Journal of Software Engineering

ISSN 1819-4311

Academic Journals Inc.

www.academicjournals.com

Journal of Software Engineering 10 (1): 170-184, 2016
ISSN 1819-4311 / DOI: 10.3923/jse.2016.170.184

Numerical Computation of the Fluctuations of Commercial Housing Prices and Disposable Incomes of Urban Residents Based on the VAR and LSE Models

[1]Kerong Zhang and [2]Wuyi Liu
[1]School of Economics and Management, Fuyang Normal University, China
[2]Department of Science and Technology, Fuyang Normal University, Qing He West Road No. 100, Fuyang 236037, People's Republic of China, China

Corresponding Author: Wuyi Liu, Department of Science and Technology, Fuyang Normal University, Qing He West Road No. 100, Fuyang 236037, People's Republic of China Tel: +86-558-2596562 Fax: +86-558-2596561

ABSTRACT

To ensure the healthy development of the national real estate industry and explore the determinant of commercial housing prices in China, the correlation between fluctuations of commercial housing prices and disposable incomes of urban residents was analyzed and predicted with annual time series data based on Vector autoregressive (VAR) model. The study dealt with the bifactorial data and found positive correlations in both the empirical VAR model and the Least Squares Estimation (LSE) model. The Granger causality test and Cointegration tests were next carried out to verify the rationality of the empirical VAR model and its settings. The present study revealed that the urban resident disposable income was one of the crucial impact factors influencing the rising commercial housing prices. There exists a positive correlation or convergence relationship between the development trends of the commercial housing prices and the disposable incomes of urban residents on the basis of a fifteen years prediction by impulse response function. Finally, some comments and suggestions were put forward relevant to the corresponding policy adjustment and decision making of commercial housing market in view of the present empirical study.

Key words: Vector autoregressive model, least squares estimation, time series data, commercial housing price, disposable income

INTRODUCTION

Housing prices are traditionally modeled using hedonic regression models following the pioneer work by Rosen (1974). Within the hedonic framework, houses are usually regarded as a bundle of attributes offering utilities to researchers and the observable housing prices are viewed as the value realization of heterogeneous bundles with varying amounts of different attributes. A inevitable issue concerned with this standard Ordinary Least Squares (OLS) specification is that errors might not be independent from one another and this case will lead to inefficiency as well as erroneous inference concerning these parameter estimates due to the spatial and temporal feature of housing transactions (Basu and Thibodeau, 1998; Pace *et al*., 1998). Later, LeSage and Pace (2004) comprehensively discussed both the theoretical and statistical reasons that would explain why data from several research fields would be prone to spatial dependence. Consequently, both the spatial and temporal dependence need to be accounted for in a specific VAR model if one aims to address the correlated errors in the traditional hedonic regression specification (Liu, 2013). Such exercises are not only beneficial to improving estimation efficiency and inference accuracy but also have crucial implications on house price prediction in the empirical application of hedonic pricing models.

In modern society, the real estate industry is regarded as the pillar for the development of national economy. It has grown up in the processes of industrialization, urbanization and modernization of contemporary national economy and promotes greatly the progress of these economic development processes in return. The real estate markets, especially the housing market, just become an important organic component of the macro social and economic system. In recent years, the commercial housing prices are on the rise in the real estate markets when the disposable incomes of urban residents rise with the rapid economic development in China. It is an important part in the process of building a moderately prosperous society in China. In the late 1980s and early 1990s, the fluctuations of real estate markets had played key roles of predictable social and economic indicators in the developed countries such as Japan and the United Kingdom in inflation. Therefore, the operation of real estate markets is a major issue in modern social and economic development and the monetary authorities and economists showed great attention and interest in the fluctuations of real estate prices. It is not only the real estate markets in China that have been examined. There are also several studies on the housing markets of the United States, the United Kingdom, Canadian (Gau, 1984, 1985; Hosios and Pesando, 1991; Clayton, 1998; Schindler, 2014). However, the economist made reports with conflicting results on the sustainability and predictability of housing markets (Schindler, 2014). Later, a comprehensive survey on further empirical findings on the sustainability and predictability was conducted by Gatzlaff and Tirtiroglu (1995). In China, the rising urban residents' disposable incomes are accompanied by the inflation of commercial housing prices in recent years. The rise of the urban residents' disposable incomes would improve the urban people's paying capacities and raise their consumption levels of real estate products. In turn, the rapid development of real estate industry and surging of commercial housing prices would very likely restrain the consumption levels of urban residents in commercial housing and result in their relatively low paying capacities. The rapid development of real estate industry and surging of commercial housing prices would very likely give rise to the consumption levels of urban residents and result in their relatively low paying capacities. Is there positive or negative correlation between these two elements in housing market? Could other possible relationships account for the phenomenon between these two elements? To our knowledge, there is yet few reports on analyzing and testing the subtle correlation between the fluctuations of commercial housing prices and disposable incomes of urban residents based on VAR models. The ideas of the present study came from the result of a wide literature review. In the study, the possible correlation between fluctuation of commercial housing prices and disposable incomes of urban residents was analyzed and predicted with annual time series data based on the VAR and LSE models. Positive correlation was found in both the VAR and LSE models. Next, the unit root tests, cointegration tests and Granger causality tests were carried out to verify the rationality of the VAR model, while the response of commercial housing prices to urban resident incomes was predicted by an impulse response function. It is shown that the impact of urban resident disposable incomes is one of the positively dominant factors to influence the rising commercial housing prices. The present study was aimed to explore and evaluate the impact of current urban resident disposable incomes on the corresponding commercial housing prices. The results obtained in the study could provide new evidences and insights into the determinant factors of commercial housing prices.

