

The Interurban House Price Gradient: Evidence from Pan-Yangtze River Delta in China

Yunlong Gong^a, Peter Boelhouwer^a and Jan de Haan^{a,b,*}

^a OTB Research Institute for the Built Environment, Delft University of Technology, the Netherlands

^b Division of Process Development, IT and Methodology, Statistics Netherlands

Abstract: Chinese cities are nowadays closely linked because of sustained regional economic integration. As a result, the housing markets of different cities are also closely linked. The main purpose of this paper is to investigate the interurban house price structure as represented by the interurban house price gradient in a region that is comprised of several hierarchical cities. Our empirical results for the Pan-Yangtze River Delta show that, except for the coastal cities, the average house price in a city declines exponentially with the distance to the central city. Subcentral cities also seem to have positive effects on other cities' house prices. Using a spline function, we are able to accurately measure the interurban house price gradient that shows the influence of central and subcentral cities in a ray-area. However, when controlling for socio-economic factors, we find that the central-city-oriented distance effect vanishes as it is capitalized into the spatial structure of income and mortgage loan size. The latter finding calls for further theoretical work.

Keywords: Interurban, house price gradient, spline function, China

1. Introduction

Whilst there is a large literature on the house price structure within a city (intracity structure) based on the idea of a trade-off between house prices and accessibility, this paper looks into the house price structure across different cities (interurban structure) which is related to the distribution of economic activities in a hierarchical urban system¹.

The famous monocentric model (Alonso, 1964; Mills, 1967, 1972; Muth, 1969) explains the urban spatial structure under the hypothesis that the increase in commuting cost to Central Business District (CBD), which provides all the job opportunities, will be offset by the decrease in house prices. The model, thus, predicts a negative house price gradient with increasing distance to CBD. Many empirical studies were carried out to verify the appropriateness of this model. Most of them indeed found negative population density, house price or land value gradients (at least for part of the sample periods) in a particular city, metropolitan area or a wider regional context with one dominating center, such as Chicago, New York, Berlin, Stockholm, Beijing and the southern part of West Norway with a dominating city (Stavanger) (McMillen, 1996; Colwell and Munneke, 1997; Atack and Margo, 1998; Ahlfeldt, 2011; Söderberg and Janssen, 2001; Zheng and Kahn, 2008; Osland *et al*, 2007).

* Corresponding Author.

E-mail Address: Y.Gong-1@tudelft.nl (Y. Gong), P.J.Boelhouwer@tudelft.nl (P. Boelhouwer), J.deHaan-1@tudelft.nl (J. de Haan)

¹ A hierarchical urban system means a cluster of many cities with different ranks, scales and functions, which are linked by a transportation system and have closely interacted economies, but where most workers tend to work in the same city they live.

Some studies, however, showed insignificant, or even positive, price gradients. As the monocentric city assumption has come under harsh criticism and subcenters, for example in Chicago, and Haifa, have been successfully identified (McMillen and McDonald, 1998; McMillen, 2001; Palut and Palut, 1998), one explanation for this finding is that modern cities are characterized by multiple centers and have more complicated structures. A few studies supported the argument that subcenters indeed play an important role in shaping the urban structure (Heikkilä *et al*, 1989; Sivitanidou, 1996; Qin and Han, 2013). Meanwhile, in a more and more dispersed city context, some scholars tested the relationship between house (land) prices and accessibility to employment measured by the so-called gravity model. The results turned out to (partially) support the positive relationship (Adair, 2000; Osland and Thorsen, 2008, 2013; Giuliano *et al*, 2010; Ahlfeldt, 2011). Noticeably, Ahlfeldt's result (2011) showed that the introduction of gravity accessibility to employment renders the negative CBD land gradient insignificant. An alternative explanation is that access to employment may be less important nowadays when people also value the accessibility to amenities, such as shopping centers, schools, and hospitals. Compelling evidence has been found to support the effects of urban amenities on price gradients (Dubin and Sung, 1987; Heikkilä *et al*, 1989; Waddell *et al*, 1993a)².

Osland *et al* (2007, 2008, 2013) extended the study area to an inter-zone context and argued that the households' decision process of residential location choices involves two spatial levels of aggregation: households firstly determine which parts of the region are in accordance with their housing market preferences, and then choose the specific residential sites within the area of interest. Motivated by this study, we argue that households are also willing to choose the city they live and work in depending on the interurban amenities and socio-economic situation in an urban system context. Thus, we think it is meaningful to study the interurban house price structure as it reflects the households' valuation of the specific city's location.

One stream of studies on the interurban or interregional housing market focussed on the determinants of the differences in interurban (regional) aggregated house prices, based on the assumption of equilibrium of housing demand and supply (Witte, 1975; Ozanne and Thibodeau, 1983; Manning, 1988; Potepan, 1996). Another stream of studies is related to the relationship between different cities' (regions') house price dynamics, like the ripple effect or diffusion effect (Cook, 2003; Jones and Leishman, 2006; Canarella *et al*, 2012). All these above interurban (regional) studies used classical economic models and treated the cities as non-spatial objectives. In other words, they did not pay attention to the effect of "location" on house prices. We should note though that a spatial-temporal model that considers both spatial and temporal dimensions to study the regional house price dynamics has been developed (Holly *et al*, 2011). Taking into account that the average house prices in central or nearby cities are usually higher than the prices in distant cities in the hierarchical urban system, this study focuses on the static structure of interurban house prices rather than the dynamic behaviour.

² For an extensive review of the theoretical and empirical studies on urban spatial structure before 1998, see Anas, Arnott and Small (1998).

The basic hypothesis in this study is that there will be a negative interurban house price gradient in a hierarchical urban system, i.e., that the average house price in a city declines with the distance to the central city, similar to the negative price gradient found in intracity studies. To our knowledge, there is no literature focusing on the house price gradient in the interurban context, so our results will contribute to the literature on the spatial structure of housing market. Our empirical work utilizes data for the Pan-Yangtze River Delta in mainland China (see fig.2). The nominal average house price in Shanghai (SH), the central city in this area, was 14400 *yuan/m²* in 2010 whereas the prices were 11854 *yuan/m²* and 6431 *yuan/m²* in Nanjing (NJ) and Hefei (HF), the capital cities of the Jiangsu and Anhui provinces, respectively³. This is a crude indication that a negative price gradient indeed exists.

After exploring the locational and socio-economic determinants of interurban house prices via an improved derived demand approach, we test the above hypothesis. First, we test three distance-price functions and choose the most appropriate one to estimate the central-city-oriented interurban house price gradient under the belief that the effects of socio-economic factors, such as income, mortgage and population, have been held constant. In this context, the influences of subcentral cities on interurban house prices structure are also discussed, and a spline function, which is expected to have a better performance, is used to fit the house price gradient of a ray-area. Then, we turn to observe how the interurban house price gradient changes after controlling the effects of socio-economic variables.

The paper is organized as follows. Section 2 gives a brief overview of the Chinese administrative division system and the housing market in urban China. In section 3, a theoretical model is discussed to seek the determinants of house price differences across cities. Section 4 introduces the study area, data and variables used in our empirical study. Section 5 presents the results. Section 6 concludes and provides some discussion.

2. The Housing Market in Urban China

Before giving an introduction into the interurban housing market in mainland China, we will provide some background information about the administrative division system in China and the corresponding urban areas. The hierarchy of the administrative division system of China consists of five different levels: central government; province; prefecture cities; county; town⁴. In this study, we focus on the housing market in the level of prefecture cities or municipalities under the central government. A prefecture city (municipality) usually consists of several sub-divisions, including city districts, county-level cities or counties. For an intuitive understanding, see figure 1. These sub-divisions are all comprised by a high densely populated urban area and surrounding rural area. By a prefecture city's housing market we mean the market associated with the urban area of the city districts. The interurban housing market refers to the integration of different cities' housing markets.

³ The average house prices come from the dataset used in this study. For a detailed calculation procedure, see section 3.

⁴ In practice, we also have other names for each level of government. We have *autonomous regions*, *municipalities directly under the central government* and *special administrative region* in province level; *autonomous prefecture* and *league* in prefecture city level; *county-level city* in county level.

