Chapter 3

Methodology

There are two models in our research. The theoretical model that is used to explain the relationship between house prices and economic growths is based on a purely stock-flow model (Wang Songtao, 2011) which was used to illustrate the common components of fluctuations in house prices. In this step we avoid making too many priori assumptions on the drivers of common fluctuations, that is, we let the common factors be latent instead of pre-specifying a number of regressors. Once an estimate of the common factors have been obtained from the statistical model, in the next step, we investigate what lies behind it, and in particular, focus on the role of economic growths. Thus, in the second step, some econometric models will be employed to study the relationship between the house price and economic growth. The theoretical and econometric models are complemented with each other to finish this study. Before going to details of the theoretical model, we introduce the data source first.

3.1 Data

This study covers annual data of house prices and economic growths in 30 provinces in People's Republic of China from 1998 to 2009. The economic growths and house prices are changed in to real terms using GDP deflator and consumer price index. All variables are in natural logarithms as LnGDP and LnHP. The logarithm form of the data can give more information. It shows when a variable change one percent, how does the other variable change.

The statistics of LnGDP and LnHP in People's Republic of China from 1998 to 2009 are shown as table 3.1. It can be seen that, there are respectively 360 observations of LnHP and LnGDP in this study. The mean value of LnGDP is 8.261517. The Median value of LnGDP is 8.344167. The maximum and minimum values of LnGDP are 10.59764 and 5.398718 respectively. And the standard deviation of LnGDP is 1.010522.

On the other hand, the mean value of LnHP is 7.697726. The median value of LnHP is 7.606432. The maximum and minimum values of LnHP are 9.599532 and 6.622061 respectively. Moreover, the standard deviation of LnHP is 0.498122.

Table 3.1 Basic statistics of LnGDP and LnHP in China, 1998-2009

	Observation	Mean	Median	Maximum	Minimum	Std. Dev.
LnGDP	360	8.261517	8.344167	10.59764	5.398718	1.010522
LnHP	360	7.697726	7.606432	9.599532	6.622061	0.498122

Source: China Statistic Bureau

Figure 3.1 shows provinces and regions of People's Republic of China. It can be seen from figure 3.1 that, the People's Republic of China is mainly divided into three parts, which are the eastern region, the central region and the western region. The most developed part is the eastern region of People's Republic of China. The major big cities, such as Beijing, Shanghai and Shenzhen are in that region. The western region is the most undeveloped part of People's Republic of China, which includes Yunnan, Qinghai, Ningxia, Xinjiang and other provinces.

Sou

rce: National Administration of Surveying, mapping and Geoinformation

Figure 3.1 China's provinces and the three regions

3.2 Theoretical model

A real estate market is often modeled with the stock-flow framework. The basic theoretical model for this study is the Wheaton (1998) stock-inflow model, which is expressed as follows:

The fundamental equations of stock-inflow housing model are as:

$$H_t^S = H_t^D\left(P_t, R_t, U_t, X_{1t}\right) \tag{3.1}$$

$$\Delta H_t^S = H_t^S - H_{t-1}^S = C_t\left(P_t, X_{2t}\right) - \delta H_{t-1}^S \tag{3.2}$$

where H_t^D is housing demand function at time t. P_t is house prices at time t, R_t is rental prices at time t, U_t is the ownership costs for housing at time t. X_{1t} are other macroeconomic and demographic variables which affect the demand for housing at time t, such as economic output (GDP_t), household permanent income (INC_t), the total population (POP_t), family structure (PS_t) and so on.

Equation (3.2) is the housing supply function. where H_t^S is the housing stock in period t. δ is the housing stock depreciation rate. C_t is the new supply of housing at time t. The main factors which influent the increment of new supply at time t are the house price P_t and other exogenous factors X_{2t} at time t, such as the construction cost (MC_t), labor costs (LC_t), various short-and long-term interest rates ($i_{s,t}$ or $i_{m,t}$), the housing vacancy rate (VAC_t) and so on.

