5-1-2012

Does School Quality Affect Real Estate Prices? The Effect of Top-Tier Elementary Schools on Property Prices in Shanghai

Wanyi Li
Macalester College, wandou1989@gmail.com

Does School Quality Affect Real Estate Prices? The Effect of Top-Tier Elementary Schools on Property Prices in Shanghai

Wanyi Li

Department of Economics

Macalester College

March 4, 2012

Abstract: This study adapted the hedonic pricing model and inspected how varying elementary school quality affects property prices in Shanghai. Because the variation of school quality appeared before the capitalization of the housing market, the obtained results suggest a causal relationship. The data set comprises top-tier elementary school information from eWOM, the yearly school attendance zones published by the government, and the Shanghai Existing Property Index, which has a sample of similarly structured apartments. Main results show that prices on average increase 41.9% more in the top-tier school districts under the standardized housing system, and range from 15.5% to 69.7% among different districts in urban Shanghai.

Table of Contents

1. Introduction

The 2000 Housing Reform in China has led to an incredible housing boom which has drawn more attention from the public in recent years. The property prices were doubled by 2005 and tripled by 2009.[1] Before the reform, apartments were not capitalized in the market but allocated to the employees as part of employee benefits or social welfare by the nearby employers and government agencies. It is not until the 2000 housing reform that the price disparities started to show up.[2] The property values vary widely not only across different cities, but also across different neighborhoods within the same city. Meanwhile, an increasing amount of existing apartments is being traded in the housing market. In recent years, traded area reaches 1 million square meters monthly. Clark and Onaka (1983) argue that family event is the most important factor in housing choices; this in the case of China, or even in Asia, is well known to be education. In the past decades, young parents have put a greater emphasis on education when considering the location of new home purchases. (Zhou, 2011, Shen, 2011) In preparation for the fierce competition children will face when applying for colleges, parents make great efforts to get their children the best education. In 1984, the government implemented a nine-year compulsory education policy, where students have to enroll in a public school unless they are admitted by a private school, and the resources for public schools coming from municipal government are allocated by the education department at the district level. According to the Shanghai Statistical Yearbook from 2004 to 2011, more than 90 percent

[1] Relevant news can be found on Renmin Ribao (www.people.com.cn) and Shanghai Securities News (english.cnstock.com)
[2] Relevant news can be found on China Index Academy (industry.soufun.com).

of primary students enroll in public schools every year. Following the *Jiu Jin Ru Xue*[3] policy, students have to enroll in the public elementary school with the specific attendance zones that cover where their *hukou*s[4] are located. This policy encourages parents to act early and purchase property in the neighborhoods corresponding to their preferred schools. According to the news reports from *Information Times* (2011), some banks even offer to loosen their mortgage policy and give discounts to support parent buyers who plan to purchase school-district housing three to five years ahead. However, house price-to-income ratio in Shanghai remains the highest of cities all over the world, which adds a lot of pressure for young parents when considering housing purchases.[5] (Numbeo Property Investment)

Although the literature usually finds that the school quality affects the housing market significantly, few have found evidence for a causal relationship between the two as most studies focused on the United States. A few that studied other countries have also suggested that this school quality effect is specific to each city because of various education and residential policies and sometimes exhibits no significant relationship. (Stadelmann, 2010) On the other hand, little research regarding China has been conducted. The limited literature on China either examines only one school, or compares the absolute differences of housing values rather than percentage changes (Wang and Ge, 2010, Wen, 2011, Song, Yang and Guo, 2010). As will be discussed in the Literature section, both cases could bias results due to the failure in controlling for important factors.

This paper is among the first to inspect how the varying quality of elementary school is capitalized into the existing property prices in the urban area of Shanghai. I used

[3] *Jiu Jin Ru Xue* (就近入学) policy is part of the nine-year compulsory education policy.
[4] *Hukou* is the household certification of citizenship, also required for public school enrollment.
[5] Reports can be found on Numbeo Property Investment (www.numbeo.com)

monthly data from the Shanghai Existing Property Index, which has a sample of apartments with similar structural characteristics. The special nature of the education and housing systems in Shanghai is advantageous to eliminate measurement biases that confound results in existing literature as will be discussed in the Literature section. The capitalization of housing in 2000 can also serve as an exogenous shock to identify a causal relationship between school quality and housing prices because it happened after the 1984 education reform. The Appendix section includes more background of the housing and education systems in Shanghai. The study contributes to the understanding of both the correlation and the causality between school quality and constant-quality housing via evaluating the magnitude of the capitalization of education quality. The results can help explain the emerging housing price disparities in Shanghai, suggest strategies for consumers on housing purchases, and explore the policy implications for narrowing the house-price gaps among different neighborhoods in the Shanghai, such as re-locating public elementary school resources.

The remainder of the paper is organized as follows. Section 2 will first discuss how the Shanghai case fits into the stream of literature. Section 3 will provide the theory background on the hedonic pricing model. Then, the empirical methods are introduced in Section 4 and 5. These methods include integrating housing data, education data and geographical data for regression, adjusting the empirical regression function and correcting for statistical biases in the regression analysis. Section 5 presents the main results and discussion, and Section 6 demonstrates robustness checks done to support main results. Finally, Section 7 contains the concluding remarks.

2. Literature

2.1 Evolution of the Pricing Models

Pricing models have been widely used in studying the endogeneity between property values and school quality. Tiebout (1956) and Oates (1969) first tried to draw attention to how local public services were capitalized into urban land markets. Lancaster (1966) developed the microeconomics foundations for pricing housing. He focused on utility-generating characteristics of the goods instead of the goods per se, using the tools of "activity analysis." His theory was not limited to the housing markets but also applied to diverse topics such as financial assets and the demand for money.

