

A Forecasting Model for the Likelihood of Delinquency, Default or Prepayment: The Case of Taiwan

Chin-Tsai Lin^a
Shih-Yu Yang^b

^a*Department of Information Management, Yuan Pei Institute of Science and Technology*

^b*Department of Finance, Chihlee Institute of Commerce, Panchiao, Taipei220, Taiwan,
R.O.C.*

ABSTRACT

In a competitive and dynamic market, financial institutions must forecast the proportion of mortgages that will become delinquent, default or prepay. This paper develops a novel forecasting model with nonstationary Markov chain and Grey forecasting, capable of predicting the likelihood of delinquency, default and prepayment. Home mortgage data, obtained by a major Taiwan financial institution from January 1, 1996 to June 30, 1998, are adopted to examine the forecasting effectiveness of the novel forecasting model and the ARIMA model. Empirical results indicate that the novel forecasting model with a low error is better than ARIMA. Thus, the novel forecasting model provides a promising means of accurately predicting the probabilities of delinquency, default and prepayment.

JEL: C60, G2, G21, O53

Keywords: Forecasting; Mortgage; Loan; Delinquency; Default; Prepayment

I. INTRODUCTION

The home mortgage sector is critical to financial institutions. However, delinquency, default and prepayment of home mortgages make regulating funds difficult for financial institutions. Managers in the home mortgage sector require the ability to forecast the proportion of mortgages that will be delinquent, defaulted or prepaid.

Many methods have been used to forecast delinquency, default or prepayment. For example, Standard & Poors (S&P) and Moody's employed accounting analytic and migration analysis to predict the probability of default. Statistical methods for forecasting default risk include linear discriminant analysis (Altman, 1968), the logistic regression model (Martin, 1977; Smith and Lawrence, 1995) and neural networks (Altman, Marco and Varetto, 1994). Meanwhile, Merton (1974) employed an option-theoretic approach to default risk. Migration analysis is the most interesting. Researchers have typically used Markov chains to perform migration analysis. Cyert, Davidson and Thompson (1962) applied stationary Markov chains to model credit accounts. However, the suitability of the stationary Markov chain remains unproven. Smith and Lawrence (1995) considered a dynamic environment and constructed a forecasting model with a Markovian structure and non-stationary transition probabilities. The forecasting model is structured as a mathematical recursion to predict the probability of a loan's being in any one of the alternative financial states annually. The calculation process is too complex to understand. Neither Cyert etc. (1962) nor Smith and Lawrence (1995) predict the probability of delinquent payment and prepayment. Smith and Lawrence (1995, 1996) also forecast losses at annual intervals, but such information is not sought by financial institutions. In a competitive and dynamic market, financial institutions must be able to predict the situation over the forthcoming months. However, previous studies have not considered this fact.

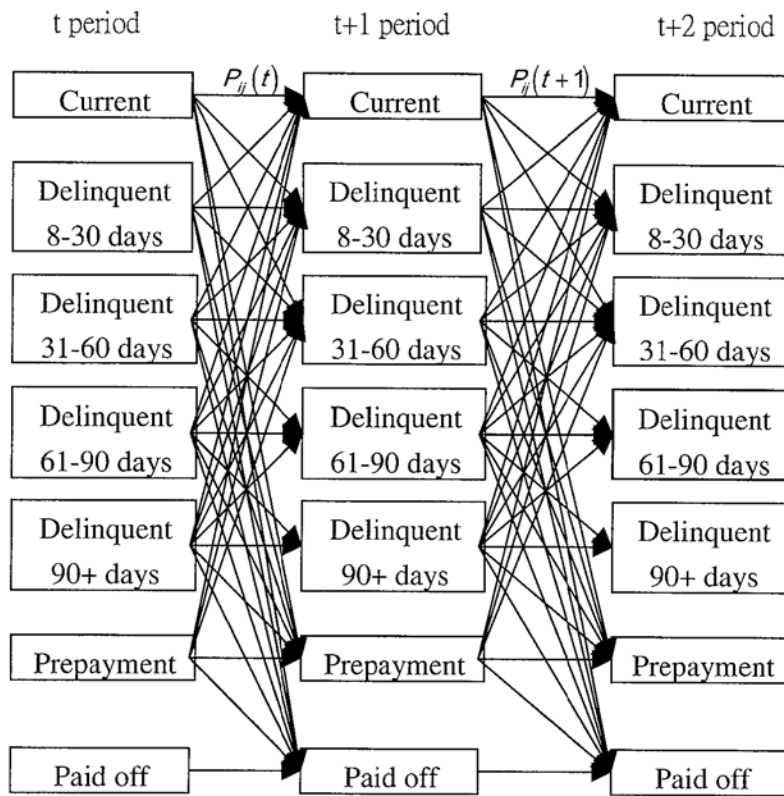
The above methods typically require many data to build the forecasting model. However, Taiwan's financial institutions normally lack such a complete database. Grey forecasting does not depend on a large amount of data. Grey forecasting is fit for Taiwan's and other financial institutions, which possess insufficient data. Furthermore, Xu and Wen (1997) applied grey forecasting to forecast accurately the passengers of international air transportation. Yi (1987) used grey forecasting to predict the number of talented persons. Their results illustrated the ability of grey forecasting to deal effectively with incomplete or uncertain information. Accordingly, this paper constructs a forecasting model with nonstationary Markov chain and grey forecasting, to forecast the likelihood of delinquency, default and prepayment of home mortgages on a monthly basis. The method can mitigate the difficulties in home mortgage risk management, facilitate financial institutions' controlling of funds, and help to establish appropriate reserves to protect against losses.