The housing reform launched in 1998 was a landmark in the history of Chinese housing reform; a mature and modern housing market is now being formed in China⁵. Three types of housing are provided to meet the housing demand of different income groups: commercial housing, government-supported affordable housing (*Jingji shiyong fang*) and government-subsidized rental housing (*Lianzu fang*) (Wang *et al*, 2012). The commercial housing market is market-oriented and provides the majority of housing. In a typical housing project development cycle, housing developers first purchase the right to use residential land parcels from local governments through public bidding or auctions⁶, and then dwellings, like apartments in most western countries, are built and sold to the households. Government-supported affordable housing, aiming to cater for low to middle-income households, enjoys privilege in obtaining the right of use of state-owned urban land and the developers' profits are limited, so this type of housing can be sold at relatively low prices. Government-subsidized social rental housing, which is designed for very low-income households, is still characterized as 'welfare-oriented'. Though affordable housing in particular has been encouraged and supported by governments in recent years, the last two types of housing still remain relatively unimportant.

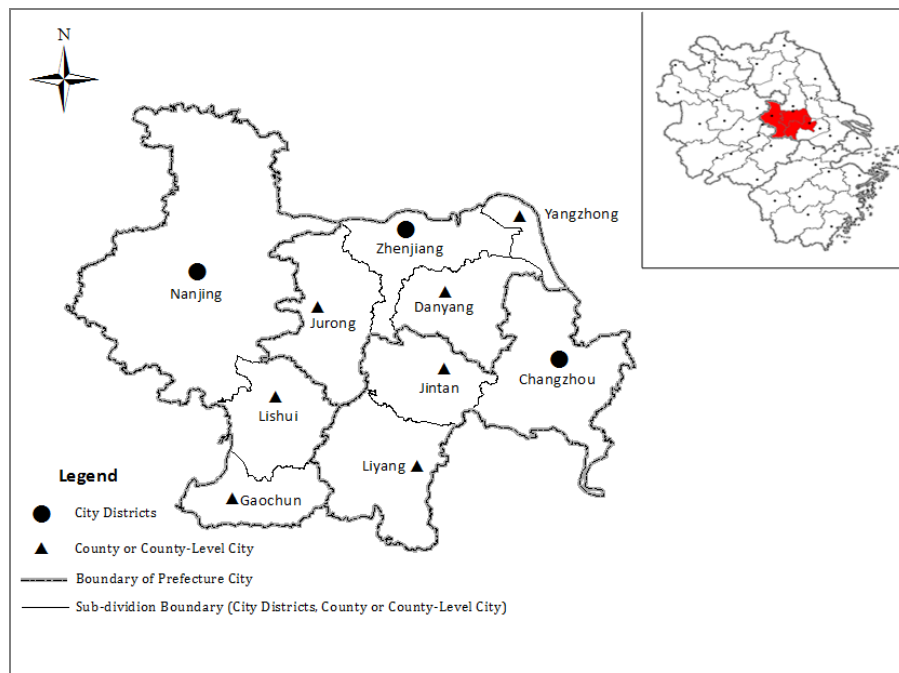


Fig.1 The sub-division of prefecture cities

In most western countries it is quite common for workers to live in one city and work in another. In China, on the other hand, most workers still live and work in the same city because of a

⁵For a longitudinal analysis of the housing market reform before 1998, refer to Huang (2004); for a housing system revolution after 1998, see Wang *et al* (2012).

⁶According to the land system in China, the State retains the ownership of urban land and the local governments are authorized to manage the use of those land. Land users can acquire the land-use right for a fixed period by bidding or auction, usually 70 years for residential uses, 50 years for industrial uses, and 40 years for commercial uses. For detailed introduction of urban land supply system and land market in China, see Wu *et al* (2012).

number of reasons, including the traditional culture, the long distance between two cities, and the high traveling costs.

3. Theoretical Approach

We will make the following three assumptions. First, a city is assumed to be comprised by both the highly dense urban area, providing the working and residential places, and the surrounding rural area. Second, the decision to purchase a home is based on maximizing the individuals' utility derived from the expenditures on housing and other goods and services. Third, capital and workers can freely travel across cities to allow individuals to choose the location of their home site both within a particular urban area and between interurban areas. However, the situation of living in one city and working in another city is ruled out.

The framework of theoretical approach comes from previous studies (Witte, 1975; Ozanne and Thibodeau, 1983; Manning, 1988; Potepan, 1996), but some changes have been made according to the housing market situation in China. House prices in a city are supposed to be determined by (long-run) housing demand and supply. Therefore, differences in average house prices among cities are caused by the differences in either the housing demand or housing supply.

3.1 Housing demand

The housing demand in a given city is determined by the aggregate of each household's willingness and ability to pay. Considering the case of one household, the willingness can be represented by the utility function $[U = U(X, Q, t)]$ which depends on the consumption of nonhousing goods and services (X), quantity of housing (Q), and the location of the city they are living in (t). Taking into account the location of city in the utility function is based on the idea that households are more likely to live in the central city or nearby cities where people will be more convenient to consume higher-grade goods and services⁷. Household's paying ability is limited by the household income (I) so that the expenditures on nonhousing goods and services ($P_x X$) and housing ($P_h Q$) subject to the income constraint $[I = P_x X + P_h Q]$ ⁸. Households are always trying to maximize their utility under the income constraint, so we have

$$Q = q(I, P_h, P_x, t) \quad (1)$$

where:

Q = Quantity of housing demand by a household;

I = Household income;

P_h = House Price;

P_x = Price of nonhousing goods and services; and

⁷ In a hierarchical urban system, the central city can provide higher-order specialized goods and services, as well as the lower-grade goods, while other cities can only provide lower-order basic goods.

⁸ In this paper, we focus on the choice of the city households are willing to live in, but not the exact home site within a city, so we incorporate the commuting costs into the term "nonhousing goods and services", which treatment is a bit different from the previous studies.

t = Location of city, represented by distance to central city.

Household income (I) consists of monetary income (I_M) and nonmonetary income (I_N) derived from the urban amenities [$I = I(I_M, I_N)$]. Monetary income (I_M) mainly depends on the household's real income. Moreover, considering that the modern financial system makes it possible for households to borrow money and spend part of their future income in advance through mortgaging activities, it is necessary to incorporate the mortgage terms in the monetary income function⁹.

$$I_M = M(w, m, i) \quad (2)$$

w = Household real income;

m = Size of mortgage; and

i = Mortgage interest rate.

The nonmonetary income (I_N) is related to the “comfort of human-living”, which depends on the urban amenities, such as the convenience of transportation system, education and medical system, environment, recreational infrastructure and so on. Combining equations (1) and (2), yields the housing demand function for a specific household. By including the population size (Z) of city t , which is assumed to be the result of the spatial equilibrium, aggregate housing demand can be expressed as

$$Q_h^D = d(P_h, w, m, i, I_N, P_x, t, Z) \quad (3)$$

3.2 Housing supply

Following the framework of Muth (1971), the housing supply in a city is a production function

$$Q_h^S = S(L, N) \quad (4)$$

L = Land inputs for housing; and

N = Other materials inputs.

Factors that influence the land inputs and other materials inputs have been fully discussed in previous studies. Apart from land, resources (N) such as construction labor, materials, capital and so on are also required for the provision of housing products. These can be expressed by the construction costs (B_c).

Land inputs (L) depend on a variety of factors, including the costs of improvements of raw land (R_c), the price of agriculture land (P_a), supply potential of agriculture land (S_p) and some government restrictions (R_g). Including these supply factors into equation (4), we obtain the housing supply function

⁹ De La Paz (2003) has fully discussed how the financial activities affect the household's accessibility to housing market, including the price/income ratio, payback/income ratio and life time of mortgage.

$$Q_h^s = s(B_c, R_c, P_a, S_p, R_g) \quad (5)$$

where:

B_c = Construction costs;

R_c = Costs of land improvements;

P_a = Price of agriculture land;

S_p = Potential of agriculture land supply; and

R_g = Policy restrictions.

3.3 Reduced form of formal model

By combining the housing demand function (3) and the housing supply function (5) and eliminating the endogenous variable, we obtain the reduced-form model

$$P_h = p(w, m, i, I_N, P_x, t, Z, B_c, R_c, P_a, S_p, R_g) \quad (6)$$

with first differences $p_1, p_2, p_4, p_7, p_8, p_9, p_{10}, p_{12} > 0$; $p_3, p_6, p_{11} < 0$ and $p_5 = ?$ (unknown in advance).