Rosen (1974) followed in this direction by focusing on characteristics, and first developed the hedonic pricing model (HPM). In this model, the implicit price of a house is a nonlinear function of its structural characteristics, the school quality in its neighborhood, and its non-school neighborhood characteristics. Although his theoretical work has been widely acknowledged, cited and employed, he didn't elaborate on how to carry out the estimation of such characteristic variables: it is hard to include all relevant parameters and comprehensive estimations of each of them. Therefore, while I followed this model closely in this study, I made several modifications to tailor it to the Shanghai case specifically.

2.2 Measuring the Structural Characteristics of Housing

The first problem researchers encountered when using this model to evaluate how school quality affects housing prices is how to obtain the "constant-quality houses." In other words, the problem is how to control for the structural characteristics in order to rule out the effects of houses per se on their prices.

Earlier studies controlled for basic housing structures including the number of bedrooms, number of bathrooms, lot size and internal square footage. (Walden, 1990, Figlio and Lucas, 2004, Din, Hoesli, and Bender, 2001) However, controlling these factors doesn't necessarily give us the constant-quality houses or reflect the extent of capitalization accurately.

For example, the housing age and bedroom facing direction[6] are important to distinguish house quality but they were not included in earlier studies in this field. Few researchers who studied the United States took apartments into consideration, and many simply failed to distinguish apartments from houses. In addition, Brasington (1999), Cheshire and Sheppard (2004), and Hilber and Mayer (2009) argue that the elasticity of the housing supply influences the extent to which school quality is reflected in the housing prices. Figlio and Lucas (2004) find a large effect of school grading in the year following grade imposition, indicating a change in elasticity of demand, but they have not explained whether these effects were permanent or transitory in their study. Cheshire and Sheppard (2004) find the school quality to be significantly and additionally discounted in areas where new construction is concentrated. They also find that researchers hardly exclude the cases when the houses or apartments in a school district are not suitable to those families with children in general and thus can obtain an omission bias or a discounting bias in the results.

Furthermore, because researchers include different combinations of structural characteristics, the results obtained by different research studies are not fully parallel and thus are not fully comparable.

[6] The bedroom facing direction is the direction toward which the bedroom window opens. It is the most important criterion in *Feng Shui* to evaluate the quality of the apartment. Most bedrooms in China face south because of this.

2.3 Measuring School Quality

Brown and Moore (1970) use the concept of "place utility" to measure the balance of residential satisfaction and dissatisfaction. According to their model, households prefer to move to locations with better amenities. In addition to ethnic composition, floor space, the number of rooms and different infrastructures, Clark and Onaka (1983) argue that family events are significant components in explaining housing choices. In Shanghai's case, this is most likely to be the availability of the best education, because society widely acknowledges that Chinese parents evaluate education to be the priority of all family events. (*Economics Information*, 2009)

When showing the existence of the school-quality premium, most previous studies use single measures for school quality. This raises the second group of concerns about the validity and efficiency of the employed measurement, and thus whether the measured aspects of school quality accurately reflect the premiums. Researchers have proposed three major categories of measures: input-based, output-based, and value-added.

Input-based measures, such as per-pupil spending and reduced-price lunch in the district, were popular among earlier studies. Among output-based measures, standardized test scores are the most popular choice. Black (1999), Weimer and Wolkoff (2001), and Downes and Zabel (2002) combine both input-based and output-based measures in their studies. However, Hanushek (1997) find no apparent impact of these inputs on student achievement. He thus claims that the input-based measures are inappropriate for education quality. His findings have led to a more prevalent use of output-based measures in recent studies.

Researchers, such as Haurin and Brasington (1996), Hayes and Taylor (1996), Bogart and Cromwell (1997), Black (1999), and Cheshire and Sheppard (2004), have all used single-subject standardized test scores. Unfortunately, these test scores usually correlate with socio-economic factors which are also taken into the regression models as controlled variables. This multicollinearity problem makes it hard for the researchers to separate the effects of school quality from those of its correlated factors, which will be discussed in detail later.

Hayes and Taylor (1996) have incorporated multiple measures by using both math test scores and student body characteristics. Weimer and Wolkoff (2001) criticize Hayes and Taylor's method, arguing that they should not leave out the peer group effects. Some researchers suggest that the more appropriate measure should be value-added measures of achievement. Hanushek (1997) states in his education production functions that these measures usually evaluate the marginal improvement in a particular cohort's performance over a period of time. However, because tracking groups over time adds sophistication to the decision-making processes of house buyers, such measures are empirically hard to construct. Moreover, Brasington (1999), Brasington and Haurin (2006) and other researchers find little support for using the value-added measures. They point out that home buyers usually favor traditional measures in their housing valuation.

2.4 Defining School Districts

In addition to the measurement problem, there are problems with the definition of school districts. Most researchers do not include either the details about how they divide school districts or the definition of the school districts in their studies. According to Cheshire and Sheppard (2004), some researchers have included the nearby schools rather

than schools with the corresponding attendance zone in the regression. They conclude that since most attendance zones studied didn't fully overlap with administrative districts, biases would be introduced and the validity of those results could be largely affected. This is consistent with what Bogart and Cromwell (2000), who found that "access to neighborhood schools has a substantial effect on house prices" in their subsequent study of redistricting in Shaker Heights following their 1997 study, implying that top private schools could also be examined in studying the school district effect on real estate prices. Unfortunately, few studies have taken such approaches so far.

When it comes to choosing what levels of schools to incorporate into the models, different researchers used different standards. Weimer and Wolkoff (2001) find small effects from high schools but significant results from elementary schools. On the other hand, Cheshire and Sheppard (2004) claim that the secondary schools have a larger effect than elementary schools, while Chiodo, Hernandez-Murillo, and Owyang (2010) do not disaggregate the potential different effects from the two; they simply integrate elementary and secondary schools by using the school-level index generated by the Missouri Department of Elementary and Secondary Education. These studies, however, share the problem that they use only math scores in their research. Researchers, such as Bifulco and Ladd (2007), Bischoff (2008), and Berends and Penaloza (2010), have all revealed that test scores are strongly correlated with socio-economic factors, such as racial compositions and gender composition.