II. METHODOLOGY

This paper applied nonstationary Markov chain and Grey forecasting to construct a forecasting model. First, a nonstationary Markov chain was used to describe the home mortgage payment transition. A loan can be categorized into seven financial states at

any given point in time, as Fig.1 illustrates

Figure 1
Home Mortgage Payment Transition Model



- State 1: current or delinquent 1-7 days
- State 2: delinquent 8-30 days
- State 3: delinquent 31-60 days
- State 4: delinquent 61-90 days
- State 5: delinquent 90+ days or default
- State 6: prepayment
- State 7: paid off

In the above home mortgage payment transition model:

$P_i(t)$ = the probability that the loan is in state i at period t

$P_{ij}(t)$ = the probability that the loan which is in state i at period t will transfer to state j for period $t+1$

Then, the probability that the loan is in state i at period $t+1$ is expressed as a Markov chain, as follows:

$$P_j(t+1) = \sum_{i=1}^6 P_i(t) P_{ij}(t) \quad (1)$$

Thus, Eq. (1) can be used to calculate the probability of delayed payment, default and prepayment per month. In Eq. (1), $P_{ij}(t)$ is unknown at period t . Meanwhile, the transition probability, $P_{ij}(t)$, is nonstationary. $P_{ij}(t)$ could not be computed via the traditional Markov chains methods. Thus, a novel method, the Grey forecasting model, is used here to predict $P_{ij}(t)$.

Deng (1982) proposed the Grey system theory to construct a Grey model for forecasting. Meanwhile, Lin and Yang (1999) applied Grey relational analysis to select home mortgage loans. Similarly, Lin and Chen (1999) used the Grey model to forecast a bank re-decreasing the required reserve ration. The Grey system theory has been verified as helpful in dealing with poor, incomplete and uncertain information. The Grey forecasting model (GM) is the core of Grey system theory. The Grey system theory treats all variables as a Grey quantity within a certain range. Grey forecasting model then collects available data to find the internal regularity. The model examines the nature of internal regularity to manage the disorganized primitive data. The model was established by transferring the arranged sequence into a differential equation. The following illustration details the method used to construct the model adopted herein by creating a sequence of one order linear moving GM (1, 1) (Deng, 1989). The first order differential equation of GM (1, 1) model is

$$\frac{dX^{(1)}}{dt} + aX^{(1)} = b \quad (2)$$

Where t denotes the variables in the system, a represents the developed coefficient, b is the Grey controlled variable, and a and b are the parameters requiring determination in the model. The variables, including $P_{ij}(1), P_{ij}(2), \dots$, and $P_{ij}(6)$, are used to construct the Grey forecasting model and predict $P_{ij}(7)$. The primitive sequence is assumed herein to be as follows:

$$X^{(0)} = (P_{ij}^{(0)}(1), P_{ij}^{(0)}(2), \dots, P_{ij}^{(0)}(6)) \quad (3)$$

When constructing a model, the Grey system must apply one order accumulated generating operation (AGO) to the primitive sequence to provide the middle message of building a model and to weaken the variation tendency. Herein, $X^{(1)}$ is defined as $X^{(0)}$'s one order AGO sequence. That is,

$$\begin{aligned} X^{(1)} &= (P_{ij}^{(1)}(1), P_{ij}^{(1)}(2), \dots, P_{ij}^{(1)}(6)) \\ &= \left(\sum_{t=1}^1 P_{ij}^{(0)}(t), \sum_{t=1}^2 P_{ij}^{(0)}(t), \dots, \sum_{t=1}^6 P_{ij}^{(0)}(t) \right) \end{aligned} \quad (4)$$

From Eqns. (2), (4) and ordinary least square method, coefficient \hat{a} becomes

$$\hat{a} = \begin{bmatrix} a \\ b \end{bmatrix} = (B^T B)^{-1} B^T Y_N \quad (5)$$

