საქართველოს მეცნიერებათა ეროვნული აკადემიის მოამბე, ტ. 18, №1, 2024
BULLETIN OF THE GEORGIAN NATIONAL ACADEMY OF SCIENCES, vol. 18, no. 1, 2024

Economics

An Alternative Approach to Modelling Default Probabilities for Macroprudential Stress Testing

Iuri Ananiashvili* **and Beka Gurgenidze***

* *Faculty of Economics and Business, Ivane Javakhishvili Tbilisi State University, Tbilisi, Georgia*

(Presented by Academy Member Vladimir Papava)

This paper addresses the challenge of estimating default probabilities in situations where loan transitions across different categories are unobservable and only aggregate proportions data is available. An alternative approach that connects default probabilities to macroeconomic variables is proposed. Through Monte Carlo simulation, we have compared the results of our approach with the benchmark method and found minimal differences. This alternative approach holds particular promise for developing countries, especially for those whose banking systems are still in the process of adopting IFRS 9 standards. It enables central banks to project conditional default probabilities in diverse macroeconomic scenarios, offering a valuable tool for enhancing financial system stress testing frameworks. By incorporating our method, central banks can gain an alternative perspective on their financial system's resilience, potentially improving risk assessment and decision-making in macroprudential policy. *© 2024 Bull. Georg. Natl. Acad. Sci.*

probability of default, Markov transition matrices, stress test, credit risk

Since the Great Financial Crisis (GFC), the importance of assessing the resilience of the banking system to adverse macroeconomic shocks has significantly increased. Central banks developed tools such as stress tests to evaluate the credit risk of financial institutions for different macroeconomic scenarios. One of the key elements of stress tests is the estimation of the effects of macroeconomic variables on default probabilities. There is a variety of historical data in most of the developed countries, including borrower-based data covering the business cycle. For such countries, there are a lot of options to model default behaviour [1]. For developing countries, where the estimation of default probabilities is related to some problems and difficulties due to data availability, it is challenging to link macroeconomic variables to credit risk. There are several types of approaches that can be applied in developing countries. One type of these methods that try to relate macroeconomic variables to credit risk, use logit transformation of non-performing loan ratio as a dependent variable and macroeconomic variables as independent variables [2]. Another type of methods try to estimate loan transition matrices using aggregate proportions data. This approach is useful when loan transitions to different categories are not observable. Based on this methodology, transition matrices can be conditioned on economic fundamentals and different estimates for contraction and expansion periods can be obtained [3, 4]. Since 2018, many countries adopted IFRS 9 standards. Due to the changes related to IFRS 9, [5] suggested the approach for credit risk modelling in top-down stress tests. The idea by [5] to model transition matrices follows the methodology developed by [6]. However, the approach suggested by [5] does not consider the problem related to transition data availability for developing countries, which are in the process of adopting IFRS 9 standards. In addition to that, the methodology described by [5] depends on the idea of a latent variable, which is assumed to follow standard normal distribution, but in practice, this assumption is rarely satisfied [7].

This paper will be trying to develop an additional method to relate default probabilities to macroeconomic variables. This approach will be particularly useful for developing countries, where loan transitions among different categories are not observable and only aggregate proportions data are available. The method presented in this paper is based on the theoretical framework suggested by Shao J., et al. [7]. However, the fundamental difference between the work done by Shao J., et al. and the approach considered in this paper is that in their work [7] use a cohort method to derive the empirical transition probability matrix, but this paper uses generalized least squares estimation (GLS) to compute transition probability matrix based on aggregate proportions data. In an ideal situation, when loan transitions between different categories are observable, the cohort approach gives more accurate results than the GLS estimation based on the aggregate proportions data. In this paper, Monte Carlo simulation will be used to generate loan contract transitions among different categories. Then the transition probability matrix estimated by the cohort approach will be used as a benchmark to compute default probabilities conditioned on macroeconomic variables. These benchmark estimates will be compared to the output when conditional default probabilities are computed using aggregate proportions data. As the results indicate, the difference between the benchmark and the GLS output is not significant, suggesting that in countries where the individual transitions among loan categories are not observable and the cohort estimation is not available, in order to link macroeconomic variables to the default probabilities, it is reasonable to use the GLS estimates based on aggregate proportions data. This indicates that central banks in developing countries can have an alternative method to conduct macroprudential stress tests to evaluate the resilience of the banking system. The approach suggested in this paper might be particularly useful for countries, where the banks are in the process of transitioning to IFRS standards and for stress testing purposes default probabilities for loan products (e.g. mortgages, consumer and business) are needed.