Other research find no significant effect of school quality on house prices in societies composed of relatively homogeneous culture background. Stadelmann (2010) finds, in Switzerland, the average distance to the next school mattered instead. He

explains that this insignificance might also be due to the small role private schools play in education in the Swiss context. These studies suggest a more careful look at different culture contexts in this area of study.

2.5 Other Concerns with Previous Empirical Methods

Researchers raise their concerns about empirically applying HPM due to the substantial interaction between neighborhood characteristics and school quality. For instance, it is hard to separate the school quality premiums from the premiums caused by non-school neighborhood characteristics. Hanushek (1997), Bischoff (2008), and Chiodo, Hernandez-Murillo, and Owyang (2010) have all pointed out that there is a significant correlation between socio-economic factors of the neighborhoods, such as racial composition and income, and school district quality. Since most literature in this field examines the school district effect in a neighborhood composed of various ethnic groups, namely cities and counties in the United States, this multicollinearity problem of the controlled variables in the models would lead to upward biases in these results. As Fischel (2001) mentions in his study, "Everything seems to be capitalized."

Black (1999) also argues that previous research using HPM has introduced an upward bias from neighborhood quality effects that are unaccounted for in the data. She instead uses an excellent method to control for non-school neighborhood characteristics by looking at houses on the boundaries of school attendance zones. This method is widely cited and praised except that it cannot separate the effect of the school zones from that of the other local public services zones if these zones overlap in a large degree. In addition, she uses single-family residences as her method in controlling for constant-quality houses. Her large standard deviation in all lot sizes indicated a large variation in

the house properties. As discussed above, the big variation of structural characteristics can potentially introduce a discounted effect since certain characteristics of the house structure might not be suitable for families with children.

On the other hand, Stadelmann (2010) find that most demographic and other socioeconomic characteristics were of minor importance in the Swiss case. This finding suggests that regions with relatively homogeneous racial compositions might help eliminate the interactions between factors discussed above when evaluating the school quality effect on real estate prices.

2.6 Research on the Shanghai Case

My research will focus on Shanghai, which offers several advantages over the existing research.

First, only a few studies have looked at the school quality effect outside U.S., such as Gibbons and Machin (2003) on England, and Stadelmann (2010) on Switzerland. Even fewer have inspected cases in Asia. This study contributes to the current literature to build a more comprehensive understanding of the school-district effect, especially given the special natures of the education and housing systems in Shanghai, with its different cultural context. Second, we can largely reduce the multi-collinearity problem in the case of Shanghai, as the cohort of citizens there has a relatively homogeneous racial composition and even levels of socio-economic factors such as income level. Third, real estate within the urban area is mainly apartment-based across all districts. These characteristics will help separate the effects of school quality from those of the controlled variables in HPM, and thus relieve the concerns from the previous researchers.

Furthermore, this study employs the Shanghai Existing Property Index. This index report was initiated in 2002. In this index, all apartments chosen share the same set of structural characteristics. In terms of the school quality of the elementary schools in the neighborhoods, I use a top-tier measurement that integrates data from education commissions, online search results, and personal connections. The department of education in each district publishes a compulsory education report every year, and all three types of measures, including input-based, output-based, and value-added, have been used to create each report upon evaluation. These reports, along with the information obtained through a private connection, will be used to confirm the top search-engine results generated in this research to help rank a public school as top-tier or not. All of these advantages help this research contribute to the current literature in urban planning, education and pricing models.

3. Theory Background

One approach to evaluating how consumers make their purchasing decisions is the well-known cost-benefit analysis. In the real estate market, the benefits one gains from buying a residence can vary according to individual preferences on structure, location and other dimensions. The most widely adopted method is the pricing model where we focus on the capitalization of the benefits into the product prices. To understand the hedonic pricing analysis, one should think of a residence as a bundle of characteristics that get capitalized into the price of the residence. Although there are many benefits one can get from the purchase and not all of them will be incorporated into the pricing model, we can study the explicit factors that influence consumer intention that might also be the dominant ones.

3.1 The Hedonic Pricing Model (HPM)

Rosen (1974) defines a differentiated product, Z, by its various characteristics, z_1, z_2, …, z_n. This well-established model can be used to quantify different effects from separated utility-bearing attributes of a certain good. Most recent studies then use Rosen's hedonic pricing model (HPM) when studying the capitalization of school quality into house prices. This research will also closely follow Rosen (1974) in applying hedonic pricing to Shanghai's existing property market. In this section, I will first introduce this model and then give an extension of hedonic price model to formalize the argument that better regional-school quality leads to the higher capitalization premium in housing prices. In addition, I explain how the Shanghai case could potentially solve the endogeneity problem of school quality, which turns out to be a central concern in previous studies of the hedonic pricing model.

3.2 Implicit Price of Attributes

Because the HPM values a good as a differentiated bundle of characteristics that embodies the good, I made a few assumptions before digging deeper into the theoretical background of hedonic pricing model, following the work from Anderson and West (2006), and Clark and Herrin (2000):

(a) There exists a large number of differentiated products and sufficient variation such that the range of product choices is essentially continuous.

(b) Housing bundles cannot be untied linearly according to consumers' tastes and preferences. The assumption of indivisibility makes arbitrage impossible. (Rosen, 1974) For example, the benefits obtained from two 500-square-foot apartments are not equivalent to one 1000-square-foot apartment (case [i]);

while a 500-square-foot residence for half a year and an 1000-square-foot residence for the other half is not the same as a 750-square-foot all year round (case [ii]).

(c) Generally, sellers either cannot or lack intention to repackage existing products as the type of repackaging or reassembling is costly or impossible.

(d) All buyers and sellers in the market have symmetric information about the bundle of attributes. In other words, the perceived characteristics that get capitalized into the prices are the same on either side of the market.