4. Data and Variables

The model will be estimated using aggregate data for the urban area of prefecture cities and municipalities (excluding county level data) in the Pan-Yangtze River Delta. The area comprises one municipality (Shanghai) and three provinces (Jiangsu, Zhejiang and Anhui) with a land area of 350 thousand square kilometers and a population of 215 million (see figure 2). The Pan-Yangtze River Delta is one of the most developed and densely populated regions in China, feeding 16% of the population and producing 23% of GDP with only 4% of the total land area. The reason for choosing this region as our study area is that it can be described as a hierarchical urban system because of the frequent interaction of labour and capital. Our data set is a panel dataset for 42 cities, including 41 prefecture-level cities and 1 municipality, covering the years 2006 to 2010. Thus, we have a total of 210 annual observations.

The dependent variable, average house price (P_h), refers to the average price of newly-built houses per square meter floor space in the urban and suburban area of a city. More specifically, the average price was calculated by dividing total sales of commercial housing by the total floor space. The data were extracted from the statistical yearbook of each city or province. According to the Chinese statistical system, property developers submit data on housing construction and sales to the statistical bureau on a monthly basis. The data are published in a monthly report and a statistical yearbook. Our measure of house price includes both finished housing and forward delivery housing¹⁰. A disadvantage is that it also covers commercial properties. However, 80%

¹⁰ In the forward delivery housing (“qifang” in Chinese) market, the purchasers buy the underlying houses that are not yet completed from the developers in the form of forward contract.

or more of total sales relates to residential properties, so we think any bias will be limited¹¹. For a few cities in some years, data on total sales and floor space was not available in the statistical yearbook, so the average house price from “the annual report of real estate” published by some real estate research institute (such as E-House China, Centaline Property Agency and World-Union Property) was used.

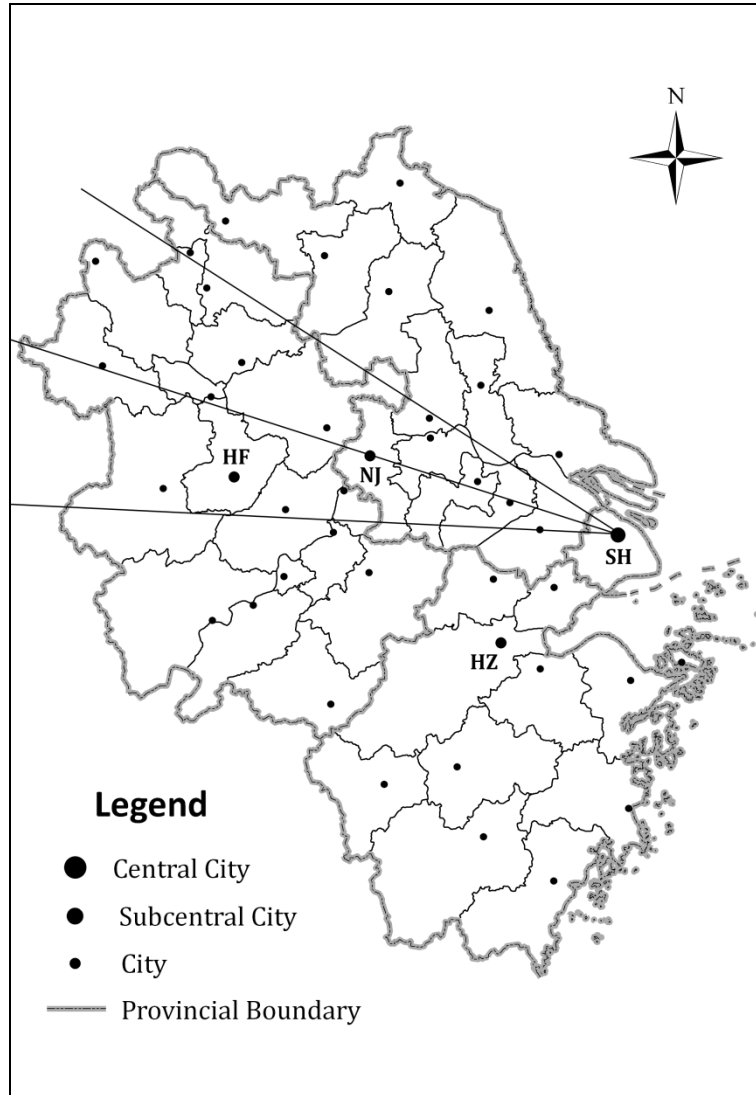


Fig.2 Pan-Yangtze River Delta

In terms of the independent variables, the location of a city (t) was measured by the Euclidean distance to central city (D_c) or to the sub-central city (D_{sc}). Shanghai was a priori defined as the central city, and Nanjing and Hangzhou as sub-central cities. The specifying criterion is

¹¹ Some cities have the statistical data of residential housing, but not all cities. According to these data, the percentage of residential housing accounting for total sold commercial housing in most cities is greater than 80%, only very few cities below this figure.

given in Section 5. A dummy variable for coastal city (CC) was also introduced, which has the value 1 if the urban area of that city is located at the coastline and the value 0 otherwise.

The data for most of the socio-economic variables came from the statistical yearbook. Monetary income (I_M) was measured by disposable income per capita (w), which can be found in the statistical yearbook. It is a measure of after-tax income and calculated by deducting income tax and other fees from total income. There was no direct indicator for the size of the mortgage associated with the purchase of the houses. As an approximation we used the loan balance per capita (m), which was calculated by dividing the total loan balance of financial institutions by the population of the urban districts. Because of the highly centralized financial system in China, there is no difference in mortgage interest rates (i) between cities, so it was not included in the estimating equation.

Nonmonetary income (I_N) was approximated by two variables: public expenditure on education per student (E_N) and the number of doctors per thousand people (D_N)¹². In order to have better education for their children, people may be willing to move house to and pay a higher price for a house in cities with high-quality education resources. We might expect that higher public expenditure on education will lead to better education facilities and higher teacher quality. The number of doctors per thousand people was chosen as an indicator of the level of medical services.

The preferred measure for population size (Z) is the number of residents who permanently live in the urban area. However, for most of the cities in our study, this data is not available. We therefore chose the population in the urban district under the prefecture government, including the registered population in urban area and rural area. Considering the very high degree of urbanization in China in the recent past, this alternative measure seems to make sense, at least to some extent.

With respect to supply factors, the price of agriculture land (P_a) could be represented by the cost of transferring agricultural land into residential construction land. The “compensation standard for agriculture land expropriation” published by the provincial government is unfortunately only available after 2010 for most of the cities, and so this variable was dropped. The potential agriculture land supply (S_p) is believed to be limited by the population density (Z_d). Cities with lower population densities are likely to have larger proportions of agriculture land and can probably more easily transfer agricultural land into land for construction purposes.

In spite of extensive search, it was really difficult to find appropriate measures for price of nonhousing goods and services (P_x), construction costs (B_c), costs of land improvements (R_c) and policy restrictions (R_g). The absence of these variables may bias the estimation results. However, as we will see below, the selected variables explain more than 86% of the interurban

¹² Considering that the educational expenditure in city level is mainly on the primary and secondary school, so this indicator was computed by dividing the number of students in primary and secondary school from the total educational expenditure.

house prices variation. Finally, considering that the dataset we use pools 5 years' cross-sectional data, we added 4 time dummy variables to control for time effects. A total of 13 variables will enter in the fully specified estimating equation; see Table 1. As mentioned above, the spatial extents of average house price and disposable income cover the urban area of the urban districts under prefecture government, while the other socio-economic variables are based on the data of the urban districts.

Table 1. Definition of variables

| <i>Variables</i> | <i>Description</i> |
|------------------|--|
| P_h | <i>Average house price</i> ; computed by dividing sold floor areas from total sales of commercial housing (Yuan/m ²). |
| D_c | <i>Distance to central city</i> ; the straight-line distance to central city (Km). |
| D_{sc} | <i>Distance to sub-central city</i> ; the shortest straight-line distance to the 3 sub-central cities (Km). |
| CC | <i>Coastal city</i> ; if the urban area next to the coastline, it will be 1, otherwise 0. |
| w | <i>Disposable income</i> ; the after-tax income in urban area (Yuan per capita). |
| m | <i>Loan balance per capita</i> ; computed by dividing the registered population in urban district from the total loan balance (Yuan per capita). |
| E_N | <i>Public expenditure on per student</i> ; computed by dividing the number of students in primary and secondary school in urban districts from the public expenditure on education (Yuan per student). |
| D_N | <i>Number of doctors per thousand capita</i> ; computed by dividing the registered population in urban districts from the number of doctors (Yuan per thousand capita). |
| Z | <i>Population size</i> ; registered population in urban district (10 ⁴ person). |
| Z_d | <i>Population density</i> ; computed by dividing the area of urban district from registered population (person/Km ²). |
| Y_t | <i>Time dummies</i> ; if the transaction occurs in the t_{th} year, it will be 1, otherwise 0 ($t=2007, 2008, 2009, 2010$). |

5. Estimation Results

A variety of specific models with different sets of explanatory variables were estimated using ordinary least squares regression. Firstly, only the location variables were used as explanatory variables, and the interurban house price gradient was estimated. The corresponding models are denoted by LM. Secondly, models SM were estimated that included the socio-economic variables but not location. Finally, the fully specified model (LSM), including both location and socio-economic variables, was estimated to measure the interurban house price gradient when controlling for the socio-economic variables.