MATERIALS AND METHODS

Data overview and pre-processing: The annual data of the commercial housing prices and the disposable incomes of urban residents were collected in time series data (1999-2013) from "China Province Statistical Yearbook" with references to "China City Statistical Yearbook". Next, the original annual data was processed and transformed into logarithms based on the scales of Chinese national Consumer Price Index (CPI) and Fixed Assets Investment Price Index (FAIPI) with necessary adjustments before subsequent analyses to eliminate possible hetero-scedasticities. The processed data were subsequently analyzed with the econometric software package Eviews 8.0 (Quantitative Micro Software Inc., USA). There might be auto-correlations of variables in the subsequent modeling process when the econometric model was used to analyze the correlation between these variables. The plots of data trends and auto-correlograms of the original variables (X, Y) and the transformed variables (logX, logY) were shown in Fig. 1a-b and 2a-b.

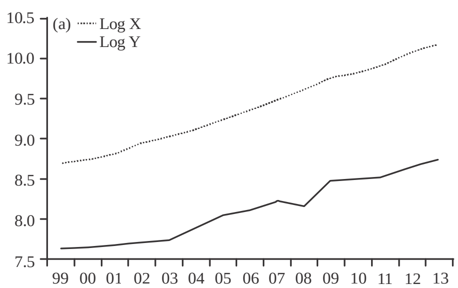
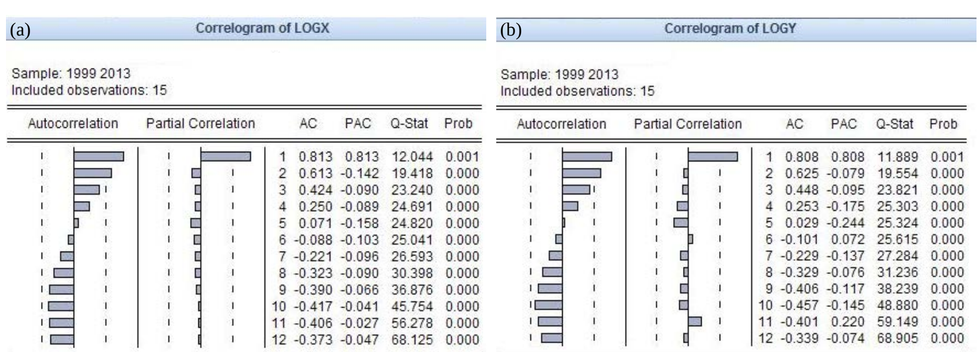


Fig. 1(a-b): Rising trend plots of the original variables (X, Y). The parameters (X, Y) on the x-axis and y-axis of Fig. 1b showed the annual data of the disposable incomes of urban residents and the commercial housing prices, while the parameters (logX, logY) on the x-axis and y-axis of Fig. 1a represented their transformed logarithm data. The units taken on the x-axis and y-axis of Fig. 1a and Fig. 1b were years (logX and X in blue on the x-axis) and the values of corresponding variables (logY and Y in red on the y-axis)



(a) Correlogram of LOGX

Sample: 1999 2013
Included observations: 15

	AC	PAC	Q-Stat	Prob
1	0.813	0.813	12.044	0.001
2	0.613	-0.142	19.418	0.000
3	0.424	-0.090	23.240	0.000
4	0.250	-0.089	24.691	0.000
5	0.071	-0.158	24.820	0.000
6	-0.088	-0.103	25.041	0.000
7	-0.221	-0.096	26.593	0.000
8	-0.323	-0.090	30.398	0.000
9	-0.390	-0.066	36.876	0.000
10	-0.417	-0.041	45.754	0.000
11	-0.406	-0.027	56.278	0.000
12	-0.373	-0.047	68.125	0.000

(b) Correlogram of LOGY

Sample: 1999 2013
Included observations: 15

	AC	PAC	Q-Stat	Prob
1	0.808	0.808	11.889	0.001
2	0.625	-0.079	19.554	0.000
3	0.448	-0.095	23.821	0.000
4	0.253	-0.175	25.303	0.000
5	0.029	-0.244	25.324	0.000
6	-0.101	0.072	25.615	0.000
7	-0.229	-0.137	27.284	0.000
8	-0.329	-0.076	31.236	0.000
9	-0.406	-0.117	38.239	0.000
10	-0.457	-0.145	48.880	0.000
11	-0.401	0.220	59.149	0.000
12	-0.339	-0.074	68.905	0.000

Fig. 2(a-b): Auto-correlogram plots of the transformed variables (a) logX and (b) logY

Econometric models: The present study carried out econometric tests mainly based on the Vector autoregressive (VAR) and Least Squares Estimation (LSE) models, unit root tests with the Augmented Dickey-Fuller (ADF) method, cointegration tests and the Granger Causality tests (Hamilton, 1994; LeSage and Pace, 2004; Gao, 2006) to make a systematic computing and modeling of the correlation between the commercial housing prices and the disposable incomes of urban residents in China.

