THE POWER OF A LEADING INDICATOR’S FLUCTUATION TREND FOR FORECASTING TAIWAN’S REAL ESTATE BUSINESS CYCLE: AN APPLICATION OF A HIDDEN MARKOV MODEL

Yun-Ling Wu1 --- Cheng-Huang Tung2 --- Chun-Chang Lee3†

1 Lecturer, Department of Real Estate Management, National Pingtung University, Taiwan
2 Associate Professor, Department of Computer Science and Information Engineering National Pingtung University, Taiwan
3 Professor, Department of Real Estate Management, National Pingtung University, Taiwan

ABSTRACT

This paper employs the discrete hidden Markov model (HMM) in order to capture information about the Markov switching model’s inner states that is not directly observable, and to pre-detect the real estate business cycle’s volatility trend. The empirical results show that this HMM can capture the asymmetry in the duration of states. Compared with the real estate leading indicator announced by the Taiwan Real Estate Research Center, this HMM yields the same results in terms of forecasting the trends of cycle fluctuations. The explanatory power of the HMM in forecasting out-of-sample data is supported both conceptually and methodologically.

Keywords: Real estate cycle, Real estate cycle leading indicator, Asymmetry, Duration, Hidden Markov model, Taiwan real estate research center.

1. INTRODUCTION

The long-term trends of the business cycle in the real estate market and how to predict fluctuations in that business cycle are subjects of great topical concern in macroeconomic analysis and decision making. Tvede et al. (2008) found that, according to a conservative estimate, construction activities account for 10% of global GDP, of which one-fifth comes from stable market fluctuations in the public sector and the rest results from significant periodic business cycle fluctuations. In other words, when activity in the real estate market decreases by 33%, there will be a 3% decrease in GDP, with that decrease not including the wealth spillover effects caused by the plunge in real estate prices. In addition, market observations over the past several decades and even centuries suggest that there are repeated business cycle fluctuations in the real estate markets around the world and that real estate market
fluctuations often lead to considerable fluctuations in raw material prices. Hence, monitoring the trend of real estate market cycle fluctuations is an important activity in the field of macroeconomic research.

Previous studies on business cycles can be categorized in two ways. First, in previous studies aimed at analyzing the real estate business cycle and its causes, researchers have applied proxy variables such as vacancy rates (Grenadier, 1995; Gordon et al., 1996) and real estate rental growth rates (Mueller, 1999) in their analyses. Second, the relationship between residential housing market price and quantity (Edelstein and Tsang, 2007) imbalances in real estate market supply and demand (Roulac, 1996) investment expectations (Huang, 2013) and the mortgage loan supply and asset price relationship (Arsenault et al., 2013) have been used to describe the causes of and changes to the real estate business cycle. However, those studies only describe the business cycle fluctuations or causes of fluctuations without comprehensively explaining the implications of the real estate business cycle. Moreover, some studies have discussed the real estate business cycle and macroeconomic relationships, or the mutual influences of relevant industries (Pritchett, 1984; Clayton, 1996; Goetzmann and Wachter, 2000; Kan et al., 2004; Leung, 2007; Huang, 2013). However, research on the composite indicators of the real estate business cycle is still relatively rare.

In order to directly quantify real estate business cycle fluctuations, recent studies have adopted the composite real estate index. Krystalogianni et al. (2004) used the leading composite index to predict the features and performance of the British commercial real estate cycle. Lee et al. (2009) applied the real estate leading composite index and Markov-switching model to explore the change in the real estate business cycle turning points. They found that the performance of the leading composite index for identifying real estate business cycle turning points is good. Although state transitions were excessively frequent and there were some disparities in terms of the durations of business recession and expansion periods, the model fitness performance and application were both found to be good. It was concluded that the composite index is superior to the single series for understanding business cycle fluctuations.

Many recent studies on macroeconomic cycles or real estate business cycles have focused on the observation of business turning points (Scott and Judge, 2000; Baum, 2001; Barras, 2005) and business duration periods. For the measurement model, the Markov switching model proposed by Hamilton (1989) has been the most widely applied, having been applied in various studies such as those of Hamilton (1989); Pelagatti (2001); Krolzig (1997) and Cruz (2005). The main advantage of the Markov-switching model is that it can capture the random changes of the business state over time through a group of unobserved state factors in the model setting mechanism. Moreover, the model can slightly modify the model pattern with different data patterns, such as the Markov model of intercept changing with state (Hamilton and Susmel, 1994) the Markov model of variance changing with state (Cai, 1994; Hamilton and Susmel, 1994; Gray, 1996) or a variety MSVAR (Markov-switching vector auto-regression models) (Krolzig, 1997). Hence, the application and analysis process of the model is more flexible and practical. The Markov-switching model is mainly dominated by a group of unobservable states and the random transition jumping mechanisms between those states. The state setting is the unobservable variable and is able to describe the features of the state random change. However, in the preset viewpoint, if the state is set as unobservable, it is not consistent with the observable data. As such, during the analysis process, it may be able to capture the state of random change but not able to observe state features, thereby losing a more accurate estimation of state path switching probability. As a result, it may lead to estimation error in estimating the trend of business cycle fluctuations.

The state random switching theory of the Markov-switching model in the framework of probability theory, namely the hidden Markov model (HMM), can be regarded as a double-embedded stochastic process (Huang et al., 2001). A complete HMM consists of two random processes: one layer is the hidden unobserved state switching series corresponding to a pure first-order Markov process, while the other layer is the observable random series in the hidden state. In the Markov-switching model, although random state switching series cannot be observed, through the observable series of state, a hidden Markov model can predict the originally unobservable state switching series.
probability. As the hidden state series is transformed into the observable state feature series, this model is known as the HMM.