5.1 Identification of Central and Subcentral City

In this study, the location of a city is represented by the distance to the central or subcentral city, so we need to define and identify those central and subcentral cities first. The central city refers to the ‘super city’ which has a global influence on the other cities in a hierarchical urban system; subcentral cities are those big cities that have local effects. Formal approaches to identify the subcenters in a metropolitan area have been presented (McMillen and McDonald, 1998; Palut and palut, 1998), and these models may be helpful in identifying the subcentral cities. However, distinguishing the central and subcentral cities is not so complicated in a hierarchical urban system, and an informal approach will suffice. Especially in China, the ranks of cities are usually determined in the country-level urban planning. According to the “Outline of National Urban System Planning (2005-2020)”, Shanghai is designed to be a nationwide-central city, while the cities of Nanjing, Hangzhou, Ningbo and Hefei are local-central cities. Consequently, it is reasonable to treat Shanghai as the central city in our study, and Nanjing, Hangzhou and Hefei, the capital of Jiangsu, Zhejiang, Anhui provinces, as the Subcentral cities. Ningbo is excluded as a subcentral city as it is under the administration of Zhejiang Province and has less influence than the capital city Hangzhou.

Table 2 provides a comparison of central, subcentral cities and the remaining cities in terms of the economic situation in 2010. It can be seen that Shanghai, the central city, has by far the largest working population, the biggest investment in fixed assets and the highest GDP. A similar result is found when comparing the subcentral cities and the remaining cities. This convinces us that the identification of central and subcentral cities is appropriate.

Table 2. Comparison of central, subcentral cities with the remaining cities in year 2010

| <i>City</i> | <i>Working population (10⁴ Person)</i> | <i>GDP (Billion Yuan)</i> | <i>Fixed asset investment (Billion Yuan)</i> |
|------------------------------------|---|-------------------------------|--|
| Shanghai | 716.74 | 1697.16 | 519.38 |
| Nanjing | 259.18 | 451.52 | 289.14 |
| Hangzhou | 333.58 | 474.08 | 218.27 |
| Hefei | 114.28 | 192.04 | 241.04 |
| Average of the remaining cities | 44.60 | 78.07 | 49.97 |

5.2 Effect of Location: Constrained Interurban House Price Gradient

Distance Effects of Central City

Intracity studies tend to recommend the semi-log (exponential) functional form to represent the relation between distance and house price, but the choice of functional form in the interurban context is still unexplored. Three functional forms were tested. The model LM1 assumes that the average house price in a city declines with the distance to the central city in a linear fashion ($P_h = \alpha + \beta d_c$), while the model LM2 uses a semi-log function and the model LM3 a log-log function ($\ln P_h = \alpha + \beta Dc$ and $\ln P_h = \alpha + \beta \ln Dc$, respectively). For all of the three models, four year dummy variables were introduced to allow the intercept to change from year to year.

The estimation results for models LM1, LM2 and LM3 are reported in Table 3. White period method was used to calculate robust standard errors and t values. All three functional forms can capture the spatial variance of interurban house prices, given the highly significant coefficients and the overall F statistics. The signs of the coefficients for the distance variables are negative, showing that the average house price in a city declines with the distance to the central city, as expected. The semi-log functional form (LM2) performs best in terms of the goodness-of-fit with an adjusted R^2 of 0.515, nearly 10% higher than the log-log function (LM3) and 20 % higher than the linear function (LM1). The estimated interurban housing price gradient is -0.0019 according to LM2, indicating that one extra kilometer away from the central city will reduce the average house price in a city by 0.19%.

Table 3 Distance Effects of Central City

| | <i>LM1</i> <i>Linear</i> | <i>LM2</i> <i>Semi-log</i> | <i>LM3</i> <i>Log-log</i> | <i>LM4</i> <i>Semi-log</i> | <i>LM5</i> <i>Semi-log</i> |
|----------------|-----------------------------|-------------------------------|------------------------------|-------------------------------|-------------------------------|
| | P_h | $\ln(P_h)$ | $\ln(P_h)$ | $\ln(P_h)$ | $\ln(P_h)$ |
| Constant | 6117.2985*** (10.30) | 8.6299*** (94.97) | 9.3306*** (26.79) | 8.3769*** (83.56) | 8.5383*** (88.42) |
| Dc | -8.9146*** (-5.12) | -0.0019*** (-7.62) | — | -0.0014*** (-5.64) | -0.0019*** (-7.45) |
| $\ln(Dc)$ | — | — | -0.2343*** (-3.77) | — | — |
| CC | — | — | — | 0.5641*** (3.89) | — |
| $Dc * CC$ | — | — | — | — | 0.0019** (2.38) |
| Y_{2007} | 611.1905*** (6.65) | 0.1535*** (10.11) | 0.1535*** (10.11) | 0.1535*** (10.11) | 0.1535*** (10.11) |
| Y_{2008} | 1102.8052*** (6.52) | 0.2629*** (12.90) | 0.2651*** (12.56) | 0.2616*** (13.22) | 0.2611*** (13.35) |
| Y_{2009} | 2093.9271*** (5.63) | 0.4353*** (16.21) | 0.4376*** (15.86) | 0.4341*** (16.57) | 0.4336*** (16.77) |
| Y_{2010} | 3212.6428*** (8.30) | 0.6323*** (25.05) | 0.6323*** (25.05) | 0.6323*** (25.05) | 0.6323*** (25.05) |
| F-statistic | 20.64 | 44.96 | 31.31 | 83.54 | 71.27 |
| Adjusted R^2 | 0.322 | 0.515 | 0.423 | 0.705 | 0.671 |
| Sample | 208 | 208 | 208 | 208 | 208 |

Note: **denotes the significance at the 5% level; *** denotes the significance at the 1% level. The t values were presented in parenthesis. Considering the existence of serial correlation, the robust standard errors corrected by White period method were used in computing the t values. When estimating Model 3, the value of Dc for central city-Shanghai-was set to 1 in order to take the natural logarithm.

However, the region under study is most likely not homogenous. In particular, average house prices in coastal cities are expected to be higher than in inland cities. A dummy variable (CC) for coastal cities was added to the semi-log model. The positive and significant coefficient in the resulting model LM4 confirms the above and increases the adjusted R^2 from 0.515 to 0.705. Furthermore, we would expect the distance effect on house prices for coastal cities to be smaller

than for inland cities. The model LM5 is an extension of LM3, which includes an interaction term for distance and the coastal-city dummy variable. The coefficient of this interaction term turned out to be significantly positive. Combining the coefficients of interaction term ($Dc \cdot CC$) and distance to central city (Dc) shows that there is almost no negative effect of distance to the central city on coastal cities' house prices. This result suggests that households are willing to pay a premium for living in coastal cities, perhaps because of the pleasant climate and environment, the diverse culture and also the typically more developed economy.