Equation (3.3) gives the housing costs of ownership which contains the marginal income tax rates ($t_{y,\,t}$), property taxes ($t_{p,\,t}$), mortgage interest rates ($i_{m,\,t}$), housing depreciation rates (ζ_t), inflations (π_t), house prices expected (P_t^e), and the credit constraints (λ_t is the credit constraints, μ_c is the unit utility for the housing) . The bigger the U_t, the lower the housing demands.

$$U_t = \left[\left(1 - t_{y,t}\right)\left(i_{m,t} + t_{p,t}\right) - \pi_t + \zeta_t - E\left(P_t^e / P_t - 1\right) + \lambda_t / \mu_t\right] \tag{3.3}$$

Di Pasquale and Wheaton further give some amendments for the basic model: First, the price adjustment is a slow process. Second, the price is expected to be myopic or rational expectations. Third, the new housing supply is constrained by the long-term housing stock level. It is slow adjusted, as the follow equations (3.4) to (3.6):

$$\Delta P_t = P_t - P_{t-1} = \gamma\left(P_t^* - P_{t-1}\right) \tag{3.4}$$

$$P_t^e = E\left[P_{t+1} \middle| I_t\right] \text{ or } P_t^e = \sum_{i=1}^{P} w_i P_{t-i} \tag{3.5}$$

$$\Delta H_t^S = C_t\left(P_t, X_{2t}\right) - \delta H_{t-1}^S = \tau\left[H_t^{S*}\left(P_t, X_{2t}\right) - H_{t-1}^S\right] - \delta H_{t-1}^S \tag{3.6}$$

In equation (3.4), γ represents the speed of actual price converge to the equilibrium price. Equation (3.6) shows that the supply of new housing stock is decided by the long-term housing stock H_t^{S*}. The relationship of H_t^{S*} and P_t is determined by the urban growth theory. The actual house price equation (3.7) can be achieved from (3.1) to (3.6) as follows:

$$P_t = P_t^* + \varepsilon_t \tag{3.7}$$

$$= P^*\left(R_t\left(t_{y,t}, t_{p,t}, i_{m,t}, \pi_t, \zeta_t, P_t^e, \lambda_t\right), \left(GDP_t, INC_t, POP_t, PS_t\right), \ldots\left(MC_t, LC_t, i_{s,t}, VAC_t\right), \delta_t, H_{t-1}^S, \ldots\right) + \varepsilon_t$$

From equation (3.7), it can be seen that the house prices, to some extent, can be affected by a variety of common factors, including the economic growths. Then, in the next step, this paper will look solely on the relationship between house prices and economic growth. This housing stock-flow model provides the theoretical background for finding the empirical relationship between house prices and economic growths in the People's Republic of China using econometric methods. The details of the econometric method are shown as the following part.

3.3 Econometric models

The methodology employed in this study has three steps: First, a panel unit root test for checking the stationary of data will be used. Second, the Pedroni (2000, 2004) test will be applied to examine the panel cointegration relationship. At last, the error correction models will be built to assess the short run and long run Granger causality between house prices and economic growths.

3.3.1 Panel unit root test

Before testing the cointegrating relationship between house prices and economic growths, some time series properties of the panel data need to be examined. Two panel unit root tests: Levin-Lin-Chu, or LLC, and Im-Pesaran-Skin (2003), or IPS, are performed to check for the presence of stationary.

The structure of the LLC analysis may be specified as follows:

$$\Delta Ln(GDP)_{i,t} = \rho_i Ln(GDP)_{i,t-1} + \alpha_{0i} + \alpha_{1i}t + \mu_{i,t} \quad (3.8)$$

$$\Delta Ln(HP)_{i,t} = \rho_i Ln(HP)_{i,t-1} + \alpha_{0i} + \alpha_{1i}t + \nu_{i,t} \quad (3.9)$$

where $Ln(GDP)_{i,t}$ and $Ln(HP)_{i,t}$ are the natural logarithms of house prices and gross domestic products, for each province *i* at time *t* respectively. ρ_i is the coefficient of one period lagged variables. $\mu_{i,t}$ and $\nu_{i,t}$ are the residuals. $\Delta Ln(GDP)_{i,t}$ and $\Delta Ln(HP)_{i,t}$ are the first difference of natural logs of $HP_{i,t}$ and $GDP_{i,t}$. It can be seen that the LLC structure has an intercept, α_{oi}, and a trend, $\alpha_{li}t$, in the equations. The hypotheses of LLC test are:

H_0: $\rho_i = \rho = 0$ the null hypothesis is that the data has a unit root. (3.10)

H_a: $\rho_i = \rho < 0$ (3.11)

If the null hypothesis is accepted, it means house prices and economic growths are nonstationary. If the null hypothesis is rejected, it means house prices and economic growths are staionary.