These assumptions lead to the nonlinear hedonic equation and also allow us to segregate housing prices into prices of the embodied attributes, which can be represented as *p(z)*, where $z = z_1, z_2, z_3, \dots, z_n$ is a vector of the perceived bundle of attributes and strictly one unit of z_i is consumed in each case. The implicit price of attribute i can thus be given by the partial derivative of *p(z)* with respect to attribute i, $p(z_i) = \frac{\partial p(Z)}{\partial (z_i)}$. In empirics, the marginal implicit prices of residence characteristics may be obtained by regressing housing value on a vector of residence characteristics.

3.3 Equilibrium in Shanghai Housing Market

Two major groups of players in the housing market in Shanghai, buyers and sellers, serve as the utility-maximizing consumers and profit-maximizing firms in Rosen's model, and determine the equilibrium hedonic price function together.

3.3.1 Demand Side Story

For buyers, namely young parents, the availability of substitutes for good public schools are limited. In Shanghai, more than 90% of the students enroll in public elementary schools (Shanghai Statistical Yearbook, 2004 to 2011). The remaining private

schools usually charge a sponsorship as an admission premium in addition to the special recruitment process, which means a higher willingness to pay from the parents doesn't necessarily guarantee a seat in the classroom for their children in the private section. On the other hand, children of those who hold *hukous* within the attendance zones of good public schools are guaranteed to be allocated to those schools according to the "Jiu Jin Ru Xue" policy[7] where there is only one public primary school available in each school attendance zone. Therefore, parents consider school quality as the major factor when choosing location for future apartment purchases. Furthermore, perceived school quality, or school reputation, is what the buyers actually use during their decision-making processes and what the sellers promote in their advertisements. This is also consistent with the assumption of symmetric information about the perceived bundle of attributes[8] of housing.

Therefore, we know that the buyers differ in terms of socioeconomic characteristics according to a vector of different socioeconomics, a and are maximizing utility,

$$U(x, z_1, z_2, z_3, \dots, z_n; a)$$

by choosing the bundle of characteristics Z and a composition good x, subject to the budget constraint,

$$y = x + p(Z)$$

where y is the consumer's income, the price of x is 1, and $p(z)$ is nonlinear as mentioned above. The first-order condition for utility maximization is satisfied by setting the

[7] This nationwide education policy started in 1984 is equivalent to the western concept of school attendance zone.
[8] Perceived bundle of attributes refers to the bundle that the consumer cares about when making a purchase.

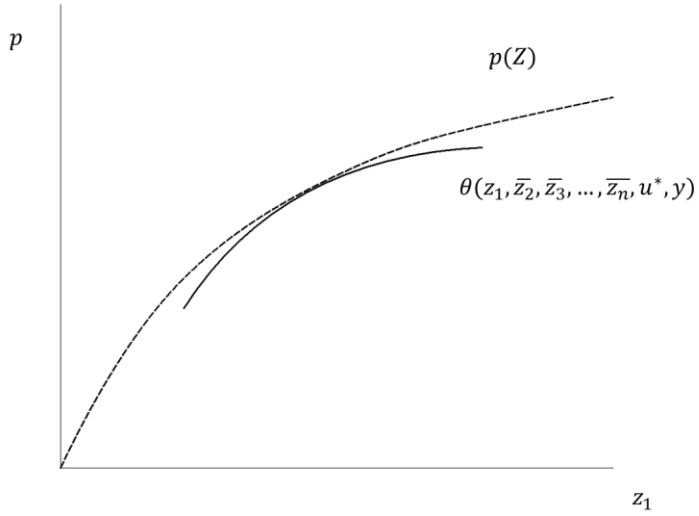
marginal rate of substitution between that characteristic and other goods in the bundle equal to the marginal price[9] of each characteristic,

$$\frac{U_{z_i}}{U_x} = \frac{\partial U/\partial z_i}{\partial U/\partial x} = \frac{\partial p(Z)}{\partial z_i} = p(z_i)$$

To express the consumer's willingness-to-pay, a bid function, $\theta(Z; u, y)$, can be defined implicitly for the particular bundle of characteristics Z, according to

$$U(y - \theta, Z; a) = u$$

at a given utility index u and level of income y. This modified utility is maximized when the marginal bid with respect to a certain characteristic equals the marginal implicit price for that characteristic. Graphically, this optimization occurs when the bid curve is tangent to the implicit price curve under the phase of each characteristic in the bundle (Anderson and West, 2006).



3.3.2 Supply Side Story

On the supply side, the housing market in Shanghai has been reaching a relatively saturated status in the past twenty years, especially within the middle ring where most top-tier elementary schools are located (Shanghai Local Chronicles Office, Shanghai Statistical Yearbook, 2004 to 2011). The supply is thus inelastic, especially within the existing property market. As a result of the high population density and city crowdedness,

[9] Here marginal price can also be interpreted as unit price if the characteristic is measured on a discrete scale.

space is limited while the existing property can hardly be improved except for interior remodeling. Meanwhile, the rapid growth of the city doubled the number of available metro lines in the urban area over the past 10 years so that taking the metro is faster and cheaper than driving, and thus becomes the major way of commuting.[10] As a matter of fact, owners of second-hand apartments can benefit from these urban structure improvements without much added costs.