Joint distance effects of Central and Subcentral Cities

To investigate the influences of subcentral cities on other cities' average house prices, we adapt the framework of Heikkila *et al* (1989), who considered the roles of subcenters in a polycentric city, to the intraurban case. Above, we defined three cities (Hangzhou, Nanjing and Hefei) as subcentral cities. It is assumed that the relationship is competitive between subcentral cities but complementary between each of the subcentral cities and the central city. Four functional forms that incorporate the influence of subcentral cities will be considered¹³:

$$\text{LM6: } P_{hi} = \alpha \cdot e^{\beta_1 Dic} \cdot e^{\beta_2 \min\{Disch, Discn, Discf\}} \cdot u_i \quad (7)$$

$$\text{LM7: } P_{hi} = \alpha \cdot e^{\beta_1 Dic} \cdot \min\{Disch, Discn, Discf\}^{\beta_2} \cdot u_i \quad (8)$$

$$\text{LM8: } P_{hi} = \alpha \cdot Dic^{\beta_1} \cdot e^{\beta_2 \min\{Disch, Discn, Discf\}} \cdot u_i \quad (9)$$

$$\text{LM9: } P_i = \alpha \cdot Dic^{\beta_1} \cdot \min\{Disch, Discn, Discf\}^{\beta_2} \cdot u_i \quad (10)$$

P_{hi} denotes the average house price in city i ; Dic refers to the distance from city i to central city (Shanghai); $Disch$, $Discn$ and $Discf$ refer to the distance from city i to subcentral cities Hangzhou, Nanjing and Hefei, respectively. The models use the distance to the closest subcentral city. Hence, two dummy variables, SCH and SCN , were included to represent the house price differences between Hangzhou, Nanjing and Hefei after controlling for the distance to central city. Note that in the models LM6 and LM7, SCN was not significant and therefore dropped. The estimation results, after taking logarithms of both sides of equations (7)-(10), are reported in Table 4.

Both the central and subcentral cities have significant negative distance effect on house prices in all the models except LM6, where the coefficient for subcentral cities is not significant, not even at the 10% level. Using the R^2 criterion, model LM7 performs best. Compared to LM4, adding

¹³ Several scenarios were presented and tested here. In the first scenario, only one of the three cities (Nanjing, Hangzhou and Hefei) was treated as subcentral city. Two of the three cities were treated as subcentral cities in the second scenario, but the relationship between subcentral cities was assumed to be complementary. The third scenario was almost the same as the second scenario except assuming the relationship between subcentral cities to be competitive. All three cities were considered as subcentral cities with complementary relationship in the fourth scenario. The fifth scenario also included all the three cities but the relationships were competitive. The fifth scenario has the best goodness-of-fit, so it was reported above.

the effect of subcentral cities raises R^2 from 0.705 to 0.767. Average house prices in Hangzhou and its neighbor cities are demonstrated to be higher than those in Nanjing and Hefei, given the positive coefficient for SCH , when controlling for the influence of Shanghai. An interesting finding is that for subcentral cities, the log-log function is more appropriate than the semi-log function. Considering that the central city may have a macro-effect that influences a larger radius and the subcentral city only has local micro-effects, it seems that the choice of functional form is sensitive to the influence sphere of the center; the log-log function performs better when the influence area is relatively small, while the semi-log function is more appropriate if the area is larger. This conclusion could be partly supported by the estimated coefficients in LM9 which estimated both the log-log forms, where the absolute coefficient of Dsc (0.0863) is a bit larger than that of Dc (0.0846) which indicates that the influence of subcentral cities would decay dramatically with the distance and therefore have a relatively small effect radius. These findings are also in line with Osland's study (Osland *et al*, 2007), in which they found that the exponential (semi-log) function performs best when the estimation is based on a large area while the power (log-log) function performs best if the data is restricted to a small area.

Table 4 Joint Effects of Central City and Subcentral City

| | <i>LM6</i> <i>Semi-log/semi-log</i> | <i>LM7</i> <i>Semi-log/log-log</i> | <i>LM8</i> <i>Log-log/semi-log</i> | <i>LM9</i> <i>Log-log/log-log</i> | <i>LM10</i> <i>Piecewise/semi-log</i> | <i>LM11</i> <i>Semi-log</i> |
|---------------|--|---------------------------------------|---------------------------------------|--------------------------------------|--|--------------------------------|
| | $\ln(P_h)$ | $\ln(P_h)$ | $\ln(P_h)$ | $\ln(P_h)$ | $\ln(P_h)$ | $\ln(P_h)$ |
| Constant | 8.2455*** (59.75) | 8.4821*** (46.95) | 8.3012*** (30.97) | 8.6148*** (27.90) | 8.9010 (531.07) | 8.6701*** (95.31) |
| Dc | -0.0009*** (-3.75) | -0.0010*** (-3.36) | — | — | — | -0.0021*** (-9.42) |
| $\ln(Dc)$ | — | — | -0.0704* (-1.93) | -0.0846** (-2.10) | — | — |
| $Dc1(spline)$ | — | — | — | — | -0.0042*** (-72.57) | — |
| $Dc2(spline)$ | — | — | — | — | 0.0121*** (47.62) | — |
| $Dc3(spline)$ | — | — | — | — | -0.0463*** (-11.74) | — |
| $Dc4(spline)$ | — | — | — | — | 0.0007 (0.69) | — |
| $Dc5(spline)$ | — | — | — | — | -0.0019*** (-2.99) | — |
| CC | 0.4850*** (3.58) | 0.4461*** (3.31) | 0.4670*** (3.29) | 0.4002*** (2.95) | — | — |
| Dsc | -0.0007 (-1.17) | — | -0.0012** (-2.02) | — | — | — |
| $\ln(Dsc)$ | — | -0.0707*** (-2.83) | — | -0.0863*** (-3.75) | — | — |
| SCH | 0.2452** (1.98) | 0.2615** (2.09) | 0.4202*** (3.20) | 0.4400*** (3.44) | — | — |
| SCN | — | — | 0.1462** (2.04) | 0.1538** (2.46) | — | — |

| | | | | | | |
|-------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Y_{2007} | 0.1535*** (10.11) | 0.1535*** (10.11) | 0.1535*** (10.11) | 0.1535*** (10.11) | 0.1465*** (7.63) | 0.1465*** (7.63) |
| Y_{2008} | 0.2599*** (13.38) | 0.2597*** (13.45) | 0.2617*** (13.18) | 0.2622*** (13.33) | 0.2380*** (10.00) | 0.2380*** (10.00) |
| Y_{2009} | 0.4324*** (16.62) | 0.4322*** (16.71) | 0.4342*** (16.47) | 0.4346*** (16.61) | 0.4285*** (14.51) | 0.4285*** (14.51) |
| Y_{2010} | 0.6323*** (25.05) | 0.6323*** (25.05) | 0.6323*** (25.05) | 0.6323*** (25.05) | 0.6397*** (20.03) | 0.6397*** (20.03) |
| F-statistic | 73.78 | 86.36 | 62.99 | 73.29 | 67.29 | 64.02 |
| Adjusted R ² | 0.738 | 0.767 | 0.729 | 0.759 | 0.939 | 0.790 |
| Sample | 208 | 208 | 208 | 208 | 85 | 85 |

Note: * denotes the significance at the 10% level; **denotes the significance at the 5% level; *** denotes the significance at the 1% level. The t values were presented in parenthesis. Considering the existence of serial correlation, the robust standard errors corrected by White period method were used in computing the t values. During the estimation of Model LM7-LM9, the value of D_c or D_{sc} was set to 1 in order to take the natural logarithm.

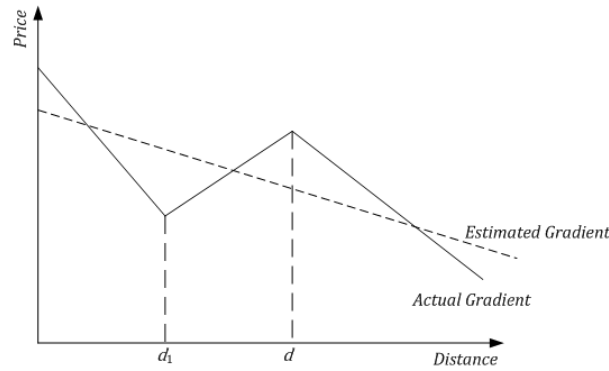


Fig.3 Actual and Estimated Housing Price Gradient

Since we have shown that subcentral cities also affect regional house prices, what does the house price gradient look like in a polynuclear context? Obviously, the influence of subcentral cities cannot be accurately captured by the traditional central city-oriented house price gradient, which is a continuously downward sloping line. In the example of figure 3, the subcentral city in location d raises the house price in that location and also affects the price in neighbor cities. The solid line shows the 'true' house price gradient, while the dashed line shows the estimated central city-oriented gradient. If the knots, such as d_1 and d in figure 3, could be correctly located, then the true price gradient could be represented by a piecewise linear spline function. Splines have been successfully applied in estimating price gradients (Dubin & Sung, 1987; Osland et al, 2007). Here we apply a piecewise linear spline to the semi-log functional form:

$$P_{hi} = \prod_{j=1}^n \alpha_j \cdot e^{\beta_j(D_{ic} - D_{j-1}^k) \delta_j} \quad (11)$$

$$\text{subject to } \alpha_{j+1} = \alpha_j \cdot e^{\beta_j(D_j^k - D_{j-1}^k)} \quad (12)$$

where the meanings of P_{hi} , Dic are the same as equation (7); D_j^{k-1} and D_j^k refer to the starting and ending knots of interval j ; δ_j is a dummy variable with the value 1 if $D_{j-1}^k \leq Dic < D_j^k$ and 0 otherwise.

