Volume 7, Number 1 May 2012

AN ANALYSIS OF RETAIL MORTGAGE PAYMENT BEHAVIOR BASED ON HIDDEN MARKOV MODELS

Chang Liu, Southern West University of Finance and Economics, China

Morsheda Hassan, Wiley College
mthassan@wileyc.edu

Raja Nassar, Louisiana Tech University
Nassar@coes.latech.edu

ABSTRACT

This paper is concerned with stochastic modeling of retail mortgage loans using hidden Markov chains (HMC). HMC are very useful in determining the unobservable variables affecting retail mortgage loans by analyzing the observable state transition behaviors of the loans. A fourth-order HMC model is presented based on the assumption that the past several periods of payment behavior have an effect on current behavior. Also, an Interactive Hidden Markov chain model is presented in order to capture the interaction between the observable states, loan transition behavior, and the unobservable underlying local macro-economic factors.

INTRODUCTION

Not all important variables affecting a process under consideration are directly observable. In such a case, a model that considers only the observable variables is often not efficient and may lead to wrong conclusions. A case in point is the modeling of credit risk.

In the modeling of credit risk, the payment behavior may be influenced by economic factors. Furthermore, in areas such as small local communities, the influence may also run in the opposite direction. This can occur if the payment behavior affects the local economy by changing the local credit environment, and thus employment, interest rate, and production.

In this paper, the effects of hidden or unobservable macroeconomic factors on the mortgage payment behavior are studied by Hidden Markov Chains. Also, a higher-order HMC (HHMC) is used to account for multiple transition lags. Furthermore, an Interactive HHMC model is used to analyze the two-way effect between the mortgage payment behavior and the local macroeconomy by using data from a local community bank.

LITERATURE REVIEW

The basic property of a Markov chain has been extended to accommodate many new applications among them are network traffic analysis, speech recognition, DNA sequence analysis, engineering designs, and inventory management. Also, new theories extending the basic Markov assumption have been developed in the past 50 years, such as High-order Markov chains, Multivariate Markov chains, and Hidden Markov chains.

Hidden Markov Chains

Although higher-order Markov chain models may provide more accurate results than first-order Markov chains, they fail to take into consideration underlying forces that may influence observed transition processes in real-world problems. Examples include speech recognition, stock market analysis, and network traffic analysis. All these problems could be solved by Hidden Markov chain models. A standard HMC model has the following elements:

(1) Hidden states, $H = \{H_1, H_2, ..., H_N\}$
(2) Observable states, $S = \{S_1, S_2, ..., S_l\}$
(3) Transition probability distribution within hidden states,
$A = \{a_{ij}\}, a_{ij} = P(H_{j,t=n} \mid H_{i,t=n-1}), 1 \le i, j \le N$
(4) Emission probabilities matrix, $B = \{b_{jk}\}$ where $b_{jk} = P(S_k \mid H_j), 1 \le j \le N, 1 \le k \le l$
(5) Initial state distribution, $\Pi = \{\pi_i\}, \pi_i = P(S_i), 1 \le i \le N.$

Thus, an HMC is completely specified by: $\Lambda = (A, B, \Pi)$. As pointed out by MacDonald and Zucchini (1997), HMC could be used to answer the following three classic problems:

(1) Given an observed sequence $S = \{S_1, S_2, ..., S_l\}$ and a model $\Lambda = (A, B, \Pi)$, how does one efficiently compute $B = \{b_{jk}\}$,
(2) Given an observed sequence $S = \{S_1, S_2, ..., S_l\}$ and a model $\Lambda = (A, B, \Pi)$, how do we choose the corresponding state sequence $A = \{a_{ij}\}$ which best explains the observations,
(3) How does one adjust the model parameters $\Lambda = (A, B, \Pi)$ to maximize $P(S / \lambda)$.

Thomas, Allen, and Kingsbury (1998) used a hidden Markov chain model for the term structure and credit risk spreads of bond prices. Their model had two dependent factors, interest rate and the credit rating of the bonds which were affected by the underlying macroeconomic factors assumed to be the unobserved hidden processes. A linear programming approach was used to solve the model and the coupon strip, assuming that there was no miss-pricing opportunity.

Also, algorithms to estimate the parameters of a hidden Markov chain have been studied by Archera and Titterington (2002). They proposed some alternatives to maximum likelihood estimates (MLE), such as EM algorithm, Zhang's mean-field approximation to the EM algorithm, and Monte Carlo simulations. Parameter estimates has been studied by Arribas-Gil, Gassiat, and Matias (2006) in pair hidden Markov models by simulation, Hobolth and Jensen (2005) for DNA

Sequences, and Knudsen and Miyamoto (2003) for Human alpha and beta-hemoglobin sequences. Many algorithms are used to efficiently solve these problems, including a forward algorithm, a backward algorithm, an EM algorithm, and a heuristic linear programming method for a higher-order HMC proposed by Ching and Ng (2006). This linear programming method is as follows:

$$(1) \qquad Min_{\lambda_i} \left\{ \left\| \sum_{j=1}^{k} \lambda_j V_j \hat{H}_i - \hat{H}_i \right\|_l \right\}, i = 1, 2,$$

$$\text{Subject to} \sum_{i=1}^{k} \lambda_i = 1, \lambda_i \geq 0$$

Where the λ_i 's are the parameters, $\hat{H}_i$ is the estimated stationary probability distribution, and V_j is the higher-order transition matrix defined as A at the beginning of this subsection.

Comparisons between the EM algorithm and the linear programming method (LPM) for different orders, provided by Ching and Ng (2006), are also presented in Tables 1 and 2.

TABLE 1. COMPARISON BETWEEN THE EM ALGORITHM AND LPM BY NUMBER OF ITERATIONS

	First-Order	Second-Order	Third-Order
Linear Programming	1381	1378	1381
EM Algorithm	1377	1375	1377

TABLE 2. COMPARISONS BETWEEN THE EM ALGORITHM AND LPM BY COMPUTATION TIME IN SECONDS

	First-Order	Second-Order	Third-Order
Linear Programming	1.16	1.98	5.05
EM Algorithm	4.02	12.88	40.15

It is seen from these tables that although there is not much difference between linear programming and the EM algorithm with regard to the number of iterations, the linear programming method is better than the EM algorithm regarding computation time, especially for a higher order. Similar approaches have also been used by Ching and Ng (2004) for parameter estimates, Ching, Ng and Fund (2003) for DNA sequence, and Ching, Ng Fund, and Siu T (2005) for a categorical data sequence.

