili (2013) provide a model for calculating fair premium rates for the deposit insurance system. Using this model, they estimate the fair premium rate for the deposit insurance system of Japan and find that it is much higher than the actual current premium rate in that country. In another study, they conclude that, to secure financial stability, Japan needs to raise the deposit insurance premium rate. It is crucial for each country to set fair premium rates to maintain financial system stability, thereby protecting depositors and ensuring an appropriate settlement of funds when financial institutions fail.

In this paper, a fair rate refers to a rate that covers the operational expenditures of an insuring agency (e.g. personnel costs and equipment costs) and provides it with sufficient funds to financially assist any failed depository financial institutions. The insuring agency is also obliged to keep adequate precautionary reserves at the end of each financial period to secure itself against further possible failures. A high premium rate reduces the capital adequacy of individual financial institutions, which endangers the stability of the financial system; a low premium rate reduces the security of the financial system.

In this paper, we expand the model first introduced by Yoshino, Taghizadeh-Hesary, and Nili (2013) for estimating one fair premium rate for the whole deposit insurance system. We conclude that many countries need to adopt a system that uses more than one fair premium rate. Depending on the soundness and stability of banks, varying premium rates should be adopted as it is unfair for all banks, irrespective of whether they are healthy or unhealthy, to pay the same premium rate to the insuring agency. Unsound and riskier banks that endanger the stability of the financial system should pay higher premiums than healthy banks and financial institutions that keep their nonperforming loans (NPLs) at adequate levels and perform well financially. Hence, it is necessary to have at least a dual fair premium rate system, which is the main argument of this paper. In Section 2 of this paper, we present a model for calculating dual fair premium rates that would allow healthier banks to pay a lower rate. For this

purpose, we need to have a mechanism for credit rating and classification of banks based on their financial soundness, which is presented in Section 3. For our empirical analysis, presented in Section 4, we use the deposit insurance system of an Asian economy that is currently in the process of establishing a deposit insurance system.

1.1 Who Pays the Deposit Insurance Premium?

As its name suggests, a deposit insurance system is intended primarily to provide for the payment of insurance claims when an insurable contingency occurs. Specifically, there are two methods of protection: the insurance payout method, whereby insurance payouts are made directly to depositors; and a method whereby the business of a failed financial institution is transferred to a different financial institution, and the deposit insurance agency or corporation (DIC) provides assistance to this second institution.

When checking the DIC websites of various countries, we typically find a sentence along the following lines: *"You [depositors] do not pay for the deposit insurance. Financial institutions that are a member of our deposit insurance system pay premiums to us."* Although member banks or financial institutions of the deposit insurance system do indeed pay the premium rate to the DIC, in practice the deposit insurance premium rate burden is divided between banks and depositors and/or banks and corporations. Figure 1 illustrates how the burden of the deposit insurance premium is shared.

Figure 1: Who Pays the Deposit Insurance Premium?
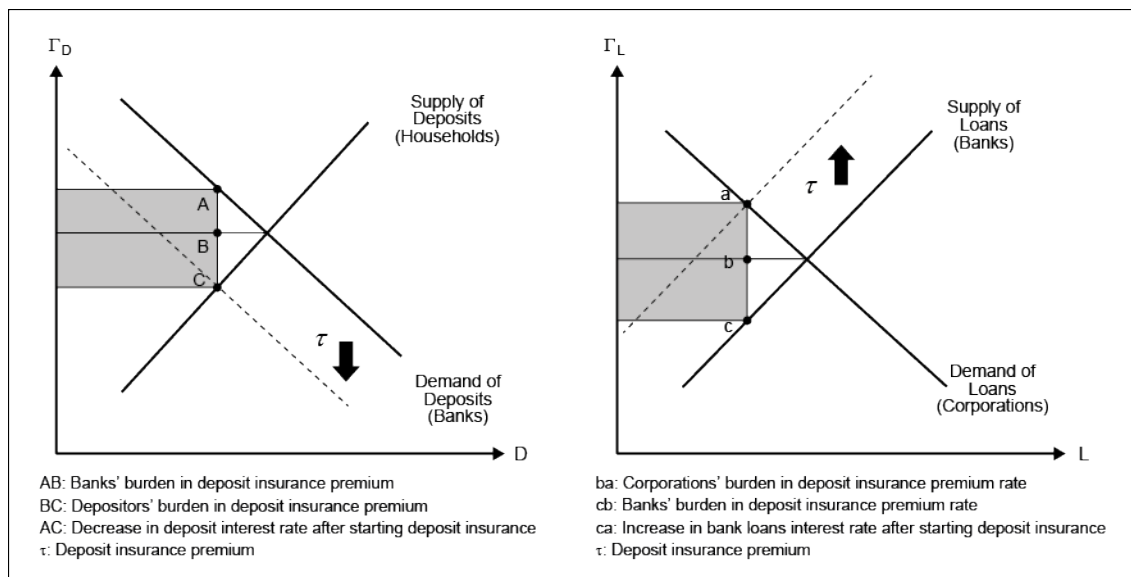
Source: Authors' compilation.

In Figure 1, τ is the deposit insurance premium banks should pay to the DIC. Paying this premium increases the banks' costs, so they have to lower the interest they pay out on customer deposits and/or raise their interest rates on loans granted. In the left-hand-side graph of Figure 1, it is assumed that banks compensate for their premium burden only by lowering the interest they pay out on deposits. In this scenario, banks are not the only parties that bear the burden of the deposit insurance cost, as it is shared between depositors and banks. As can be seen in the figure, the higher costs incurred by banks due to the launch of a deposit insurance system decrease the demand for deposits and consequently the demand curve shifts to the left. The result is

a decrease in interest rates on deposits. "AB" signifies the share of the burden of the deposit insurance premium borne by the banks, "BC" is the depositors' share of the burden of the deposit insurance premium, and "AC" is the total decrease in the interest paid out on deposits, which is equal to τ. In the right-hand-side graph of Figure 1, a scenario in which the premium burden of banks is compensated only by raising their lending rates on loans to corporations is presented. In this case, the increase in cost incurred by the banks due to the launch of a deposit insurance system results in a decrease in the provision of loans to customers and consequently the supply curve shifts to the left. As a result, banks' lending rates for loans rise. This tends to divide the premium burden between banks and corporations that are demanders of loans. The corporations' share is depicted by "ba," the banks' share by "cb," and "ca" depicts the total change in the banks' lending rate as a result of paying premiums to the DIC, which is equal to τ.

2. MODEL

In this paper we present two models—the first one is for estimating dual premium rates of deposit insurance; the second is for forecasting NPLs for each group of banks, which is a requirement for estimating the premium rates of deposit insurance. In Section 2.1 we define the dual premium rate model, and in Section 2.2 we explain how to forecast banks' NPLs using our model.

2.1 Dual Premium Rate Model

In the development of our model we were inspired by Yoshino, Taghizadeh-Hesary, and Nili (2013). They provide a model using a discounted present value mechanism to calculate a single fair premium rate for deposit insurance systems through which they estimate the fair premium rate for the deposit insurance system of Japan. Their model enables us to calculate a single premium rate for all financial institutions, which is what many countries use. However, in many other countries the monetary authorities prefer to use dual or multiple premium rates for their deposit insurance system, which means healthier financial institutions pay a lower premium to the DIC. This gives financial institutions an incentive to improve their soundness, so they can attain higher credit rating levels for paying lower premiums.

In this paper, we expand the Yoshino, Taghizadeh-Hesary and Nili (2013) model, and use a discounted present-value mechanism to calculate dual fair premium rates for the deposit insurance system. It can be expanded further to calculate multi-premium rates, should a particular DIC wish to use more than two premium rates.

Figure 2 shows the general outline of our new model for calculating the different premium rates for each group of banks:

As Figure 2 shows, according to our model, the premium income the DIC earns from each group of banks (A, B) has to be equal to the total amount of financial assistance the DIC provides to each group in the case of a banking default in that group, operational expenditures incurred by the DIC for each group, and precautionary future reserves kept by the DIC for each group separately.

Figure 2: Income and Expenditure of DIC in the Case of Dual Premium Rates
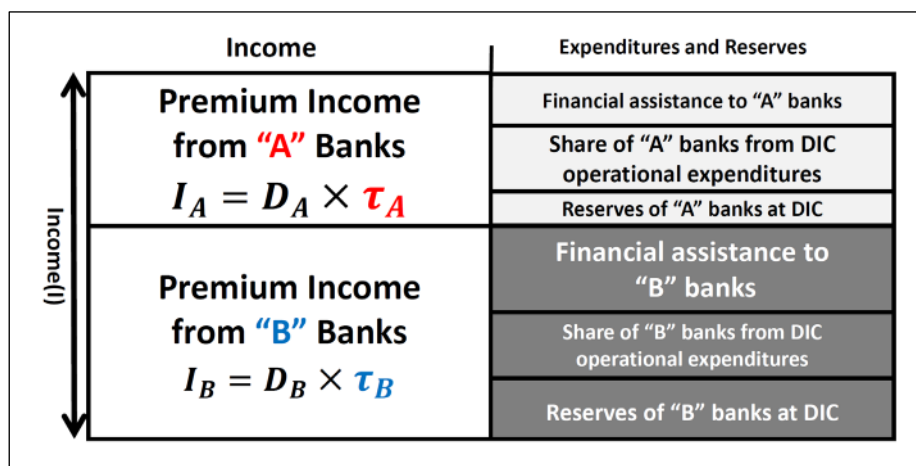
DIC = deposit insurance agency/corporation.