The HMM was first applied in the late 1970s for the identification of sound signal fluctuations (Baum and Petrie, 1966) and is now widely used in engineering, genetics and other fields, such as communications audio classification and speech recognition, with considerable research achievements. According to Hassan and Nath (2005) the HMM model has the following advantages: (1) a strong statistical basis; (2) the ability to handle new information robustly; (3) the capacity to calculate and forecast similarity modes more efficiently. In recent years, the HMM model has been applied in economic, financial, and management fields, including economics in general (Leng and Wang, 2014) and the study of financial series fluctuations more specifically (Gregoir and Lenglart, 2000; Hassan and Nath, 2005; Korolkiewicz and Elliott, 2008; Oguz and Gurgen, 2008; Liu, 2010; Roamn et al., 2010; Zhu and Cheng, 2013). In the management field, it has been used to study customer relations management (Bouchaffra and Tan, 2004; Shen and Zhao, 2006; Netzer et al., 2008; Sepideh and Aaghaie, 2011) and online purchasing behaviors (Wu et al., 2005).

By following the current literature on the forecasting of real estate business cycle fluctuations and measurement models, based on the developed Markov-switching model, this study uses the real estate business cycle leading indicator composite index (footnotes 1) released by the Taiwan Real Estate Research Center to establish a discrete hidden Markov-switching model (discrete HMM; HMM with discrete output observations), in order to capture unobservable state implications. It is expected that the model will be able to predict the trends and changes in the real estate business cycle, as well as detect business turning points and state duration periods.

The remainder of this paper is organized as follows. Section 2 introduces the HMM theory and model parameter estimation; Section 3 explains the data description and analysis, as well as the features extracted from the economic cycle leading indicators series to comply with HMM model setting principles and ideas; Section 4 discusses the empirical results analysis; Section 5 offers conclusions.

2. HIDDEN MARKOV MODEL AND MODEL PARAMETER ESTIMATION

In previous studies, economists have developed a series of non-linear econometric models by describing the changing process of linear model parameters. The state transition model (regime switching model) can process multiple jumping processes of the parameters (footnotes 2). In order to observe the real estate business cycle leading indicator’s trend characteristics and the predictability of series fluctuations, this paper uses a non-linear HMM model as the measurement research tool. The model and parameter estimation results are as explained in the sections that follow.

2.1. HMM

In terms of the overall parameters and random concepts, the HMM can be divided into two parts. The first part can be described as a Markov chain to generate hidden state random series; the second part of the random process is described by the distribution of the observation variable probability in the state. The basic elements are as shown below (Huang et al., 2001; Koskinen and Öller, 2004):

1). Hidden state set: \( S = \{s_1, s_2, ..., s_N\} \), where \( N \) is the state number and \( q_{t} \) is the state at time point \( t \).

2). In-state output observation series set: \( O = \{o_1, o_2, ..., o_M\} \), where \( M \) is the number of observations in a state. (footnotes 3)

3). State transition probability distribution: \( A = \{a_{ij}\} \).
where \( a_{ij} = P(q_{t+1} = S_j | q_t = S_i) \), \( 1 \leq i, j \leq N \), and \( a_{ij} \geq 0 \). \( \sum_{i=1}^{N} a_{ij} = 1 \) represent the probability of switching from \( S_i \) to \( S_j \) and from \( t \) to \( t+1 \).

4). Under the conditions of state \( S_i \), the output observation variable probability distribution is: \( B = \{ b_i(o_m) \} \), where \( o_m \) is the output observation spatial sample in the state, and \( 1 \leq i \leq N \), \( o_m \in O \); \( b_i(o_m) = f(O_t = o | q_t = S_i) \). \( O_t \) is the output observation random variable at time \( t \), which can be a univariate variable or multiple variable. This paper sets the output observation value as a univariate discrete number.

5). Model initial state probability distribution: \( \Pi = \{ \pi_i \} \), \( 1 \leq i \leq N \), where \( \pi_i = P(q_1 = S_i) \)

To sum up, the parameters needed to describe a complete HMM model are \( \{S, A, B, \pi\} \), which is simplified in general literature in the form of \( \lambda = \{A, B, \pi\} \). In other words, the HMM model can be described by three setting parameters including the initial state probability distribution \( \{\Pi\} \), the hidden state transition probability distribution \( \{A\} \) and the in-state output observable series probability distribution \( \{B\} \).

The model makes two major assumptions. The first is the first-order Markov assumption, which assumes that the inter-state switching probability is related to the initial probability and the last term probability. The current term probability is subject to the influence of the previous term probability. The inter-state switching probability is non-time-varying as shown in Eqs. (1) and (2):

\[
P(Q_{\lambda}) = P(q_1, ..., q_t, ..., q_T | \lambda) = P(q_1 | \lambda) \prod_{t=2}^{T} P(q_t | q_{t-1}, \lambda) \tag{1}
\]

\[
a_{ij} = P(q_{t+1} = j | q_t = i), \ 1 \leq i, j \leq N \tag{2}
\]

The second assumption is the in-state observations \( o_m \) (output-independent) assumption, which assumes that observations and state \( q_t \) are dependent but that the observations are mutually independent. The conditional equation is as shown in Eq. (3).

\[
P(O | Q, \lambda) = P(o_1, ..., o_m, ..., o_M | q_1, ..., q_T, \lambda) = \prod_{t=1}^{T} b_{q_t}(o_m) \tag{3}
\]

In parameter setting, the HMM model is flexible and varying in pattern (footnotes 4). The proposed model pattern settings are as shown in Figure 1. There are three states including \( S_1 \), \( S_2 \), and \( S_3 \), and the in-state observations are divided into seven types by feature extraction method into \( O_1,...,O_7 \). The inter-state switching
probability $\alpha_{ij}$ can only delay the current term or switch forward over time to be set as $\alpha_{11}, \alpha_{12}, \alpha_{22}, \alpha_{33}$ and $\alpha_{31}$. The HMM pattern is a left-to-right pattern. By switching probability, the corresponding different business state duration periods can be inferred. With the three-state settings as an example, the switching probability is a $3 \times 3$ matrix and the diagonal estimate is the probability of the current term state remaining as the previous term state $(i = 1, 2, 3)$, respectively, being $\alpha_{11}, \alpha_{22}$ and $\alpha_{33}$. For this reason, it is inferred that the state duration periods are $1/(1 - \hat{\alpha}_{11})$, $1/(1 - \hat{\alpha}_{22})$ and $1/(1 - \hat{\alpha}_{33})$.