MODELS

The ultimate purpose of the HMC is to better understand and predict the transition probabilities between the observable states by analyzing the underlying forces that have influence on the observable behavior. Generally speaking, what people are really interested in are the observable states. However, to better simulate or estimate the true pattern of the state transition under different prevailing underlying situations, underlying forces must be taken into account in the model. Empirically speaking, as more information is built into the model, more accurate results could be expected, which is the general idea of the higher-order HMC. From the linear programming scheme proposed by Raftery (1985), which was extended by Ching and Ng (2006) by allowing for non-stationary transition intensity ($Q_i, i = 1,2,...,T$) overtime, one can avoid the problem of having to estimate too many parameters in a higher-order Markov model. In addition, the higher-order model could be further improved by assuming that the observable states could also have influences on the unobservable or hidden states. As a result, an HMC will allow for the interaction between these two types of states and might produce even more accurate prediction results.

In the Markov chain model, let S_j be a state of past due corresponding to the days of past due. The loan normally requires monthly payment. Based on monthly payments, Table 3 defines the different states of the Markov chain representing mortgage loan payments. According to the Basel accord II, Basel Committee on Banking Supervision (1997), the definition of default is more than 90 days past due, which is represented by S_3. However, there have been cases where the obligations on a loan, which have already been more than 90 days past due, has been paid off. As a result, the definition of default is modified to be the state of default that is triggered by a permanent force, such as death or an application of chapter 7 or chapter 13 bankruptcy protections.

Referring to Table 3, Let R_k be the default state and let S_{-j} be the state of a prepaid period defined as $S_{-j} = (X_i - Y_i)/Y_i$, where X_i is the actual payment at month i and Y_i is the scheduled payment at month i. One can see that state S_{-j} is defined as the extra payment over the scheduled payment, which measures how many future monthly payments have been made.

For example, if a loan monthly payment (or scheduled amount) is \$1500 and the payment for that month is \$6000, then the loan is three months prepaid (6000-1500)/1500 = 3) and the state is S_{-2}. Likewise, a past due state is obtained as (due amount – scheduled amount) / scheduled amount. Furthermore, if a loan’s prepayment is less than 50% of the remaining balance, the loan stays in the S states and is classified according to the rules above. If a loan is prepaid more than 50% and less than 75% of its remaining balance, then it belongs to R3. When the prepayment is 75% or more, the loan is in state R4.

TABLE 3. DEFINITIONS OF THE DIFFERENT STATE OF THE MARKOV CHAIN WITH REGARD TO PAYMENTS ON A MORTGAGE LOAN

Past Due and Prepayment States $S_j, j=-3,-2,-1,0,1,2,3$		**Default States** R_k $R_k, k=1,2,3,4$	
S_{-3}	Prepaid More than 91 days	R_1	Sold by Bank
S_{-2}	Prepaid 61 days – 90 days	R_2	All others
S_{-1}	Prepaid 31 days – 60 days	R_3	Prepayment is more than 50% and less than 75% of the remaining balance
S_0	No more than 30 days past due	R_4	Prepayment is 75% or more of the remaining balance
S_1	31 days – 60 days past due		
S_2	61 days – 90 days past due		
S_3	More than 91 days past due		

Hidden Markov chain model

In most cases, an observable phenomenon is veiled by invisible forces. In this case, these hidden forces are crucial to understanding the perceivable pattern. In this subsection, a simple Hidden Markov Model is introduced to track and predict the transition probabilities of payment states in retail mortgage loans by taking local macroeconomic situations into consideration. The macro-economic environment is the main factor influencing business development. It is desirable to have a measurement which could track hidden macroeconomic transition processes that have a close relationship with the financial industry. One good candidate is the state space model concerning the business industry industrial production index by Liu et al. (2007). The model is given as:

$$y_t = 0.4096y_{t-2} + 0.0835Ir_{t-2} - 0.6258Un_{t-2} - 0.0619In_{t-2} - 0.0236Dp_{t-2} - 0.987529Ir_{t-1} + 0.26377In_{t-1} + 0.002143Dp_{t-1} \quad (2)$$

where, y_t is the industrial production index at time t, Ir_t is interest rate, Un_t is unemployment, In_t is inflation, and Dp_t is disposable personal income at time lags. We define an economic environment to be positive if the industrial production index is at least 100 at that period and negative otherwise. Thus, we have 2 hidden states. From time to time, the hidden state transits from good to bad or from bad to good. Without loss of generality, we assume that the probability of the industrial production index being positive is α, and the probability of it being negative is $1-\alpha$. Also, we follow the definition of observable retail mortgage states in Table 3. By the definition of hidden states, we can observe the steady state probability distribution (under different hidden states), $O^t_{i,s}, i=1,2,\ \ S\in(S_j,R_k), j=-3,-2,-1,0,1,2,3;$ k = 1,2,3,4, which are defined as:

$$(3)\quad O^t_{i,s}=\begin{cases} O^t_{1,S}\text{, if observed under a positive economic environment at time } t \\ \\ O^t_{2,S}\text{, if observed under a negative economic environment at time } t \end{cases}$$

A new method for estimating the parameter α has been introduced by Ching and Ng (2006). Following their method, one needs to define a probability distribution at steady state. Unfortunately, in a dynamic economic environment, a steady state does not exist. One way we can bypass this dilemma is as follows: Let X_S be the S^{th} element of the steady state probability distribution vector X, $S\in(S_j,R_k), j=-3,-2,-1,0,1,2,3, k=1,2,3,4$ in terms of an average value, X_S can be expressed as:

$$(4)\qquad X_S=\frac{\sum_{i=1}^{2}\sum_{t=1}^{n}O^t_{i,S}}{n}, S\in(S_j,R_k), j=-3,-2,-1,0,1,2,3, k=1,2,3,4, n=16$$

Thus, the steady state probability distribution is approximated by averaging all the observed distributions over the intersections, where n is the number of switches (or transitions from positive economy to negative economy) in the available time series data. Thus, to estimate α in the hidden Markov chain, we use Equation (5) as suggested by Ching and Ng (2006). Equation (5) minimizes the sum of squared deviations between $\hat{P}_S$ and X_S.

$$(5)\qquad \begin{aligned} &Min_{0\le\alpha\le1}\{\psi\}=\{\left\|\hat{P}_S-X_S\right\|_2\} \\ &S\in(S_i,R_k), i=-3,-2,-1,0,1,2,3, k=1,2,3,4 \end{aligned}$$

$\hat{P}_S$ is given by the following matrix manipulation. Let P be so defined such that

$$(6)\qquad P=\begin{pmatrix} 0 & H_{2\times11} \\ P'_{11\times2} & 0 \end{pmatrix}_{13\times13}$$

where, $H_{2\times11}=\begin{pmatrix} \alpha & \dots & \alpha \\ 1-\alpha & \dots & 1-\alpha \end{pmatrix}_{2\times11}$, and $P'_{11\times2}=\begin{pmatrix} O^T_{S|1} & O^T_{S|2} \end{pmatrix}_{11\times2}$.

