EXPLORING THE IMPACT OF REAL ESTATE POLICY ON REAL ESTATE TRADING USING THE TIME SERIES ANALYSIS

Shiwei Shao*, Xin Huang, Lixia Xiao, Hui Liu

Wuhan Land Resource and Urban Planning Information Center, Wuhan, China, shaoshiwei@163.com

KEY WORDS: Time series analysis, Real estate, Policy analysis, Turning point detection

ABSTRACT:

Housing price is a major issue affecting people's lives, but also closely related to the interests of the people themselves. Housing prices are affected by various factors, such as economic factors, population size factors, social factors, national policy factors, the internal factors of real estate and environmental factors. With the deepening of urbanization and the agglomeration of urban population in China, housing prices have been further accelerated. The Chinese government has also introduced a series of policies to limit real estate transactions and affect property prices. This paper also aims to explore a time series analysis method to analyse the impact of real estate policies on real estate prices. Firstly, the article searches for policy factors related to real estate through government official channels such as state, Prefecture and city, and analyses key words related to policy by means of natural language processing. Then, the real estate registration volume, transaction volume and transaction house price data which are arranged into time series are modelled using ARIMA time series model, and the data are processed according to scatter plot, autocorrelation function and partial autocorrelation function graph of the model to identify its stationarity. Finally, the LPPL (logarithmic periodic power) model and MPGA (multi-population genetic algorithm) are used to fit and detect turning points of real estate registration data, and the time series detection algorithm is used to obtain the inflection time nodes of the sequence, and then the relationship between real estate policy and real estate transactions is analysed. Taking the real estate registration data in Wuhan as an example, this paper validates the above time series analysis method. The results show that some real estate policies (such as purchase restriction policy, public rental policy, etc.) have a certain impact on real estate transactions in a short time. Part of the real estate policy (such as graduate security, settlement policy, etc.) does not have a significant impact on real estate transactions. To sum up, the government's brutal blockade of macro-control of the housing market cannot fundamentally solve the housing difficulties of the people, but also standardize the real estate market trading mechanism, innovate the market trading mode, so as to promote the long-term development of the housing market.

1. INTRODUCTION

Housing price is a major issue affecting people's lives, but also closely related to the interests of the people themselves. Housing prices are affected by various factors, such as economic factors, population size factors, social factors, national policy factors, the internal factors of real estate and environmental factors. With the development of urbanization and the agglomeration of urban population in China, housing prices have been further accelerated. The Chinese government has also introduced a series of policies to limit real estate transactions and regulate the real estate market in order to make the real estate market develop smoothly and orderly.

At present, the research on real estate is mainly to study the cyclical fluctuation of prices. Cheng et al. optimized the LPPL model parameters based on a multi-population genetic algorithm and effectively extracted the time series turning points to study the periodicity of oil prices and real estate prices [3]. Smirnov et al. established a reference chronology for the Russian economic cycle from the 1980s to 2015. In order to test the peaks and valleys of the economic cycle, three methods (local minimum/maximum, Bayesian model and Markov conversion model) are used to determine the periodic turning point of the time series data [2]. Giusto et al. proposed a simple machine learning algorithm called Learning Vector Quantization (LVQ) to quickly and accurately assess the turning point in the US real estate cycle over the past five years [4]. Greble and Burns compared the total real estate, public buildings, private buildings and residential buildings in the United States from 1950 to 1978. It was found that during the 28 years, the house experienced about 6 cycles, non-residential experience has 4 cycles and the real estate cycle lags behind the national economic cycle by about 11 months [4]. Hendershott and Kane cited data from 30 cities in the United States in the 1980s and came up with the relationship between office vacancy rates and economic losses. In the case of a 20% vacancy rate, the office building can bring about 13 billion yuan in losses per year to the United States [4]. Baldi pointed out that it was very important to analyze the change of real estate price from the perspective of credit funds. We should pay attention to the impact of changes in credit conditions such as the proportion of down payment on the real estate market and real estate prices [6]. Adalid and Detken analyzed the 18 member states of the Organization for Economic Co-operation and Development. The study showed that credit funds have a significant impact on the changes in asset prices. They pointed out that due to excessive support for credit funds in the real estate market, real estate prices have risen sharply [7]. Guido Baldi adds real estate industry-related variables to the New...
Keynesian dynamic stochastic equilibrium model to study the impact of bank credit changes on housing prices and macroeconomics in the process of rising housing prices\cite{8}.