Furthermore, accumulated matrix B is

$$B = \begin{bmatrix} -\frac{1}{2} [P_{ij}^{(1)}(1) + P_{ij}^{(1)}(2)] & 1 \\ -\frac{1}{2} [P_{ij}^{(1)}(2) + P_{ij}^{(1)}(3)] & 1 \\ \vdots & \vdots \\ -\frac{1}{2} [P_{ij}^{(1)}(5) + P_{ij}^{(1)}(6)] & 1 \end{bmatrix}$$

Meanwhile, the constant vector Y_N is

$$Y_N = [P_{ij}^{(0)}(2), P_{ij}^{(0)}(3), \dots, P_{ij}^{(0)}(6)]^T$$

The approximate relationship can be obtained as follows by substituting \hat{a} obtained in the differential equation, and solving Eq. (2):

$$\hat{P}_{ij}^{(1)}(t+1) = \left(P_{ij}^{(0)}(1) - \frac{b}{a} \right) e^{-at} + \frac{b}{a} \quad (6)$$

where $\hat{P}_{ij}^{(1)}(1) = \hat{P}_{ij}^{(0)}(1)$, the acquired sequence one order inverse-accumulated

generating operation (IAGO) is acquired and the sequence that must be reduced as Eq. (7) is obtainable.

$$\hat{P}_{ij}^{(0)}(t+1) = \hat{P}_{ij}^{(1)}(t+1) - \hat{P}_{ij}^{(1)}(t) \quad (7)$$

Given $t = 1, 2, \dots, 6$, the sequence of reduction is obtained as follows:

$$\hat{X}^{(0)} = (\hat{P}_{ij}^{(0)}(1), \hat{P}_{ij}^{(0)}(2), \dots, \hat{P}_{ij}^{(0)}(7))$$

Here $\hat{P}_{ij}^{(0)}(7)$ is the Grey elementary predicting value of $P_{ij}(7)$. To conform with the rules of Markov chains, this study uses Eq. (8) to revise.

$$\tilde{P}_{ij}^{(0)}(t) = \frac{\hat{P}_{ij}^{(0)}(t)}{\hat{P}_{i1}^{(0)}(t) + \hat{P}_{i2}^{(0)}(t) + \dots + \hat{P}_{i7}^{(0)}(t)} \quad i=1,2,\dots,6; \quad j=1,2,\dots,7; \quad t=7 \quad (8)$$

From Eq. (8), $\tilde{P}_{ij}^{(0)}(7)$ is obtainable. So $\tilde{P}_{ij}^{(0)}(7)$ is the predicting value of $P_{ij}(7)$ in the Grey forecasting model. $P_i(7)$ is a known number. Then Eq. (1) can be used to compute the predicting probability, $\tilde{P}_j(8)$, of delay payment, default and prepayment.

III. DATA

This paper adopts data regarding home mortgages loaned by a major Taiwanese financial institution. The data refers to 4359 home mortgage loans and includes borrowers' characteristics, mortgage and property characteristics, and loan status. Borrower characteristics include income, education, sex, occupational risk ranking, and years of service to the current company. Mortgage and property characteristics include loan-to-value ratio, loan amount, and loan source. The loan status of individual loans refers to the payment record during the research period after the loan was obtained. Loan status records were used to determine the transition probabilities. The research period is from January 1, 1996 to June 30, 1998, giving 29 months of transition probabilities.

IV. RESULTS

This paper develops a forecasting model using a nonstationary Markov chain and grey forecasting, and to accurately predict the probabilities of delinquency, default and prepayment of home mortgage payment. The transition probability, $P_{ij}(t)$, is

nonstationary in the Markov chain model. It now applies a novel method, the Grey forecasting model, to predict $P_{ij}(t)$. The known transition probabilities of the first, second, third, fourth, fifth and sixth months are used to forecast the transition probability of the seventh month and uses the same rules to accurately predict the transition probability of the eighth month, ninth month, and so on. Then, it obtains the $\tilde{P}_{ij}^{(0)}(t)$, the predicting value of $P_{ij}(t)$, in the Grey forecasting model. It uses Eqn. (1) to calculate the likelihood of delay payment, default and prepayment for a certain month.