Linking Macroeconomic Variables with Default Probabilities

This section follows the idea of [7] applied to 3 loan categories and is extended to the case when loan transitions are not observable. It is assumed that the latent variable X is given in the following form:

$$X = V'\beta + Z, \tag{1}$$

where V is the vector of macroeconomic variables representing the business cycle, β is the vector of parameters to be estimated and Z is the idiosyncratic component which is standard, normally distributed. In this paper, K' denotes the transpose of K. It is further assumed that Z and V are mutually independent. For the borrower, who is initially in the grade i there exist thresholds in the support of the random variable X. Namely, if X_t^i is located at $(x_{j-1}^i, x_j^i]$, then it means that the borrower i will be in the loan category j at the end of time period t. Since there are only 3 loan categories, i and j take values from $C = \{1, 2, 3\}$, for which the numbers 1, 2, and 3 denote standard, watch and non-performing loan categories, respectively.

Therefore, for the borrower, who is initially in the loan category number 1 the thresholds $(x_0^1, x_1^1]$, (x_1^1, x_2^1) and $[x_2^1, x_3^1)$ mean that, based on the location of X_t^1, this borrower will probably stay in loan category one, move to loan category 2 or move to loan category 3, respectively. Most interestingly, the assumption about the distribution of X is not needed. To illustrate the idea in terms of probabilities for the borrower located in loan category number 1, consider probabilities conditioned on macroeconomic variables:

$$P(X_t^1 > x_1^1 \mid V = v) = P(Z > x_1^1 - v'\beta) = \sum_{j=2}^{3} p_{1jt}(v)\,, \tag{2}$$

where $p_{1jt}(v)$ denotes conditional transition probability from state 1 to state j. Since it is assumed that Z follows the standard normal distribution, from (2) it is easily derived that:

$$1 - \sum_{j=2}^{3} p_{1jt}(v) = P(Z < x_1^1 - v'\beta) = F(x_1^1 - v'\beta). \tag{3}$$

In (3), F represents the cumulative distribution function of a random variable Z. From (3) follows that:

$$x_1^1 - v'\beta = F^{-1}\left(1 - \sum_{j=2}^{3} p_{1jt}(v)\right). \tag{4}$$

If $p_{1jt}(v)$ is replaced by the estimator $\widetilde{p_{1jt}}(v)$, it can be obtained:

$$x_1^1 - v'\beta + e_t = F^{-1}\left(1 - \sum_{j=2}^{3} \widetilde{p_{1jt}}(v)\right), \tag{5}$$

where e_t is the error term caused by the replacement. In other words, (5) represents a multivariate regression model, which can be estimated using different econometric methods. Equation (5) is shown for the borrower who is in category number 1, but the idea is the same for all types of borrowers. After estimating x_1^1 and β, it is straightforward to have fitted values and to obtain estimates for transition probabilities. Equation (5) allows to project transition probabilities for different macroeconomic scenarios.

It is important to mention, that in order to obtain the estimator $\widetilde{p_{1jt}}(v)$ in (5), Gross M., et al. use the cohort method [5]. However, in case the transitions to loan categories are not observable, this paper employs the GLS estimation method based on aggregate proportions data, which is described in detail in the next part. Moreover, the next section describes a simulation by which transitions to different loan categories and aggregate proportions data are generated. This simulation allows us to compare the conditional transition probabilities calculated using cohort and GLS methods.

Estimation of Markov Transition Matrices Using Proportions Data

This section describes a discrete Markov chain process. Assume that there is a sequence of random variables: $\{X_t\}_{t=1}^{T}$, where X_t represents the borrower's state in time t, taking values from $C = \{1, 2, 3\}$. To describe the Markov process it is convenient to write transition probabilities in a matrix form:

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}. \tag{6}$$

where p_{ij} denotes the transition probability from state i to state j and $\widehat{p_{ij}} = n_{ij} \Big/ \sum_{j=1}^{3} n_{ij}$. When borrowers' transitions to different categories are observable, estimation of the matrix P is straightforward, namely, using cohort method $\widehat{p_{ij}} = n_{ij} \Big/ \sum_{j=1}^{3} n_{ij}$ for time t, where n_{ij} is the number of transitions from state i to state j and $\sum_{j=1}^{3} n_{ij} = 1$ is the total number of transitions from state. However, the situation is different, when the transitions are not observable, but only proportions data are available in each category. In this case, a simple idea can be applied. In particular, from the theoretical point of view, using the total probability rule, for any j and i:

$$P(X_t = j) = \sum_{i=1}^{3} P(X_{t-1} = i) P(X_t = j \mid (X_{t-1} = i). \tag{7}$$