In model (11), the average house price is a piece-wise function of distance to central city, where the parameters can vary between different intervals. The purpose of equation (12) is to ensure that two adjacent intervals have the same price at the boundary (knot location). Substituting equation (12) into equation (11) yields

$$\text{LM10: } P_i = \alpha_1 \cdot \left[\prod_{j=1}^n e^{\beta_j (Dic - D_{j-1}^k) \cdot \delta_j} \right] \cdot \left[\prod_{j=1}^n e^{\beta_j (D_j^k - D_{j-1}^k) \cdot \tau_j} \right] \quad (13)$$

where the dummy variable τ_i is 1 if $Dic \geq D_j^k$, and 0 otherwise. It should be emphasized that model LM10 is highly sensitive to the heterogeneity of the region. That is, the estimated price gradient will depend not only on the distance to the central city but also on the direction. To obtain an accurate estimator, an area divided by two rays¹⁴ was chosen to estimate model LM10 (see figure 2). In this area, there are two subcentral cities (Nanjing & Hefei), so we would expect to have four knots in the spline function.

For comparison reasons we also estimated the traditional semi-log model for the ray area, LM11. The OLS results for this model in Tabel 4 show that the central city has a significant negative effect on the average house price in the cities in the ray-area. Model LM11 explains the regional house price structure reasonably well with an adjusted R^2 of 0.79. The piecewise-linear spline function, however, has much more explanatory power, with an adjusted R^2 of 0.93. The spline function approximates the 'true' house price gradient very well, the two subcentral cities being correctly located at the peak points. Note that the location of the knots were determined using a searching procedure by maximizing the adjusted R-squared¹⁵.

Figure 4 shows that the central city (Shanghai) initially has a much sharper diminishing effect than was estimated by the traditional exponential function. After 200 kilometers, an upward effect is observed due to the influence of subcentral city Nanjing. Yet Nanjing has a very limited affecting radius. This can be explained by two reasons. First, for the cities on the left of Nanjing, Shanghai is more attractive and these cities are more willing to keep close contact with Shanghai. Second, the cities located on the right of Nanjing are under the management of another province (Anhui), which also limits the influence of Nanjing, as the administrative demarcation has seriously limit the socio-economic links between cities in different provinces. Another subcentral city, Hefei, is less developed itself and its average house price is still staying at a low level, so the influence on the other cities is also not strong. The weakly upward gradient is not significant on

¹⁴ The two rays are derived by rotating the line crossing Shanghai (central city) and Nanjing (subcentral city) approximately 15 degrees, so the two rays are 30 degrees apart. There are totally 17 cities in this ray area, and all the cities are inland cities except the central city Shanghai.

¹⁵ It is assumed that the four knots can be any four locations of 17 cities. The combination which had the best performance was selected at last.

the left hand and the slope of negative gradient on the right hand is even flatter than that of exponential function.

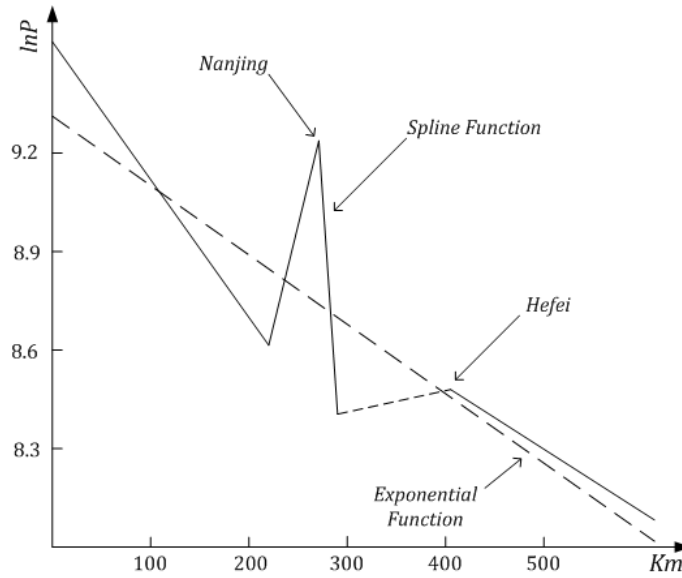


Fig.4 Empirical Results of Exponential Function and Spline Function

5.3 Effect of Socio-economic Factors

The theoretical model suggests that differences in the socio-economic circumstances across cities, such as the income level, mortgage size, and public infrastructure, will also affect the interurban house price variation. In this section, we investigate the effects of socio-economic factors. All the socio-economic variables were transformed into natural logarithms so that the coefficients can be interpreted as elasticities.

The OLS results are reported in the first two columns in Table 5. The model SM1 contains all the socio-economic variables. This model performs quite satisfactory. It explains almost 86% of the variance in interurban house prices, which is much more than the 76.7% explained by location variables. Three variables, *disposable income*, *loan balance per capita* and *population size*, have significant effects on house prices at the 1% level, and the signs of the coefficients are according to expectations. In contrast, the signs of the coefficients for the other three socio-economic variables are counterintuitive, of which the coefficient for *population density* is significant at the 10% level. Multicollinearity seems to have been a problem here. The collinearity diagnostic indicator VIF was calculated for each variable and shows that the VIF of w (5.42) and m (5.71) are greater than 5 which means the existence of multicollinearity in model SM1.

Table 5 Estimation results of the model SMs and LSMs

| <i>SM1</i> <i>OLS</i> | <i>SM2</i> <i>Stepwise</i> <i>OLS</i> | <i>LSM1</i> <i>OLS</i> | <i>LSM2</i> <i>OLS</i> | <i>LSM3</i> <i>OLS</i> | <i>LSM4</i> <i>OLS</i> |
|--------------------------|---|---------------------------|---------------------------|---------------------------|---------------------------|
| $\ln(P)$ | $\ln(P)$ | $\ln(P)$ | $\ln(P)$ | $\ln(P)$ | $\ln(P)$ |

| | | | | | | |
|-------------------------|---------------------|---------------------|---------------------|-----------------------|---------------------|---------------------|
| Constant | -1.1191 (-0.64) | -1.7647* (-1.75) | -1.4099 (-1.03) | 8.0332*** 23.15 | -2.7834* (-1.89) | 4.8658*** (7.26) |
| Dc | — | — | 0.0002 (0.64) | -0.0014*** (-5.83) | -0.0001 (-0.29) | -0.0001 (-0.31) |
| CC | — | — | 0.2088*** (2.76) | 0.5296*** (3.79) | 0.2717*** (3.05) | 0.2793*** (2.65) |
| $\ln(w)$ | 0.8236*** (3.26) | 0.8177*** (5.98) | 0.7939*** (5.53) | — | 1.1390*** (7.73) | — |
| $\ln(m)$ | 0.2682*** (4.32) | 0.1705*** (5.05) | 0.1527*** (4.12) | — | — | 0.3219*** (5.65) |
| $\ln(E_N)$ | -0.1333 (-1.17) | — | — | — | — | — |
| $\ln(D_N)$ | -0.0289 (-0.35) | — | — | — | — | — |
| $\ln(Z)$ | 0.1206*** (3.30) | 0.0678* (1.70) | 0.0604 (1.58) | 0.1113 (1.61) | — | — |
| $\ln(Z_d)$ | -0.1106* (-1.88) | — | — | — | — | — |
| Y_{2007} | 0.0436 (1.10) | — | — | — | — | — |
| Y_{2008} | 0.0532 (0.95) | — | — | — | — | — |
| Y_{2009} | 0.0963 (1.45) | — | — | — | — | — |
| Y_{2010} | 0.1901** (2.41) | 0.1139*** (3.45) | 0.1288*** (2.96) | 0.4157*** (23.79) | 0.1159*** (2.83) | 0.2571*** (7.48) |
| F-statistic | 121.37 | 279.91 | 223.93 | 94.63 | 271.71 | 205.87 |
| Adjusted R ² | 0.856 | 0.846 | 0.868 | 0.647 | 0.840 | 0.798 |
| Sample | 204 | 204 | 204 | 205 | 207 | 208 |

Note: * denotes the significance at the 10% level; **denotes the significance at the 5% level; *** denotes the significance at the 1% level. The t values were presented in parenthesis. The robust standard errors corrected by White period method were used in computing the t values. In SM2, the variable $\ln(Z_d)$ was selected into the model according to the stepwise procedure, but its robust t value is not significant, so it was dropped out.