To reveal the efficiency of the novel forecasting model, this study takes the ARIMA model as a comparison. This paper uses Eq. (9) to compute the mean absolute deviation (MAD) of above methods. Table 1 summarizes those results.

$$MAD = \frac{1}{n} \sum_{t=1}^n |P_j(t) - \tilde{P}_j(t)| \quad j = 1, 2, \dots, 7 \quad (9)$$

Table 1
The average error in financial states from different methods

Forecasting methods	A forecasting model with Markov			
	MAD	chain and Grey forecasting	ARIMA	No forecast model used
State 1 (current)		0.9717%	1.0600%	5.64%
State 2 (delinquent 8-30 days)		0.7048%	0.8817%	2.92%
State 3 (delinquent 31-60 days)		0.5365%	0.5052%	2.42%
State 4 (delinquent 61-90 days)		0.1926%	0.1991%	1.17%
State 5 (delinquent 90+ days or default)		0.1391%	0.1608%	0.78%
State 6 (prepayment)		0.3330%	0.3743%	1.82%
State 7 (paid off)		0.2726%	0.2983%	1.82%

Table 1 reveals that the proposed forecasting model has fewer mean absolute deviations than the ARIMA model. For example, the mean absolute deviation of state 1 (current) in the new forecasting model in research period is 0.9717%. However, the mean absolute deviation of state1 in the ARIMA model during the research period is 1.0600%. If no other model is adopted to predict the probability of state 1 (current), the mean absolute deviation will reach 5.64%. Similarly, state 2 and the financial state share the same situation. Clearly, the new forecasting model has better predicting validity than the ARIMA model for financial states except state 3. Compared with the ARIMA model the new forecasting model seems a promising way of predicting the probabilities of delinquency, default and prepayment.

V. CONCLUSION

This paper has developed a novel means of accurately predicting the likelihood of delinquency, default and prepayment of home mortgages. The new forecasting model with a nonstationary Markov chain and grey forecasting has become the conventional means of predicting probability. Furthermore, the novel forecasting model considers a dynamic environment and forecasts the probabilities of all financial states. Earlier studies were not as powerful. According to the results presented here, our forecasting model has higher prediction validity than the ARIMA model. Clearly, the new forecasting model is a viable alternative. Small financial institutions in Taiwan often have too few professionals and insufficient data to construct a predictive model. Hence, such institutions commonly assess the likelihood of delinquency, default and prepayment by past experiences. However, the novel forecasting model needs neither professional involvement nor many data. For this reason, financial institutions can use this novel forecasting model to predict the likelihood of delinquency, default and prepayment over forthcoming months. Forecasted probabilities of delinquency, default and prepayment can used to determine the funds that must be maintained by financial institutions over the present and future months.

REFERENCES

- Altman, E. I. (1968). "Financial Ratio Discriminant Analysis and the Prediction of Corporate Bankruptcy," *Journal of Finance*, 23(3), 589-609.
- Altman, E. I., Marco, G., and Varetto, F. (1994). "Corporate Distress Diagnosis: Comparisons Using Linear Discriminant Analysis and Neural Networks (The Italian Experience)," *Journal of Banking and Finance*, 18, 505-529.
- Cyert, R. M., Davidson, H.J. and Thompson, G. L. (1962). "Estimation of the Allowance for Doubtful Accounts by Markov Chains," *Management Sciences*, 8, 287-303.
- Deng, J. (1982). "Control Problems of Grey System," *Systems & Letter*, 5, 288-294.
- Deng, J. (1989). "Introduction Grey System Theory," *Journal of Grey System*, 1, 1-24.
- Green, R. K. and LaCour-Little, M. (1999). "Some Truths about Ostriches: Who Doesn't Prepay Their Mortgages and Why They Don't," *Journal of Housing Economics*, 8(3), 233-248.

- Lin, C. T., and Yang, S. Y. (1999). "Selection of Home Mortgage Loans Using Grey Relational Analysis," *Journal of Grey System*, 11(4), 359-368.
- Lin, C. T. and Chen, L. H. (1999). "A Grey Analysis of Bank Re-decreasing the Required Reserve Ration," *Journal of Grey System*, 11(2), 119-132.
- Martin, D. (1977). "Early warning of Bank Failure: A Logic Regression Approach," *Journal of Banking and Finance*, 1, 249-276.
- Merton, R. C. (1974). "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance*, 29, 449-470.
- Smith, L. D., and Lawrence, E. C. (1995). "Forecasting Losses on a Liquidating Long-term Loan Portfolio," *Journal of Banking and Finance*, 19, 959-985.
- Smith, L. D., Sanchez, S. M., and Lawrence, E. C. (1996). "A Comprehensive Model for Managing Credit Risk on Home Mortgage Portfolios," *Decision Science*, 27, 291-317.
- Xu, Q. Y. and Wen, Y. H. (1997). "The Application of Grey Model on the Forecast of Passenger of International Air Transportation," *Transportation Planning Journal Quarterly*, 26(3), 525-555.
- Yi, D. S. (1987). "Grey Model and Forecast of Talented Persons," *System Engineering*, 1, 36-43.