In equation (7), total probabilities can be replaced by observable proportions, namely:

$$y_j(t) = \sum_{i=1}^{3} y_i(t-1) p_{ij} + u_i(t), \tag{8}$$

where $y_i(t)$ denotes the portion of borrowers in category j and $u_i(t)$ is an error term due to the replacement. It is convenient to write (8) in the matrix form for $j = 1,2,3$:

$$\begin{bmatrix} y_j(2) \\ \vdots \\ y_j(t-1) \\ \vdots \\ y_j(T) \end{bmatrix} = \begin{bmatrix} y_1(1) & y_2(1) & y_3(1) \\ \vdots & \vdots & \vdots \\ y_1(t-1) & y_2(t-1) & y_3(t-1) \\ \vdots & \vdots & \vdots \\ y_1(T-1) & y_2(T-1) & y_3(T-1) \end{bmatrix} \begin{bmatrix} p_{1j} \\ p_{2j} \\ p_{3j} \end{bmatrix} + \begin{bmatrix} u_j(2) \\ \vdots \\ u_j(t-1) \\ \vdots \\ u_j(T) \end{bmatrix}. \tag{9}$$

In order to write the whole system of equations, the left side vector of (9) will be denoted by y_j, the first matrix of the right side of (9) will be denoted by x_j, the probability vector will be denoted by p_j and the error term vector will be denoted by u_j. Based on these notations the whole system of equations can be written as follows:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_1 & 0 & 0 \\ 0 & x_2 & 0 \\ 0 & 0 & x_3 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \tag{10}$$

or

$$Y = Xp + U. \tag{11}$$

For (11) it is assumed that the error term is normally distributed and $E(U) = 0$, $E(UU') = \Sigma$. Considering the fact that the parameters to be estimated are probabilities and the sum of each row in (6) has to equal 1, it is convenient to remove the equations for y_3 from (10) and (11) and obtain a reduced system of equations:

$$Y^* = X^* p^* + U^*. \tag{12}$$

After the estimation of (12), it is easy to obtain the values for the last column of (6) by using the formula: $p_{i3} = 1 - \sum_{j=1}^{2} p_{ij}$, $i = 1,2,3$. In order to estimate the transition probabilities from (12), the iterative generalized least squares technique will be employed suggested by [8]. The first step of the procedure is to estimate the transition probabilities using constrained least squares, then use these estimates to compute the covariance matrix of the error term. The estimated covariance matrix will be used in the GLS estimation. This procedure will be continued until the last two estimates of the transition probability matrix will not be close to each other. Formally the procedure is written as follows:

$$\min_{p^*}(Y^* - \breve{X}p^*)'\hat{\Sigma}^{-1}(Y^* - \breve{X}p^*), \quad Gp^* \le \mu, \quad p^* \ge 0\,, \tag{13}$$

where

$$Y = [Y_2', Y_3', \dots, Y_T']', \qquad Y_t = [y_{1t}, y_{2t}, y_{3t}]' \;,$$

$$\breve{X} = [\breve{X}_1', \breve{X}_2, \dots, \breve{X}_{T-1}']', \qquad \breve{X}_t = I_3 \otimes X_t' \,,$$

$$X_t = [y_{1t}, y_{2t}, y_{3t}, y_{4t}]' \;,$$

$$\hat{\Sigma} = diag(\hat{\Sigma}_1, \hat{\Sigma}_2, \dots, \hat{\Sigma}_{T-1}),$$

$$\hat{\Sigma} = diag(P^{*\prime} X_t) - P^{*\prime} diag(X_t) P^*.$$

In the expression (13), $G = [I_1, I_2]$, where $I_1 = I_2$ they represent identity matrices with dimensions 3 by 3. I_3 is an identity matrix with a dimension of 2 by 2. μ is an identity vector of size 3, $\otimes$ – represents Kronecker product, P^* is the matrix (6) but excluded the last column and the operator $diag$ – makes diagonal matrix from the elements inside the operator. In the purpose of the simulation discussed below, it is assumed that all the elements of μ consist of 1's. However, in practical applications, if it is assumed that the third category is absorbing, the last element of μ can be equal to 0, meaning that $\sum_{j=1}^{2} p_{3j} \le 0$ in the constraint in (13).