Note: "A" banks are healthier banks than "B" banks. τ_A is the premium rate for Group "A," τ_B is the premium rate for Group "B" of banks. I_A, I_B are the premium income of the DIC from Group A and Group B of banks, respectively. D_A, D_B are the cumulative deposit of Group A and Group B of banks, respectively.

Source: Authors.

According to our model, the discounted cumulative amounts of these variables are important, meaning:

> Discounted cumulative premium income of the DIC from each group of banks (including future expected income) = Discounted cumulative operational expenditures of the DIC toward each group of banks (including future expected operational expenditures) + Discounted cumulative financial assistance of the DIC to failed financial institutions of each groups (including future expected financial assistance) + discounted precautionary future reserves of the DIC at the end of the period for each group of banks.

Below, in Equations 1–8, we present each of these elements:

Present value of income (including future income) of the DIC from Group A banks:

$$PVI^A = \frac{\tau_A D_0^A}{(1+r_0)^0} + \frac{\tau_A D_1^A}{(1+r_1)^1} + \frac{\tau_A D_2^A}{(1+r_2)^2} + ... + \frac{\tau_A D_n^A}{(1+r_n)^n} \tag{1}$$

Present value of income (including future income) of the DIC from Group B banks:

$$PVI^B = \frac{\tau_B D_0^B}{(1+r_0)^0} + \frac{\tau_B D_1^B}{(1+r_1)^1} + \frac{\tau_B D_2^B}{(1+r_2)^2} + ... + \frac{\tau_B D_n^B}{(1+r_n)^n} \tag{2}$$

where PVI^A and PVI^B denote the present value of income (including future income) of the DIC from Group A and Group B banks, respectively; D_i^A and D_i^B are the cumulative amount of eligible deposits of Group A and Group B banks, respectively, in each year; τ_A and τ_B are the deposit insurance premium rates for Group A and Group B banks, respectively; and τ_i stands for the average long-term interest rate used for discounting values in each.

Present value of operational expenditures (including future expenditures) of the DIC:

$$PVE = \frac{E_0}{(1+r_0)^0} + \frac{E_1}{(1+r_1)^1} + \frac{E_2}{(1+r_2)^2} + ... + \frac{E_n}{(1+r_n)^n} \tag{3}$$

PVE stands for the present value of the operational expenditures (including future expenditures) of the DIC (e.g. personnel costs and equipment costs) and E_i denotes the operational expenditures of the DIC in each year.

Present value of financial assistance (including future financial assistance) of the DIC to Group A banks:

$$PVF^A = \frac{F_0^A}{(1+r_0)^0} + \frac{F_1^A}{(1+r_1)^1} + \frac{F_2^A}{(1+r_2)^2} + ... + \frac{F_n^A}{(1+r_n)^n} \tag{4}$$

Present value of financial assistance (including future financial assistance) of the DIC to Group B banks:

$$PVF^B = \frac{F_0^B}{(1+r_0)^0} + \frac{F_1^B}{(1+r_1)^1} + \frac{F_2^B}{(1+r_2)^2} + ... + \frac{F_n^B}{(1+r_n)^n} \tag{5}$$

PVF^A and PVF^B are the present value of financial assistance (including future financial assistance) of the DIC to Group A and Group B banks, respectively; and F_i^A and F_i^B are financial assistance of the DIC to Group A and Group B banks in each year, respectively. As the current year[1] is 0, for (1, 2, 3, … , n) years, this is the anticipated amount of financial assistance, which will be forecast using our second model (in Section 2.2 below).

Present value of future desired precautionary reserves of the DIC:

$$PVRES = \frac{RES_n}{(1+r_n)^n} \tag{6}$$

where $PVRES$ is the present value of the desired precautionary reserves of the DIC at the end of year n; and RES_n is desired future reserves of the DIC at the end of year n, also for precautionary purposes.

The dual premium rate model is as follows:

$$\begin{cases} PVI^A = \alpha PVE + PVF_A + \beta PVRES \\ PVI^B = (1-\alpha)PVE + PVF_B + (1-\beta)PVRES \end{cases} \tag{7}$$

---

[1] The first year is the year the deposit insurance system is launched in the country of our estimation. This could also be the year that the banking system had a structural change or experienced a financial crisis and then was supposed to calculate the optimal premium rate following that specific year.

where α and β are shares of Group A banks from the present value of the operational expenditures (including future expenditures) of the DIC and from the present value of the desired precautionary reserves of the DIC at the end of year n, respectively.

Substituting Eqs. 1–6 in Eq. 7 results in the following final dual premium rates model:

$$
\begin{cases}
\dfrac{\tau_A D_0^A}{(1+r_0)^0} + \dfrac{\tau_A D_1^A}{(1+r_1)^1} + \dfrac{\tau_A D_2^A}{(1+r_2)^2} + ... + \dfrac{\tau_A D_n^A}{(1+r_n)^n} = \alpha \left[\dfrac{E_0}{(1+r_0)^0} + \dfrac{E_1}{(1+r_1)^1} + \dfrac{E_2}{(1+r_2)^2} + ... + \dfrac{E_n}{(1+r_n)^n} \right] + \\[2em]
\dfrac{F_0^A}{(1+r_0)^0} + \dfrac{F_1^A}{(1+r_1)^1} + \dfrac{F_2^A}{(1+r_2)^2} + ... + \dfrac{F_n^A}{(1+r_n)^n} + \beta \left[\dfrac{RES_n}{(1+r_n)^n} \right] \\[2em]
\dfrac{\tau_B D_0^B}{(1+r_0)^0} + \dfrac{\tau_B D_1^B}{(1+r_1)^1} + \dfrac{\tau_B D_2^B}{(1+r_2)^2} + ... + \dfrac{\tau_B D_n^B}{(1+r_n)^n} = (1-\alpha) \left[\dfrac{E_0}{(1+r_0)^0} + \dfrac{E_1}{(1+r_1)^1} + \dfrac{E_2}{(1+r_2)^2} + ... + \dfrac{E_n}{(1+r_n)^n} \right] + \\[2em]
\dfrac{F_0^B}{(1+r_0)^0} + \dfrac{F_1^B}{(1+r_1)^1} + \dfrac{F_2^B}{(1+r_2)^2} + ... + \dfrac{F_n^B}{(1+r_n)^n} + (1-\beta) \left[\dfrac{RES_n}{(1+r_n)^n} \right]
\end{cases}
\tag{8}
$$

2.2 Forecasting Banks' Nonperforming Loans: Macro Shocks versus Idiosyncratic Shocks

Financial assistance from the DIC is mainly related to the amount of NPLs—the larger the amount of NPLs, the higher the default risk, which means the DIC needs to provide greater financial assistance (Yoshino and Hirano 2011, 2013). Hence, we need to develop a model for forecasting NPLs.

To establish what variables have an impact on future amounts of NPLs, we refer to two earlier studies, Yoshino and Hirano (2011) and Yoshino, Taghizadeh-Hesary, and Nili (2013). They use macroeconomic variables to forecast NPLs and financial assistance from the DIC. We are inspired by this earlier research, but we expand these papers' models. Our model for forecasting NPLs is as follows:

$$L_{NPL} = \rho(Y, P_S, P_L, i_B, Z_i) \times L \tag{9}$$

where L_{NPL} denotes NPLs and Y, P_S, P_L, i_B are the gross domestic product (GDP), price of stock, price of land, and government bond interest rate (safe asset interest rate), respectively. ρ is the expected share of loans that results in default. Z_i is the financial profile of all banks and is a variable to capture idiosyncratic shocks to banks. L is the total amount of loans of banks. In Yoshino and Hirano's model, the amount of NPLs (L_{NPL}) depends on the various economic factors mentioned above (Y, P_S, P_L, i_B). When land prices increase, collateral value increases as well, so default risk ρ declines. When business conditions improve, increases in GDP growth and stock prices cause a reduction in default risk ρ, and when the government bond interest rate, one of the safest asset interest rates, is raised, banks tend to invest more in safe assets that reduce default risks. The four macro variables can capture macro shocks, but some banks can fail even if the macro financial system is sound. So we need additional variables that can capture idiosyncratic uncertainty in the economy. This is the reason we insert Z_i in our model, i.e. to capture micro shocks to each bank or to each group of banks. Z_i denotes the banks' financial profile, which we further explain below. Hence, our model has the ability to capture macro and micro shocks.

3. AN ANALYSIS OF BANKS' CREDIT RATING

In our dual premium rates model, healthier banks pay lower premium rates. To enable us to identify the healthier group of banks, classification or credit rating is needed.

Credit ratings are opinions expressed in terms of ordinal measures, reflecting the current financial creditworthiness of issuers such as governments, firms, or financial institutions. These ratings are conferred by ratings agencies, such as Fitch Ratings, Moody's, and Standard and Poor's (S&P), and may be regarded as comprehensive evaluations of an issuer's ability to meet their financial obligations in full and on time. Hence, they play a crucial role by providing participants in financial markets with useful information for financial planning. To conduct rating assessments of banks, agencies resort to a broad range of financial and nonfinancial pieces of information, including domain experts' expectations. Ratings agencies usually provide general guidelines on their rating decision process, but detailed descriptions of the rating criteria and of the determinants of banks' ratings are generally not provided (Orsenigo and Vercellis 2013). In search of more objective assessments of the creditworthiness of financial institutions, there is a growing body of research into the development of reliable quantitative methods for the automatic classification of banks according to their financial strength, and the solutions for raising the creditworthiness of banks and the financial system. Cooper (2011) analyzes the off-balance sheet (OBS) behavior of a sample of small commercial banks in the USA in 2006. In particular, the paper aims to study the impact that monitoring intensity has on bank OBS usage. The results lend support to the argument of stronger regulation in the banking industry since monitoring impacts both bank management behavior and decision-making.