In order to avoid the influence of multicollinearity, a stepwise regression was employed and shown as SM2 in table 5. Note that the variable *population density* was entered into the equation as the last step, but its robust t value was not significant at the 10% level and so it was excluded from the final equation. The coefficients for the included three variables have expected signs. Together with one time dummy variable, 84.6% of the interurban house price variance was explained, which is only 1% less than that of model SM1.

Looking at the results of SM2, *disposable income* has a highly significant positive effect on house prices, similar to what was found in previous studies (Witte, 1975; Manning, 1998 and Potepan, 1996). The income elasticity in our study is 0.8177, indicating that a 1% increase in average disposable income in a particular city will result in a 0.82% increase in the average house price of that city. Potepan (1996) estimated an income elasticity of 0.64 on intermetropolitan data for the USA. Our slightly higher figure can be explained by the shortage of housing supply and the traditional culture in China. Chinese households seem to be willing to spend a larger part of their disposable income on housing than households in Western

countries. The shortage of supply and the high demand for owner-occupied dwellings will drive up house prices.

De La Paz (2003) pointed out that the accessibility of households to the housing market also affects house prices. Specialization and maturation of the financial system is likely to increase the possibility for households to obtain a mortgage to finance home purchases. A higher loan balance per capita may thus increase accessibility and lead to higher house prices. Our results support this: a 1 % increase in *loan balance per capita* results in a 0.17% increase in house prices.

The coefficient for *population size* has the expected sign, though it is only significant at the 10% level. Its value of 0.0678 is approximately half the value of 0.14 estimated by Potepan (1996). The reason for this low elasticity may be the data we used. The *population size* in this study refers to the population in the urban district of a city, which includes many rural residents also, so it is not a 'true' urban area we studied. Put differently, not all households belonging to the statistical population have access to the housing market so that we underestimate the effect of the 'true' urban population size.

As mentioned earlier, *public expenditure on education per student*, *the number of doctors per thousand residents* and *population density* were excluded from the final equation in the stepwise regression. Apart from statistical reasons, i.e. the high collinearity with the other (included) variables, this could also mean that households mainly consider "monetary" factors when they wish to buy a house and choose the city to live in, and care much less about "nonmonetary" factors, such as education opportunities, the medical system and so on.

5.4 House Price Gradient When Controlling for the Socio-economic Variables

As we have shown, location and socio-economic factors significantly affect the average house price in a city. The questions addressed in this section are what the joint effect of location and the socio-economic variables is, and how the interurban house price gradient will change when controlling for the socio-economic variables. Following Söderberg and Janssen (2001) in their intracity study, who use the same functional forms as us, the full specification of equation 6 will be tested in this section. The model contains the three socio-economic variables of model SM2 (in logarithmic form), two location variables (Dc and CC) and one year dummy variable (Y_{2010}) as explanatory variables.

The OLS results are reported in the third column of Table 5 (LSM1). The performance of the full specification model in terms of the adjusted R squared is slightly better than the models that include only socioeconomic variables or location. The small improvement suggests that there may be an interaction between the location and socio-economic variables. The most striking finding is that the coefficient for the distance to central city is now insignificant and positive. Waddell, Berry and Hoch (1993b) argued that the value associated with the relative location in *intracity* house prices is capitalized into the structural characteristics of the houses and into neighborhood quality. Similarly, it may be that the *interurban* house price gradient has been capitalized into the spatial structure of income and mortgage size. This is supported by the

results of models LSM2, LSM3 and LSM4 in Table 5. In model LSM2, only one socio-economic variable, the *population size* $\ln(Z)$, is included, and a negative and significant interurban house price gradient is observed. However, when we use another, either $\ln(w)$, as in LSM3, or $\ln(m)$, as in model LSM4, the (absolute) negative coefficient for the distance to central city decreases from 0.0014 to 0.0001 and is no longer significant. The *LR* test results also told us that D_c is the redundant variable in LSM3 ($LR=0.39$, $P=0.53$) and LSM4 ($LR=0.53$, $P=0.46$) respectively.

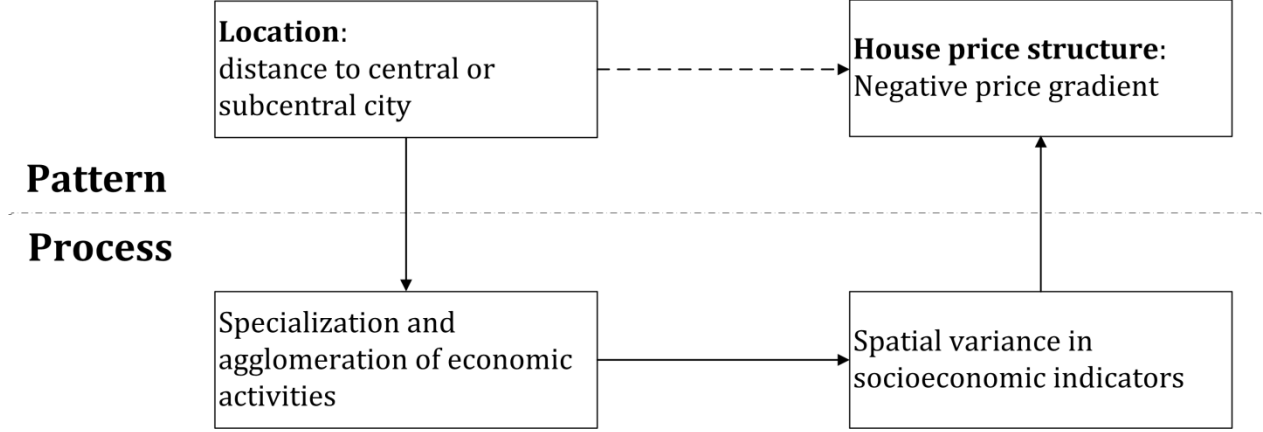


Fig.5 The capitalization process of location effects

From a statistical perspective, the capitalization is caused by the high correlation of the three variables, $\ln(w)$, $\ln(m)$ and D_c ; the distance effect is overshadowed by the effects of income and mortgage size. Theoretically, this result could be related to the evolution of interurban economic activities (see figure 5). Because of spatial heterogeneity and scale economies, different economic activities will be concentrated in different locations. The process of specialization and agglomeration will render spatial differences in socio-economic development in terms of income, loan, population and so on, which in turn will affect the spatial structure of interurban house prices. In other words, the effect of location on the interurban house prices structure is *indirect* and has been generated by a chain process. If we control for the direct “bridge” variables, such as income, then the indirect location effect vanishes.

6. Conclusion and Discussion

Based on the aggregated data of the urban area of prefecture cities in Pan-Yangtze River Delta, in this study we estimated the interurban house price gradient in a hierarchical urban system and investigated how this gradient changes when considering the impacts of socio-economic factors. A significant negative interurban house price gradient was found, indicating that the urban average house price declines exponentially with the distance to the central city (Shanghai) except for the coastal cities. Adding the distance effects of three subcentral cities (Nanjing, Hangzhou and Hefei) further increased the performance of the model. A spline function was employed to fit the ‘true’ price gradient for a ray-area containing the central and subcentral cities. The results demonstrated that the central and subcentral cities both have distance effects on other cities’ house prices, but with a different radius and intensity. Regarding the socio-

economic factors, we find that the most three important variables affecting the interurban housing price are income, mortgage size and population, which is in accordance with previous studies. Finally, when considering the effects of those socio-economic variables in isolation, the central-city-oriented house price gradient vanishes, suggesting that the indirect distance effects on interurban house prices were capitalized into the direct effects imposed by income and mortgage.