Therefore, in the optimization problem, sellers choose a desired bundle of characteristics of the housing good, Z, and the number of housing good offered, M, at the given technology factor, b, to have the corresponding level of cost[11],

$$C(M, z_1, z_2, z_3, \dots, z_n; b)$$

so as to maximize the profits, which is simply revenues minus cost,

$$\underset{choose\ M,Z}{max}\ \pi = M \cdot p(Z) - C(M, Z; b)$$

Again, the first-order conditions are satisfied by setting the marginal cost of each characteristic per unit sold equal to the marginal implicit price of that characteristic,

$$C_{z_i}/M = \frac{\partial C/\partial z_i}{M} = \frac{\partial p(Z)}{\partial z_i} = p(z_i)$$

Then, similar to the demand side story, an offer function, $\varphi(Z; \pi, b)$, can be defined to represent the seller's willingness to sell at constant profit π and indifferent technology level b, which then gives the modified first-order condition,

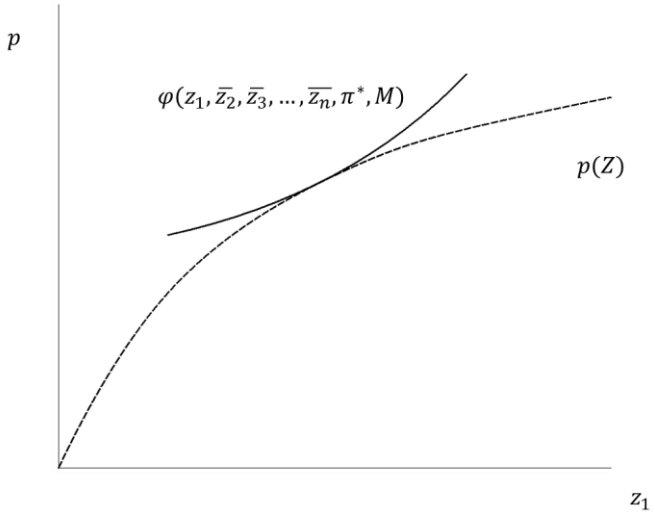
$$\varphi = C_M(M, Z)$$

[10] In Shanghai, neighborhoods are regulated to meet the facility standards, including the bus system. Therefore, even the bus is still one of the most important forms of transportation, the conditions are relatively homogeneous across neighborhoods within the urban area, and thus are not utilized to control for transportation accessibility in this paper.

[11] The cost for sellers here refers to the amount sellers paid at purchase in the existing property market.

This modified function helps us to solve for φ in terms of Z, π, and b by eliminating the number M.

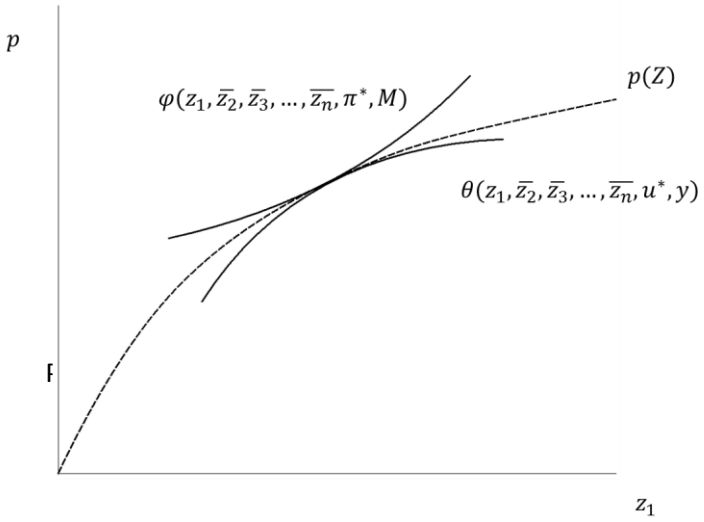
Graphically, the marginal condition of these offer curves that trace out a set of indifferent surfaces can be represented by their tangency with an implicit price curve at each characteristic dimension.



3.3.3 Market Equilibrium

Noticing that $p(Z)$ is the smallest amount a consumer would need to pay in the market while the maximum price accepted by the seller, both players can reach their optimums at $p(Z)$. For the market to clear, buyers' willingness to pay must meet sellers' willingness to offer, in other words, both curves are tangent to the implicit price curve while at the same time, the two curves are tangent with each other. If buyers are identical, $p(Z)$ is everywhere identical with pay functions, and thus identifies the structure of demand. On the other hand, if sellers are identical, $p(Z)$ is everywhere identical with offer functions, and thus identifies the structure of supply.

The solution function for the price of a given bundle of only one attribute can be



obtained using Euler's equation solution (Rosen, 1974). The equilibrium price function is generally thought to be nonlinear because the cost of arbitrage activity that repackages bundles of attributes once a house is built is assumed prohibitive (Clark and Herrin, 2000).[12] Here is a commonly used functional form that consists of all three major categories of house characteristics:

$$\ln(p_{ijt}) = \alpha + \beta_1 S_{it} + \beta_2 Q_{jt} + \beta_3 N_{jt} + \varepsilon_{ijt}$$

where p_{ijt} is the price of residence *i* in neighborhood *j* at time *t*. The vector Q_{jt} is the major interest, school quality in neighborhood *j* at time *t*. The vector S_{it} includes all utility-generating structural characteristics of house *i* at time *t*, such as number of bedrooms and bathrooms, while N_{jt} denotes the vector of all neighborhood characteristics at that time. Since all characteristics are utility-generating, I expect positive signs and significant values for all coefficients in the equation, including β_2 for school quality. This is consistent with the hypothesis that school quality is capitalized into housing prices as one utility-generating characteristic when everything else is fixed. The transformation for the Shanghai case is presented at the beginning of the *Empirical Results*, which in turns leads to the guiding equation for empirical analysis of a multi-characteristic model.

4. Summary Statistics

4.1 Housing Data in Shanghai

Property prices are given by the Shanghai Existing Property Index, retrieved from the Shanghai Existing Property Index Office. This index was initiated in 2002 and provides monthly prices available on 51 control points from 2004 to 2011, where each control point takes its own price in 2002 as index 1.

[12] Please refer to assumption c in section 3.2.

The sampling standards for the Shanghai Existing Property Index are:

1. The residential building was 15 years old in 2002.
2. The residential building has a brick masonry structure.
3. The apartment is located on the third floor of the building that has six to eight levels.
4. The apartment has average interior decoration with all bedrooms facing south, one living room, one bathroom, and one kitchen.
5. The area of the apartment is close to 35 m^2 for a single-bedroom structure, 66 m^2 for a double-bedroom structure, and 88 m^2 for a triple-bedroom structure.