$O^T_{S|i}, i=1,2,\ \ S\in(S_i,R_k), i=-3,-2,-1,0,1,2,3, k=1,2,3,4$. Thus,

$$(7)\quad P^2=\begin{pmatrix} 0 & H_{2\times11} \\ P'_{11\times2} & 0 \end{pmatrix}\times\begin{pmatrix} 0 & H_{2\times11} \\ P'_{11\times2} & 0 \end{pmatrix}=\begin{pmatrix} H_{2\times11}\times P'_{11\times2} & 0 \\ 0 & P'_{11\times2}\times H_{2\times11} \end{pmatrix}_{13\times13}$$

Therefore, $\hat{P}_S$, the probability distribution taking hidden states into consideration with α known, is defined as:

$$(8) \quad \hat{P}_S = P'_{11\times2} \times H_{2\times11} \times 1_{11\times1},$$

$$\text{where } 1_{1x11} = (1,1,...,1)^T$$

Based on the assumption that $\hat{P}_S$ is a stationary probability distribution, we can build a Markov prediction model to approximate the probability distribution in the next period under the consideration of a hidden process. The model is given as:

$$(9) \quad \begin{cases} Min_\lambda \{\psi\} = \{\left\|\lambda V_S \hat{P}^t_S - \hat{P}^{t+1}_S\right\|_l\}, l = 1,2,\infty \\ \text{subject to } \lambda > 0 \\ S \in (S_i, R_k), i = -3,-2,-1,0,1,2,3, k = 1,2,3,4 \end{cases}$$

where V_S is the transition intensities. Once we find the parameter λ, we can use the probability distribution observed at time $t-1$ to predict that at time t. A higher-order Markov prediction model for hidden processes will be presented in the next subsection.

Heuristic Method for the Higher-Order HMC (HHMC)

Given observed states, a higher-order HMC can be used to address the following three problems: (1) prediction of the probability distribution of observed states $P(O|\Lambda), \Lambda = (A,B,\Pi)$, (2) determining the optimal hidden states that best explain the observed behaviors, and (3) estimating the model parameters, $\Lambda = (A,B,\Pi)$. In the real economic world, we seldom have the capability to choose underlying factors affecting the observable behavior of a process. Thus, problem (2) is irrelevant to our case. To solve problems (1) and (3) by conventional methods require tedious recursive algorithms such as the forward algorithm for problem (1), and the EM algorithm for problem (3). Detailed discussion of the forward and EM algorithms could be found in MacDonald and Zucchini (1997).

In this subsection, we will present a Heuristic method proposed by Ching and Ng (2006) for a fourth-order HMC based on the assumption that the emission probabilities matrix, $B = \{b_{S|j}\}$, where $b_{jk} = P(S_k | H_j), 1 \le j \le N, 1 \le k \le i$ could be observed, which is generally the case. Let $\{\hat{h}_i\} \in \hat{H}, i = 1,2$ be the stationary probability distribution for the hidden states, and $\{\hat{v}_{i,t}\} \in \hat{V}_t, t = 1,2,3,4, i = 1,2$ be the transition intensities between the hidden states with different time lags. An equation for estimating λ_i in a fourth-order hidden Markov model is given as:

$$(10) \qquad Min_{\lambda_i}\{\left\|\sum_{j=1}^{k}\lambda_j\hat{V}_k\hat{H}-\hat{H}_i\right\|_l\}, i=1,2, k=1,2,3,4,$$

$$\text{subject to} \sum_{i=1}^{k}\lambda_i = 1, \lambda_i \geq 0$$

For practical reasons, we choose $l = 1$ in the vector norm $\| \, \|_l$. Thus, the more applicable version of Equation (10) that could be solved by the Excel ***Solve()*** function is:

$$(11) \qquad Min_{\lambda}\sum_{l=1}^{4} w_l, \text{subject to}$$

$$\begin{cases} \begin{pmatrix} w_1 \\ w_2 \\ \cdots \\ w_i \end{pmatrix} \geq H - [V_1H \mid V_2H \ldots \mid V_kH] \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \cdots \\ \lambda_k \end{pmatrix} \\ \begin{pmatrix} w_1 \\ w_2 \\ \cdots \\ w_i \end{pmatrix} \leq H + [V_1H \mid V_2H \ldots \mid V_kH] \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \cdots \\ \lambda_k \end{pmatrix} \end{cases}$$

Here, $\hat{H}_i$, the hidden stationary probability distribution, must be approximated since it cannot be observed directly. Ching and Ng (2006) proposed a method to calculate $\hat{H}_i$ from the observed probability distribution, $O_{i,S}$:

$$(12) \qquad \begin{matrix} \left\|\hat{O}_{i,S} - B\hat{H}_i\right\|_l, l = 1,2,\infty, i = 1,2 \\ S \in (S_j, R_k), j = -3,-2,-1,0,1,2,3, k = 1,2,3,4 \end{matrix},$$

where B is the emission probability matrix, $B = \{b_{S|i}\}$, $b_{S|i} = P(S_S \mid H_i), i = 1,2$, and $\hat{O}_S$ is the observed probability distribution. For the accuracy of the model, we choose $l = 2$ and Equation (12) given in matrix form becomes:

$$(13) \qquad \begin{matrix} Min\left\|\{o_{i,S}\}_{11\times1} - \{b_{S|i}\}_{11\times2}\{h_i\}_{2\times1}\right\|_2, i = 1,2 \\ S \in (S_j, R_k), j = -3,-2,-1,0,1,2,3, k = 1,2,3,4 \end{matrix}$$

Also, in need of estimation are the transition intensities among the hidden states, $\{\hat{v}_{i,t}\} \in \hat{V}_t, t = 1,2,3,4, i = 1,2$. As pointed by Ching and Ng (2006), $\hat{H}_i$, the hidden stationary probability distribution estimated by Equation (13) could be used to estimate the first-order transition intensity matrix for hidden states:

$$\hat{H}_{2\times 1} = \begin{pmatrix} \hat{h}_1 \\ \hat{h}_2 \end{pmatrix} \Rightarrow \hat{V}_1 = \begin{pmatrix} \hat{h}_1 & \hat{h}_2 \\ \hat{h}_2 & \hat{h}_1 \end{pmatrix}$$

Thus, as the transition intensity matrix is assumed to be stationary, the second, third, and fourth order could be estimated by the following procedure:

$$\begin{aligned} &\hat{V}_2 = \hat{V}_1 \times \hat{V}_1, \\ (14) \qquad &\hat{V}_3 = \hat{V}_1 \times \hat{V}_1 \times \hat{V}_1, \\ &\hat{V}_4 = \hat{V}_1 \times \hat{V}_1 \times \hat{V}_1 \times \hat{V}_1, \end{aligned}$$

As such, the above estimation provides us a stable method to approximate different orders of transition intensities.

The following is a summary of the above steps for a higher-order HMC:

Step 1: Use Equation (12) to find the stationary probability distribution for the hidden states, where $b_{S|i}$ is the emission transition from hidden states to observed states given by $b_{S|i} = P(S_{i,k} \mid H_i), i = 1, 2$;

Step 2: find the transition intensities for various orders by Equation (14); and

Step 3: Use Equation (13) to estimate model parameters $\lambda_i, i = 1, 2, 3, 4$ for a fourth-order HMC.