The regression model of vector autoregressive: The vector autoregression (VAR) is commonly used for forecasting systems of interrelated time series and for analyzing the dynamic impact of random disturbances on the system of variables (Johansen, 1991, 1995). In the study, the method of VAR model is adopted to estimate the effect of the fluctuation of urban residents' disposable incomes on commercial housing prices in China. The usage of this approach let us recognize the cumulative effect taking into account the dynamic response between the fluctuations of urban residents' disposable incomes and commercial housing prices. In time series analysis, if the differential variables for both the original endogenous and exogenous variables are stated in stationary process and cointegration (Note: the tests of stationary process and cointegration for both the endogenous and exogenous will be showed subsequently. However, time series dynamics tend to switch between multiple stationary regimes in practice. In small samples, the distributions of the coefficients may be improved in econometric studies by the estimation of VAR model in the 1st or 2nd differences (Hamilton, 1994). The vector autoregressive (VAR) process is capable of scaling to a large number of noisy time-series and the appropriate differential is significant because the estimations of most algorithms fail when time series are not stationary. It could be deduced the following VAR model to estimate and forecast the long-run equilibrium or relationship between the two variables mentioned above. The A VAR model consists of a set of K endogenous variables. Thus, a VAR process can be defined as shown in Eq. 1:

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + A_3 Y_{t-3} + ... + A_p Y_{t-p} + BX_t + \mu_t \tag{1}$$

where, Y_t is a K vector of endogenous variables, X_t is a vector of exogenous variables, A_p and B Matrices of coefficients to be estimated and μ_t is a vector of innovations that may be contemporaneously correlated but are uncorrelated with their own lagged values and uncorrelated with all of the right-hand side variables. Specifically, A_p is a K×K coefficient matrix and μ_t is a K dimensional zero-mean in the covariance stationary innovation process, i.e., $E[\mu_t] = 0$ and $E[\mu_y \mu_t^T] = \sum_\mu$. Since only lagged values of the endogenous variables appear on the right-hand side of the VAR equation, simultaneity is not an issue and Ordinary Least Squares (OLS) yields consistent estimates with t-statistics and the statistics R^2.

The regression model of Least Squares Estimation (LSE): The LSE model of the commercial housing prices and the urban residents' disposable incomes could be expressed in the following Eq. 2 (Table 1):

$$\log Y = \alpha + \beta \log X \tag{2}$$

where, α is the constant and β is represents the least squares regression coefficient for variable logX. It should be noted that other impacting factors involved in the fluctuation of commercial housing prices were looked as minor factors and ignored in subsequent analyses.

Furthermore, the long-run equilibrium fluctuations and correlation between the commercial housing prices and the disposable incomes of urban residents was finally revealed and verified by the VAR and LSE models. However, economic theory scarcely provides some guidance for which variables appear to have a stochastic trend and when these trends are common among the examined variables as well. Therefore, the variables should be tested with the statistical methods of unit root, cointegration and the Granger Causality theory to explore whether they were suitable to be applied in the econometric data modeling, when the VAR and LSE models were executed and computed. The next steps were carried out following the subsequent definitions.

Unit root test: A series is regarded to be (weakly or covariance) stationary if the mean and autocovariances of the series do not depend on time. Any series that is not stationary is called being nonstationary. Standard inference procedures do not apply to regressions which contain an integrated dependent variable or integrated regressors. Therefore, it is important to check whether a series is stationary or not before using it in a regression. The formal method to test the stationary of a series is the unit root test. Nevertheless, the data of stationary sequences often appears in fluctuations around a mean value and its tendency to the mean value is clear and easy to test. A difference stationary series is said to be integrated and is denoted as I(d) where d is the order of integration. The order of integration is the number of unit roots contained in the series, or the number of differencing operations it takes to make the series stationary. If a variable is stated in stationary sequence, it can be represented as I(0). If its 1st order differential variable is stated in stationary sequence, it can be denoted as I(1). Similarly, if its 2nd order differential variable is stated in stationary sequence, it can be denoted as I(2). In the unit root tests, the null hypothesis is that a unit root exists and the sequenced variable or time series data is non-stationary. Therefore, these types of tests are all called unit root tests. In the study, the Augmented Dickey-Fuller (ADF) (Dickey and Fuller, 1979) unit root test is used for the estimation of individual time series with intention to provide evidence for when the variables are integrated which is usually followed by a multivariate cointegration analysis. All of these tests are available here as a view of a series. For simplicity, the general form of the ADF test is defined as shown in Eq. 3:

$$\Delta x_t = a + bt + rx_{t-1} + \sum_{i=1}^{p} r_i \Delta x_{t-i} + \varepsilon_t \tag{3}$$

where, the "a" was a constant, "t" was a variable of time trend and "p" meant the number of lags. The null hypothesis (H_0) of the test was that r = 0 and the alternative hypothesis (H_1) was that r ≠ 0. If the null hypothesis was accepted, then there existed a unit root within the variable tested with t-statistics, i.e. non-stationary. Otherwise, the alternative hypothesis should be accepted and there was no unit root within the variable tested with t-statistics, i.e., stationary.

Johansen cointegration test: The finding that many macro time series may contain a unit root has spurred the development of the theory of non-stationary time series analysis. The concept of cointegration was suggested by Granger (Engle and Granger, 1987) to deal with error correction models in empirical estimations. Engle and Granger (1987) pointed out that a linear combination of two or more non-stationary series may be stationary. The main idea of cointegration test was to test the hypothesis whether there existed a long-term regression (or correlation) relationship between the two time series variables. The common methods of the cointegration test are the

Engle-Granger two-step method and the Johansen cointegration test. The Engle-Granger's concept of cointegration is usually a residual-based testing for simple unit root tests, while the later is aimed to determine the causality direction in the data during selected years. Moreover, the asymptotic distributions of the Engle-Granger statistics are probable non-standard and depend on the deterministic regressors specification, while its critical values for the statistics are obtained from simulation results. The Johansen test may be performed using a group object or an estimated VAR object. Its residual tests may be computed using a group object or an equation object estimated using nonstationary regression methods. The EViews 8.0 supports VAR based cointegration tests using the methodology developed in Johansen (1991, 1995) performed using a group object or an estimated VAR object.