However, some researches find that higher creditworthiness of the financial system does not necessarily come from tighter supervision and a regulated environment. Abildgren (2016) considers the case of Denmark and investigates whether the absence of banking crises is due to the robustness of the banking sector's customers rather than tight regulation. The paper finds that the Danish household sector in the 1950s had a high debt payment ability and was very robust to even large income shocks. The results indicate that the stability of the Danish financial sector is not only due to tight regulation but also reflects the high credit quality of the banking sector's loan portfolio.

Extensive empirical research devoted to analyzing the stability and soundness of financial institutions dates back to the 1960s. Ravi Kumar and Ravi (2007) provide a comprehensive survey of the application of statistical and intelligent techniques for predicting the default of banks and firms. Despite the obvious relevance, however, the development of reliable quantitative methods for the prediction of banks' credit rating has only recently begun to attract strong interest. These studies are mainly conducted within two broad research strands focusing on statistical and machine learning techniques, and may address both feature selection and classification. Poon, Firth and Fung (1999) develop logistic regression models for predicting financial strength ratings assigned by Moody's, using bank-specific accounting variables and financial data. Factor analysis is applied to reduce the number of independent variables and retain the most relevant explanatory factors. The authors show that loan provision information, and risk and profitability indicators, add the greatest predictive value in explaining Moody's ratings. Huang et al. (2004) compare support vector machines and back propagation neural networks to forecast the rating of financial institutions operating in markets in the United States and Taipei,China, respectively. In both cases five rating categories are considered, based on information released by S&P and TRC. Analysis of variance is used to discard noninformative features. In this study, support vector machines and neural networks achieve comparable classification results. However, the

authors find that the relative importance of the financial variables used as inputs by the optimal models is quite different between the two markets. A study by Orsenigo and Vercellis (2013) presents an empirical evaluation of two linear and nonlinear techniques—principal component analysis (PCA) and double-bounded tree-connected Isomap (dbt–Isomap)—to assess their effectiveness for dimensionality reduction in bank credit rating prediction, and to identify the key financial variables endowed with the greatest explanatory power. Extensive computational tests concerning the classification of six banks' ratings data sets show that the use of dimensionality reduction accomplished by nonlinear projections often induces an improvement in the classification accuracy, and that dbt-Isomap outperforms PCA by consistently providing more accurate predictions (Yoshino and Taghizadeh-Hesary, 2014a).

In our present research on credit ratings of banks, we employ the statistical techniques used by Yoshino and Taghizadeh-Hesary (2014b) for credit rating and classification of small and medium-sized enterprises (SMEs). They use PCA and cluster analysis and apply various financial variables of 1,363 SMEs in Asia. In our present paper, we assign credit ratings to 32 banks of an Asian economy and classify all of them into two groups, and in our empirical analysis we calculate the premium rate for each group of banks.

To be able to do so and to ensure our results are credible, we need to select variables that capture all the relevant characteristics of the banks that are the subject of our examination.

3.1 Selection of Variables

It is widely known that ratings are directly affected by the financial performance of banks. Based on this assumption, we focus on banks' financial profiles and employ eight financial variables that describe all general characteristics of banks. These variables are listed in Table 1:

Table 1: Variables Examined

No.	Symbol	Definition
1	L–D	Total loans/total deposits
2	PR–L	Properties/total loans
3	(SD+LD)–D	(Saving deposits + long-term deposits)/total deposits
4	A–L	Total assets/total loans
5	SC–L	Securities/total loans
6	CA–D	Cash/total deposits
7	CBR–D	Accounts receivable from central bank/total deposits
8	OBR–D	Accounts receivable from other banks/total deposits

Note: Properties are land, buildings, and other hard assets owned by banks. Securities include shares of corporate stock or mutual funds, bonds issued by corporations or governmental agencies, limited partnership units, and various other formal investment instruments that are negotiable and fungible. Accounts receivable from the central banks include reserve requirement (or cash reserve ratio) and other sums that are normally in the form of cash stored physically in a bank vault (vault cash) or deposits made with a central bank. Accounts receivable from other banks are sums loaned to other banks.

Loans, properties, securities, cash, accounts receivable from the central bank, and accounts receivable from other banks are components of a financial institution's assets. The higher these variables, the more stable and sound a particular financial institution tends to be. At the next stage, two statistical techniques are used: principal component analysis (PCA) and cluster analysis. The underlying logic of both techniques is dimension reduction (i.e. summarizing information on numerous variables in just a few variables), but they achieve this in different ways. PCA reduces the number of variables into components (or factors), whereas cluster analysis reduces the number of banks by placing them in small clusters. In this survey, we use components (factors), which are the result of PCA, and subsequently carry out a cluster analysis to classify the banks.

3.2 Principal Component Analysis

PCA is a standard data reduction technique that extracts data, removes redundant information, highlights hidden features, and visualizes the main relationships that exist between observations.[2] PCA is a technique for simplifying a data set, by reducing multidimensional data sets to lower dimensions for analysis. Unlike other linear transform methods, PCA does not have a fixed set of basis vectors. Its basis vectors depend on the data set, and PCA has the additional advantage of indicating what is similar and different about the various models created (Ho and Wu 2009). Through this method we reduce the eight variables listed in Table 2 to determine the minimum number of components that can account for the correlated variance among the banks.

To examine the suitability of these data for factor analysis, we perform the Kaiser–Meyer–Olkin (KMO) test and Bartlett's test of sphericity. KMO is a measure of sampling adequacy to indicate the proportion of common variance that might be caused by underlying factors. High KMO values (higher than 0.6) generally indicate that factor analysis may be useful, which is the case in this study: KMO = 0.61. If the KMO value is lower than 0.5, factor analysis is not useful. Bartlett's test of sphericity reveals whether the correlation matrix is an identity matrix, indicating that variables are unrelated. A level lower than 0.05 indicates that there are significant relationships among the variables, which is the case in this study: the significance of Bartlett's test is < 0.0 (Yoshino and Taghizadeh-Hesary 2015).

Next, we determine how many factors to use in our analysis. Table 2 reports the estimated factors and their eigenvalues. Only those factors accounting for more than 10% of the variance (eigenvalues > 1) are kept in the analysis, which means only the first three factors are retained (Table 2).

Taken together, Z_1 through Z_3 explain 82.421% of the total variance of the financial ratios.

In running the PCA, we use direct oblimin rotation. Direct oblimin rotation is the standard method to obtain a nonorthogonal (oblique) solution, i.e. one in which the factors are allowed to be correlated. To interpret the revealed PCA information, the pattern matrix must subsequently be studied. Table 3 presents the pattern matrix of factor loadings using the direct oblimin rotation method, where variables with large loadings—absolute value (> 0.5) for a given factor—are highlighted in bold.

---

[2] PCA can also be called the Karhunen–Loève transform (KLT), named after Kari Karhunen and Michel Loève (Yoshino et al. 2016)

Table 2: Total Variance Explained

Component	Eigenvalues	% of Variance	Cumulative Variance %
Z1	3.685	46.065	46.065
Z2	1.765	22.057	68.122
Z3	1.144	14.299	82.421
Z4	0.639	7.992	90.413
Z5	0.540	6.752	97.165
Z6	0.198	2.469	99.634
Z7	0.022	0.275	99.909
Z8	0.007	0.091	100

Table 3: Factor Loadings of Financial Variables after Direct Oblimin Rotation

Variables (Financial Ratios of Banks)	Component		
	Z1	Z2	Z3
L–D	(0.238)	(0.912)	(0.143)
PR–L	0.042	0.190	0.780
(SD+LD)–D	(0.287)	0.819	(0.123)
A–L	0.987	0.083	0.130
SC–L	(0.096)	(0.140)	0.875
CA–D	0.379	(0.536)	0.039
CBR–D	0.954	(0.104)	(0.102)
OBR–D	0.981	(0.011)	(0.117)

( ) = negative value.

Note: The extraction method is principal component analysis. The rotation method is direct oblimin rotation with Kaiser normalization. For definitions of the variables, please refer to Table 1.

As can be seen in Table 3, the first component, Z1, has three variables with an absolute value (> 0.5), which are all positive—(i) total assets/total loans, (ii) accounts receivable from central bank/total deposits, and (iii) accounts receivable from other banks/total deposits. For Z1, the variables with large loadings are mainly assets, hence Z1 generally reflects the assets of the examined banks. As this factor explains the greatest variance in the data, it is the most informative indicator of a bank's overall financial health. Z2 represents deposits and this component has three major loading variables: (i) total loans/total deposits, which is negative; (ii) (saving deposits + long-term deposits)/total deposits, which is positive; and (iii) cash/total deposits. If the amount of deposits increases, Z2 increases. Z3 has two major loadings, which are (i) properties/total loans, (ii) securities/total loans, so it reflects 1/total loans. The larger the amount of loans, the smaller the Z3.

Table 4 presents the correlation matrix of the components and shows there is no correlation between these three components. This means we can use a regular orthogonal rotation approach to force an orthogonal rotation. But in this survey we use an oblique rotation method, which still provides an orthogonal rotation factor solution, because these three components are not correlated with each other and are distinct entities.

Table 4: Component Correlation Matrix

Component	Z1	Z2	Z3
Z1	1	(0.282)	0.059
Z2	(0.282)	1	0.162
Z3	0.059	0.162	1

(…) = negative value.