An Interactive Higher-Order Hidden Markov Model (IHHMC)

The interactive HMC is different from the regular HMC in the sense that hidden states of an interactive HMC are affected by previous hidden states and by observable states. In case of retail mortgage analysis, not only local macro-economic factors can affect the mortgage payments, but the payment behavior also determine the collection policy deployed by the banks such as high mortgage rate to cover the foreseeable credit risks of the unusual payment patterns, which, in turn, affect the local businesses in many ways. Therefore, an interactive higher-order HMC seems to be a good candidate for capturing the mechanism in this system. Let $O_{S,i}$ be the observed probability distributions under different hidden states such that:

$$(15) \quad O_{S,i} = \begin{cases} O_{S,1}\text{, if observed under a positive economic environment} \\ O_{S,2}\text{, if observed under a negative economic environment} \end{cases}$$

We define $\alpha_{S,} S \in (S_i, R_k), i = -3, -2, -1, 0, 1, 2, 3, k = 1, 2, 3, 4$ to be the probability of the hidden state being positive, given the observable states in S. Thus, the transition matrix is given as:

(16) $$P = \left(\begin{array}{c|c} 0 & O_{2\times 11} \\ \hline A_{11\times 2} & 0 \end{array} \right)_{13\times 13}, A_{11\times 2} = \{a_{i,s}\}$$

where, $A = \begin{pmatrix} \alpha_1 & \cdots & \alpha_{11} \\ 1-\alpha_1 & \cdots & 1-\alpha_{11} \end{pmatrix}^T, O = \begin{pmatrix} o_{1,1} & \cdots & o_{1,11} \\ o_{2,1} & \cdots & o_{2,11} \end{pmatrix}$ $S \in (S_i, R_k), i = -3,-2,-1,0,1,$ $2,3, k = 1,2,3,4$. Thus,

$$P^2 = \begin{pmatrix} 0 & O_{2\times 11} \\ A_{11\times 2} & 0 \end{pmatrix} \times \begin{pmatrix} 0 & O_{2\times 11} \\ A_{11\times 2} & 0 \end{pmatrix}$$

(17) $$= \begin{pmatrix} O_{2\times 11} \times A_{11\times 2} & 0 \\ 0 & \hat{P}_S \end{pmatrix}_{13\times 13}$$

$$= \begin{pmatrix} O_{2\times 11} \times A_{11\times 2} & 0 \\ 0 & A_{11\times 2} \times O_{2\times 11} \end{pmatrix}_{13\times 13}$$

where, $\hat{P}_S$ the probability distribution under hidden states, is defined as

(18) $$\hat{P}_S = A_{11\times 2} \times O_{2\times 11} \times 1_{11\times 1},$$

where $1_{11\times 1} = (1,1,...,1)^T$.

To estimate the parameters α_S, we need the steady state one-step transition probability matrix which could be approximated by $\tilde{P}_{11\times 11} = \{\tilde{p}_S\}_{11\times 11}, S \in (S_i, R_k)$ $, i = -3,-2,-1,$ $0,1,2,3, k = 1,2,3,4$. Letting $c_{ik}, i = -3,-2,-1,0,1,2,4, k = 1,2,3,4$ be the transition frequency between state i and state k, the calculation of $\tilde{p}_S$ is given as:

(19) $$C_{ik} = \begin{pmatrix} c_{-3,-3} & \cdots & c_{-3,4} \\ \vdots & \ddots & \vdots \\ c_{4,-3} & \cdots & c_{4,-3} \end{pmatrix}_{11\times 11} \quad \tilde{P}_S = \begin{pmatrix} \tilde{p}_{-3,-3} & \cdots & \tilde{p}_{-3,4} \\ \vdots & \ddots & \vdots \\ \tilde{p}_{4,-3} & \cdots & \tilde{p}_{4,-3} \end{pmatrix}_{11\times 11}$$

$$\tilde{p}_{i,k} = \begin{cases} \dfrac{c_{i,k}}{\sum_{i=S_{-3}}^{R_4} c_{i,k}}, & \text{if } \sum_{i=S_{-3}}^{R_4} c_{i,k} \neq 0 \\ 0, & \text{Otherwise} \end{cases}$$

We define the Frobenius norm as $\|A_{n\times n}\|_F^2 = \sum_{j=1}^{n}\sum_{i=1}^{n} A_{ij}^2$. Thus, the parameters α_S could be approximated by minimizing the Frobenius norm given as:

(20) $$Min_{\alpha_i} \left\| \hat{P}_S - \tilde{P}_S \right\|_F^2$$

Therefore, the above minimizing algorithm could also be expressed as:

(21)
$$\begin{aligned}
&(1)\alpha_1 : Min_{0\le\alpha_1\le1}\{(\tilde{p}_{-3,-3}-\hat{p}_{-3,-3})^2+...+(\tilde{p}_{-3,4}-\hat{p}_{-3,4})^2\};\\
&(2)\alpha_2 : Min_{0\le\alpha_2\le1}\{(\tilde{p}_{-2,-3}-\hat{p}_{-2,-3})^2+...+(\tilde{p}_{-2,4}-\hat{p}_{-2,4})^2\};\\
&\vdots \qquad\qquad\qquad \vdots \qquad\qquad\qquad \vdots\\
&(11)\alpha_{11} : Min_{0\le\alpha_{11}\le1}\{(\tilde{p}_{4,-3}-\hat{p}_{4,-3})^2+...+(\tilde{p}_{4,4}-\hat{p}_{4,4})^2\};
\end{aligned}$$

The equation to estimate λ_i in a fourth order hidden Markov model is given as

(22)
$$Min_{\lambda_i}\{\left\|\sum_{j=1}^{k}\lambda_j V_j \hat{P}_S - \hat{P}_S\right\|_l\}, i=1,2,$$
$$\text{subject to} \sum_{i=1}^{k}\lambda_i = 1, \lambda_i \ge 0,$$

where $\hat{P}_S$, the hidden stationary probability distribution, is given by Equation (18). Finally, the transition intensities among hidden states could be estimated by exactly the same idea of Equation (14). The only difference is the fact that the transition intensities are 11×11 matrices to capture the effects between observed processes and hidden processes. Thus, from Ching and Ng (2006), the higher-order interactive transition intensities can be calculated as follows:

Let $\{\hat{p}_S\}\in\hat{P}_S,\ S\in(S_i,R_k), i=-3,\ -2,-1,0,1,2,3, k=1,2,3,4$:

(23)
$$\hat{V}_1 = \begin{pmatrix} \hat{p}_{-3} & \hat{p}_{-2} & \cdots & \hat{p}_4 \\ \hat{p}_{-2} & \hat{p}_{-3} & \hat{p}_4 & \vdots \\ \vdots & \hat{p}_4 & \ddots & \hat{p}_{-2} \\ \hat{p}_4 & \cdots & \hat{p}_{-2} & \hat{p}_{-3} \end{pmatrix}_{11\times11}$$
$$\hat{V}_2 = \hat{V}_1\times\hat{V}_1,$$
$$\hat{V}_3 = \hat{V}_1\times\hat{V}_1\times\hat{V}_1,$$
$$\hat{V}_4 = \hat{V}_1\times\hat{V}_1\times\hat{V}_1\times\hat{V}_1,$$

The whole algorithm for an Interactive Higher-Order HMC is as follows:

Step 1: Use Equation (21) to find the stationary probability distribution for hidden states, where $b_{S|i}$ is the emission transition from hidden states to observed states given by:
$$b_{S|i} = P(S_{i,k}\mid H_i), i=1,2, S\in(S_i,R_k), i=-3,-2,-1,0,1,2,3, k=1,2,3,4$$

Step 2: determine the transition intensities by Equation (23); and

Step 3: Use Equation (22) to estimate model parameters $\lambda_i, i=1,2,3,4$ for a fourth-order HMC.

APPLICATIONS

A bank, providing the retail mortgage services, never operates in a vacuum because the transitions of its mortgage payment behavior and its credit asset quality are affected by many macroeconomic factors. In general, the transition pattern of the mortgage payment behavior varies under different macro-economic environments which, in turn, are presented by a group of indices or factors. HMCs, however, could provide a way to unveil more accurate transition processes and therefore provides a probability distribution for mortgage payment states closer to the real prevailing macro-economic situation.

In this section, 18 consecutive months of monthly paid retail mortgage data, provided by an Ohio local bank, will be analyzed by the hidden Markov model. This includes the basic first-order HMC given in Equation (9), a higher-order HMC given in Equation (11), and finally, an interactive HMC in Equation (21).

The computations performed in this paper are accurate and are not affected by round-off error. The Solver procedure in Excel can do calculations up to 8 digits in accuracy. Also, Matlab accuracy is up to 26 digits in Floating point computation.

HMC for Unobservable Factors in Retail Mortgages

In this section, a basic first-order HMC is used to analyze and predict the probability distribution among states considering the effects of underlying macro-economic factors. Due to the lack of an industrial production index in the local Ohio area where the bank data were obtained, we estimated the index from Equation (2) by using macro-economic data for Ohio from February 2005 to September 2006. The macro-economic data for Ohio from Feb 2005 to Sep 2006 are presented in Table 4.

TABLE 4. MACRO-ECONOMIC DATA AND INDEX FOR OHIO

Year Month	Un	Ir	In	Dp	Index
2005 02	5.78	5.93	3.52	5.23	11.53
2005 03	5.80	5.87	4.20	5.08	12.20
.					
.					
.					
2006 07	5.80	6.76	5.47	2.50	16.64
2006 08	5.40	6.52	2.99	2.10	13.94
2006 09	5.00	6.40	5.74	2.10	15.10

In this Table, Ir is interest rate, Un is unemployment, In is inflation, Dp is disposable personal income at different times. For the purpose of this analysis, we refer to the industrial production index from the model in Equation (2) as the macro-economic situation in Ohio. The hidden Markov index sequence is presented in Figure 1.

The average index from Table 4 is 14.023. If we let a year takes a value of 1 or 0 depending on whether the index for that year is larger or smaller than 14.023, respectively, we obtain the hidden transition sequence in Table 5.

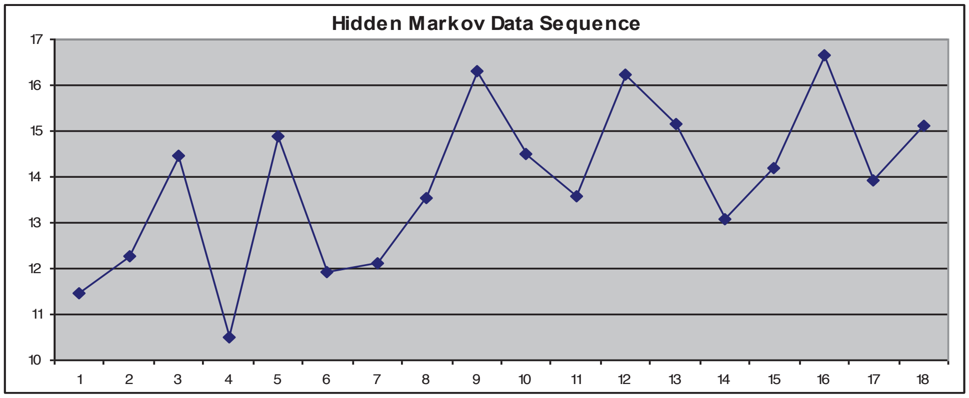

FIGURE 1. HIDDEN MARKOV DATA SEQUENCE

Table 5. Hidden transition sequence:

$$t: 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9,\ 10,\ 11,\ 12,\ 13,\ 14,\ 15,\ 16$$

$$H_t: 0,\ 0,\ 1,\ 0,\ 1,\ 0,\ 0,\ 0,\ 1,\ 1,\ 0,\ 1,\ 1,\ 0,\ 1,\ 1,\ 0$$

From the data sequence in Table 5, one can estimate the emission probability matrix, $B = \{b_{S|j}\}$. We define the steady state probability distribution for the positive hidden states (1's in Table 5) as: $O_{S|1} = \frac{\sum_t o_S}{8}, t = 3,5,9,10,12,13,15,16$. Similarly, the steady state probability distribution for negative (0's) hidden states as: $O_{S|2} = \frac{\sum_t o_S}{9},\ t = 1,2,4,\ 6,7,8,11,14,17$. From MathCAD, we obtained the probability distribution as shown in Table 5.

TABLE 5. STEADY STATE PROBABILITY DISTRIBUTIONS

$$O_{S|1} = (0.0052\quad 0.0094\quad 0.0578\quad 0.9452\quad 0.0547\quad 0.0412\quad 0.0224\quad 0.0001\quad 0.0378\quad 0.0028\quad 0.0014)^T$$

$$O_{S|2} = (0.0021\quad 0.0023\quad 0.0098\quad 0.7380\quad 0.0531\quad 0.1078\quad 0.0009\quad 0.0300\quad 0.0424\quad 0.0015\quad 0.0021)^T$$

$$X_S = (0.0038\quad 0.0087\quad 0.0187\quad 0.8012\quad 0.0947\quad 0.0094\quad 0.0145\quad 0.0300\quad 0.0147\quad 0.0024\quad 0.0019)^T$$

where X_S is given by Equation (4). We let $H_{2\times 11} = \begin{pmatrix} \alpha & \ldots & \alpha \\ 1-\alpha & \ldots & 1-\alpha \end{pmatrix}_{2\times 11}$ and $P'_{11\times 2} = \begin{pmatrix} O^T_{S|1} & O^T_{S|2} \end{pmatrix}_{11\times 2}, O^T_{S,i}, i = 1,2$. Here, α is the probability of the hidden state being positive and 1- α the probability of being negative. Thus, the parameter α, could be calculated by Equation (5) or the following algorithm by letting $l = 2$:

$$(24)\qquad \begin{cases} Min_{\alpha}\{\sum_{S}(\hat{P}_S - X_S)^2\} \\ subject \text{ to } 0 \le \alpha \le 1 \end{cases}$$

where $\hat{P}_S$ is given by: $\hat{P}_S = \begin{pmatrix} 0.0073\alpha + 0.0021 & 0.0117\alpha + 0.0023 & \ldots & 0.0035\alpha + 0.0021 \end{pmatrix}_{1\times 11}^T$