Pairwise granger causality test: The pairwise Granger causality tests are used to determine whether an endogenous variable can be treated as exogenous (Engle and Granger, 1987), i.e., whether one of the time series is useful in forecasting another. For each equation in the VAR, the output displays χ^2 (Wald) statistics for the joint significance of each of the other lagged endogenous variables in that equation. The statistic in the last row is the χ^2 (Wald) statistic for joint significance of all other lagged endogenous variables in the equation. Generally speaking, a time series A is said to Granger-cause B if it can be shown that those values provide statistically significant information about future values of b, usually through a series of t-tests and F-tests on the lagged values of A. If a time series data is in stationary process, the Granger causality tests are performed using the level values of two or more variables. If the variables are non-stationary, the Granger causality tests are thus made using the first or higher order differences (i.e., 1st and 2nd lags). The number of lags to be included is usually chosen using an information criterion.

Analysis of the predicted data by impulse response function: To obtain the impulse response function, a VAR should be first estimated. A shock to the i-th variable would directly affect the i-th variable and transmit it to all of the other endogenous variables through the dynamic (lag) structure of the VAR too. An impulse response function can trace the effect of a one-time shock to one of the innovations on current and future values of the endogenous variables. If the innovation ε_t is contemporaneously uncorrelated, interpretation of the impulse response is straightforward. The i-th innovation $\varepsilon_{i,t}$ is simply a shock to the i-th endogenous variable ($Y_{i,t}$). However, innovations are usually correlated and may be viewed as having a common component which cannot be associated with a specific variable. In order to interpret the impulses, it is common to apply a transformation P to the innovations so that they become uncorrelated, as shown in Eq. 4.

$$V_t = P\varepsilon_t \approx (0, Q) \tag{4}$$

where, D is a diagonal covariance matrix.

RESULTS AND DISCUSSION

Description of the empirical data: The income data are nominal per capita disposable incomes of urban residents given in Table 1. The income variable X_t is the ratio of the income series divided by the same urban Consumer Price Index (CPI). For commercial housing, the price data are variable Y_t is the ratio of the above price series divided by the same national Fixed Assets Investment Price Index (FAIPI). These empirical data collected were further processed and

Table 1: The descripted data of the commercial housing prices and the disposable incomes of urban residents

Time (year)	CPI (consumer price index) for urban residents (percentage of the same period last year)	Urban residents' nominal per capita disposable income	Urban residents' actual per capita disposable Income (X)	FAIPI (fixed assets investment price index) for urban house investment (percentage of the same period last year)	Nominal commercial housing price	Actual commercial housing price (Y)
1999	98.7	5854.0	5931.1	99.6	2053.0	2061.2
2000	100.8	6280.0	6230.1	101.1	2112.0	2089
2001	100.7	6859.6	6811.9	100.4	2170.0	2161.4
2002	99.0	7702.8	7780.6	100.2	2250.0	2245.5
2003	100.9	8472.2	8396.6	102.2	2359.0	2308.2
2004	103.3	9421.6	9120.6	105.6	2778.0	2630.7
2005	101.6	10493.0	10327.8	101.6	3168.0	3118.1
2006	101.5	11759.5	11585.7	101.5	3367.0	3317.2
2007	104.5	13785.8	13192.2	103.9	3864.0	3719
2008	105.6	15780.8	14943.9	108.9	3800.0	3489.4
2009	99.1	17174.7	17330.6	97.6	4681.0	4796.1
2010	103.2	19109.4	18516.9	103.6	5032.0	4857.1
2011	105.3	21809.8	20712.0	106.6	5357.0	5025.3
2012	102.7	24564.7	23918.9	101.1	5791.0	5728
2013	102.6	26955.1	26272.0	100.2	6237.0	6224.6

Data sources: China statistical yearbooks (2000-2014)

Table 2: Descriptive statistics of the variables logY and logX

Variables	Mean	Median	Maximum	Minimum	Standard deviation	Skewness	Kurtosis
logY	8.113056	8.106876	8.736264	7.631044	0.389923	0.1829	1.607179
logX	9.389879	9.357527	10.17626	8.687965	0.494913	0.109249	1.711552

transformed into logarithms so as to eliminate the possible hetero-scedasticities. The statistical characteristics of the variables logY and logX were showed in Fig. 2 (auto-correlogram plots) and Table 2 (descriptive statistics).

Estimation of the econometric models

The estimated VAR model: The VAR model has a few important merits, such as the application of Granger causality test. Wheaton (1990) suggested that capital turnover and house prices were jointly determined by reducing sales time and increasing seller reservations, which predicts that capital turnover Granger causes house price changes, while Stein (1995) suggested that house price changes should Granger cause capital turnover due to equity constraints and loss aversion of households. However, all the above theoretical insights have not been tested using large cross-sectional time series data. The VAR model can allow us to directly test the Granger causality between commercial housing prices and urban resident incomes. Moreover, this approach also enables us to better understand the determination of commercial housing prices and their fluctuations. By building the bivariate VAR model, it was assumed that the equilibrium level of commercial housing prices is a function of the variables of urban resident incomes, exogenous variables and lagged endogenous variables. After verifying that the logarithms of the studied variables are cointegrated, these variables were computed and estimated with the VAR and LSE model. Actually, the theoretical basis of cointegration test is also combined with the dynamic autoregressive distributed lag model. In order to effectively model the long-term equilibrium relationship between the two variable series (logX, logY), the following VAR models were established and are represented in Eq. 5, 6, 7 and 8:

$$\text{LogX} = C(1,1)\times \text{LogX}(-1)+C(1,2)\times \text{LogX}(-2)+C(1,3)\times \text{LogY}(-1)+C(1,4)\times \text{LogY}(-2)+C(1,5) \quad (5)$$

$$\text{LogY} = C(2,1)\times \text{LogX}(-1)+C(2,2)\times \text{LogX}(-2)+C(2,3)\times \text{LogY}(-1)+C(2,4)\times \text{LogY}(-2)+C(2,5) \quad (6)$$