Im-Pesaran-Skin (2003) extends the Levin and Lin framework to allow for heterogeneity in the value of the autoregressive coefficient under the alternative hypothesis. The structure of the IPS analysis may be specified as follows:

$$\Delta Ln(GDP)_{i,t} = \rho_i Ln(GDP)_{i,t-1} + \sum_{L=1}^{p_i} \phi_{iL} \Delta Ln(GDP)_{i,t-L} + z'_{i,t}\gamma + u_{i,t} \quad (3.12)$$

$$\Delta Ln(HP)_{i,t} = \rho_i Ln(HP)_{i,t-1} + \sum_{L=1}^{p_i} \phi_{iL} \Delta Ln(HP)_{i,t-L} + z'_{i,t}\gamma + \nu_{i,t} \quad (3.13)$$

The null hypothesis of IPS test can be written as $H_0 = \rho_i = 0 \; \forall i$. While the alternative hypothesis is given by:

$$H_a : \begin{cases} \rho_i = 0, \; for \; i = 1,2,...,N_1 \\ \rho_i < 0, \; for \; i = N_1,...,N \end{cases}$$ where N is the number of observations. (3.14)

The alternative hypothesis of IPS allows some coefficients of individuals equal zero, while others not equal zero. The IPS method permits the heterogeneity of the data. Given the short span of panel data in this study, the more powerful IPS panel test results, are more acceptable and reliable.

3.3.2 Panel cointegration test

Given that the variables appear to be non-stationary, this study will proceed to test whether house prices and economic growths are cointegrated or having a long-run dynamics relationship. Pedroni (2000, 2004) method is used to perform the cointegration test.

The Pedroni method is based on Engle-Granger (1987) two-step approach. In the first step, a generalized least square (GLS) method was used to estimate the cointegration regressions. Then, the residuals obtained from the regression functions will be checked for stationary. If the residuals are stationary, it means the variables are cointegrated at *I(1)*.

For checking the cointegration relationship between house prices and economic growths, the following cointegration regressions can be set as:

$$Ln(GDP)_{i,t} = \alpha_{1i} + \beta_{1l} Ln(HP)_{i,t} + \mu_{i,t} \quad (3.15)$$

$$Ln(HP)_{i,t} = \alpha_{2i} + \beta_{2l} Ln(GDP)_{i,t} + \nu_{i,t} \quad (3.16)$$

where $\mu_{i,t}$ and $\nu_{i,t}$ are the residuals respectively.

Then, the residuals will be checked for stationary by unit root test. If the residuals are stationary, it means that house prices and economic growths are cointegrated at *I(1)*.

3.3.3 Error correction models and Granger causality test

The following panel error correction models for examining the causality between house price and economic growth can be set as:

$$\Delta Ln(GDP)_{i,t} = \alpha_{1j} + \sum_{k=1}^{m} \alpha_{1k} \Delta Ln(GDP)_{i,t-k} + \sum_{l=1}^{n} \beta_{1l} \Delta Ln(HP)_{i,t-l} + \lambda_{1i} \mu_{i,t-1} + \varsigma_{i,t} \quad (3.17)$$

$$\Delta Ln(HP)_{i,t} = \alpha_{2j} + \sum_{k=1}^{p} \alpha_{2k} \Delta Ln(HP)_{i,t-k} + \sum_{l=1}^{q} \beta_{2l} \Delta Ln(GDP)_{i,t-l} + \lambda_{2i} \nu_{i,t-1} + \xi_{i,t} \quad (3.18)$$

where $\mu_{i,t-1}$ and $\nu_{i,t-1}$ are one period lagged values of residuals from cointegration regressions (3.15) and (3.16), respectively. The optimal lag lengths *m, n, p,* and *q* are decided based on the Akaike Information Criteria (AIC).

Equation (3.17) is employed to detect the causation from the house prices to economic growths, whereas equation (3.18) is used to test the causation from the economic growths to house prices. If the estimated coefficients of $\Delta Ln(HP)_{i,t-l}$ in equation (3.17) are significant, then there is evidence that changes in economic growths are Granger caused by changes in house prices. Similarly, the significances of coefficients of $\Delta Ln(GDP)_{i,t-l}$ in equation (3.18) indicate a causality relationship running from the economic growths to house prices.

Importantly, the ECM approach also enables us to analyze the market's dynamic convergence process, as it incorporates the effect of the last period's equilibrium error on the magnitude and direction of subsequent prices changes into its framework through ECTs, the coefficient of error correction term λ_{1i} in equation (3.17), measures the speed of markets adjustments toward its equilibrium and also indicates the efficiency of economy. Similarly, the house market's efficiency is represented by the speed of error correction λ_{2i} in equation (3.18).