By the Excel ***Solver()*** function, we estimate α to be 0.9143, which means that 91.43% of the time between Apr 2005 to Sep 2006 the macro-economic environment would stay in a positive state. As a result, the estimated probability distribution affected by the hidden macro-economic factors is given as:

$$\hat{P}_S = \begin{pmatrix} 0.0045 & 0.0083 & 0.0520 & 0.8010 & 0.0455 & 0.0284 & 0.0204 & 0.00248 & 0.0031 & 0.0021 & 0.0011 \end{pmatrix}_{1\times 11}^T$$

In the next section, we will apply a Higher-order HMC to the retail mortgage data.

A Higher-Order HMC

In this section, we will use a higher-order HMC to track and predict the hidden transition process. We approximate the steady state hidden probability distribution by a modified version of Equation (13) that could be solved directly by the Excel ***Solver()***. This gives:

$$(25)\qquad \begin{cases} Min_{h_i}\{\sum_{S}(\{o_S\}_{11\times 1} - \{b_{S|i}\}_{11\times 2}\{h_i\}_{2\times 1})^2\}, i = 1,2 \\ subject \text{ to } 0 \le h_i \le 1, \sum_{i=1}^{2} h_i = 1 \\ S \in (S_i, R_k), i = -3,-2,-1,0,1,2,3, k = 1,2,3,4 \end{cases}$$

where $\{b_{S|i}\} \in B_{S|i}$, the emission probabilities, represent the probability distribution vectors under hidden states 1 and 2, respectively. These are given in Table 6.

TABLE 6. INPUT VARIABLES FOR EQUATION (13)

$$b_{S|1} = \begin{pmatrix} 0.0052 & 0.0094 & 0.0578 & 0.9452 & 0.0547 & 0.0412 & 0.0224 & 0.0001 & 0.0378 & 0.0028 & 0.0014 \end{pmatrix}^T$$
$$b_{S|2} = \begin{pmatrix} 0.0021 & 0.0023 & 0.0098 & 0.7380 & 0.0531 & 0.1078 & 0.0009 & 0.0300 & 0.0424 & 0.0015 & 0.0021 \end{pmatrix}^T$$
$$O_S = \begin{pmatrix} 0.0038 & 0.0087 & 0.0187 & 0.8012 & 0.0947 & 0.0094 & 0.0145 & 0.0300 & 0.0147 & 0.0024 & 0.0019 \end{pmatrix}^T$$

From the solution to Equation (25) we have $\hat{H} = \{\hat{h}_1, \hat{h}_2\} = \{0.4033, 0.5967\}$

In the next step, we will approximate the transition intensities for different orders by Equation (23). Note that the first-order transition intensity matrix is given by:

$\hat{H}_{2\times 1} = \begin{pmatrix} \hat{h}_1 \\ \hat{h}_2 \end{pmatrix} \Rightarrow \hat{V}_1 = \begin{pmatrix} \hat{h}_1 & \hat{h}_2 \\ \hat{h}_2 & \hat{h}_1 \end{pmatrix}$. Thus, the transition intensities for four orders are estimated from MathCAD to give:

$$\text{(26)} \quad \hat{V}_1 = \begin{pmatrix} 0.4033 & 0.5967 \\ 0.5967 & 0.4033 \end{pmatrix} \quad \hat{V}_2 = \begin{pmatrix} 0.5187 & 0.4813 \\ 0.4813 & 0.5187 \end{pmatrix}$$
$$\hat{V}_3 = \begin{pmatrix} 0.5007 & 0.4993 \\ 0.4993 & 0.5007 \end{pmatrix} \quad \hat{V}_4 = \begin{pmatrix} 0.4964 & 0.5036 \\ 0.5036 & 0.4964 \end{pmatrix}$$

The method to estimate the parameters $\lambda_i, i = 1,2,3,4$ for is given by Equation (11). The linear programming scheme is as follows:

$$\hat{H} = (0.4033 \quad 0.5967)^T$$
$$V_1\hat{H} = (0.5187 \quad 0.4813)^T$$
$$V_2\hat{H} = (0.4964 \quad 0.5036)^T$$
$$V_3\hat{H} = (0.5007 \quad 0.4993)^T$$
$$V_4\hat{H} = (0.4999 \quad 0.5001)^T$$

$$\begin{cases} Min_{\lambda_1,\lambda_2,\lambda_3,\lambda_4} (w_1 + w_2 + w_3 + w_4) \\ w_1 \geq 0.4033 - 0.5187\lambda_1 - 0.4964\lambda_2 - 0.5007\lambda_3 - 0.4999\lambda_4 \\ w_2 \geq 0.5967 - 0.4813\lambda_1 - 0.5036\lambda_2 - 0.4993\lambda_3 - 0.5001\lambda_4 \\ w_1 \geq -0.4033 + 0.5187\lambda_1 + 0.4964\lambda_2 + 0.5007\lambda_3 + 0.4999\lambda_4 \\ w_2 \geq -0.5967 + 0.4813\lambda_1 + 0.5036\lambda_2 + 0.4993\lambda_3 + 0.5001\lambda_4 \\ w_1, w_2, w_3, w_4 \geq 0, \\ \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1, \lambda_1, \lambda_2, \lambda_3, \lambda_4 \geq 0 \end{cases}$$

Subject to: (the constraints above)

Applying the above scheme to the Excel ***Solver(),*** the parameters for the higher-order Markov chain model are given by $\lambda = (0.1876 \quad 0.8125 \quad 0 \quad 0)$. As a result, the HHMC is given as

$$\text{(27)} \quad \hat{H}_{t+1} = 0.1876V_1\hat{H}_t + 0.8125V_2\hat{H}_{t-1}$$

Equation (27) implies that the probability distribution of the hidden states at $t = n+1$ are dependent on only those at $t = n$ and at $t = n-1$.