Note: The extraction method is principal component analysis. The rotation method is direct oblimin rotation with Kaiser Normalization.
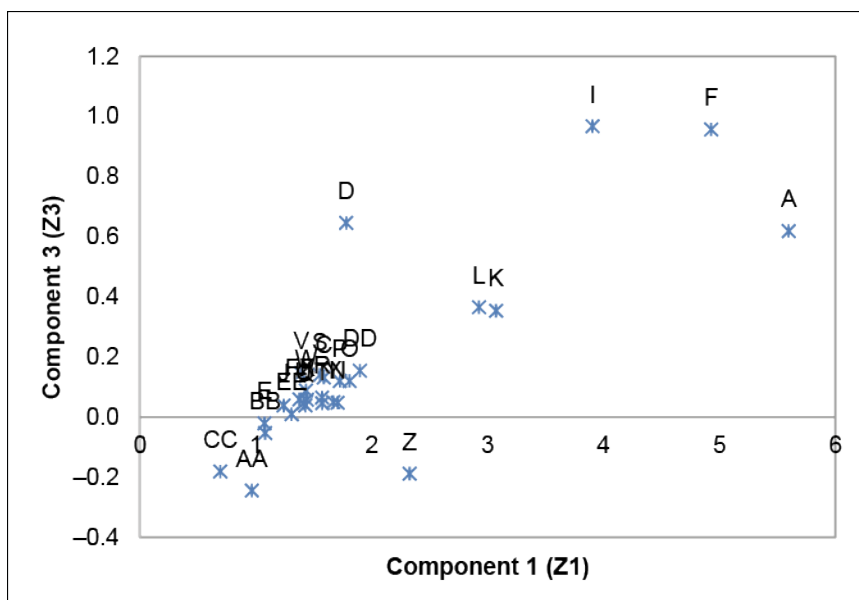
Figure 3 shows the distribution of the three components (Z1, Z2, and Z3) for 28 out of a total of 32 banks of an Asian country.

Figure 3: Distribution of Factors for 28 Banks

11
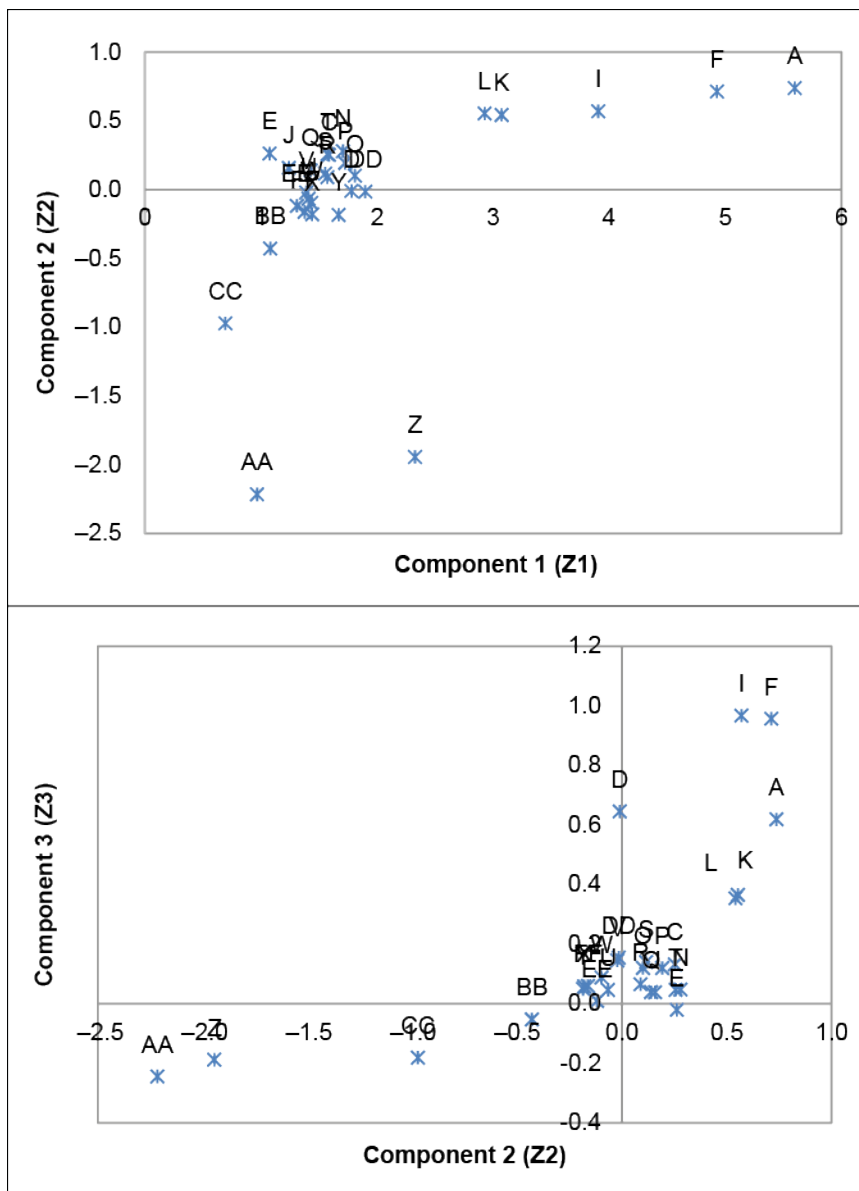
Figure 3 *continued*

Note: Each star represents one bank; these are named alphabetically, A, B, C, …, Z, AA, BB, CC, DD, EE, and FF, for 32 banks of an Asian economy. Four banks (banks B, G, H, and M) are outliers in positive parts of the graphs and are not visible in the above graphs.

3.3 Cluster Analysis

In this section, we take the three components obtained in the previous section and identify those banks that have similar traits. We then generate clusters and place the banks in distinct groups. To do this, we employ cluster analysis, which organizes a set of data into groups so that observations from a group with similar characteristics can be compared with those from a different group (Martinez and Martinez 2005). In this case, banks are organized into distinct groups according to the three components derived from the PCA obtained in the previous section. Cluster analysis techniques can themselves be broadly grouped into three classes: hierarchical clustering, optimization clustering,[3] and model-based clustering. We use the method most prevalent in the literature—hierarchical clustering. This produces a nested sequence of partitions by merging (or dividing) clusters. At each stage of the sequence, a new partition is optimally merged with (or separated from) the previous partition according to some adequacy criterion. The sequence of partitions ranges from a single cluster containing all the individual banks to a number of clusters (*n*) containing a single bank. The series can be described by a tree display called a "dendrogram" (Figure 4). Agglomerative hierarchical clustering proceeds by means of a series of successive fusions of the *n* objects into groups. By contrast, divisive hierarchical methods divide the *n* individuals into progressively finer groups. Divisive methods are not commonly used because of the computational problems they pose (see Everitt et al. [2001] and Landau and

---

[3] The main difference between the hierarchical and optimization techniques is that in hierarchical clustering the number of clusters is not known beforehand. The process consists of a sequence of steps in which two groups are either merged (agglomerative) or divided (divisive) according to the level of similarity. Eventually, each cluster can be subsumed as a member of a larger cluster at a higher level of similarity. The hierarchical merging process is repeated until all subgroups are fused into a single cluster (Martinez and Martinez 2005). Optimization methods, on the other hand, do not necessarily form hierarchical classifications of the data as they produce a partition of the data into a specified or predetermined number of groups by either minimizing or maximizing some numerical criteria (Feger and Asafu-Adjaye 2014).

Chis Ster [2010]). Below, we use the average linkage method, which is a hierarchical clustering technique.

Figure 4 shows the dendrogram that results from this hierarchical clustering:

Figure 4: Dendrogram Using Average Linkage
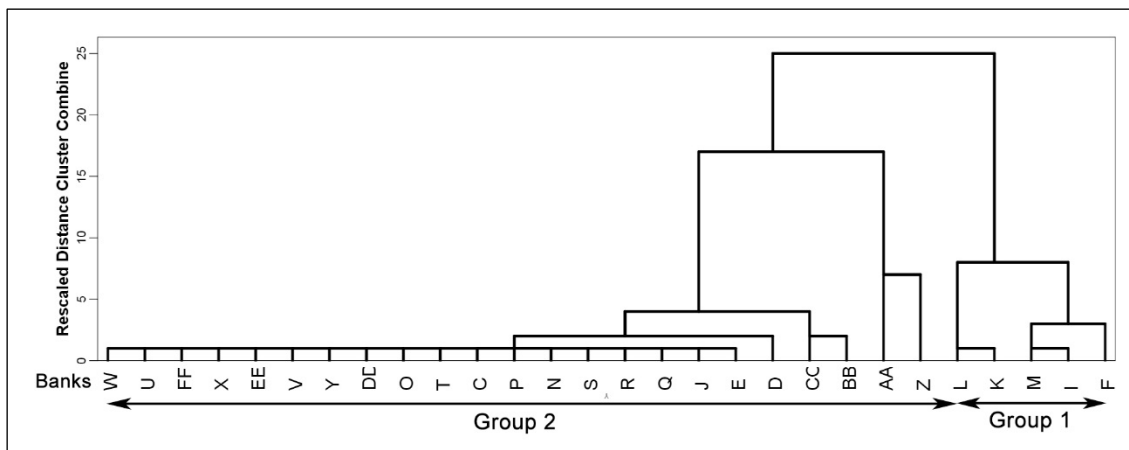
The resulting dendrogram (hierarchical average linkage cluster tree) provides a basis for determining the number of clusters by sight. In the dendrogram shown in Figure 4, the horizontal axis shows 28 banks, which are named alphabetically. As mentioned above, 32 banks are the subject of our examination. However, four banks have outlying positive data that are far removed from the data for the other 28 banks. We do not include these four banks in our cluster analysis as our result is not a rational clustering. This is the reason Figure 4 shows only 28 banks on the horizontal axis.