These data originally came with different report dates in two different forms: original prices and price indices. The original prices were reported either by apartment or by unit area.[13] A sample of the original reports is included in the appendix as Table 1. I combined all data into an index-only dataset where the value in 2002 serves as base 1 for each control point index, and thus obtained 82 or 83 observations for each apartment with three sets of indices (oneBedroom, twoBedroom, threeBedroom) whose standard sizes are given in Table 2.

4.2 Using eWOM Messages to Measure School Quality

Shanghai used to have "key elementary schools" that were widely acknowledged as top-tier schools with the best education quality, before the open policy in 1978. Some educationalists, including Dongchang He, questioned this system as parents and students focused too much on the college entrance rate of young children.[14] In 1982, when

[13] According to the official description, each by-unit-area price value was calculated by dividing the total price of the apartment by the mean area of its apartment structure. For example, if a single-bedroom apartment has a 500,000 RMB value on the report day, the by-unit price is 500,000 (RMB) / 35 (m^2).
[14] He also stated that the system led to a higher rate of nearsightedness among young kids in Shanghai.

Dongchang He was appointed to the head of the Ministry of Education, he encouraged local government to cancel this "key school" system in major cities (He, 2009; Jiang, 2009).

The ideal school-quality data should come from this "key school" information. Although the district education commissions in Shanghai have lists of district-acknowledged "key schools," these lists are kept private and are not for public use in order to maintain the perception of equality in public education resources.[15]

Although China is a developing country, it has the largest group of internet users in the world at more than 513 million people; this number covers about 32.7 percent of China's population in 2011 according to International Telecommunication Union (*Wikipedia*).[16] Online consumer discussion forums are extremely popular and provide a virtual avenue for users to share their consumer opinions via the Internet (Cheung, Luo, Sia and Chen, 2009). This results in a new prevalence in electronic Word-of-Mouth (namely, eWOM) evaluations for products. Numerous studies have shown the influential role eWOM plays in consumer intentions (e.g., Cheung, Luo, Sia and Chen, 2009, Hennig-Thurau and Walsh, 2003, Zhang, Craciun and Shin, 2010). These studies show that the information provided by online opinion platforms can exert a strong influence on consumer buying and communication behavior. Cross-cultural studies also state that this influence varies across different cultures and is stronger in collectivist cultures such as China, where people care more of the opinions from others (Cheung, Luo, Sia and Chen, 2009, Morahan-Martin, 2004). When it comes to the different behaviors of consumers in online forums, low-involvement consumers are affected by review quantity rather than

[15] This is because the district education commissions allocate resource and funding in a rather arbitrary way.
[16] http://en.wikipedia.org/wiki/List_of_countries_by_number_of_Internet_users

quality – as their major motive in using eWOM messages is to save time in their decision-making processes (Park, Lee and Han, 2007). In a fast-pace city with an enormous amount of information like Shanghai, most consumers act as low-involvement. They tend to use search engines, especially Baidu.com, in obtaining information as fast as possible. Page rank thus comes into play where pages with higher search engine ranks get more visits and become more popular (Cho and Roy, 2004). Especially for service purchases like elementary-school education, the WOM messages in reality and virtual contexts become one of the major approaches consumers take in their decision-making processes.

The proxy for school quality in this research comes from the lists of "top-tier schools" that are widely referred to in the eWOM messages, as most consumers in China are low-involvement consumers who care more of the quantity of the online forum messages when making purchases (Park, Lee and Han, 2007). I used the major search engine in China, Baidu. According to Forbes in 2011, Baidu has earned the moniker "Google of China" with a dominant 75% share of the market. The keywords used for district A in year B are "key school," "district A," and "year B."[17] A school is then identified as top-tier if it appears frequently in the search results; otherwise it is non-top-tier. In general, the page rank effect on the popularity of web pages indicates that internet users visit those pages that show up in the top search results more (e.g., Cho and Roy, 2004), and thus the threshold of frequent appearances for the main regression is set at 50% of the first 50 results. Other thresholds will be tested in the robustness checks. Since most parents in China either go to the major online forums and school information sessions or

[17] For example, the keywords for top-tier school in Xuhui district in 2006 are "key elementary school", "Xuhui district", and "2006"; or "重点小学 徐汇区 2006".

learn through acquaintances to find the "key school" information and advice on good education resources,[18] these "top-tier" data, instead of other publicized school resources, are what buyers actually refer to when making purchase decisions and thus are the real utility-generating characteristics of the housing goods. Although "key elementary school" data are not all publicly available, I obtained one list for the Xuhui ("xh") district through personal connection. Thus, this single list is used to verify the results obtained from eWOM described above. In fact, the schools on that list agree with those top-tier schools published on major online education forums, which provides support for the validity of the top-tier measure under the Shanghai education system.[19]

4.3 Matching Housing Data and Education Data

School attendance zones for elementary schools in Shanghai are available on some but not all websites of the district education commissions. This paper included six districts that have complete data published from 2004 to 2011. A sample of the administrative map, control points for the housing index, and a sample year of school attendance zone for one district ("cn" district from the six districts in total) are included in the appendix (Figure 1, 2, and 3). No attendance zones overlap and each neighborhood is assigned to only one public school in each district. Private schools have the rest of the student body, which is no more than ten percent every year (Shanghai Statistical Yearbook, 2004 to 2011). Some private schools openly recruit from the entire city while others can only admit students within the district. The private schools will thus not be

[18] Numerous articles and reports can be found in major newspapers, online news sections, and the education channels of major websites, such as edu.ifeng.com (*Feng Huang Wang*, 凤凰网), edu.sina.com.cn (*Xin Lang Wang*, 新浪网), etc.