The left-to-right pattern is a type of generalized HMM pattern. The generalized HMM model pattern means that the inter-state switching probability can switch randomly without limitation. The main difference between the left-to-right pattern and the generalized pattern is that the inter-state switching probability can delay the current term or switch forward only (i.e., it cannot switch backward). The concept of such a setting mainly comes from the notion of moving forward in time and the idea that the state switching path will naturally jump forward. As shown in the above HMM model, the observation output sample space and probability distribution in hidden space may have different setting patterns according to the literature. This paper emphasizes the capability of the HMM model to capture the inter-state switching. Hence, it is set as a univariate discrete observation data pattern.

![Figure 1](image)

**Source:** Compiled by this study.

### 2.2. HMM Model Parameter Estimation Method

There are three basic problems to solve by using the HMM model to obtain the optimal parameters, namely, model training, hidden state optimal path estimation, and the computation of the maximum likelihood estimate (Baum and Petrie, 1966). The estimation method is briefly described as follows:

1). Computation of maximum likelihood estimate

With a given model optimization parameter $\hat{\lambda}_{ML} = \{A, B, \pi\}$, the computation of the maximum likelihood value of the observation series in the state $O = \{o_1, \ldots, o_m\}$ is used to compute the likelihood function $f(O|\hat{\lambda})$ value or logarithm likelihood function $\ln f(O|\hat{\lambda})$ value; indicating the model parameter $\hat{\lambda}_{ML} = \{A, B, \pi\}$ simulates the
observation series \( O = \{o_1, \ldots, o_m\} \) accuracy. This study uses the forward-backward procedure/algorithm to conduct the optimization of the HMM model to obtain the optimal solution.

2). Hidden state optimal path estimation

With a given model parameter \( \lambda_{ML} = \{A, B, \pi\} \) and observation series \( O = \{o_1, \ldots, o_m\} \), the estimation of the most possible hidden state optimal path \( Q = \{q_1, \ldots, q_T\} \) is used to estimate the optimal possible state path of the observation series. This paper uses a Viterbi algorithm for estimations.

3). Model learning/training

Model training is required for the solution of the model parameter estimation. After setting the HMM initial model and the observable sample series for the computation state \( O = \{o_1, \ldots, o_m\} \), computing the model parameter \( \lambda = \{A, B, \pi\} \) simulates the in-state observable sample series to obtain \( \hat{\lambda}_{ML} = \arg \max_{\lambda} f(O|\lambda) \) which determines the optimization model parameter: \( \hat{\lambda}_{ML} = \{A, B, \pi\} \). The EM algorithm or the Baum-Welch algorithm can be used for the estimations. The Baum-Welch algorithm is used herein.

In the evaluation of the model predictability, this paper uses expected loss function principles of turning point error (TPE) and mean squared error (MSE) to evaluate the forecasting errors between the actual observation values of the leading indicator and the estimated predicted values by using the HMM model to illustrate the performance of applying the HMM model in predicting the trends of the real estate business cycle leading indicator fluctuations. The two inspections can simply and directly compute the expected loss or expected cost error between the predicted values and the actual values.

3. DATA SOURCE AND ANALYSIS

The data used in this study were sourced from the real estate business cycle leading indicator composite index values from 1971 Q1 to 2009 Q4 released by the Taiwan Real Estate Research Center, a period including 156 samples of seasonal data. The Center releases data regarding the trends of the real estate business cycle via quarterly reports and elaborates on the compilation of economic cycle leading indicator series in the appendices of those quarterly reports. As the series data contained in the economic cycle leading indicators have different characteristics, after the verification of the quarterly and trend statistics method, all the contained series have been de-trend adjusted, while some data are seasonally adjusted by X12 software. The construction stock price index representing investments, the construction loan credit balance change representing production, and the CPI representing transactions have been adjusted according to the trend. In addition, the GDP and monetary supply values in the investment dimension have also been revised according to the de-trend adjustment and seasonal adjustment.

According to the previous literature, when analyzing a relevant time series by using an empirical model, the data should consist of a stationary series (footnotes 5) for parameter estimation and statistical inference. This paper conducts the ADF (augmented Dickey-Fuller test) unit root test of economic cycle leading indicators in advance to determine whether the series data are stationary. On the assumption of null hypothesis with a unit root, the ADF test confirms that the real estate business cycle leading indicator has a unit root, suggesting that series data are non-stationary. After the first differentiation of the data, this study conducts the ADF unit root, finding that the economic cycle leading indicator series at the 1% significance level rejects the null hypothesis after the first differentiation. In
other words, after the first differentiation, the economic cycle leading indicator series is stationary. Therefore, this study adopts the first differentiation series of leading indicator for analysis.

Regarding the setting of the HMM model, the number of states should be set in the beginning. According to the purpose and pre-concept, the number of states can be increased or decreased independently. Most of the previous studies on the overall business cycle or real estate business cycles have described business cycle fluctuations in terms of two states (Lee et al., 2009). Some research has used three states. For example, Cruz (2005) pointed out that the expansion and shrinkage stages of the three-state business cycle model described in previous studies was relatively similar to the turning points released by the NBER (National Bureau of Economic Research). The proposed model sets three business states, namely, market recession, market of no significant change, and market expansion. The setting of two states may result in excessive range to capture the feature information in the HMM model. In the estimation of the state switching probability, inter-state switching will be too slow to lead to an excessively lengthy duration period for each state. As a result, it is difficult for the HMM model to capture in-state features.