The dendrogram classifies the banks into two main clusters (Group 1 and Group 2), but it does not show which of these two clusters contain the financially healthier banks, so we have to take one further step. By comparing the classification resulting from cluster analysis and the distributions of factors in Figure 3, we can conclude that the sequence of banks on the horizontal axis of our dendrogram is based on their soundness. Among these 28 banks, bank "F" has the highest stability and soundness, whereas bank "W" has the lowest.

3.4 Robustness Check of Banks' Credit Rating

For robustness, we check the rankings of three banks out of the 28 banks for all eight examined financial variables. We randomly pick one bank from Group 1 and one from Group 2, and the bank that is in the middle of the credit ranking selected. The results are summarized in Table 5:

Table 5: Robustness Check for Three Sample Banks

Bank	Credit Rank	Rank of L–D	Rank of PR–L	Rank of (SD+LD) –D	Rank of A–L	Rank of SC–L	Rank of CA–D	Rank of CBR–D	Rank of OBR–D
I	2	24	1	16	3	5	8	21	2
R	14	14	17	12	15	9	11	9	7
W	28	11	20	22	20	6	10	3	18

Note: Credit rank is the ranking shown by our dendrogram—the lower this number, the healthier the bank. For definitions of the variables, please refer to Table 1.

The first randomly picked bank from Group 1 is bank I. Bank I is the second most sound and stable bank according to our credit rating result, and as is clear from Table 5, the robustness check supports this result. This bank shows a fairly stable and healthy status in most of our eight financial variables. It is the top bank for PR–L (properties/loans), meaning this bank has a relatively large amount of properties compared with the amount of loans, which means it is stable. It ranks second for OBR–D (accounts receivable from other banks/total deposits), fifth for SC–L (securities/loans), and third for A–L (assets/loans)—these results indicate that this bank has sufficient assets, which favors its stability and soundness. Although it has one of the lowest ranks for L–D (loans/deposits), this suggests this bank is trusted by depositors, and therefore the amount of deposits is large compared with loans. The second bank in our robustness check is bank R, which can be found in the middle of the horizontal axis of our dendrogram with a credit rank of 14, which is close to the middle of these 28 banks. When considering bank R's ranking in terms of the eight variables, for most of these variables it appears in the middle of the ranking. If we take a simple average of the rank of this bank in our eight variables, the result is almost 12, which is close to the credit rank of 14 suggested by our method. The third bank in our robustness check is bank W, a bank we pick randomly from Group 2. Bank W has the lowest soundness and stability in this group and among all 28 banks. When considering the ranking of this bank in our eight variables in Table 5, it is apparent that this bank is not sound. It has very low rankings for PR–L (properties/loans), (SD+LD)/D ((saving deposits + long-term deposits)/total deposits), A–L (assets/loans), and OBR–D (accounts receivable from other banks/total deposits), which suggests this bank is unsound and unstable—it has the lowest credit rank of the banks examined.

4. EMPIRICAL ANALYSIS

In Section 4.1, we first use the model developed in Section 2.2 to forecast NPLs for each group of banks. We then use the results of the estimations obtained in Section 4.1 to calculate a fair deposit insurance premium rate for each group of banks, using the model we develop in Section 2.1 of this paper.

4.1 Forecasting Banks' Nonperforming Loans

There are several recent works on forecasting banks' nonperforming loans and stress testing. For instance, Kahlert and Wagner (2017) stress tested Eurozone banks of systemic importance by applying a historical simulation approach. To forecast each group of banks' NPLs, we run regressions using the vector autoregression or vector error correction (VAR/VEC) model. As mentioned above, for our empirical analysis in this paper we use data from an Asian economy, so we use macroeconomic data and all 32 banks' financial profiles to forecast the NPLs for each group of banks (Group 1 and Group 2). As per Eq. 9, we need to use macroeconomic variables (real GDP, price of land, price of stock, government bond interest rate) and Zi, which represents the financial profile of banks and captures idiosyncratic shocks, to forecast NPLs. In our empirical analysis, for the macroeconomic variables we employ real GDP, and instead of the price of stock and price of land, due to lack of data, we use the consumer price index (CPI), which is the best representative for the price level in an economy and can be used as a substitute for these two price levels. In this study, using the government bond interest rate is not practical since the selected Asian country has implemented Islamic banking and fiscal rules, which are quite different from conventional rules. And as interest rates are affected by monetary policy, instead of the real interest rate, we

use another monetary variable, M1, which has a high correlation with the interest rate, as shown in many previous studies.

Eq. 9 has two categories of variables for forecasting NPLs—the first category consists of the macroeconomic variables described above; the second element is Zi, reflecting the financial profile of banks. The latter category is made up of three significant components—Z1, Z2, and Z3—obtained using principal component analysis in Section 3.2 with their factor loadings presented in Table 3. Using the loadings of all eight financial ratios (Table 3), we obtain Z1, Z2, and Z3 for each group of banks (Group 1 and Group 2), and since those eight financial ratios of banks are time series variables, Z1, Z2, and Z3 will also be time series variables. For our empirical analysis, we use monthly data from 2011M1 to 2013M12 from the Central Bank.

Since we have two groups of banks, we should run two regressions—one for each group. The left-hand side of Eq. 9 for each group's regression is the sum of NPLs of that group/total loans of that group of banks; the right-hand side of Eq. 9 is the macroeconomic variables and Z1, Z2, and Z3 for that group of banks.

4.1.1 Data Analysis

To evaluate the stationarity of all series, we use an augmented Dickey–Fuller (ADF) test. The results we obtain imply that all variables are nonstationary. These variables include GDP growth rate; CPI inflation rate (inflation rate of each month compared to the same month of the previous year); M1 growth rate (growth rate of M1 in each month compared with the same month of the previous year—the original quarterly data are converted to monthly data); sum of NPLs/sum of total loans for Group 1 and Group 2 of the banks; and Z1, Z2, and Z3 for each group of banks. However, when we apply the unit root test to their first differences, we are able to reject the null hypothesis of unit roots for each of the variables. These results suggest that all variables each contain a unit root. When we perform the unit root test and discover that the variables are nonstationary in level and stationary at the first difference level, they are integrated of order one. The next step is to conduct a cointegration analysis to examine whether a long-run relationship exists among these variables.

4.1.2 Cointegration Analysis

We conduct a cointegration analysis using Johansen's technique by assuming a linear deterministic trend and for two cases—with intercept, and with intercept and trend. Given the short period of our data, the Akaike information criterion (AIC) suggests using variables with one lag. The results of the cointegration rank test using trace are presented in Table 6.

As is clear from Table 6, the above test rejects the null hypothesis of noncointegrating variables for Group 1 and Group 2. This means that all variables are cointegrated and there is a long-run association among variables, or, in other words, in the long run, these seven variables (NPL/L, GDP growth rate, CPI inflation rate, M1 growth rate, Z1, Z2, and Z3) for each group of banks move together. Hence, we should run a vector error correction model (VECM). The AIC results of our linear deterministic VEC model indicate that estimating the model by including trend and intercept is slightly better than including just intercept for both bank groups, so we also retain this finding.

Table 6: Cointegration Rank Test (Trace)

	Group 1 of Banks					
	Intercept			Intercept and Trend		
Hypothesized no. of CEs	Eigenvalue	Trace Statistic	Prob.	Eigenvalue	Trace Statistic	Prob.
None	0.80	192.62*	0.00	0.80	217.14*	0.00
At most 1	0.75	136.33*	0.00	0.75	160.38*	0.00
At most 2	0.61	87.91*	0.00	0.62	111.82*	0.00
At most 3	0.53	55.01*	0.01	0.55	77.80*	0.00
At most 4	0.39	28.35	0.07	0.51	49.89*	0.01
At most 5	0.25	11.06	0.21	0.35	24.98	0.06
At most 6	0.02	0.86	0.35	0.25	10.10	0.12
	Group 2 of Banks					
	Intercept			Intercept and Trend		
Hypothesized no. of CEs	Eigenvalue	Trace Statistic	Prob.	Eigenvalue	Trace Statistic	Prob.
None	0.80	167.96*	0.00	0.81	200.61*	0.00
At most 1	0.75	112.06*	0.00	0.80	141.91*	0.00
At most 2	0.48	64.19	0.13	0.58	86.33	0.07
At most 3	0.46	41.23	0.18	0.47	55.63	0.20
At most 4	0.24	19.41	0.46	0.38	33.63	0.31
At most 5	0.21	9.58	0.31	0.24	16.82	0.43
At most 6	0.03	1.17	0.28	0.19	7.34	0.31

CE = cointegrating equation; prob. = probability.

Note: * denotes rejection of the noncointegrating hypothesis at the 5% level.

Prob. shows MacKinnon–Haug–Michelis p-values.

4.1.3 Vector Error Correction Model (VECM)

We estimate Model 9 in a VECM setting including the seven variables—NPL/L, GDP growth rate, CPI inflation rate, M1 growth rate, Z1, Z2, and Z3—for each group. The VECM can be defined as follows (see Yoshino et al. 2014):

$$dV_t = A(O)dV_t + \Pi V_{t-1} + \varepsilon_t \tag{10}$$

for

$$V = \left(\text{NPL/L}, \text{gdp}, \text{cpi}, \text{m1}, Z1, Z2, Z3\right) \tag{11}$$

where d denotes the first differences, O is the lag operator, and ε is an error term. Π can be written as $\Pi = \alpha\beta'$, where α and β are $p \times r$ matrices, and p is the number of variables in V. gdp is GDP growth rate, cpi is CPI inflation rate, and m1 is M1 growth rate. β is a vector of the cointegrating relationship and α is a loading matrix defining the adjustment speed of the variables in V to the long-run equilibrium defined by the cointegrating relationship. The rank of Π is denoted by r. As mentioned above, the AIC standard suggests one lag.