Interactive Effects Analysis for Retail Mortgages

The observable probability distributions, $O_{2\times 11}$, under both positive and negative states, are given as:

$$\text{(28)} \quad O_S = \begin{pmatrix} 0.0052 & 0.0094 & 0.0578 & 0.9452 & 0.0547 & 0.0412 & 0.0224 & 0.0001 & 0.0378 & 0.0028 & 0.0014 \\ 0.0021 & 0.0023 & 0.0098 & 0.7380 & 0.0531 & 0.1078 & 0.0009 & 0.0300 & 0.0424 & 0.0015 & 0.0021 \end{pmatrix}$$

Therefore, $\hat{P}_S$, the probability distribution under hidden states, is given by Equation (17) as $\hat{P}_S = A_{11\times 2} \times O_{2\times 11} 1_{11\times 1}, 1_{1\times 11} = (1,1,\ldots,1)^T$ where $A = \begin{pmatrix} \alpha_1 & \ldots & \alpha_{11} \\ 1-\alpha_1 & \ldots & 1-\alpha_{11} \end{pmatrix}^T$. Thus, $\hat{P}_S$ is given as

$$\begin{pmatrix}
1+0.0031\alpha_1 & 1+0.0071\alpha_1 & 1+0.048\alpha_1 & 1+0.2072\alpha_1 & 1+0.0016\alpha_1 & 1-0.0666\alpha_1 & 1+0.0215\alpha_1 & 1-0.0299\alpha_1 & 1-0.0046\alpha_1 & 1+0.0013\alpha_1 & 1-0.0007\alpha_1 \\
1+0.0031\alpha_2 & 1+0.0071\alpha_2 & 1+0.048\alpha_2 & 1+0.2072\alpha_2 & 1+0.0016\alpha_2 & 1-0.0666\alpha_2 & 1+0.0215\alpha_2 & 1-0.0299\alpha_2 & 1-0.0046\alpha_2 & 1+0.0013\alpha_2 & 1-0.0007\alpha_2 \\
1+0.0031\alpha_3 & 1+0.0071\alpha_3 & 1+0.048\alpha_3 & 1+0.2072\alpha_3 & 1+0.0016\alpha_3 & 1-0.0666\alpha_3 & 1+0.0215\alpha_3 & 1-0.0299\alpha_3 & 1-0.0046\alpha_3 & 1+0.0013\alpha_3 & 1-0.0007\alpha_3 \\
1+0.0031\alpha_4 & 1+0.0071\alpha_4 & 1+0.048\alpha_4 & 1+0.2072\alpha_4 & 1+0.0016\alpha_4 & 1-0.0666\alpha_4 & 1+0.0215\alpha_4 & 1-0.0299\alpha_4 & 1-0.0046\alpha_4 & 1+0.0013\alpha_4 & 1-0.0007\alpha_4 \\
1+0.0031\alpha_5 & 1+0.0071\alpha_5 & 1+0.048\alpha_5 & 1+0.2072\alpha_5 & 1+0.0016\alpha_5 & 1-0.0666\alpha_5 & 1+0.0215\alpha_5 & 1-0.0299\alpha_5 & 1-0.0046\alpha_5 & 1+0.0013\alpha_5 & 1-0.0007\alpha_5 \\
1+0.0031\alpha_6 & 1+0.0071\alpha_6 & 1+0.048\alpha_6 & 1+0.2072\alpha_6 & 1+0.0016\alpha_6 & 1-0.0666\alpha_6 & 1+0.0215\alpha_6 & 1-0.0299\alpha_6 & 1-0.0046\alpha_6 & 1+0.0013\alpha_6 & 1-0.0007\alpha_6 \\
1+0.0031\alpha_7 & 1+0.0071\alpha_7 & 1+0.048\alpha_7 & 1+0.2072\alpha_7 & 1+0.0016\alpha_7 & 1-0.0666\alpha_7 & 1+0.0215\alpha_7 & 1-0.0299\alpha_7 & 1-0.0046\alpha_7 & 1+0.0013\alpha_7 & 1-0.0007\alpha_7 \\
1+0.0031\alpha_8 & 1+0.0071\alpha_8 & 1+0.048\alpha_8 & 1+0.2072\alpha_8 & 1+0.0016\alpha_8 & 1-0.0666\alpha_8 & 1+0.0215\alpha_8 & 1-0.0299\alpha_8 & 1-0.0046\alpha_8 & 1+0.0013\alpha_8 & 1-0.0007\alpha_8 \\
1+0.0031\alpha_9 & 1+0.0071\alpha_9 & 1+0.048\alpha_9 & 1+0.2072\alpha_9 & 1+0.0016\alpha_9 & 1-0.0666\alpha_9 & 1+0.0215\alpha_9 & 1-0.0299\alpha_9 & 1-0.0046\alpha_9 & 1+0.0013\alpha_9 & 1-0.0007\alpha_9 \\
1+0.0031\alpha_{10} & 1+0.0071\alpha_{10} & 1+0.048\alpha_{10} & 1+0.2072\alpha_{10} & 1+0.0016\alpha_{10} & 1-0.0666\alpha_{10} & 1+0.0215\alpha_{10} & 1-0.0299\alpha_{10} & 1-0.0046\alpha_{10} & 1+0.0013\alpha_{10} & 1-0.0007\alpha_{10} \\
1+0.0031\alpha_{11} & 1+0.0071\alpha_{11} & 1+0.048\alpha_{11} & 1+0.2072\alpha_{11} & 1+0.0016\alpha_{11} & 1-0.0666\alpha_{11} & 1+0.0215\alpha_{11} & 1-0.0299\alpha_{11} & 1-0.0046\alpha_{11} & 1+0.0013\alpha_{11} & 1-0.0007\alpha_{11}
\end{pmatrix}$$

Also, the observed one-step transition intensity matrix, calculated from Equation (19) is $\tilde{P}_S$ =

	R_1	R_2	R_3	R_4	S_{-3}	S_{-2}	S_{-1}	S_0	S_1	S_2	S_3
R_1	1	0.	0	0	0	0	0	0	0	0	0
R_2	0	1	0	0	0	0	0	0	0	0	0
R_3	0	0	1	0	0	0	0	0	0	0	0
R_4	0	0	0	1	0	0	0	0	0	0	0
S_{-3}	0.0195	0	0	0	0.5759	0.0996	0.1646	0.1329	0	0	0
S_{-2}	0.0105	0	0	0	0.0365	0.6137	0.1563	0.1825	0.0005	0	0
S_{-1}	0.0101	0	0	0	0.02756	0.0341	0.6774	0.2399	0.0015	0	0
S_0	0.02901	0	0	0	0.0041	0.0097	0.0491	0.8920	0.0109	0	0
S_1	0.0133	0.1523	0	0.2090	0.0091	0.0089	0.0140	0.2078	0.2063	0.0552	0
S_2	0	0.0905	0.1456	0.1221	0	0.0031	0.0237	0.0347	0.0853	0.1747	0.3184
S_3	0	0.1053	0.1505	0.1305	0	0	0.0021	0.0632	0.0952	0.1510	0.2931