Second, with regard to the observable information value of each in-state type and the corresponding univariate business leading composite index, the univariate variable is set as discrete. This study uses the feature extraction method to process by dividing the information of the economic cycle leading indicators series by change into (significant increase, general increase, modest increase, small change, modest decrease, general decrease, significant decrease) output signals. Therefore, from 1971 Q2 to 2009 Q4, the leading indicators’ discrete observable data were extracted. This paper applies the feature extraction method in a manner similar to the concept of compiling business countermeasure signals (footnotes 6) to categorize series into a number of features by change. In addition to the emphasis on the application of the HMM model in the forecasting of the trends of business cycle fluctuations, it is expected to highlight the market state fluctuations’ degree and data switching features including business data expansion, recession, and others. Huang et al. (2001) extracted Dow Jones industrial average index features to represent the bull market, bear market, and fluctuations market of no change. Unlike the five types of business monitoring indicators, that study extracted seven types of trend signals of features fluctuations, expecting to further categorize the fluctuation degree of the leading indicator to observe more characteristics of the trends of the business cycle fluctuations. As mentioned above, Huang et al. (2001) argued that the continuous data contain a considerable degree of information despite the feature extraction of discrete data. However, if they are further divided into nine types or more, the value of each change is too small and the distinguishing of features will be too insignificant. As a result, it may result in excessively frequent jumping of state switching path probability. Therefore, this paper sets seven types of features.

To measure the forecasting capabilities of the HMM model against the trends of fluctuations, this paper divides the business leading composite index series (1971 Q2~2009 Q4) into two groups. One group consists of the in-sample (1971 Q2~2008 Q4) samples; this group, with a total number of 152 samples, is used as the training data for parameter estimation. Another group is the out-of-sample (2009 Q1~2009 Q4) observation data for 4-step-ahead forecasting. In addition, by using the rolling window sampling method (footnotes 7), with 4 seasons as a unit time length, the sample data are sorted out by time. The one-term lag data start with 1971 Q2 and develop at the interval of one season to 2008 Q4 to result in a total of 148 pattern training samples. The rolling window for the preprocessing of the observation samples can increase the number of samples to improve the accuracy of the parameter estimation. Meanwhile, it is expected to simulate the similar fluctuation path pattern series by using the in-sample samples as the four-term pattern samples in the forecasting of out-of-sample 4-step-ahead forecasts to improve forecasting accuracy.
4. ANALYSIS OF EMPIRICAL RESULTS

This section contains two parts. The first part is the description and analysis of the HMM model estimation and 4-step-ahead forecasts. The second part further elaborates on the forecasting capability of the application of the model in the trends of the economic cycle leading indicators out-of-sample fluctuations.

4.1. Analysis of HMM Model Estimation Results

The settings of the model initial parameters are shown in Table 1. The initial settings are the same as the model pattern shown in Figure 1: left-to-right pattern and state switches with time in a forward and irreversible away. Hence, regarding the inter-state switching probability of $a_{13}$, $a_{21}$ and $a_{32}$, before parameter estimation, the probability is set as 0, the rest ($a_{11}$, $a_{12}$, $a_{22}$, $a_{23}$, $a_{31}$, and $a_{33}$) are preset as 0.5. Regarding initial probability $\pi$, before training, the built-in software program uses an automatic setting that is equal to the $a_{ij}$ probability. The settings of the overall state switching probability $a_{ij}$ are similar to the assumption conditions in Eqs. (1) and (2), that is, the first-order Markov assumption and the assumption of non-time-varying inter-state switching probability. As shown in Table 1, the initial setting of each in-state observation value $b_i(o_m)$ probability is $1/7$, that is, the probability of each in-state observation value before parameter estimation is the same, which is consistent with the concept of the random appearance of samples.

For model training and program computation, this paper uses the built-in HMM program of MATLAB 7.0 software to obtain the parameter optimal solutions by iteration computation of the logarithm maximum likelihood value. Figure 2 illustrates the iteration time’s curve of the logarithm maximum likelihood during the optimal parameter estimation process. When the tolerance rate is 0.000001, the model parameter estimation converges when the iteration times are 432 during the model parameter estimation, and the logarithm maximum likelihood values are -1048.6.

Table 1. Initial parameter settings of the model before training ($\hat{\lambda}$)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Initial setting (before training)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_i$</td>
<td>State 1</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
</tr>
<tr>
<td>$a_{ij}$</td>
<td></td>
</tr>
<tr>
<td>State 1</td>
<td>0.5</td>
</tr>
<tr>
<td>State 2</td>
<td>0</td>
</tr>
<tr>
<td>State 3</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>Significant increase</td>
</tr>
<tr>
<td>$b_i(o_m)$</td>
<td>State 1</td>
</tr>
<tr>
<td>State 2</td>
<td>0.14286</td>
</tr>
<tr>
<td>State 3</td>
<td>0.14286</td>
</tr>
</tbody>
</table>

Source: Compiled by this study.