Compared with the study of foreign real estate cycle theory, the research on real estate in China started late and the related theoretical research is still immature. Since the establishment of the market economic system in 1993, real estate in China has entered the period of recovery and development. So far, the research on the real estate market in China has gradually started. Today, Chinese scholars have also made some achievements in the theory of the fluctuation of the real estate cycle, mainly from the following aspects.

Hu Kelin evaluate the Beijing real estate regulation and control policies, such as land-stopping policies, credit policies, financial policies, and safeguard policies by the PSR model (pressure-state-response) based on the monthly data from 2001 to 2010. The study selected six indicators to test, including the infrastructure investment amount, land price, population quantity, urban resident’s disposable income, CPI, M2. House prices are significantly affected by real estate control policies and have a three-month lag period \cite{9}. Kuang studied the interaction between real estate investment, real estate credit and economic growth based on the selected data from 1996 to 2007 in 35 large and medium-sized cities in China. The results of the study show that real estate prices have a greater impact on bank credit than interest rates and economic growth \cite{10}. Zheng Zhonghua and Zhang Yu established a dynamic stochastic general equilibrium model of multiple economic entities. The banking system and the real estate market constructed the main transmission factors of economic fluctuations in the model and simulated the impact on other factors in the macroeconomics. The results of the study show that the introduction of credit funds will lead to the flow of funds in the market to high-yield industries, thus pulling the price of real estate\cite{11}. Based on the quarterly data of the national real estate market between 2003 and 2013, Wu established the ARDL model to study the long-term and short-term relationship between real estate prices and bank credits. It is found real estate prices and bank credits are mutually causal in the long run and house price fluctuations will drive bank credit in the short term. Based on the annual panel data of 35 cities in 2003-2013, the SYS-GMM method is used to estimate the relationship between house price fluctuations and bank credits in the eastern, central and western regions and it is found that the relationship between the two has obvious regional characteristics. The credit elasticity in the western region is relatively large and the central region is the smallest \cite{12}.

In order to further improve the efficiency of real estate registration, explore the value of the real estate data and assist in the decision-making of real estate business, we intend to carry out research on the spatial-temporal analysis and visualization of real estate data. The logarithmic periodic power law (LPPL) model is a timed and quantitative analysis method for the asset bubble of logarithmic periodic power law bubble theory. The construction of the model mainly consists of two parts. One is that the noise traders can influence each other and the influence is continuously strengthened to reach a peak and the investors will collectively sell assets to make the market collapse at the same time. Second, the market crash may occur before the asset bubble breaks. The LPPL model can be used to judge the state of the real estate market bubble to predict the future trend of real estate and can also be used to judge the bursting point of the real estate market bubble.

## 2. Time Series Turning Points Detection Based on LPPL Model Optimized by MPGA

### 2.1 MPGA Algorithm

Genetic algorithms must be highly adaptable to changes in complex dynamic environments. The genetic algorithm must be able to closely track the change of the solution in dynamic environment until the best solution is obtained that is an important difference between dynamic optimization and traditional evolutionary algorithms. The goal of traditional evolutionary algorithms is to gradually converge to a satisfactory solution which makes the population lose diversity that is the necessary condition to effectively explore the entire feasible space. Traditional evolutionary algorithms lose ability to adapt to environmental changes later in evolution which is the main challenge for the application of evolutionary algorithms in dynamic environments. In recent years, many scholars find some methods to solve this problem. Multi-population genetic algorithm is a fairly good algorithm to further extend the genetic algorithm. The multi-population genetic algorithm can divide the entire population into small populations, several of the populations track the current extreme point and the others continue to search for new extreme points. Compared with traditional algorithms, multi-group genetic algorithms pay more attention to ensuring the diversity of populations which is a necessary condition for algorithms to adapt to environmental changes.

Taking the double population as an example, the chromosome migration process between the populations is shown in the Figure 1.

![Figure 1. Map of chromosome migration between populations](image)

### 2.2 MPGA Algorithm

The LPPL model is an algorithm commonly used in the financial field to detect the bursting of the economic bubble. It is also used to detect the oil price transition point. LPPL can be regarded as a method for detecting the turning point of time series. Therefore, the LPPL model is adopted as the research method of time series turning point detection in this study.