In fact, the Bayesian VAR Estimates were modeled and computed too (data not showed). After solving the equation above, an unbiased estimation of coefficients of the VAR matrix for the variable series (logX, logY) was obtained. Finally, the estimated VAR models (with substituted coefficients) could be represented as:

$$\text{LogX} = 0.52910375127\times \text{LogX}(-1)+2.87757778188\times \text{LogX}(-2)-2.93593645116\times \text{LogY}(-1)+0.616511301308\times \text{LogY}(-2)-0.354958226303 \quad (7)$$

$$\text{LogY} = 1.0080209995\times \text{LogX}(-1)-0.840035003456\times \text{LogX}(-2)+0.99197900428\times \text{LogY}(-1)-0.159965000402\times \text{LogY}(-2)+3.61914420259e\text{-}10 \quad (8)$$

In Eq. 7 and 8, the statistics R^2, Adjusted R^2, Log likelihood statistics and F statistics for the variables (logX, logY) were estimated as (0.9973, 1), (0.9960, 1), (31.07094, 311.8643) and (739.5855, 4.35E+21) respectively. The log likelihood statistic for the VAR matrix was estimated as 342.9353

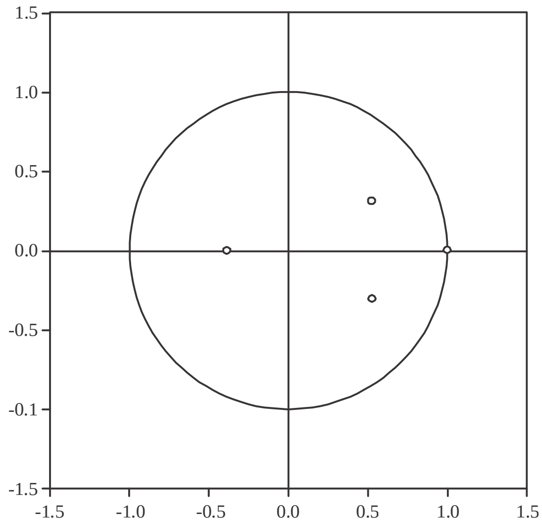


Fig. 3: Inverse roots of AR characteristic polynomial for log VAR model, The Fig. revealed these inverse roots (four points in blue) fell in a reasonable scope of AR (Vector autoregressive) values. The parameters on the x-axis and y-axis of Fig. 3 showed the AR characteristic polynomial for log VAR model

Table 3: Vector autoregression estimates for the variables (logX, logY). Vector autoregression estimates (sample: 1999-2013) Included observations: 13 after adjustments, standard errors in (?) and t-statistics in [?]

Variables	LOGX	LOGY
LOGX(-1)	0.529104	1.008021
	(0.40755)	(1.7E-10)
	[1.29826]	[5.9e+09]
LOGX(-2)	2.877578	-0.840035
	(2.95620)	(1.2E-09)
	[0.97341]	[-6.8e+08]
LOGY(-1)	-2.935936	0.991979
	(2.91440)	(1.2E-09)
	[-1.00739]	[8.2e+08]
LOGY(-2)	0.616511	-0.159965
	(0.43174)	(1.8E-10)
	[1.42797]	[-8.9e+08]
C	-0.354958	3.62E-10
	(0.35545)	(1.5E-10)
	[-0.99863]	[2.44552]
R-squared	0.997303	1.000000
Adj. R-squared	0.995955	1.000000
Sum sq. resids	0.006390	1.11E-21
S.E. equation	0.028261	1.18E-11
F-statistic	739.5855	4.35E+21
Log likelihood	31.07094	311.8643
Akaike AIC	-4.010913	-47.20990
Schwarz SC	-3.793625	-46.99261
Mean dependent	9.494082	9.483027
S.D. dependent	0.444339	0.447956
Determinant resid covariance (dof adj)		1.11E-25
Determinant resid covariance		4.19E-26
Log likelihood		342.9353
Akaike information criterion		-51.22081
Schwarz criterion		-50.78623

(very significant). The inverse roots of AR characteristic polynomial was showed (Fig. 3). Detailed results and data were also input into Table 3 with standard errors and t-statistics.

The estimated LSE model: For the sake of making comparison with the testing result of VAR model, the LSE model was used in the regression of the variables logY and logX too. The estimated LSE model was expressed as follows:

$$\log Y = 0.0772+0.99111 \log X \tag{9}$$

where the statistics for the variable 1ogX were computed as following:

t-statistic = 60.23699
Probability = 0.0000
R^2 = 0.9964
SE = 3628.495

For another, the statistics for the constant α were simultaneously computed too. From the estimated LSE model, it could be deduced that when the disposable incomes of urban residents increase by 1%, the commercial housing prices would rise by 77. 86% on average. This statistics showed that the impact of resident incomes on housing prices is positive (0.0772) and significant. This case agreed well with that of the estimation of VAR model. However, it could only account that the disposable incomes of urban residents would determine the commercial housing prices theoretically. Whether the evidences from realistic society and the real estate market are in favour of the estimated VAR and LSE models above, is depend on the subsequent analyses, such as the unit root tests, Johansen cointegration tests and Granger Causality tests.