Model 12 shows our VECM for Group 1 with four cointegrating equations and one lag for each variable:

$$
\begin{aligned}
d(NPL_1/L_1) = &\; \Phi_1[Z_{1,1}(-1) - 47.45\ NPL_1/L_1(-1) - 33.89\ P(-1) \\
&+ 1.82\ Y(-1) + 0.34\ trend - 12.36] + \Phi_2[Z_{1,2}(-1) - 8.83\ NPL_1/L_1(-1) \\
&- 5.43\ P(-1) + 0.75\ Y(-1) + 0.05\ trend - 1.55] + \Phi_3[Z_{1,3}(-1) \\
&- 23.10\ NPL_1/L_1(-1) - 17.63\ P(-1) + 6.89\ Y(-1) + 0.24\ trend - 9.12] \\
&+ \Phi_4[\ M(-1) - 0.92\ NPL_1/L_1(-1) - 2.17\ P(-1) + 2.35\ Y(-1) \\
&+ 0.03\ trend - 1.59] + \Phi_5\ d[Z_{1,1}(-1)] + \Phi_6\ d[Z_{1,2}(-1)] + \Phi_7 d[Z_{1,3}(-1)] \\
&+ \Phi_8 d[M(-1)] + \Phi_9 d[NPL_1/L_1(-1)] + \Phi_{10} d[P(-1)] + \Phi_{11} d[Y(-1)] + \Phi_{12}
\end{aligned}
\tag{12}
$$

where NPL_1/L_1 is the ratio of NPLs over total loans for Group 1; $Z_{1,1}$ denotes the first component, $Z_{1,2}$ is the second component, and $Z_{1,3}$ is the third component, all three for Group 1; $d(Z_{1,1})$, $d(Z_{1,2})$, $d(Z_{1,3})$, $d(M)$, $d(NPL_1/L_1)$, $d(P)$, and $d(Y)$ are first differences of the first component, the second component, the third component (all three for Group 1), M1 growth rate, NPLs over total loans for Group 1, CPI inflation rate, and GDP growth rate, respectively. In this VECM, trend is also included, since we calculate the cointegration with intercept and trend. Φ_1, Φ_2, Φ_3, and Φ_4 are the coefficients of the four cointegrating equations; $\Phi_5 \dots \Phi_{11}$ are the coefficients of the lagged variable for the seven variables of our model; and Φ_{12} is a constant.

Model 13 shows our VECM for Group 2 with one cointegrating equation and one lag for each variable:

$$
\begin{aligned}
d(NPL_2/L_2) = &\; \Phi_{13}[Z_{2,1}(-1) + 0.67\ Z_{2,2}(-1) - 3.90\ Z_{2,3}(-1) + 0.03\ M(-1) \\
&- 2.04\ NPL_2/L_2(-1) - 1.11\ P(-1) - 0.04\ Y(-1) + 0.008\ trend - 0.97] \\
&+ \Phi_{14}\ d[Z_{2,1}(-1)] + \Phi_{15}\ d[Z_{2,2}(-1)] + \Phi_{16}\ d[Z_{2,3}(-1)] + \Phi_{17} d[M(-1)] \\
&+ \Phi_{18} d[NPL_2/L_2(-1)] + \Phi_{19} d[P(-1)] + \Phi_{20} d[Y(-1)] + \Phi_{21}
\end{aligned}
\tag{13}
$$

where NPL_2/L_2 is the ratio of NPLs over total loans for Group 2; $Z_{2,1}$ denotes the first component, $Z_{2,2}$ is the second component, and $Z_{2,3}$ is the third component, all three for Group 2; $d(Z_{2,1})$, $d(Z_{2,2})$, $d(Z_{2,3})$, $d(M)$, $d(NPL_2/L_2)$, $d(P)$, and $d(Y)$ are first differences of the first component, the second component, the third component (all three for Group 2), M1 growth rate, NPLs over total loans for Group 2, CPI inflation rate, and GDP growth rate, respectively. In this VECM, trend is also included since we calculate the cointegration with intercept and trend. Φ_{13} is the coefficient of the cointegrating equation; $\Phi_{14} \dots \Phi_{20}$ are the coefficients of the lagged variable for the seven variables of our model; and Φ_{21} is a constant.

We use models 12 and 13 to forecast the NPL/L for each group of banks. To do so, we need some assumptions. As mentioned above, in developing our VECM we use monthly data from 2011M1 to 2013M12. We assume real GDP growth of 2.8%, year on year, for 2014 and 2.9% for 2015. We assume a CPI inflation rate of 23%, year on year, for 2014 and for 2015 we expect 18%. As for the M1 growth rate, the selected country under a new governor continues to pursue tightening monetary policies, as it did in 2013, to control the high inflation rate. Hence, we assume that in 2014 and 2015, M1 grows at the same rate as in 2013M09–2013M12. Also for NPL/L and the three components for each group of banks for 2014 and 2015, we assume they stay on the same growth path as in 2013M09–2013M12. Using these assumptions, we forecast the NPL/L for each group and use these to calculate the premium rates for each group of banks (which are presented in Section 4.2).

4.1.4 Impulse Response Analysis

In this section, we conduct impulse response (IR) analysis to provide further evidence of the dynamic response of NPL/L to macro and idiosyncratic innovations. (For more information on IR analysis, see Yoshino and Taghizadeh-Hesary [2014c].)

The accumulated response of NPL/L to macro and idiosyncratic innovations for Group 1 of the banks is shown in Figure 5.

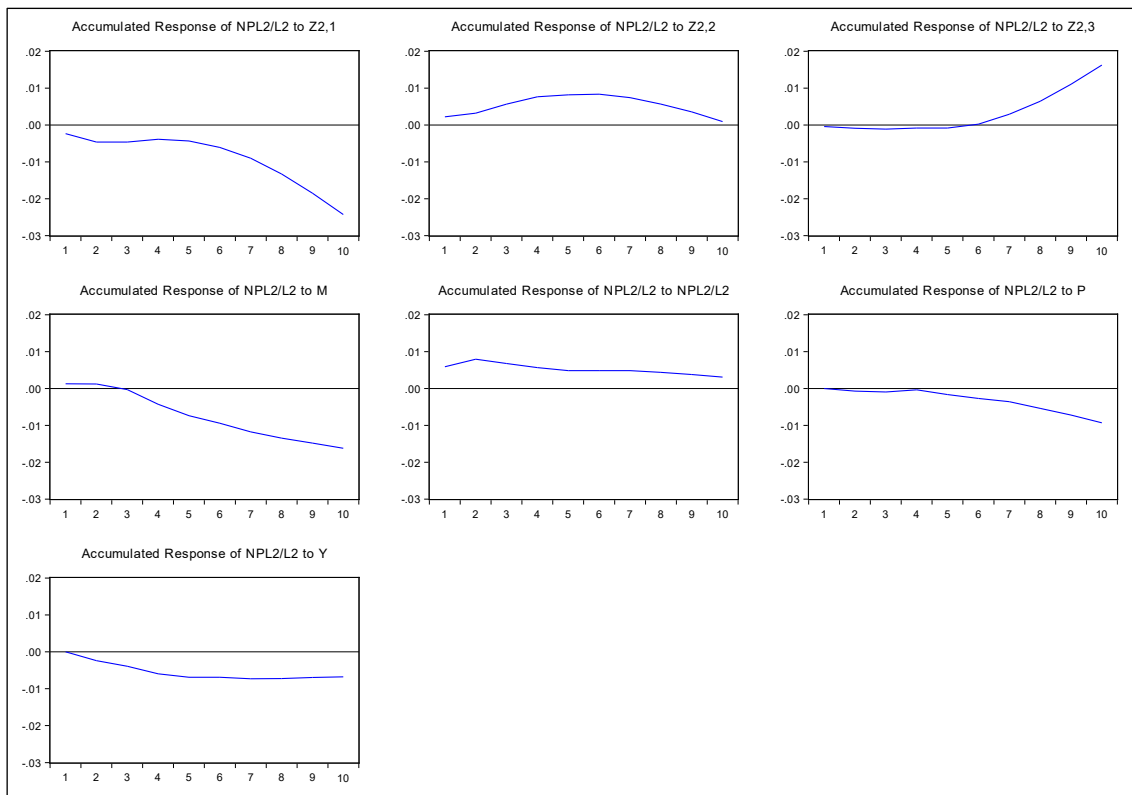**Figure 5: Response of NPL/L to Innovations (Group 1 of Banks)**
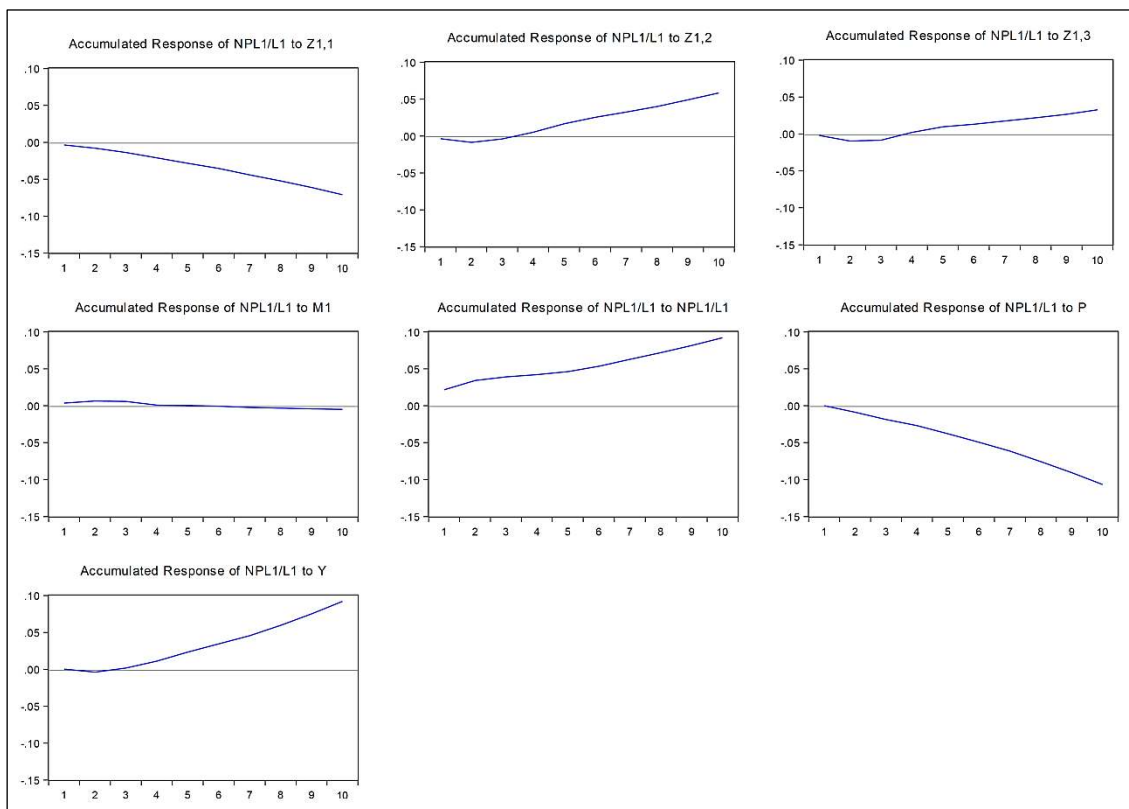
Note: Accumulated response to Cholesky one-standard deviation innovations. NPL_1/L_1 is the ratio of NPLs over total loans for Group 1 of the banks; $Z_{1,1}$ denotes the first component, $Z_{1,2}$ the second component, and $Z_{1,3}$ the third component, all three for Group 1; M1 denotes M1 growth rate, P denotes CPI inflation rate, and Y denotes GDP growth rate.