The LPPL model formula is shown in Equation (1).

$$y(t) = A + B(t_c - t)^x + C(t_c - t)^y \cos(\omega \ln(t_c - t) + \phi)$$

\(y(t)\) is the statistical feature quantity at time \(t\), \(A, B, C\) and \(\phi\) are the structural parameters of the formula, \(\alpha\) is the index, \(I_c\) is the corresponding time series turning point, \(\omega\) is the amplitude
of the time series. To some extent, \((t_i - t)^\alpha\) as a power law part describes how fast the feature quantity grows.

Time series turning point test

In order to verify the results of the time-breaking points obtained by the LPPL model, the spectral analysis method is used to test the model results. The spectrum analysis method has the following parts:

1. A series of frequencies set in advance (length is \(M\))

2. Calculate the frequency spectral density for each frequency value, and the calculation formula is shown in Equation (2), the formula for calculating the parameters in Equation 2 are shown in Equation (3), Equation (4), Equation (5), Equation (6).

\[
p(\omega) \approx \frac{1}{2\pi} \left\{ \frac{\sum_j x_j \cos(2\pi f_j(t_i - t))}{\sum_j \cos(2\pi f_j(t_i - t))} - \frac{\sum_j x_j \sin(2\pi f_j(t_i - t))}{\sum_j \sin(2\pi f_j(t_i - t))} \right\}
\]

\[
\sigma^2 = \frac{1}{M-1} \sum_{j=1}^{M} (x_j - \bar{x})^2
\]

\[
r = \frac{1}{4\pi f} \arctan \frac{\sum_{j=1}^{M} \sin(4\pi f_j)}{\sum_{j=1}^{M} \cos(4\pi f_j)}
\]

(3) The highest frequency is removed because it may be generated by a random sequence. The frequency sequence statistic is calculated as Equation (7).

\[Z = -\ln(1 - (1 - p)^1/M)\]

Remove the frequency value where the P value is less than the statistic z.

4. Select the frequency with the largest \(P\) value as the result of spectral analysis test.

2.3 Algorithms and processes

We mainly use the LPPL model to fit historical real estate data to derive potential turning points of time series. The LPPL model has 7 parameters, including 3 linear parameters \((A, B, C)\) and 4 nonlinear parameters \(\{t, \omega, \varphi, \alpha\}\). This study focuses on the optimization of four nonlinear parameters in the LPPL model. At the same time, linear parameters are solved according to the principle of least squares to achieve global optimization of the LPPL model. The time series turning point detection model includes the LPPL model and the multi-group genetic algorithm (MPGA) which form a research framework. There are several steps:

1. The four nonlinear parameters of LPPL are binary coded which is divided into four parts corresponding to four nonlinear parameters. The value range of \(t\) is \([0, 0.6]\) and occupies 6 bits in binary encoding. The value range of \(\omega\) is \([0, 2\pi]\) that takes the accuracy to 0.1 and occupies 6 bits in binary encoding. The value range of \(\varphi\) is \([0, 0.4]\) that takes the accuracy to 0.01 and occupies 13 bits in binary encoding. The value range of \(\alpha\) is \([0.1, 1]\) that takes the accuracy to 0.1 and occupies 3 bits in binary encoding. The total code length is 28 bits.

2. The initial set population number is 10 and the number of individuals in each population is 100. The chromosomes of each individual are randomly selected from 0, 1.

3. Fitness function is the standard for survival of the fittest in genetic algorithm. In this study, the fitness function value is the mean square error obtained by substituting the parameters obtained by each individual into the LPPL model. In this project, the linear parameters need to be calculated by least squares under the condition that the nonlinear parameters of each individual are calculated and the fitness function is solved in the LPPL. Substitute the parameters into the LPPL to solve the fitness function. The smaller the value of the function, the better the LPPL model parameters are.

4. Population selection is based on the classic roulette algorithm. The fitness of each individual in the population needs to be calculated first. The smaller the fitness, the better the individual, the greater the probability of being selected to enter the next generation of reproduction. The roulette algorithm is used to determine the final selected individual.