In practice, observable proportions in (8) can be calculated based on the number of contracts in each category. However, if these contracts are not observable, or there are data issues related to this, instead of using the number of contracts, loan amount in each category can be employed to compute proportions in (8). In the following section, the proportions in (8) are calculated based on the number of contracts in each category created by the simulation.

Simulation and Results

As a first step of the simulation, the "true" transition probability matrix (6) was constructed in each time period, 36 periods in total. Initial values of (6) were arbitrarily created and in each time period, these initial values were linked to GDP growth. The idea behind the creation of a time-varying probability transition matrix that is linked to GDP growth is to construct the data set in which the movement of borrowers to different loan categories will also be related to GDP growth. It is important to mention again that initial values of (6) are arbitrary and their selection does not affect the results. The next step is to generate transitions of borrowers to different credit categories. It is assumed that the portfolio consists of 100 borrowers and initially, they are randomly assigned to loan categories, namely 80 percent of borrowers are initially in state 1 and the remaining 20 percent in states 2 and 3. In the next period, a borrower n, for $n = 1, 2, \dots, 100$ will remain in the same loan category or move to another category. The movement of a borrower n will depend on the random draw from uniform distribution $U_n \sim U[0,1]$. In particular, if at time period t, a borrower n is in category 1, in the next period he/she will be in the same category, move to category 2 or 3, if random draw from U_n falls in $[0, p_{11,t}]$, $(p_{11,t}, p_{11,t} + p_{12,t}]$ or $(p_{11,t} + p_{12,t}, 1]$, respectively. It has to be mentioned, that the probabilities above are already given in the "true" transition probability matrix. The procedure of Monte Carlo simulation can be summarized as follows:

$$X_{n,t+1} = \begin{cases} 1: if\ U_{n,t+1} \in (0, p_{i1,t}] \\ 2: if\ U_{n,t+1} \in (p_{i1,t}, p_{i1,t} + p_{i2,t}] \\ 3: if\ U_{n,t+1} \in (p_{i1,t} + p_{i2,t}, 1] \end{cases}, \quad (14)$$

where, as stated in the previous part, $X_{n,t+1}$ shows the state of the nth borrower in period $t+1$, given that initially the borrower was in loan category $i = 1,2,3$.

Monte Carlo simulation allowed us to generate transitions of 100 borrowers to loan categories for 36 periods. In addition to that, it was straightforward to compute the proportions of borrowers in each category for each time period. Based on the data, on the one hand, $\widetilde{p_{1jt}}(\nu)$ in (5) was calculated using the cohort method and, on the other hand, it was computed using the GLS algorithm described in the previous section. Applying the idea discussed in the first part of the paper the result for the conditional transition probability from category 1 to category 3 (default probability) is given below for the two methods.

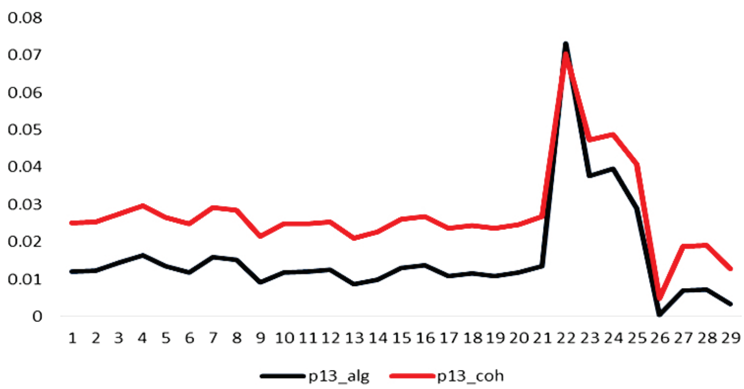

Fig. Conditional transition probabilities.
Source: Authors' calculations.

In Figure p13_alg and p13_coh show conditional probabilities estimated using the GLS algorithm and cohort method, respectively. As the results indicate, the path of the two probabilities are similar and the estimates are close to each other. The results presented in Figure are based on a single Monte Carlo simulation. However, conducting several simulations does not give significantly different results. Estimations using the cohort method are more accurate in general than the iterative GLS estimations based on aggregate proportions data. However, since the estimates of conditional default probabilities given by these two methods do not differ significantly and both reflect the GDP growth similarly, it can be concluded that, in general, estimation of conditional default probabilities based on aggregate proportions data is a good solution, when individual transitions data are unavailable.