Note: (1) $\pi_i$ is the model initial probability; $a_{ij}$ is the inter-state switching probability; $b_i(o_m)$ is the probability of observation value $o_m$ at state $i$, $1 \leq i, j \leq 3$. $1 \leq m \leq 7$.
(3) In-state observation value $o_m$: significant increase, general increase, modest increase, small change, modest decrease, general decrease, and significant decrease.
After training, the optimization model parameters are as shown in Table 2. The state initial probability $\hat{\pi}_i$ suggests that the probability of the model initial state in state 1 (recession market state) is the highest at 0.8089. This can be verified by the trends of the real estate leading composite index as shown in Figure 3. As shown there, the duration of the market recession is not lengthier than the duration of the market expansion or market of no significant fluctuations. Obviously, the estimation of the probability of the initial state in state 1 is reasonable. Moreover, the Table 2 inter-state switching probability value $\hat{a}_{ij}$ suggests that the probability of the previous term state of 3 and the current term state of 1 is the highest, followed by the probability of the previous term state of 1 and the current term state of 1. The probability of the previous term state of 2 and the current term state of state 3 is the third rank. Based on the switching probability of the current term of state 1 and the following term of state 1 and the current term of state 2 and the following term of state 2, the degree of state continuation is high. However, the probability of the current term of state 3 and the following term of state 3 ($\hat{a}_{33}$) is almost 0, suggesting that the continuation rate for state 3 is very low. This can be verified by the long-term trends of the leading indicator as shown in Figure 3. When the market is in the expansion state of state 3, in most cases, it will switch to the recession market state of state 1. In other words, there are more downward turning points than upward turning points, indicating that the probability of remaining in state 3 is low. All of the above verifies the reasonability of $\hat{a}_{33}$ estimation.

By using the switching probability estimate $\hat{a}_{ij}$, this paper estimates the business state’s average duration period (footnotes 8). The probability of the current term business being in the market recession state (state 1) and the following term business being in the recession state is $\hat{a}_{11}=0.8089$. The probability of the current term business state being in the unchanged market state (state 2) and the following term business being in the unchanged market state is $\hat{a}_{22}=0.40991$, while the probability of the current term business being in the market expansion state (state 3) and the following term being in the expansion state $\hat{a}_{33}$ is almost 0. The average duration period of business in the state of
recession (state 1) is 5.23 seasons, the average duration period of business in the state of market of no change (state 2) is 1.69 seasons, and the average duration period of business in the expansion state (state 3) is very short. From the perspective of state 1 and state 2 switching probability, the state is continuous, which confirms the characteristic of Taiwan’s real estate market business state being difficult to change and supports the viewpoint of continuous business state. The state 3 (expansion market state) probability estimation suggests that it is difficult to remain in the same state. This paper infers that the setting of the HMM model shifts from left to right. As a result, the probability shifts from state 3 to state 1. In addition, the leading indicator trends in Figure 3 suggest that the trends in the expansion stage are steep and that the downward fluctuation sections are frequent. Hence, the trend deepness in the expansion state is relatively less steep as compared with the deepness in the state recession period. Therefore, it is relatively easier to shift to the recession period. The probability value of state 3 suggests that it is unlikely to remain in state 3. The estimation verifies that the period of business recession is longer than the expansion period.

The above estimation results of the business average duration period suggest that the real estate business cycle’s recession and expansion periods are asymmetric (footnotes 9). In this respect, the findings of this study are consistent with most of the available empirical results for the overall business cycle or the real estate business cycle, indicating that expansion periods and recession periods have asymmetric trends. The real business cycle model is considered to be subject to the influence of the same random variables in business expansion period and recession period states, and thus they have the same degree of fluctuations and the same dynamic characteristics. However, Figure 3 suggests that the trend of increasing real estate business cycle leading indicator data during the period from 1971 to 2009 is steeper and shorter in duration. On the contrary, the trends of the decreases in most indicators are less steep and last longer. The long-term trends of time series data are consistent with the empirical results.

According to the \( \hat{\beta}_i(o_m) \) estimation results, when the market is in the recession state, the total probability of observations of decreases in leading indicators (i.e., modest decreases, general decreases, and significant decreases) is 0.2065. When the market is in a state of no significant change, the probability of the trend change observation (small change) is close to 50%, and up to 0.4136. When the market is in a state of expansion, the feature value of the in-state observation value general increase is 0.5213. Roughly speaking, the in-state feature value probability distribution meets our expectations, as the features in the expanding market state (state 3) (significant increase, general increase and modest increase) may concentrate on the observation feature of the modest increase. The probability of the occurrence of the two observation features is almost 0. This may be caused by the steep trends of the real estate business cycle leading indicator in the expansion period, and the in-state feature closer to the boundaries of other states can be more easily captured by the small change features of state 2 in probability. As a result, it is difficult for the model to capture in the expanding market state the probability of two features including significant increase and general increase but concentrate on the boundary section of state 2 features. Another possible reason is that the samples are not categorized by smaller features and thus the abnormalities of each in-state feature can easily affect the bias of the probability estimation.

In 4-step out-of-sample forecasting, through the known in-sample observation data, the discrete HMM model is established to allow the known out-of-sample forecasting observations and known in-sample observation data to have a similar pattern and similar statistical characteristics. The forecasting results error rate is as shown in Table 3. According to the absolute error \( \varepsilon_i \) and mean absolute error \( \bar{\varepsilon} \) ratio of the simple state estimates, the forecasting accuracy of the out-of-sample data is not satisfactory, which is as expected. According to real estate business cycle leading indicator trends as shown in Figure 3, the indicator has been considerably decreasing since 2008. There is a possibility of series structural change.
Table 2. After HMM training, model parameter \( \hat{\lambda} \) parameter

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Optimal parameter (after training)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\pi}_i )</td>
<td>State 1: 0.8089, State 2: 0.1911, State 3: 0</td>
</tr>
<tr>
<td>( \hat{a}_{ij} )</td>
<td>State 1: 0.8089, State 2: 0, State 3: 1</td>
</tr>
</tbody>
</table>

\[
\hat{b}_i(\mathbf{o}_m) = \begin{cases} 
0.1371 & \text{significant increase} \\
0.1176 & \text{general increase} \\
0.1252 & \text{modest increase} \\
0.4136 & \text{small change} \\
0.1283 & \text{modest decrease} \\
0.0782 & \text{general decrease} \\
0.0000 & \text{significant decrease} \\
\end{cases}
\]

<table>
<thead>
<tr>
<th>State</th>
<th>Significant increase</th>
<th>General increase</th>
<th>Modest increase</th>
<th>Small change</th>
<th>Modest decrease</th>
<th>General decrease</th>
<th>Significant decrease</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1371</td>
<td>0.1176</td>
<td>0.1252</td>
<td>0.4136</td>
<td>0.1283</td>
<td>0.0782</td>
<td>0.0000</td>
</tr>
<tr>
<td>2</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.2588</td>
<td>0.0000</td>
<td>0.0478</td>
<td>0.3628</td>
<td>0.3307</td>
</tr>
<tr>
<td>3</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.4787</td>
<td>0.5213</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Source: Compiled by this study.