5. Crossover is the core of genetic algorithms. Set the crossover coefficients \((0.7-0.9)\) randomly to determine the number of crossed individuals in each population. Perform a single point crossover operation on selected individuals in each population that randomly select a gene position and cross-interchange all the gene positions in the two chromosomes before and after the gene position to obtain two new individuals.

6. Mutation operation is an important means for genetic algorithms to jump out of local best. The coefficient of variation ranges from 0.001 to 0.05 and is set randomly. The number of variant gene positions is determined and the corresponding gene positions are randomly selected from all individual gene position sets in the population. Mutations vary from 0 to 1 or from 1 to 0.

7. Inter-population migration is the biggest difference between multi-population genetic algorithms and standard genetic algorithms. Inter-population transfer passes the optimal value of one population to the next population to replace the worst performing individual in the next population to speed up the convergence and prevent some populations from falling into local optimum prematurely.

8. Finally, the optimization parameters obtained after multiple iterations are substituted into the LPPL model, and the significance and effectiveness of the model are verified by spectral analysis.

The overall algorithm flow is shown in the Figure 2.
3. EXPERIMENTS AND RESULTS

3.1 Experimental data

3.1.1 Real estate registration data

The research data of this project is the real estate registration data of Wuhan from December 2016 to October 2018. We analyze Wuhan real estate registration transaction volume data by time series analysis method.

3.1.2 Data related to real estate policy

Policy release is determined by several government departments and can be divided into national, provincial, and municipal policies. Since policies are issued by different departments, we collect policy data from the official websites of 11 agencies. The agency information is shown in the Figure 3.

We collected a total of 123 real estate-related policies from October 1, 2016 to October 1, 2018, including various types, such as housing purchase policies, real estate market macro-control policies, housing leasing policies, settlement policies, real estate convenience policy, rural land policy, provident fund policy, etc. Policy data information is shown in the Table 1.

| Key Points                  | Name of Document                                                                 | Policy Collection Department                                      | Public Date |
|-----------------------------|----------------------------------------------------------------------------------|-------------------------------------------------------------------|-------------|
| Regulate the Real Estate Market | Circular of the Ministry of Housing, Urban and Rural Construction on Implementing the Asset Assessment Law and Regulating the Management of Real Estate Valuation Industry | Ministry of Housing and Urban-Rural Development of the People’s Republic of China | 2016/12/6   |
| Guarantee family rental allowance | Guidance of the Ministry of Finance of the Ministry of Housing, Urban and Rural Construction on the Work of Urban Housing Security and Family Rental Subsidies | Ministry of Housing and Urban-Rural Development of the People’s Republic of China | 2016/12/8   |
| Convenient for Real estate | WuHAn Natural Resources and Planning Bureau Optimizing and Integrating for the People and Benefiting the People: New Measures for the People Promoted by the East-West Lake Real Estate Registration Center | WuHAn Natural Resources and Planning Bureau | 2016/8/26   |
| Remote district Limited Purchase | Notice of the Wuhan Housing Administration on Expanding the Limitation of Housing Purchase | Wuhan public accumulation fund management center | 2016/12/21  |
| Construction of Small Towns | Notice on transmitting the notice of the Ministry of housing and urban rural development of China Construction Bank on promoting commercial finance to support the construction of small towns | Department of Housing and Urban-Rural Development of Hubei Province | 2016/12/30  |
| Distribution of public | Notice on Effectively Strengthening the | Department of Housing | 2017/1/8    |

This contribution has been peer-reviewed. 
https://doi.org/10.5194/isprs-archives-XLII-3-W10-1281-2020 | © Authors 2020. CC BY 4.0 License.
Firstly, the data preprocessing of Wuhan real estate registration data includes a series of processing such as data regularization and abnormal data detection. Then, policy text data is processed by word segmentation, feature selection and weighting methods to eliminate the impact of data irregularities on the results.

3.2 Results

3.2.1 Policy data theme extraction

We assume that each month from December 2016 to October 2018 is a time turning point. Taking into account the delays in the impact of policies on real estate, we assume that every month from December 2016 to October 2018 is a time turning point. Since the impact of policies on real estate is delayed, we conducted data mining on the policies in the three months before each turning point and obtained the results of policy interpretation.