[19] In 2009 and 2011, there are also posts on private blogs that compared schools that had the most students awarded in major math and English competitions with the best schools listed in the forums and found the two correspond to each other. This gives additional support for using school reputation to proxy for school quality in Shanghai.

considered in the main regressions because of the small share they have in the Shanghai education system, but some top-tier ones will be examined in the robustness tests.

Combining top-tier school data, the administrative map data and the school attendance zone data, a school quality variable is then generated for each housing control point in the dataset. The values of this variable are indicated as a dummy variable, where 1 indicates that the corresponding public school for this neighborhood is top-tier and 0 indicates non-top-tier. In my robustness check, I present a more continuous measure, namely segregating the school quality into more than 2 tiers.

The integrated data set contains the monthly housing index of three structures for thirty control points of the housing index across six districts in the urban area of Shanghai from 2004 to 2011. Each observation has the following variables associated:

1. "topTier" is the top-tier school indicator for the availability of a top-tier primary in the neighborhood;
2. Non-school neighborhood characteristics including "shopping" for whether the neighborhood lies within a major shopping area, and "metro" for the availability of the metro stations nearby;
3. Geographical information including "neighbor" for the neighborhood and a set of 6 "district" dummy variables for the district that neighborhood lies in;
4. Macro level indicators including "interest" for interest rate[20] and "cpiadjusted" for adjusted monthly CPI level in the given month of the given year to proxy for inflation level, which uses the values from the same month in 2003 as a base 100;

[20] Interest rate data are short-term interest rates (per cent per annum) published on Organisation for Economic Co-operation and Development (OECD) website.

5. Month data represented as "dateYM."

A sample plot of single-bedroom indices over time for one of the six districts is given in the appendix (Figure 6). These preliminary results suggest that most areas with top-tier schools available in the neighborhood have faster growth in prices even when compared to the shopping areas that don't have top-tier schools. I will present a more careful analysis on empirical methodology used and results obtained in both the empirical section and the robustness section.

Other neighborhood characteristics are also collected, including transportation accessibility and Center Business Division. The metro is the most widely used way to commute in Shanghai, as it is the convenient public transportation system (versus heavy highway traffic and expensive taxi fare). Current Metro lines are categorized into major lines (line 1 to line 4) that were put into operation before 2004, and another 7 lines (line 5 to line 11) that were built after 2004.[21] Availability of metro station and major line metro station are used as proxies for transportation accessibility. Shanghai has 6 Major Central Business Districts (CBD) and 7 Major Shopping Areas. Whether a neighborhood is within any of these areas, along with the number of metro stops needed to get to People's Square, will be used as measures for distance to the center of the city. Interest rates and CPI are also collected in this paper to adjust for macroeconomic factors over years in the panel data.

The full panel sample consists of 6 districts in Shanghai over 8 years from 2004 to 2011. These are the 30 controlled points in total from the Shanghai Existing Property Index, where school attendance zone information is available. The full summary statistics

[21] Line 13 is Expo line which was only used during 2010 EXPO

are given in Table 3. Figure 1 is a sample district with neighborhood, controlled points and school attendance zone labeled respectively in Figure 2 and 3.

5. Empirical Analysis

5.1 Functional Form Specification

Recall that the model obtained from the hedonic pricing model is in the following form:

$$\ln(p_{ijt}) = \alpha + \beta_1 S_{it} + \beta_2 Q_{jt} + \beta_3 N_{jt} + \varepsilon_{ijt} \quad (2)$$

In this form, α is the constant term, ε_{ijt} is the error term. Recall that p_{ijt} is the real price of the real estate for the residence *i* in neighborhood *j* at time *t*, thus the price index can be a measure to interpret $\ln(p_{ijt})$ as it gives the percentage change in real price, p_{ijt}. S_{it} is a vector of the structural characteristics that are utility-generating for the resident *i* at time *t*. The school-quality variable Q_{jt} represents the accessibility to a top-tier elementary school in the neighborhood *j* at time *t*. N_{jt} is the vector of all non-school neighborhood characteristics that generate utility by purchasing the resident *i* at time *t*.

To further tailor the empirical regression equation to the Shanghai case, I re-write the empirical equation as below:

$$index_{it} = \alpha + \beta_{1it} q_{it} + \beta_{2it} s_{it} + \beta_{3it} m_{it} + \varepsilon_{it} \quad (3)$$

The summation term of the structural characteristics is eliminated and a time subscript *t* is remained because the sample apartments collected in the Shanghai Existing Property Index report share the same major structural characteristics at different times, and thus S_{it} is already controlled on the left-hand-side of the equation. The index of this apartment *j* at time *t* is thus the ratio of the current price over its base price. In addition, among all

control points chosen in this index, no two apartments selected belong to the same neighborhood. In other words, all control points have distinct neighborhood information. Therefore, the index serves as a measure for $\ln(p_{ij})$ in the original form. This form represents how much of each characteristic is capitalized in the housing index as one unit change in index is equivalent to a 100% change in the original value of the specific property. The neighborhoods in the urban area of Shanghai are usually considered relatively similar as mentioned in the Shanghai Housing System section because of the standardization of neighborhood facilities in Shanghai. Therefore, it is reasonable to assume the homogeneity of neighborhoods in Shanghai, as previous researchers have done. (Chen and Hao, 2004) Therefore, only two additional neighborhood vectors, shopping area (s_{it}) and metro station (m_{it}) are introduced as non-school neighborhood characteristics in the Shanghai case to proxy for distance to the center of business and transportation accessibility respectively. The shopping area dummy, s_{it} is 1 if the neighborhood *i* is within the shopping area, while the metro dummy, m_{it} is 1 if the neighborhood has access to at least one metro station. A positive sign is expected for the coefficient on each of the neighborhood characteristics as of their utility-generating nature. Besides these two factors,

5.2 Regression Methodology

5.2.1 Macroeconomics Factors

Though the dataset in this study appears to be panel data, this study focuses on the cross-sectional variation of the effects a top-tier school has on the residence prices in the neighborhood. Since the time span for the monthly data set here is about 8 years long, macro indicators, namely interest rate, r_t, and city CPI, p_t are included in the model to

control for loan policy changes and inflation respectively. These two macro indicators can help fix the variation over time and to allow us observe the variation of property prices among different neighborhoods in Shanghai regarding the availability of the best schools at a given time. This paper also includes the district variable d_{it} to control for other variations across districts when conducting the regressions.