An issue in this study is whether the gradient is a proper measure for describing the structure of finite discrete “points” (we treat each city’s housing market as a point in the study area) while all the house price gradients in intracity studies are based on a continuous area. However, we are not trying to formulate a surface of house prices in a regional context, but merely to reveal the relationship between cities’ house prices and their location. From this perspective, the gradient seems to be a useful tool. It should also be emphasized that the existence of an interurban house price gradient can only be expected in a hierarchical urban system or in a region where the cities are closely linked. The Pan-Yangtze River Delta is a good example as there is a close interaction between the cities in terms of capital, labor and information due to the industrial division and the national economic development plan. Another potential weakness of our study is the choice of independent variables. Appropriate measures for several factors could not be found, which may result in omitted-variables bias. Also, the measures of non-monetary income we used were rather crude. Finally, we did not consider the impact of spatial effects. These effects have been discussed by, for example, Yu *et al* (2007), McMillen (2010) and Osland (2010). We are intending to address the spatial effects in interurban housing market by spatial analysis.

The intracity house price gradient is formed by the trade-off between commuting costs and housing costs. It can be argued that this trade-off also plays a role in the interurban house price structure: residents have to travel to central cities to purchase higher-order products and services which are not available outside the central city. This behavior between cities is similar to the commuting behavior in the intracity. But that is probably not the main part of the story. The capitalization of distance effects into the spatial variation of income and mortgage indicates that the interurban housing market structure is more likely to relate to the distribution of economic activities. The emergence of New Economic Geography, which explains how the economic activities aggregate to formulate a “core-periphery” structure and then evolve to a hierarchical urban system based on the interaction of “centralization” and “centrifugation” forces (Krugman, 1991; Fujita *et al*, 1999), provides us with a new perspective to understand the spatial structure of interurban house prices. Yet, more theoretical work is needed to analyse the linkages between interurban housing market and interurban economic activities.

References

Adair, A., McGreal, S., Austin S., *et al*. House prices and accessibility: the testing of relationship within the Belfast Area. *Housing Studies* 15(5), 699-716.

- Ahlfeldt, G., 2011. If Alonso was right: modeling accessibility and explaining the residential land gradient. *Journal of Regional Science* 51(2), 318-338.
- Alonso, W., 1964. *Location and land use: toward a general theory of land rent*. Cambridge MA: Harvard University Press.
- Anas, A., Arnott, R., Small, K. A., 1998. Urban spatial structure. *Journal of Economic Literature* 36(3), 1426-1464.
- Atack, J., Margo, R. A., 1998. "Location, location, location!" The price gradient for vacant urban land: New York, 1835 to 1900. *Journal of Real Estate Finance and Economics* 16(2), 151-172.
- Colwell, P.F., Munneke, H.J., 1997. The structure of urban land prices. *Journal of Urban Economics* 41(3), 321-336.
- Cook, A. 2003. The convergence of regional house prices in the UK. *Urban Studies* 40(11), 2285-2294.
- De La Paz, P. T. 2003. Determinants of housing prices in Spanish cities. *Journal of property investment and finance* 21(2), 109-135.
- Dubin, R. A., Sung, C. 1987. Spatial variation in the price of housing: rent gradients in non-monocentric cities. *Urban Studies* 24(3), 193-204.
- Fujita, M., Krugman, P., Venables, A. J. 1999. *The spatial economy: cities, regions, and international trade*. Cambridge MA: The MIT Press.
- Giorgio, C., Miller, S., Pollard, S. 2012. Unit roots and structural change: an application to US house price indices. *Urban Studies* 49(4), 757-776.
- Giuliano, G., Gordon, P., Pan, Q., *et al.* 2010. Accessibility and residential land values: some tests with new measures. *Urban Studies* 47(14), 3103-3130.
- Heikkila, E., Gordon, P., Kim, J. I., *et al.* 1989. What happened to the CBD-distance gradient?: land values in a polycentric city. *Environment and Planning A* 21(2), 221-232.
- Holly, S., Pesaran, M. H., Yamagata, T. 2011. The spatial and temporal diffusion of house prices in the UK. *Journal of Urban Economics* 69(1), 2-23.
- Huang, Y.Q., 2004. The road to homeownership: a longitudinal analysis of tenure transition in urban China (1949-1994). *International Journal of Urban and Regional Research* 28(4), 774-795.
- Jones, C., Leishman, C. 2006. Spatial dynamics of the housing market: an interurban perspective. *Urban Studies* 43(7), 1041-1059.
- Krugman, P. 1991. Increasing returns and economic geography. *Journal of Political Economy* 99(3), 483-499.

- Manning, C. A., 1988. The determinants of intercity home building site price differences. *Land Economics* 64(1), 1-14.
- McMillen, D. P., 1996. One hundred fifty years of land values in Chicago: a nonparametric approach. *Journal of Urban Economics* 40(1), 100-124.
- McMillen, D.P., 2001. Nonparametric employment subcenter identification. *Journal of Urban Economics* 50(3), 448-473.
- McMillen, D. P. 2010. Issues in spatial data analysis. *Journal of Regional Science* 50(1), 119-141.
- McMillen, D.P., McDonald, J.F., 1998. Suburban subcenters and employment density in metropolitan Chicago. *Journal of Urban Economics* 43(2), 157-180.
- Mills, E. S., 1967. An aggregation model of resource allocation in a metropolitan area. *American Economic Review* 57(2), 197-210.
- Mills, E.S., 1972. *Studies in the structure of the Urban Economy*. Baltimore: Johns Hopkins Press.
- Muth, R. F., 1969. *Cities and housing: the spatial pattern of urban residential land use*. Chicago: University of Chicago Press.
- Muth, R. F., 1971. The derived demand for urban residential land. *Urban Studies* 8(3) , 243-254.
- Osland, L. 2010. An application of spatial econometrics in relation to hedonic house price modelling. *Journal of Real Estate Research* 32(3), 289-320.
- Osland, L., Thorsen, I., Gitlesen, J.P. 2007. Housing price gradients in a region with one dominating center. *Journal of Real Estate Research* 29(3), 321-346.
- Osland, L., Thorsen, I. 2008. Effects on housing prices of urban attraction and labor-market accessibility. *Environment and planning A* 40(10), 2490-2509.
- Osland, L., Thorsen, I. 2013. Spatial impacts, local labour market characteristics and housing prices. *Urban Studies*, doi: 10.1177/0042098012474699.
- Ozanne, L., Thibodeau, T., 1983. Explaining metropolitan housing price differences. *Journal of Urban Economics* 13(1), 51-66.
- Palut, P.O., Palut, S.E., 1998. Endogenous identification of multiple housing centers in metropolitan area. *Journal of Housing Economics* 7(3), 219-245.
- Potepan, M.J., 1996. Explaining intermetropolitan variation in housing prices, rents and land prices. *Real Estate Economics* 24(2), 219-245.
- Qin, B., Han, S. S. 2013. Emerging polycentricity in Beijing: evidence from housing price variations, 2001-05. *Urban Studies*, doi: 10.1177/0042098012471979.

- Sivitanidou, R., 1996. Do office-commercial firms value access to service employment centers? A hedonic value analysis within polycentric Los Angeles. *Journal of Urban Economics* 40(2), 125-149.
- Söderberg, B., Janssen C., 2001. Estimating distance gradients for apartment properties. *Urban Studies* 38(1), 61-79.
- Waddell, P., Berry, B. J. L., Hoch, H. 1993a. Residential property values in a multinodal urban area: new evidence on the implicit price of location. *Journal of Real Estate Finance and Economics* 7(2), 117-141.
- Waddell, P., Berry, B. J. L., Hoch, I. 1993b. Housing price gradients: the intersection of space and built form. *Geographical Analysis* 25(1), 5-19.
- Wang, Y.P., Shao, L., Murie, A. *et al*, 2012. The maturation of the neo-liberal housing market in urban China. *Housing Studies* 27(3), 343-359.
- Witte, A. D. 1975. The determination of interurban residential site price differences: a derived demand model with empirical testing. *Journal of Regional Science* 15(3), 351-364.
- Wu, J., Gyourko, J., Deng, Y. H., 2012. Evaluating conditions in major Chinese housing markets. *Regional Science and Urban Economics* 42(3), 531-543.
- Yu, D., Wei, Y. D., Wu, C. 2007. Modelling spatial dimensions of housing prices in Milwaukee, WI. *Environment and Planning B: Planning and Design* 34(6), 1085-1102.
- Zheng, S.Q., Kahn, M. E. 2008. Land and residential property markets in a booming economy: new evidence from Beijing. *Journal of Urban Economics* 63(2), 743-757.