The following picture shows all the themes and key words from the corresponding policies in each month from December 2016 to October 2018.

| Theme | Key Words |
|-------|-----------|
| Settlers policy | Register, settle, population, apply, legitimate, land use, city proper |
| Rental market policy | Lease, housing, public, distribution, check in, guarantee, register |
| Real estate market regulation | The real estate, enterprise, development, housing, business, real-estate market |
| Urban affordable housing Purchase restriction loan | Housing, supportability, town, lease, guarantee, security department, regulations, housing loan, purchase, city, limited loan, restrictions, apply, resident family |

3.2.2 Turning point of real estate time series trading data

We use MPGA-optimized LPPL model to perform turning point detection on real estate time series trading data of Wuhan and all administrative regions. The results are as follows.

| Turning Point Date | Policy Theme |
|--------------------|--------------|
| 2016/12            | Purchase restriction loan, Regulate the Real Estate Market…… |
| 2017/6             | Convenience of Real property registration, Real estate inventory…… |
| 2018/6             | Priority to first needed family, Regulate the Real Estate Market…… |

3.2.3 Visualization of the theme data from real estate policy

The following picture shows cloud map visualization of policy keywords mined from the corresponding policies of each month from December 2016 to October 2018. The policy theme at each month can be clearly shown from the word cloud map.

Cloud map visualization of policy keywords from October 2016 to December 2018 was as the word cloud map for December 2016. We got four topics, including restrictions on purchases and loans, regulating the real estate market, building smart cities and rural land related policies.

Restrictions on purchases and loans and regulating the real estate market have significant effects on real estate registration, real estate transaction volume and transaction price, so the two themes are selected to do time series analysis.

There are several policy themes during the period from December 2016 to October 2018: convenient real estate registration, regulating the real estate market, restrictions on purchases and loans, rural land related policies, building smart cities, priority purchase for newly needed families, personal provident fund loan related policies, settlement policies, points to households, urban affordable housing policy, college graduate housing security policy, housing ticket policy, building a characteristic town.
Figure 6. Cloud map visualization of policy keywords in June 2018, the pictures indicate Priority to Just needed family, Regulate Real Estate market, Housing rental market policy, Convenience of Real property registration and Housing provident fund personal loan policy.

4. DISCUSSION

4.1 Policy interpretation of real estate price time series turning point

From the perspective of the whole city of Wuhan, the house price has shown a rising trend from January to February 2017. In September-December 2016, Wuhan City issued a series of restrictions on purchase and loan, which led to the low price of house prices in early 2017. The rise in 2017 may be affected by the release of some settlement policies and the policies of university students staying in Wuhan. Until April and May of 2018, there is a large upward fluctuation, and then it tends to stabilize or even decline. This is due to the regulation of the real estate market by the state and local governments.

From the perspective of various administrative districts in Wuhan, it is consistent with the overall trend of Wuhan. Among them, Jiang'an District, Hannan District and Xinzhou District have experienced a relatively large sustained growth since January-February 2017, and have been fluctuating up and down since July and August. Qingshan District, Qiaokou District, Dongxihu District, Hannan District, Xinzhou District have a peak turning point from November 2017 to January 2018, which may be affected by the policies of 20% discount for college students to purchase house after November 2017. Hongshan District, Dongxihu District and Jiangxia District also had an obvious peak turning point around July 2018. It was the result of the policy of personal provident fund loan, the need for family-first purchase and other rent-related policies.

4.2 Timeliness Analysis of Policy on Real Estate Transactions

From the real estate related policy data and transaction data of Wuhan. At the end of 2016, a number of restrictions on purchases and loans were issued, which led to a trough in February 2017. The timeliness of such policies will continue to affect the real estate transactions within two to three months. There was also a trough in March 2018, which may be the result of a short-term one-month policy for urban security family housing in February 2018. From the perspective of various administrative districts in Wuhan, there are lows in January-February 2017 and March-April 2018, which may be affected by the policies of restricting purchases and loans at the end of 2016 and the urban security family housing issued in February 2018. This is consistent with the overall situation in Wuhan. Jiang'an District, Qingshan District, Hongshan District, Dongxihu District, Caidian District and Jiangxia District also had a significant peak in May-July 2018, which may be affected by the April personal provident fund loan policy and the need for family priority housing purchase policy in March. Policy has had an impact within 2-3 months.