By the Frobenius norm defined in Equation (20), the 11 linear programming schemes are given as:

$$\begin{cases} Min_{\alpha_1}\{(1+0.0031\alpha_1-1)^2+(1+0.0071\alpha_1)^2+(1+0.2072\alpha_1)^2+\ldots+(1-0.0007\alpha_1)^2\} \\ subject\ to: 0 \le \alpha_1 \le 1 \end{cases}$$

$$\begin{cases} Min_{\alpha_2}\{(1+0.0031\alpha_2)^2+(1+0.0071\alpha_2-1)^2+(1+0.2072\alpha_2)^2+\ldots+(1-0.0007\alpha_2)^2\} \\ subject\ to: 0 \le \alpha_2 \le 1 \end{cases}$$

$$\vdots$$

$$\begin{cases} Min_{\alpha_{11}}\{(1+0.0031\alpha_{11})^2+(1+0.0071\alpha_{11}-0.1053)^2+(1+0.2072\alpha_{11}-0.1505)^2 \\ \qquad +\ldots+(1-0.0007\alpha_{11}-2931)^2\} \end{cases}$$

$$subject\ to: 0 \le \alpha_{11} \le 1$$

$$\text{Letting } \hat{P}_S = \begin{cases} 0, & if\ \hat{P} \le 0 \\ 1, & if\ \hat{P} \ge 1 \\ \hat{p}_S, & \text{otherwise} \end{cases}$$

the probability distribution under the hidden states is given as:

$$A = \{\alpha_S\} = \begin{pmatrix} 0.0001 & 0.0001 & 0.0001 & 0.0001 & 0.0047 & 0.0004 & 0.0001 & 0.0008 & 0.0002 & 0.0008 & 0.0001 \end{pmatrix}^T$$

$$\begin{array}{r} \\ R_1 \\ R_2 \\ R_3 \\ R_4 \\ S_{-3} \\ \hat{P}_S = S_{-2} \\ S_{-1} \\ S_0 \\ S_1 \\ S_2 \\ S_3 \end{array} \begin{array}{c} \begin{array}{ccccccccccc} R_1 & R_2 & R_3 & R_4 & S_{-3} & S_{-2} & S_{-1} & S_0 & S_1 & S_2 & S_3 \end{array} \\ \left[\begin{array}{ccccccccccc} 1 & 0. & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.8202 & 0.1899 & 0.0324 & 0.0034 & 0 & 0 & 0 \\ 0.0208 & 0 & 0 & 0 & 0.0717 & 0.8508 & 0.2881 & 0.0231 & 0.0011 & 0 & 0 \\ 0.0218 & 0 & 0 & 0 & 0.0544 & 0.0671 & 0.8959 & 0.0021 & 0.0029 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.0081 & 0.0195 & 0.0958 & 0.9883 & 0.0217 & 0 & 0 \\ 0.3545 & 0.2814 & 0 & 0.3743 & 0.0181 & 0.0179 & 0.0279 & 0.0012 & 0.1285 & 0.1074 & 0 \\ 0 & 0.1728 & 0.1254 & 0.0234 & 0 & 0.0064 & 0.0468 & 0.0682 & 0.1633 & 0.3189 & 0.5355 \\ 0 & 0.1993 & 0.2784 & 0.2440 & 0 & 0 & 0 & 0.1224 & 0.1814 & 0.2793 & 0.5004 \end{array}\right] \end{array}$$

The above matrix is the transition intensities between observable states with the assumption of an interaction between the local macro-economic situation and retail mortgage payments. Because elements of the probability vector, $A = \{\alpha_S\}$, are small, we can conclude that retail mortgage payment behaviors of a single local bank have little to do with the local macroeconomic factors.

CONCLUSION

The models presented in this paper are used to further analyze the relationship between local macro-economic factors and the payment pattern for a local bank's retail mortgages. From this analysis we conclude the following:

(1) Based on a first-order HMC, the probability of stay in a positive macro-economic state is 0.9143;

(2) For the period from April 2005 to September 2006, the estimated steady state probability distribution of the hidden macro-economic states is given as

$$\hat{H} = \begin{pmatrix} \hat{h}_1 \\ \hat{h}_2 \end{pmatrix} = \begin{pmatrix} 0.4033 \\ 0.5967 \end{pmatrix} \quad ; \text{and}$$

(3) The effect of the macro-economic states on retail mortgage loans is strong as indicated by the relatively large differences between corresponding observation $O_{S,1}$ and $O_{S,2}$ in rows 1 and 2 of the O_S matrix in Equation (28).

REFERENCES

Archera, G., and Titterington, D. (2002). Parameter estimation for hidden Markov chains. *Journal of Statistical Planning and Inference*, 108 (1-2), 1 November 2002, pp. 365-390.

Arribas-Gil, A., Gassiat, E., and Matias, C. (2006*)*. Parameter estimation in pair-hidden Markov models. *Scandinavian Journal of Statistics*, 2006, 33 (4), pp. 651-671.

Basel Committee on Banking Supervision. *International Convergence of Capital Measurement and Capital Standards*, Ch-4002, Basel 1997.

Ching, W., and Ng, M. (2004). Building Simple Hidden Markov Models. *International Journal of Mathematical Education in Science and Engineering*, 35, pp. 294-299.

Ching, W., and Ng, M. (2006). *Markov Chains: Models, algorithms and applications*, Springer.

Ching, W., Ng, M., and Fung, E. (2003). *Higher-order hidden Markov models with application to DNA sequence.* Lecture Notes in Computer Science, Springer.

Ching, W., Ng, M., Fung, E., and Siu, T. (2005). An Interactive Hidden Markov Model for Categorical Data Sequence, Working Paper.

Hobolth, A., and Jensen, J. (2005). Applications of hidden Markov models for characterization of homologous DNA sequences with a common gene. *Journal of Computational Biology*, 12 (2), 1 March 2005, pp. 186-203, doi:10.1089/cmb.2005.12.186.

Knudsen, B., and Miyamoto, M. (2003). Sequence alignments and pair hidden Markov models using evolutionary history. *Journal of Molecular Biology*, 33 (2), 17 October 2003, pp. 453-460 (8).

Liu, C., Hassan, M., and Nassar, R. (2007). Time series Analysis of the sensitivity of net incomes of industrial sections to macroeconomics factors. *Journal of International Business Disciplines*, 2 (1), November 2007.

MacDonald, I., and Zucchini, W. (1997). *Hidden Markov and other models for discrete-valued time series.* Chapman & Hall, CRC 1997.

Raftery, A. (1985). A model for high-order Markov chains. *Journal of the Royal Statistical Society*, B, 47 (3).

Thomas, L., Allen, D., and Kingsbury, N. (1998). A hidden Markov chain model for the term structure of bond credit risk spreads. Available at SSRN: http://ssrn.com/abstract=92568 or DOI: 10.2139/ssrn.92568.

JOURNAL OF INTERNATIONAL BUSINESS DISCIPLINES

Volume 7, Number 1 May 2012

Published By:
International Academy of Business Disciplines and Frostburg State University
All rights reserved

ISSN 1934-1822 WWW.JIBD.ORG