The three graphs in the first row of Figure 5 show accumulated responses of NPL/L to an unanticipated positive shock to Z1, Z2, and Z3 for Group 1 of the banks. The response of NPL/L to Z1 is statistically negative and very persistent. This means a positive shock to Z1, which mainly represents assets, decreases the NPL/L of Group 1. An unanticipated positive shock to Z2, which represents deposits, has a statistically negative effect on the NPL/L of Group 1 and builds up over the first 3 months, after which it becomes insignificant, meaning an unanticipated increase in deposits reduces the NPL/L for Group 1. An unanticipated positive shock to Z3, which represents 1/loans, has a statistically negative effect on NPL/L of Group 1 and builds up over the first 3 months, after which it becomes insignificant. The four other graphs in Figure 5 show accumulated responses of the NPL/L of Group 1 of the banks to positive shocks to macro variables and to lagged NPL/L. The response of NPL/L to M1 growth rate shocks is statistically positive and builds up over the first 5 months, after which it becomes insignificant. An unanticipated positive shock to P (CPI inflation) has a

statistically negative and persistent effect on the NPL/L of Group 1, which is consistent with Yoshino and Hirano (2011, 2013). An unanticipated positive shock to Y (GDP growth rate) has a statistically negative effect on the NPL/L of Group 1 and builds up over the first 2 months, after which time it becomes insignificant.

Figure 6 depicts the accumulated responses of NPL/L to macro and idiosyncratic innovations for Group 2 of the banks.

Figure 6: Response of NPL/L to Innovations (Group 2 of Banks)

Note: Accumulated response to Cholesky one-standard deviation innovations. NPL_2/L_2 is the ratio of NPLs over total loans for Group 2 of the banks; $Z_{2,1}$ denotes the first component, $Z_{2,2}$ the second component, and $Z_{2,3}$ the third component, all three for Group 2; M1 denotes the M1 growth rate, P the CPI inflation rate, and Y the GDP growth rate.

Group 2 shows similar responses to innovations to macro variables. It means that focusing only on a model based on macro variables for forecasting NPLs of different groups of banks leads to misinterpretation as it is possible that under good economic conditions some banks show a negative financial performance and have a high default risk.

The responses of the NPL/L of Group 2 of the banks to an unanticipated positive shock to Z1 and Z3 are similar to Group 1's responses, but for shocks to Z2 the responses differ. The response of the NPL/L of Group 2 to positive shocks to Z2 is statistically positive and persistent, which goes against our finding for Group 1. This means that increasing deposits, which are good news for banks, tend to result in an increase in NPL/L for Group 2. This shows that Group 2 does not manage their NPL/L well—by expanding their business and accepting more deposits the NPL/L ratio increases, which indicates that Group 2 is not as sound as Group 1.

These results confirm our findings in the previous sections of this paper. Moreover, they back up our suggestion that macro variables are not sufficient in an NPL forecasting model for different groups of banks. The model also needs to have the capability to capture idiosyncratic shocks, as does our Model 9 above.

4.2 Fair Deposit Insurance Premium Rate for Each Group of Banks

In this paper, a fair premium rate is defined as a rate that covers the operational expenditures of an insuring agency (e.g. personnel costs and equipment costs), provides it with sufficient funds to enable it to pay a certain percentage of deposit amounts to depositors in the case of a banking default, and provides it with sufficient funds as precautionary reserves to secure itself against further failures. High premium rates reduce the capital adequacy of individual financial institutions, which can in turn endanger the stability of the financial system. Low premium rates reduce the overall safety of the financial system.

Figure 7 shows a bank's balance sheet in the case of default. In order to calculate the fair deposit insurance premium rate, we need to calculate the financial assistance of the deposit insurance.

Figure 7: Financial Assistance of the Deposit Insurance Corporation/Agency in a Failed Bank's Balance Sheet
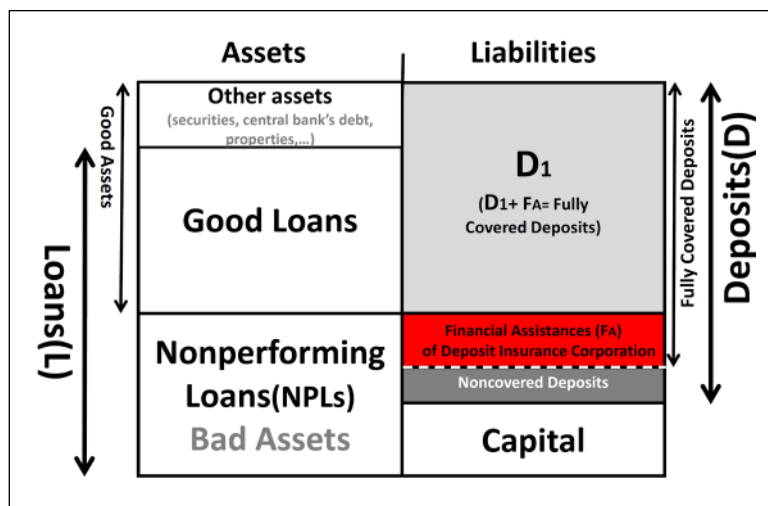
Source: Authors' compilation.

To estimate fair premium rates for each group of banks, we need to make some assumptions regarding: the percentage share of insurance coverage for each type of deposit; the level of the insuring agency's operational expenditures; the estimated default ratio of NPLs; and the percentage share of excess over the forecasted financial assistance from the DIC and operational expenditures that need to be kept by this organization as precautionary reserves.

To calculate the fair premium rate for each group of banks in this case, we make the following assumptions (though these assumptions can be modified to take account of decisions by policymakers and monetary authorities):

i.   All banks pay the same membership fee to the DIC, as we do not have information about the DIC's operational expenditures. This membership fee is the only source of financing operation expenditures (e.g. personnel costs and equipment costs) of the insuring organization and premium income of the DIC is the only source of financial assistance from the DIC in the case of bank failure.

ii. Insurance coverage for long-term deposits is 80%, and other deposits, including savings deposits, short-term deposits, current accounts, and other deposits, are fully covered, meaning in the case of bank default that 100% of the latter is refunded by the DIC to depositors.

iii. The default ratio of NPLs is 100%.

iv. For all assets except loans ("other assets" in Figure 5), the liquidity percentage is 90%, meaning that in the case of bank default, 90% of other assets is converted into cash and the remainder goes to default.

v.   Assets and liabilities grow every year in accordance with the CPI inflation rate.

vi. The DIC will keep 10% in excess of forecasted financial assistance as precautionary reserves.

Based on these assumptions and our earlier forecast for NPLs, we estimate the present value of FA, and are thus able to obtain fair deposit insurance premium rates for each group of banks. For Group 1, which is the group with higher soundness and stability, the calculated premium rate is 0.64%, and for Group 2, which has lower soundness and stability, the rate is 0.86%. To calculate the fair premium income from each group of banks, these two rates need to be multiplied by the amount of eligible deposits. The effective fair premium rate, which is the weighted average of the two estimated premium rates, is 0.83%. This rate can be used by the DIC if it decides to adopt a single premium rate policy.