By associating policy interpretation information with real estate transaction time series, policy information is semantically attached to the real estate transaction time series turning point. It realizes the interpretation of the policy interpretation of the time turning point of real estate transaction in Wuhan and all administrative districts. The specific results are as follows.

The overall transaction of Wuhan real estate has shown a steady trend from January 2017 to October 2018. But after the trough in April 2017 and April 2018, it began to rise again. Due to the impact of the public rental housing policy and the relaxation of settlement policies, a time series turning point was formed in April 2017. The regulation policy for the real estate market and town security family housing policy for housing have resulted in a time turning point in April 2018.

Due to the impact of policies such as restrictions on purchases and loans and government regulation of the market, the real estate transaction number in Jiang'an District formed a time series turning point in February-March 2017.

Figure 7. March 2017 Wuhan real estate transaction timing turning point policy interpretation

Figure 8. April 2018 Wuhan real estate transaction timing turning point policy interpretation

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLII-3-W10-1281-2020 | © Authors 2020. CC BY 4.0 License.
1286
5. SUMMARY

This project conducts time series analysis from three aspects of real estate registration, transaction number and transaction price. It constructs time series for the statistical characteristics of multiple spatial scales of Wuhan city, administrative district, cadastral district and cadastral sub-district. The ARIMA model is used to model the time series after the statistics, and decompose the original sequence into periodic components, trend components, and residual components. On this basis, the log-periodic power law (LPPL) model under the optimization of multi-group genetic algorithm (MPGA) is used as the trend turning point detection of time series. At the same time, the project collects real estate-related policy data from the web portals of major real estate related departments since the second half of 2016, and mines the topic information and keywords in the policy information through the natural language processing theme model. Finally, the policy mining information is combined with the real estate statistical time series, and the policy is used to interpret the turning trend of real estate time series and visualize the display.

REFERENCES

Cheng F, Fan T, Fan D, et al. The prediction of oil price turning points with log-periodic power law and multi-population genetic algorithm[J]. Energy Economics, 2018,72:341-355.

Giusto A, Piger J. Identifying business cycle turning points in real time with vector quantization [J]. International Journal of Forecasting, 2017,33(1):174-184.

Grebler L, Burns L S. Construction Cycles in the United States Since World War II*[J]. Real Estate Economics, 1982,10(2):123-151.

Hendershott P H, Kane E J. CAUSES AND CONSEQUENCES OF THE 1980s COMMERCIAL CONSTRUCTION BOOM[J]. Journal of Applied Corporate Finance, 1992,5(1):61-70.

Baldi G. The economic effects of a central bank reacting to house price inflation[J]. Journal of Housing Economics, 2014,26:119-125.

Adalid R, Detken C. Liquidity Shocks and Asset Price Boom/Bust Cycles[J]. Working Paper, 2007,27(6):697-711.

Goodhart C, Hofmann B. House prices, money, credit, and the macroeconomy[J]. Oxford Review of Economic Policy, 2008,24(1):180-205.

Li Ling, Zhu Daolin, Hu Kelin. Application of PSR model to the effects of real estate regulation policy on house price: a case of Beijing,2012,34(4):787-793.

Kuang Weida. Real estate investment, real estate loan and economic growth in China, 2011,V(1):59-68.

Zheng Zhonghua, Zhang Yu. The real estate market, the banking system and the fluctuation of economic in China: a multiple departments’ analysis based on dynamic stochastic general equilibrium, 2015,V33(2):53-69.

Wu Chengchen. Study on the regional layering effects of the relation between the real estate price fluctuation and bank credit, 2015.

Cheng Xiangjun, He Zhenhuan, Yang Zhaoxia. Machine learning traffic signal control approach based on genetic algorithm, 2004,24(8):130-135.

Yuan Wanghuang, You Xiaoming, Liu Sheng,et al. Adaptive simulated annealing ant colony algorithm for solving TSP problem,2018,35(2):261-266.

Tian Dongping Chi Hongqin. Hybrid genetic algorithm and simulated annealing, 2006,42(22):63-65.

Smirnov S V, Kondrashov N V, Petronovich A V. Dating Cyclic Turning Points for Russia: Formal Methods and Informal Choices[J]. Journal of Business Cycle Research, 2017,13(1):53-73.