Note: (1) \( \hat{\pi}_i \) is the model initial probability; \( \hat{a}_{ij} \) is the inter-state switching probability; \( \hat{b}_i(\mathbf{o}_m) \) represents the probability of observation value \( \mathbf{o}_m \) at state \( \hat{\pi} \).

1 \( \leq i, j \leq 3 \), 1 \( \leq m \leq 7 \)


(3) In-state observation value \( \mathbf{o}_m \): significant increase, general increase, modest increase, small change, modest decrease, general decrease, and significant decrease.

Figure 3. Trends of the real estate business cycle leading indicator (1971Q1 ~ 2009Q4)

Source: Compiled by this study.

Table 3. HMM model 4-step out-of-sample forecasting results - mean absolute error rate

<table>
<thead>
<tr>
<th>Forecasting period</th>
<th>Actual state value ( y_i )</th>
<th>Predicted state value ( \hat{y}_i )</th>
<th>Absolute error ( \epsilon_i )</th>
<th>Mean absolute error ( \bar{\epsilon} ) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009 Q1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>100%</td>
</tr>
<tr>
<td>2009 Q2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2009 Q3</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2009 Q4</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Source: Compiled by this study.

Notes: (1) In this study, the business state is divided into three types: State 1: recession market. State 2: unchanged market. State 3: expansion market.

(2) Forecasting period: 4-step-ahead forecasts.

(3) The absolute error in the table \( \epsilon_i \) is set as \( y_i = \hat{y}_i \), then \( \epsilon_i = 0 \), and the rest \( \epsilon_i \) is 1.

(4) Mean absolute error \( \bar{\epsilon} = \frac{1}{n} \sum_{t=1}^{n} |y_t - \hat{y}_t| \)

This paper further examines the model-estimated data of each year, and deletes the samples of 2008-2029 suspected of structural changes. The total sample HMM model is then applied in the estimation of parameters, and the
means absolute error rate is applied in the detection of total sample of each year in four-step forecasting as shown in Table 4. The results are consistent with those of Wu (2009). The mean estimation error rate of the total sample is 44.90%; the accuracy of forecasting samples out of the four-step forecasting (2007Q1-2007Q4) is up to 75%. It is concluded that the HMM model has good accuracy in forecasting fluctuations of the business cycle leading indicator. The results validate the inference of this study. In the period of 2009 Q1 to Q4, there are structural changes in Taiwan’s real estate business cycle leading indicator. Previous literature (Rau et al., 2001; Chen, 2006) on the overall business cycle of Taiwan indicates that structural change occurred during the 1990s. In the field of real estate business cycle research, many articles on the measurement of the real estate business cycle by housing price, such as those by Chen (2003) and by Peng et al. (2004) argue that housing prices did undergo structural changes. However, to avoid divergence of research topics and to simplify the model to highlight the characteristics of the HMM model for capturing the jumping state, the topics of structural change are not considered and explored in this study.

### Table 4. HMM model forecasting results of total sample in each year-mean absolute error rate detection table

<table>
<thead>
<tr>
<th>Period sample forecasting</th>
<th>Mean absolute error rate (%)</th>
<th>Period of sample forecasting</th>
<th>Mean absolute error rate (%)</th>
<th>Period of sample forecasting</th>
<th>Mean absolute error rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>60Q2-Q4</td>
<td>0</td>
<td>73Q1-Q4</td>
<td>75</td>
<td>86Q1-Q4</td>
<td>0</td>
</tr>
<tr>
<td>61Q1-Q4</td>
<td>25</td>
<td>74Q1-Q4</td>
<td>25</td>
<td>87Q1-Q4</td>
<td>75</td>
</tr>
<tr>
<td>62Q1-Q4</td>
<td>75</td>
<td>75Q1-Q4</td>
<td>25</td>
<td>88Q1-Q4</td>
<td>75</td>
</tr>
<tr>
<td>63Q1-Q4</td>
<td>25</td>
<td>76Q1-Q4</td>
<td>0</td>
<td>89Q1-Q4</td>
<td>50</td>
</tr>
<tr>
<td>64Q1-Q4</td>
<td>100</td>
<td>77Q1-Q4</td>
<td>25</td>
<td>90Q1-Q4</td>
<td>25</td>
</tr>
<tr>
<td>65Q1-Q4</td>
<td>50</td>
<td>78Q1-Q4</td>
<td>50</td>
<td>91Q1-Q4</td>
<td>25</td>
</tr>
<tr>
<td>66Q1-Q4</td>
<td>75</td>
<td>79Q1-Q4</td>
<td>50</td>
<td>92Q1-Q4</td>
<td>75</td>
</tr>
<tr>
<td>67Q1-Q4</td>
<td>75</td>
<td>80Q1-Q4</td>
<td>25</td>
<td>93Q1-Q4</td>
<td>50</td>
</tr>
<tr>
<td>68Q1-Q4</td>
<td>50</td>
<td>81Q1-Q4</td>
<td>50</td>
<td>94Q1-Q4</td>
<td>75</td>
</tr>
<tr>
<td>69Q1-Q4</td>
<td>75</td>
<td>82Q1-Q4</td>
<td>0</td>
<td>95Q1-Q4</td>
<td>50</td>
</tr>
<tr>
<td>70Q1-Q4</td>
<td>75</td>
<td>83Q1-Q4</td>
<td>25</td>
<td>96Q1-Q4</td>
<td>25</td>
</tr>
<tr>
<td>71Q1-Q4</td>
<td>25</td>
<td>84Q1-Q4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>72Q1-Q4</td>
<td>50</td>
<td>85Q1-Q4</td>
<td>25</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Compiled by this study