5. CONCLUSION

Since the start of the recent global financial crisis, triggered by the collapse of Lehman Brothers in September 2008, there has been an ongoing international debate about the reform of financial regulation and supervision intended to prevent the recurrence of a similar crisis. Strengthening deposit insurance systems is one of the fundamental steps in this reform. It is crucial for each country to select "fair" deposit insurance premium rates to maintain financial system stability, thereby protecting depositors and ensuring the settlement of funds related to failed financial institutions. A fair rate refers to a rate that covers the operational expenditures of an insuring agency (e.g. personnel costs and equipment costs) and provides sufficient funds to the insuring agency to enable it to financially assist any failed depository financial institutions. This insurance agency also keeps an appropriate amount of precautionary reserves at the end of each financial period to secure itself against further failures. A high premium rate reduces the capital adequacy of individual financial institutions, which endangers the stability of the financial system. A low premium rate reduces the security of the financial system.

It is unfair for all banks, healthy or unhealthy, to pay the same premium rate to the insuring agency. Unsound and riskier financial institutions that jeopardize the stability of the financial system should pay higher premiums than sound financial institutions that keep their nonperforming loans (NPLs) at reasonable levels and demonstrate good financial performance. Hence, it is necessary to have at least a dual fair premium rate system, which is the main argument of this paper.

For classification and credit rating of financial institutions, we use two statistical techniques on various financial variables taken from banks' statements. The underlying logic of both techniques—principal component analysis and cluster analysis—is dimension reduction (i.e. summarizing information on numerous variables in just a few variables), but they achieve this in different ways. These techniques enable us to classify our sample of banks, made up of all banks, into two clusters—one cluster has higher soundness and stability than the other, and should, based on the aforementioned logic, pay a lower deposit insurance premium rate.

The level of financial assistance from the deposit insuring agency (DIC) in the case of banking failure is mainly based on the amount of NPLs—the larger the amount of NPLs, the higher the default risk, so the greater the financial assistance the DIC provides. Hence, we develop a model for forecasting NPLs (Model 9). Our model has the capability to capture macroeconomic shocks as well as idiosyncratic shocks. The macroeconomic variables we use in our model are GDP, price of stock, price of land, and government bond interest rate. When land prices increase, collateral value increases as well, so the default risk of loans declines. When business conditions improve, increases in GDP growth and stock prices cause a reduction in default risk. When the government bond interest rate, one of the safest asset interest rates, rises, banks tend to invest more in safe assets that reduce default risks. Four macro variables can capture macro shocks, but some banks fail even if the macro financial system is sound and healthy. Hence, we add additional variables that can capture idiosyncratic uncertainty in the economy. To obtain these variables that can capture idiosyncratic shocks, we take eight financial variables from all banks' statements, which explain all financial characteristics of these banks. To summarize information on these eight variables in just a few variables, we use principal component analysis, which reduces them to three components that can capture idiosyncratic shocks. We subsequently use these three components in our model for forecasting NPLs.

To forecast NPLs, we run our model in a vector error correction setting. We conduct impulse response (IR) analysis to provide further evidence of the dynamic response of NPLs to macro and idiosyncratic innovations. IR analysis backs up our suggestion that macro variables are not sufficient in an NPL forecasting model for different groups of banks. The model needs to have the capability to capture idiosyncratic shocks as well. Our empirical analysis for two groups of banks reveals that NPL/loans for both bank groups respond similarly to unanticipated macroeconomic shocks, but their response to idiosyncratic shocks varies.

Finally, using our results of forecasting NPLs and employing Model 8, we calculate the fair premium rates for both groups of banks and also the effective premium rate, which is the weighted average of these two premium rates. The effective rate can be used by the DIC if it decides to adopt a single premium rate policy. Our calculated fair premium rate for the group that has higher soundness is lower than that for the other group.

REFERENCES

Abildgren, Kim. 2016. "Household Micro-data, Regulation and Financial Stability: The Case of Denmark in the 1950s." *Studies in Economics and Finance* 33(2): 320–335. doi: 10.1108/SEF-07-2015-0176

Acharya, S., and J. Dreyfus. 1989. "Optimal Bank Reorganization Policies and the Pricing of Federal Deposit Insurance." *Journal of Finance* 44(5): 1313–1333.

Bartholdy, J., G. W. Boyle, and R. D. Stover. 2003. "Deposit Insurance and the Risk Premium in Bank Deposit Rates." *Journal of Banking & Finance* 27(4): 699–717.

Buser, S. A., A. H. Chen, and E. J. Kane. 1981. "Federal Deposit Insurance, Regulatory Policy, and Optimal Bank Capital." *Journal of Finance* 36(1): 51–60.

Cooper, Elizabeth W. 2011. "Determinants of Off–balance Sheet Usage in Private Banks," *Studies in Economics and Finance* 28(4): 248–259, doi: 10.1108/10867371111171528

Everitt, B. S., S. Landau, and M. Leese. 2001. *Cluster Analysis*. 4th ed. London: Edward Arnold Publishers.

Feger, T., and J. Asafu-Adjaye. 2014. "Tax Effort Performance in Sub-Sahara Africa and the Role of Colonialism." *Economic Modelling* 38: 163–174.

Ho, C. B., and D. D. Wu. 2009. "Online Banking Performance Evaluation Using Data Envelopment Analysis and Principal Component Analysis." *Computers & Operations Research* 36: 1835–1842.

Horvitz, P. M. 1983. "The Case against Risk-related Deposit Insurance Premiums." *Housing Finance Review* 2(3): 253–263.

Huang, Z., H. Chen, C. J. Hsu, W. H. Chen, and S. Wu. 2004. "Credit Rating Analysis with Support Vector Machines and Neural Networks: A Market Comparative Study." *Decision Support Systems* 37: 543–558.

Hwang, D., C. F. Lee, and K. T Liaw. 1997. "Forecasting Bank Failures and Deposit Insurance Premium." *International Review of Economics & Finance* 6(3): 317–334.

Inakura, N., and S. Shimizutani. 2010. "Deposit Insurance and Depositor Discipline: Direct Evidence on Bank Switching Behaviour in Japan." *Applied Economics* 42(26): 3401–3415.

Kahlert, D. and Wagner, N. 2017. "Are Systemically Important Eurozone Banks Undercapitalized? A Stress Testing Approach" (April 2017). Available at SSRN: https://ssrn.com/abstract=2568614 or http://dx.doi.org/10.2139/ssrn.2568614

Kane, E. J. 1986. "Appearance and Reality in Deposit Insurance: The Case for Reform." *Journal of Banking & Finance* 10(2): 175–188.

Landau, S., and I. Chis Ster. 2010. "Cluster Analysis: Overview." In: *International Encyclopedia of Education*. 3rd ed. Oxford, UK: Elsevier. pp. 72–83.

Martinez, W. L., and A. R. Martinez. 2005. *Exploratory Data Analysis with Matlab*. Boca Raton, FL: Chapman and Hall/CRC Press.

Orsenigo, C., and C. Vercellis. 2013. "Linear versus Nonlinear Dimensionality Reduction for Banks' Credit Rating Prediction." *Knowledge-Based Systems* 47: 14–22.

Poon, W. P. H., M. Firth, and H. G. Fung. 1999. "A Multivariate Analysis of the Determinants of Moody's Bank Financial Strength Ratings." *Journal of International Financial Markets, Institutions and Money* 9: 267–283.

Ravi Kumar, P., and V. Ravi. 2007. "Bankruptcy Prediction in Banks and Firms via Statistical and Intelligent Techniques–A Review." *European Journal of Operational Research* 180: 1–28.

Yoshino, N., and T. Hirano. 2011. "Pro-cyclicality of the Basel Capital Requirement Ratio and its Impact on Banks." *Asian Economic Papers* 10(2): 22–36.

———. 2013. "Counter-cyclical Buffer of the Basel Capital Requirement and its Empirical Analysis." In: *Current Developments in Monetary and Financial Law.* Volume 6. pp. 417–429. Washington: International Monetary Fund.

Yoshino, N., and F. Taghizadeh-Hesary. 2014a. Hometown Investment Trust Funds: An Analysis of Credit Risk. *ADBI Working Paper* No. 505. Tokyo: Asian Development Bank Institute.

———. 2014b. "Analytical Framework on Credit Risks for Financing SMEs in Asia." *Asia-Pacific Development Journal* 21(2): 1–21.

———. 2014c. "Monetary Policies and Oil Price Fluctuations Following the Subprime Mortgage Crisis." *International Journal of Monetary Economics and Finance* 7(3): 157–174.

———. 2015. "Analysis of Credit Risk for Small and Medium-Sized Enterprises: Evidence from Asia." *Asian Development Review* 32(2): 18–37.

Yoshino, N., F. Taghizadeh-Hesary, A. Hassanzadeh, and A. Danu Prasetyo. 2014. "Response of Stock Markets to Monetary Policy: The Tehran Stock Market Perspective." *Journal of Comparative Asian Development* 13(3). doi:10.1080/15339114.2014.985458

Yoshino, N., F. Taghizadeh-Hesary, and F. Nili. 2013. "Optimal Premiums for the Deposit Insurance System: An Empirical Work on the Deposit Insurance System of Japan." *Keio Economic Studies Discussion Paper*. KESDP No. 13–1. Tokyo: Keio University.

Yoshino, N., F. Taghizadeh-Hesary, P. Charoensivakorn, and B. Niraula. 2016. "Small and Medium-sized Enterprise (SME) Credit Risk Analysis Using Bank Lending Data: An Analysis of Thai SMEs." *Journal of Comparative Asian Development* 15(3): 383–406.