Stationary tests

The result of unit root tests: In order to avoid "false return" and test the order of integration of the variables, the standard unit root test was used, namely the Augmented Dickey-Fuller (ADF) test proposed by Dickey and Fuller (1979). The selection of test methods used from general to specific, determine the optimal lag order number using the standards of Akaike info criterion, Schwarz criterion, Hannan-Quinn criterion and Durbin-Watson statistics. Firstly, the Stationary of the variables (logX, logY) was tested and then the 1st and 2nd differential variables (e.g., ΔlogX and ΔΔlogX) for both logX and logY. Table 4 presents unit root test statistics for all these data series. Results are reported for the augmented Dickey and Fuller (1979) tests in Table 4. The results of the ADF unit root test are reported in Table 4 indicating that the model must be estimated in differentiation levels of the variables. The Dickey-Fuller tests are consistent with the

Table 4: The resulted data of unit root tests for log VAR model

Variables	t-Statistic for ADF test	t-Statistic for thresholds 10 (%)	5 (%)	1 (%)	p value (probability)	Durbin-Watson statistic	Remark
logX	1.239747	-2.690439	-3.098896	-4.004425	0.9964	2.086869	nonstationary
ΔlogX	-4.143489**	-2.701103	-3.119910	-4.057910	0.0086	2.298400	stationary
ΔΔlogX	-5.823570**	-2.728985	-3.175352	-4.200056	0.0010	2.285862	stationary
logY	1.822134	-2.690439	-3.098896	-4.004425	0.9991	2.075683	nonstationary
ΔlogY	-2.649955	-2.701103	-3.119910	-4.057910	0.1086	1.671400	nonstationary
ΔΔlogY	-4.643059**	-2.728985	-3.175352	-4.200056	0.0051	2.155758	stationary

Marks * and ** for relevant variables meant significant at the levels of 5 and 1%, respectively. The lag order p values were determined in accordance with the AIC rules hypothesis of a unit root type of non-Stationary for the transformed data. Although it was indicated that the variables (logX and logY) themselves were non-stationary in Table 4, their first-order or second-order differences were stationary and the previous analyses of the VAR and LSE model were credible and applicable.

The result of Johansen cointegration test for the VAR model: The dominant approach in testing for cointegration in Gaussian vector autoregressions (VAR) is Johansen' s maximum likelihood approach. EViews 8.0 supports VAR-based cointegration tests using the methodology developed by Johansen (1991, 1995). The Johansen cointegration test was initially discussed in Johansen (1991, 1995) system maximum likelihood approach to cointegration analysis and testing, which is supported using VAR objects. The main idea of Johansen cointegration test was to test the hypothesis whether there existed long-term regression or possible correlation relationships between two time series variables. In brief, if each element of a vector of time series X_t (vector or matrix) first achieves Stationary after differencing but a linear combination $\alpha' X_t$ is already stationary and the time series X_t are said to be cointegrated with cointegrating vector α. There may be several such cointegrating vectors (α). Interpreting $\alpha' X_t = 0$ as a long run equilibrium, cointegration implies that deviations from equilibrium are stationary, with finite variance, even though the time series themselves are non-stationary and have infinite variance. Table 5 shows the result of Johansen cointegration test based on VAR model. It could be inferred that logY and logX were non-stationary but their 1st differential variables were stationary with a probability of 0.0001 and a Log likelihood value of 295.5561 in the cointegrating equation. The resulted data and statistics of Johansen

Table 5: The resulted data of Johansen cointegration test for log VAR model Johansen cointegration test (sample: 1999-2013) included observations: 13 after adjustments trend assumption: Linear deterministic trend Series: LOGX LOGY Lags interval (in first differences): 1-1 unrestricted cointegration rank test (trace)

Hypothesized No. of CE (s)	Eigenvalue	Trace statistic	0.05 Critical Value	Probability**
None *	1.000000	400.5934	15.49471	0.0001
At most 1	0.161511	2.290001	3.841466	0.1302
Trace test indicates 1 cointegrating eqn(s) at the 0.05 level Unrestricted Cointegration Rank Test (Maximum Eigenvalue)				
Hypothesized No. of CE(s)	Eigenvalue	Max-eigen statistic	0.05 Critical Value	Probability**
None *	1.000000	398.3034	14.26460	0.0001
At most 1	0.161511	2.290001	3.841466	0.1302
Max-eigenvalue test indicates 1 cointegrating eqn (s) at the 0.05 level unrestricted cointegrating coefficients (normalized by b: *S11*b: I)				
LOGX				LOGY
115.4441				-115.4441
319.4218				-310.4542
Unrestricted Adjustment Coefficients (alpha):				
D(LOGX)		-0.006075		0.009730
D(LOGY)		0.001455		2.22E-16
1 Cointegrating Eq.):		Log likelihood		295.5561
Normalized cointegrating coefficients (standard error in parentheses)				
LOGX				LOGY
1.000000				-1.000000
				(NA)
Adjustment coefficients (standard error in parentheses)				
D(LOGX)				-0.701345
				(0.93169)
D(LOGY)				0.167986
				(NA)

The mark * denotes rejections of the hypothesis at the 0.05 levels, while the significances denoted by the mark ** were measured with one-sided p-values of Mackinnon (Ericsson and MacKinnon, 2002) cointegration test showed that there were stable cointegration relationships between both the unrestrictedly adjusted variables (logX, logY) and their 1st differential variables (D (logY), D (logY)). With the long-term and stable cointegration relationships, the linear combination of these two variables and the cointegrating coefficients were computed or deduced too.