This paper further uses the MSE criterion to estimate the accuracy of estimation by determining the degree of error as indicated by the difference between the actual value and the forecast value of the business indicator. The total MSE result is 1.0148, as shown in Table 5, where \( \hat{y}_i \) is the leading indicator value deduced from the optimal path by the estimation of the HMM model. Before the application of the HMM model, this study has extracted the features of the economic cycle leading indicators for intermittent classification. In the process, to extract features, the category average number is estimated. Hence, \( \hat{y}_i \) is the average feature category deduced from the optimal path of forecasting.

### Table 5. HMM model 4-step-ahead forecasts-MSE criterion test

<table>
<thead>
<tr>
<th>Forecasting period</th>
<th>Actual leading value (( y_i ))</th>
<th>Estimated leading indicator value (( \hat{y}_i ))</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009 Q1</td>
<td>91.97</td>
<td>91.9682</td>
<td>1.0148</td>
</tr>
<tr>
<td>2009 Q2</td>
<td>92.57</td>
<td>92.5682</td>
<td></td>
</tr>
<tr>
<td>2009 Q3</td>
<td>92.58</td>
<td>93.9798</td>
<td></td>
</tr>
<tr>
<td>2009 Q4</td>
<td>94.63</td>
<td>93.1809</td>
<td></td>
</tr>
</tbody>
</table>

Source: Compiled by this study.

Note: (1) \( y_i \) is the real leading indicator value; \( \hat{y}_i \) is the estimated leading indicator value.

\[
MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2
\]

(2) MSE criterion testing value

© 2017 AESS Publications. All Rights Reserved.
4.2. Analytic Comparison of HMM Model Forecasting Results and Conditions Released by Taiwan Real Estate Research Center

For the business duration period, the estimated average duration period of business in the state of recession (state 1) is 5.23 seasons, the estimated average duration period of business in the state of market of no change (state 2) is 1.69 seasons, and the estimated average duration period of the business in the state of expansion (state 3) is very short. These estimations are in line with the current condition that business recession durations are longer than the durations of the expansion periods. According to the trends of economic cycle leading indicators released by the Real Estate Research Center as shown in Figure 4, the state of recession is more frequent and longer than expansion in the business cycle. However, the results on the average of the long and short durations are different from those released by the Taiwan Real Estate Research Center (footnotes 10). This is caused by the difference in the number of states. As some state paths of the model are affected by state 2 of market of no change, thus the probability of state 1 and state 3 estimated paths is decreased. The possible reason is due to HMM model left-to-right pattern and initial state switching probability setting. The regression computation of different initial settings has a considerable impact on the parameter estimation of the optimal HMM model. In addition, the overall parameter of \( \hat{\alpha}_{31} \) is consistent with the settings of the HMM model. The cause of the low duration of state 3 may lead to the difference between the estimation of business duration and asymmetry and the results released by the Real Estate Research Center. According to the leading indicator trends released by the Real Estate Research Center (Figure 4), compared to recession market states, the probability of the current term being in the expansion state and the following term being in the expansion state is not very high. As seen in the figure, the market often switches to the recession market state after it enters into the expansion state. This is consistent with the estimated switching probability of \( \hat{\alpha}_{31} \).

Furthermore, the HMM model estimation results of switching probability suggest that there is an asymmetric relationship between the recession and expansion periods of the business cycle. Previous empirical research into the overall business cycle and real estate business cycles suggests that there are asymmetric trends for the expansion periods and recession periods. As shown in Figure 4, from 1971 to 2009, Taiwan’s real estate business cycle leading indicator data had steep increases of shorter durations. On the contrary, the trends of decreases for most indicators tended to be less steep and last for longer periods of time. The long-term trends of the time series data can verify the asymmetric estimation results.

![Figure 4. Real estate business cycle leading and coincident indicator composite index trends](Source: Architecture and Building Research Institute, Ministry of the Interior, Taiwan Real Estate Research Center, National Chengchi University (Quarterly report of Taiwan’s real estate business cycle trends, December 2009))
In order to observe the changes of the real estate business cycle leading indicator and HMM model forecasting capabilities by following the practice of the 4-step out-of-sample forecasting method, this study uses the average change rate of the features of the state optimal path to estimate the real estate business cycle leading indicator. The model estimated trend forecasting is as shown in Figure 5. Between 1972 Q3 and 1974 Q4, the model estimation has apparent lagging trends. The trends of the upward and downward fluctuations of the model forecasting are in the same direction with those released by the Taiwan Real Estate Research Center. Wherein, there are five decreasing forecasting values, which are consistent with the real estate leading indicators released by Taiwan Real Estate Research Center, in particular, those for the period of 1971 Q2~Q4 and the period of 2008 Q1~Q4. As far as the business cycle fluctuations deepness is concerned, the HMM-estimated leading indicator is deeper than the real indicator in the state of recession. Changes in forecasting average fluctuations suggest significant differences. These differences are caused by the different number of states.

Figure 5. predicted trends of real estate business cycle leading indicator

Source: Compiled by this study.