Conclusion

This paper proposes a possible way to link macroeconomic variables with default probabilities, when individual transitions data are unavailable and only aggregate proportions data are observable. According to the results of the simulation, the computations based on the suggested approach do not differ significantly from the estimates obtained from the benchmark (cohort) method. It can be concluded, that in developing countries, where there are data limitations, it is reasonable to apply the approach described in this study. By implementing this method, central banks will be able to project conditional default probabilities in different macroeconomic scenarios and incorporate the results into their stress testing frameworks to have an alternative assessment of the financial system's resilience.

ეკონომიკა

დეფოლტის ალბათობის მოდელირების ალტერნატიული მიდგომა მაკროპრუდენციული სტრესტესტისათვის

ი. ანანიაშვილი* და ზ. გურგენიძე*

* *ივანე ჯავახიშვილის სახ. თბილისის სახელმწიფო უნივერსიტეტი, ეკონომიკისა და ბიზნესის ფაკულტეტი, თბილისი, საქართველო*

(წარმოდგენილია აკადემიის წევრის ვ. პაპავას მიერ)

ნაშრომში განხილულია დეფოლტის ალბათობების შეფასების ალტერნატიული მიდგომა სიტუაციისათვის, როდესაც სესხების სხვადასხვა კატეგორიაში გადასვლები დაკვირვებადი არ არის, მაგრამ არსებობს აგრეგირებული ფარდობითი მონაცემები. პრობლემის გადასაწყვეტად შემოთავაზებული ალტერნატიული მიდგომა დეფოლტის ალბათობებს მაკროეკონომიკურ ცვლადებთან აკავშირებს. მონტე-კარლოს სიმულაციაზე დაყრდნობით ნაჩვენებია, რომ შემოთავაზებული მეთოდის შედეგები მცირედით განსხვავდება პრაქტიკაში აპრობირებული საორიენტაციო მიდგომის შედეგებისგან. ეს გარემოება გვაფიქრებინებს, რომ შემოთავაზებული ალტერნატიული მეთოდი, განსაკუთრებით, გამოსადეგი იქნება იმ განვითარებადი ქვეყნებისათვის, რომელთა საბანკო სისტემაც ჯერ კიდევ ფინანსური აღრიცხვის საერთაშორისო სტანდარტებზე (IFRS 9) გადასვლის პროცესშია. სხვადასხვა მაკროეკონომიკური სცენარის დროს, აღნიშნული ქვეყნების ცენტრალურ ბანკებს მეთოდის გამოყენება დაეხმარება საკუთარი სტრეს-ტესტის ჩარჩოსთვის საჭირო დეფოლტის პირობითი ალბათობების შეფასებასა და მაკროპრუდენციული პოლიტიკის გატარებაში.

REFERENCES

1. Simons D. and Rolwes F. (2009) Macroeconomic default modeling and stress testing, *International Journal of Central Banking,* 18:177-204.
2. Wezel T., Canta M., Luy M. (2014) A practical example of the nonperforming loans projection approach to stress testing, chapter 30, in: a guide to IMF stress testing. Methods and Models (ed.), L. Ong, International Monetary Fund.
3. Jones M.T. (2005) Estimating Markov transition matrices using proportions data: an application to credit risk, IMF Working Papers 2005/219, available at: https://www.imf.org/en/Publications/WP/Issues.
4. Washle D. (2016) Estimating transition probabilities from aggregate mortgage arrears data in Ireland, available at SSRN: https://ssrn.com/abstract=2817825.
5. Gross M., Laliotis D., Leika M., Lukyantsau P. (2020) Expected credit loss modeling from a top-down stress testing perspective, IMF Working Papers 20/111, available at: https://www.imf.org/en
6. Belkin B. and Suchower S. (1998) A one-parameter representation of credit risk and transition matrices. *CreditMetrics Monitor,* 46-56.
7. Shao J., Li S. and Li, Y. (2016) Estimation and prediction of credit risk based on rating transition systems, available at: https://arxiv.org/abs/1607.00448.
8. MacRae E.C. (1977) Estimation of time-varying Markov processes with aggregate data, *Econometrika,* **45** (1): 183–98.

Received September, 2023