Pairwise granger causality tests: The result of pairwise Granger causality tests are reported in Table 6. The cointegration relationship between two variables and diverse Stationary tests could not fully explain and prove the dynamic correlations in both short-term and long-term between the tested variables. Therefore, pairwise Granger causality tests are essential to determine whether there are Granger causes between the tested variables in a long-term or short-term period (Engle and Granger, 1987). For the Engle-Granger statistic, it could be found that these P values were not significant for the null hypothesis "LogY does not Granger Cause logX", while those were very significant for the null hypothesis "LogX does not Granger Cause logY" (Table 6). Thus, the former could not be rejected and the latter should be rejected, i.e. the hypothesis “LogX might Granger Cause logY” is acceptable and reasonable. With the result of pairwise Granger causality test, it can be concluded that the disposable incomes of urban residents is probably one of the crucial impact factors for Chinese commercial housing prices.

The predicted result of Impulse response function: The impulse-response functions were used to provide a more intuitive description of how shocks in exogenous variables generate the co-movements of the commercial housing prices and urban resident incomes in the market. The relationships between the commercial housing prices and the disposable incomes of urban residents are simulated and computed by the impulse response function of 15 years prediction (Fig. 4). It was shown that the impact of urban resident incomes on commercial housing prices was obviously in a positive trend in the long run (Fig. 4a) but the response of urban resident incomes to the impulse

Table 6: Result of Granger causality tests for log VAR model

Lags for the variables tested	Null hypothesis	df	F-Statistic	p value (probability)
1	LogY does not Granger Cause logX	14	0.32393	0.5807
1	LogX does not Granger Cause logY	14	558.681	9.E-11
2	LogY does not Granger Cause logX	13	1.03026	0.3998
2	LogX does not Granger Cause logY	13	1.4E+20	7.E-79

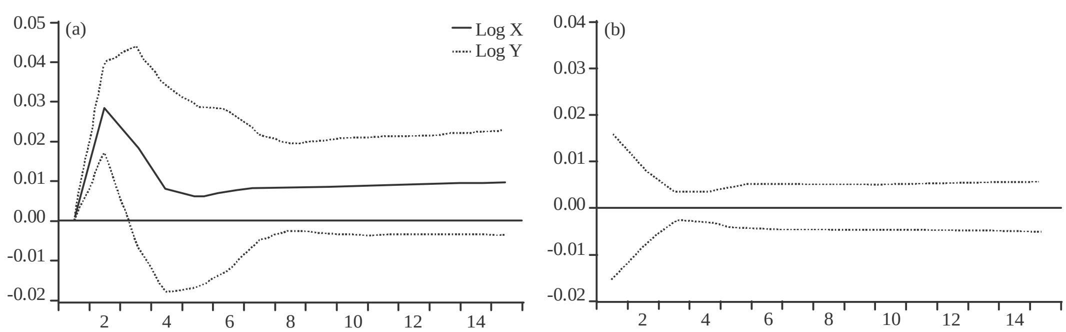


Fig. 4(a-b): Responses of commercial housing prices and urban resident incomes to exogenous shocks. The parameters (logX, logY) on the x-axis and y-axis of Fig. 4a and 4b showed the responses of logY (commercial housing prices) to the exogenous shock from logX (urban resident incomes) in Fig. 4a and the responses of logX (urban resident incomes) to the exogenous shock from logY (commercial housing prices) in Fig. 4b impact of commercial housing prices was non-active (Fig. 4b). During a predicted period of 15 years by impulse response function, in response to the fluctuations of urban resident incomes, commercial housing prices will be rising. Generally, commercial housing prices will be promoted rapidly from the first 1-2.5 years and then come into a process of decline from the next 3-8 year. Finally, commercial house prices will rise slowly and fall since the 9th year. On the contrary, the impact of commercial house prices on urban resident incomes is very limited and n response to the fluctuations of commercial housing prices, the corresponding response of urban resident incomes will be low too (Fig. 4).

Variance decomposition for the VAR model: While the impulse response function traces the effect of each shock to one endogenous variable on to the other variable in the VAR model, variance decomposition separates the variation in an endogenous variable into the component shocks based on the VAR model. Thus, the variance decomposition was obtained which provides new information about the relative importance of each random innovation in affecting the variables. The dynamic characteristics of the VAR model could be well studied with variance decomposition.

With the evidences from realistic society and the results of the estimated VAR and LSE models and subsequent econometric analyses and prediction of Impulse response function above, it can be inferred that the disposable incomes of urban residents is one of the positive determinants of the housing demand and commercial housing prices. This conclusion agrees well with previous studies, although there are presently a few reports of commercial housing price fluctuations relevant to the present work in China. Some academic literatures studied the determination of nominal housing price and found that disposable incomes or monetary factors were really at work (Shen, 2012; Ren *et al*., 2012; Zhang *et al*., 2012; Zhang, 2013). For instance, Ren *et al*. (2012) found that the economic theory of demand and supply could explain the relative price of urban housing in China satisfactorily. Shen (2012), Zhang *et al*. (2012) and Zhang (2013) also found that housing price was relatively reasonable because the affordability in China was high due to higher growth rate and low interest rates and disposable incomes of urban residents was vital.