5. CONCLUSION

Based on the time series data of the real estate business cycle composite leading indicator jointly compiled and regularly released by the Architecture and Building Research Institute, as well as the Taiwan Real Estate Research Center of National Chengchi University, this paper uses the HMM to capture the optimal path of state transition to observe the trends of fluctuations of out-of-sample data. Regarding the identification and application of the economic cycle leading indicators and HMM model in the trends of business cycle fluctuations, the results confirm that trends of the real estate business cycle fluctuations are asymmetric and that the average duration of recession periods is longer than that of expansion periods.

By overcoming the unobservable limitation of using the Markov-switching model to capture state series setting, this study applies the univariate discrete HMM model and feature extraction method to extract signals from the unobservable state to increase data simulation. Overall, the in-state observation value probability is consistent with the market fluctuations that the HMM model extracts for the in-state observation values to observe the unobservable state of the Markov-switching model. In terms of the out-of-sample 4-step-ahead forecasts, except for the sections of the fluctuations of structural change, the results of the mean absolute error rate of the estimation accuracy suggest that the average accuracy of using the HMM model to forecast 4-step out-of-sample state is up to 55.41%. This indicates that the proposed model is advantageous in model estimation.

However, the proposed model has not considered the differences caused by the fluctuation trend path estimation due to structural change. Therefore, it cannot effectively master the fluctuation path of total samples in the current
term. Previous studies have suggested that the leading indicator is not only useful for setting and selection but can also serve as the leading reminder of macroeconomic phenomena. Therefore, future studies can use the improved dual-layer built-in HMM measurement models based on the leading indicator series compiled by the Taiwan Real Estate Research Center, such as the multivariate Markov vector auto-regression model for describing the common fluctuation characteristics between series (Krolzig, 1997) the switching probability time-varying Markov-switching model (Peersman and Smets, 2001) the duration dependent Markov-switching model of switching probability with duration dependent characteristics (Pelagatti, 2001) or the continuous hidden Markov-switching model, to further reduce forecasting error.

Note

Note 1: The leading indicator composite index is compiled by the Taiwan Real Estate Research Center, College of Social Sciences, National Chengchi University, at the delegation of Construction and Planning Agency, Ministry of the Interior. Starting with the first quarter of 1971, it initially consisted of relevant macroeconomics indicators including GDP, monetary supply, and CPI. In 1981 and 1989, indicators relating to the real estate industry and the construction stock index as well as the construction credit balance change were used to form the current real estate business cycle leading composite index.

Note 2: The state transition model is a non-linear model. It can be divided into two types by the state transition observation. If the model’s state transition process is determined by the observable variables, it is known as the threshold model, in which case state change is determined by whether the observable variable is above the threshold value. If it is above threshold value, the state transition occurs. For another type of the model, the state transition process is determined by the unobservable variables. Therefore, the model should define the state transition process. The Markov switching model is such a model. The Markov switching model considers that data come from different parent matrices. By setting a group of self-regression equations and using a Markov chain to understand the inter-state switching process, the current term state will be subject to the influence of the previous state, so that the data of various periods will be continuous and relevant. The probability theory can be applied to estimate the non-linear shifting of state transitions.

Note 3: In the HMM model, there are a number of patterns of hidden in-state observation output sample spaces and probability distributions such as univariate discrete observations, continuous observations, and mixed continuous observations. This paper sets the output observations as type of univariate discrete data distribution.

Note 4: There are a number of general patterns for the HMM model. The main change is that inter-state switching probability can switch in between states without limitation. Second, the pattern of the in-state observable value can be univariate discrete data, univariate continuous data, multivariate discrete data, or multivariate continuous data. To summarize, the two assumptions of this study should be satisfied, that is, in-state observation value \( O_m \) and state \( q_t \) are mutually dependent and mutually independent with \( O_m \). The first-order Markov assumption and the current term switching probability are subject to the influence of the current term and previous term switching probability only.

Note 5: The term “stationary series” means that the statistical characteristics of the time series data generation process, including average, variance, and covariance, should be limited constants. It expects the time series variable’s important statistical characteristics including average, variance… to be non-time-varying
to facilitate statistical inference and parameter estimation. On the other hand, if it is misused, it can easily lead to estimation bias. The most well-known example of such bias is proposed by Granger and Newbold (1974) when they note that spurious regression often occurs in between non-stationary variables.

Note 6: In Taiwan’s real estate business cycle trend quarterly reports, the scores are given in terms of changes in the indicators. The real estate business cycle trend signals are categorized into five lights, specifically, red, yellow-red, green, yellow-blue, and blue lights, to represent business indicators of ranging from a significant increase to a significant decrease including overheated business, business boom, business stability, poor business, and business recession.

Note 7: The term “rolling window” refers to the sampling method of a given sample length and sampling period range shifting by term.

Note 8: The estimation of the business duration period is determined by the computation of inter-state switching probability, that is, $1/(1 - \hat{\alpha}_{11})$, $1/(1 - \hat{\alpha}_{22})$ and $1/(1 - \hat{\alpha}_{33})$ represent the average duration periods of the market recession, market of no change, and market expansion.

Note 9: The term “business cycle asymmetry” refers to the inconsistency in the durations of expansion and recession periods in business cycles. Sichel (1993) proposed two features relating to business cycle asymmetric fluctuations: deepness and steepness. In that study, deepness is described as the inconsistency in the distance from the trend values in the case of valley and peak of the business cycle in fluctuations. In addition, steepness is mainly used to describe the inconsistency in the slope of the valley and peak movements in the business cycle, that is, the speeds at which rebounds from the peak and the valley occur are not the same.

Note 10: According to the data released by the Taiwan Real Estate Research Center, the current average expansion period lasts 9 seasons and the average recession period lasts 24 seasons, indicating an apparent asymmetry of longer business recessions than expansions.

Funding: This study received no specific financial support.

Competing Interests: The authors declare that they have no competing interests.

Contributors/Acknowledgement: All authors contributed equally to the conception and design of the study.

REFERENCES