Furthermore, the social implication and practical aspects are developed as following. In recent years, the realistic incomes of urban residents are significantly increased and urban living conditions are relatively improved in China. It also raises or promotes the demand of urban residents for commercial housing, while the urban residents can afford modern commercial houses. According to the market mechanism of demand and supply for commercial housing, the different aims of the commercial house buyers can be inferred and analyzed as following. As for the residential demand of urban residents, the increasing incomes of urban residents' income enhanced greatly their desires and raised their purchasing power for commodity, houses and other social products. On the other hand, the gradually springing up housing products of green houses, health houses and ecological residential houses has greatly improved the quality and level of commercial housing consumption, which also accelerated strongly the urban residents' purchasing and consumpting action of commercial houses. Moreover, the urban displacement and resettlement and other relevant urbanization actions also improve and increase the urban residents' demands for the commercial houses and other real estate products in China. In return, the raised urban residents' demands for real estate products will promote the continuous and steady development of social economy. After analyzing the demand and supply of commercial housing in market, it is also found that there is a dynamic interaction between the urban resident incomes and commercial housing prices. With the improvement of the urban residents' disposable incomes and the augment of private capital flow, private investment and funds become increased and strong. A lot of private capital is to be invested and many investment channels are high-risk. Therefore, commercial housing and other real estate products have become the first choice of investment because of relatively low risks. Furthermore, recent stock market crash and inflation made urban residents put more money into the real estate industry, which increased the demand for commercial housing and greatly promoted the rise of Chinese urban commercial housing prices. In summary, the phenomena of fluctuations of commercial housing prices should be observed and thought of reasoning in view of the market mechanism of demand and supply, urban housing issues, residential requirements and their risk awareness to constantly optimize and improve the market regulatory relationship and mechanism of demand and supply, relevant policies and management regimes.

CONCLUSION

In this study, the market economic theory and mechanism and econometric models were applied to explain the demand for and the supply of urban residential housing in China. The inclusion of urban resident incomes in the commercial housing market could be motivated by the market economic theory and regulatory mechanism of demand and supply. The following conclusion can be inferred that the disposable incomes of urban residents are positively one of the impact factors on the increasing housing demand and commercial housing prices. Specifically, it was found that the positive correlation could almost completely be explained by both the empirical analysis using VAR and LSE models. Furthermore, strong evidence was found that the urban resident incomes Granger cause the commercial housing prices from a series of econometric tests in the present study.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous referees for constructive suggestions. This work is supported by National Statistical Scientific Foundation (No. 2013LY051), Anhui Provincial Natural Soft Science Research Project (No. 1502052051) Anhui Provincial Educational Commission's Key Projects (No. SK2014A069, No. 2014SK006, No. 2014SK020), Fuyang City Social Scientific Project (FSK201413) and Fuyang Normal University Educational Study Project (No. 2013 JYXM24).

REFERENCES

Basu, S. and T.G. Thibodeau, 1998. Analysis of spatial autocorrelation in house prices. J. Real Estate Finance Econ., 17: 61-85.

Clayton, J., 1998. Further evidence on real estate market efficiency. J. Real Estate Res., 15: 41-57.

Dickey, D.A. and W.A. Fuller, 1979. Distribution of the estimators for autoregressive time series with a unit root. J. Am. Stat. Assoc., 74: 427-431.

Engle, R.F. and C.W.J. Granger, 1987. Co-integration and error correction: Representation, estimation and testing. Econometrica, 55: 251-276.

Ericsson, N.R. and J.G. MacKinnon, 2002. Distributions of error correction tests for cointegration. Econometrics J., 5: 285-318.

Gao, T.M., 2006. Econometric Analysis and Modeling: Reviews Applications and Examples. Tsinghua University Press, Beijing, China.

Gatzlaff, D.H. and D. Tirtiroglu, 1995. Real estate market efficiency: Issues and evidence. J. Real Estate Literature, 3: 157-189.

Gau, G.W., 1984. Weak form tests of the efficiency of real estate investment markets. Financial Rev., 19: 301-320.

Gau, G.W., 1985. Public Information and abnormal returns in Real Estate investment. Real Estate Econ., 13: 15-31.

Hamilton, J.D., 1994. Time Series Analysis. Princeton University Press, Princeton, USA., ISBN: 9780691042893, Pages: 799.

Hosios, A.J. and J.E. Pesando, 1991. Measuring prices in resale housing markets in Canada: Evidence and implications. J. Housing Econ., 1: 303-317.

Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. Econometrica, 59: 1551-1580.

Johansen, S., 1995. Likelihood-based Inference in Cointegrated Vector Autoregressive Models. Oxford University Press, New York, ISBN: 9780198774501, Pages: 267.

LeSage, J.P. and R.K. Pace, 2004. Introduction. In: Spatial and Spatiotemporal Econometrics: Advances in Econometrics, LeSage, J.P. and R.K. Pace (Eds.). 1st Edn., Elsevier, Oxford, ISBN-13: 978-0762311484, pp: 1-32.

Liu, X., 2013. Spatial and temporal dependence in house price prediction. J. Real Estate Finance Econ., 47: 341-369.

Pace, R.K., R. Barry, J.M. Clapp and M. Rodriquez, 1998. Spatiotemporal autoregressive models of neighborhood effects. J. Real Estate Finance Econ., 17: 15-33.

Ren, Y., C. Xiong and Y. Yuan, 2012. House price bubbles in China. China Econ. Rev., 23: 786-800.

Rosen, S., 1974. Hedonic prices and implicit markets: Product differentiation in pure competition. J. Polit. Econ., 82: 34-55.

Schindler, F., 2014. Persistence and predictability in UK house price movements. J. Real Estate Finance Econ., 48: 132-163.

Shen, L., 2012. Are house prices too high in China? China Econ. Rev., 23: 1206-1210.

Stein, J.C., 1995. Prices and trading volume in the housing market: A model with down-payment effects. Q. J. Econ., 110: 379-406.

Wheaton, W.C., 1990. Vacancy, search and prices in a housing market matching model. J. Polit. Econ., 98: 1270-1292.

Zhang, C., 2013. Money, housing and inflation in China. J. Policy Modell., 35: 75-87.

Zhang, Y., X. Hua and L. Zhao, 2012. Exploring determinants of housing prices: A case study of Chinese experience in 1999-2010. Econ. Modell., 29: 2349-2361.