Therefore, I obtained the governing equation for regression tailored to the Shanghai case:

$$index_{it} = \alpha + \beta_{1it} q_{it} + \beta_{2it} s_{it} + \beta_{3it} m_{it} + \beta_{4t} r_t + \beta_{5t} p_t + \beta_{6it} d_{it} + \varepsilon_{it} \quad (4)$$

I expect a negative sign for β_{4t} on r_t as higher interest rates discourages consumers' incentives to obtain mortgages. Otherwise, I expect a positive sign for all other β's because of their utility-generating nature.

5.2.2 Model Specification

A multicollinearity test is conducted on all explanatory variables and the results are given in Table 4. No explanatory variable is correlated with school quality variable. This might due to the nature of the Shanghai education system that is not largely influenced by macro policies. Except for the time variable ("dateYM", the record date of an observation) and the inflation variable ("cpiadjusted", the adjusted CPI level at the given time), no other multicollinearity problem exists in these model specifications. However, these two correlated variables are not included in any model specifications at the same time in the research, and thus would not cause biases to the results.

Using simple Ordinary Least Squared (OLS), I tested five different model specifications in Table 5. Model V attempted to capture the joint effect of all six districts in the model by using "areg" command with "absorb" option in Stata, which gives the F-

statistic value for the six districts as a group. The F-statistic shows a significant correlation between district ("districtDum") and the dependent variable, housing price index, which re-assures the addition of district variable d_{it} in the main equation.

Among all five different specifications, the final form is the best fit with the specified model and indeed has the highest adjusted R-squared. This model has all coefficients statistically significant with expected signs except for the "metro" which represents the transportation accessibility. I will discuss possible explanations for this in Results section.

5.2.3 Statistical Correction

I test for heteroskedasticity and serial correlation problems using a Breusch-Pagan test and an autocorrelation test provided by David Drukker (2003) in Stata respectively. Literature suggests when both problems exist, one can use Generalized Least Squared (GLS) corrected for ARMA errors to correct both problems if the sample size here is large enough and the models exhibit autocorrelation and/or moving average errors at the same time (Yaffee, 2003; STATA, 2003; Sayrs, 1989). Therefore, GLS regression is used instead of simple OLS in this study.[22]

Again, the primary goal of this paper is to understand the cross-spatial variation of school district effect on housing prices. Therefore, Table 6 presents a comparison between between-effect regression, and GLS regression after correcting for heteroskedasticity and autocorrelation.[23] All coefficients of interest have the expected signs in between-effect inspection and they are statistically significant with similar

[22] We use the options "corr(ar1)" and "p(h)" under "xtgls" regression in Stata to correct for heteroskedasticity and autocorrelation respectively.

[23] A Hausman Test was performed on the random-effect results and between-effect results and it failed to reject the null hypothesis where there is systematic difference between the two. Therefore, between-effect results are presented and discussed here.

magnitudes in the corrected GLS regressions except for the interest-rate variable. This might result from the multicollinearity problem macroeconomics indicators have, as mentioned above. A Hausman test was also conducted but failed to reject the null that the difference in coefficients is not systematic between random-effect and between-effect regressions. Therefore, it supports that both are consistent, and thus allows us to focus on the GLS regression results.

5.3 Results and Discussion

As discussed in the previous section, we will focus on the GLS regression results given in Table 7 for one-bedroom, two-bedroom, and three-bedroom apartment index reports, which are mostly significant at the one-percent level. The following subsections will present discussion on the capitalization of school quality, the causal relationship suggested, and other interesting findings, respectively.

5.3.1 Capitalization

The coefficients on the variable, top-tier represent the absolute percentage price difference between having access to top-tier elementary schools and non-top-tier for the default district, district 1 (cn). To get the absolute percentage price difference for the other five districts in the sample, we can add top-tier coefficients on those of each district. The reformatted results are given in Table 8.

The minimum and maximum values are highlighted in green and red respectively for each district. The wide range from 15.5 percent to 69.7 percent indicates large variation of price disparities across different districts in urban area of Shanghai. The average values by district show that the wide range is mainly captured when we break

down results by district, while the effects are similar among categories when we break down by apartment types.

On the average, having a top-tier elementary school can increase the housing price by 41.9 percent, which is significantly larger than what most studies have found. The availability of the top-tier elementary school indeed has the largest and dominant impact among all utility-generating characteristics in this model. This agrees with the expectation within the Chinese society that education plays the dominant role in housing purchase decisions.

5.3.2 Causality

This study uses the 2000 housing capitalization as an exogenous shock to suggest a causal relationship between elementary-school quality and house-price disparities. On one hand, the variation of elementary-school quality appeared when housing resources were still employee benefits rather than market products, and stayed relatively stable over the years. On the other hand, it is not until the housing reform in 2000 that the larger disparities in housing market among different neighborhood started to show up. The significant coefficient of the school quality variable in this study thus suggests a one-directional causal relation from school quality to house prices, which can contribute to the literature in solving the concerns about endogeneity.

5.3.3 Other Findings

In this model, the adjusted CPI[24] exhibits substantial positive coefficients. However, one should notice that the range of CPI data is from 1.00 to 1.25 over the past 8 years, this result should be interpreted as a one-percent increase in CPI capitalized into the housing index by about seven percent. In other words, the inflation is captured and

[24] Adjust CPI: for each month, the base 100 is valued at the same month in 2003

positively amplified by about seven times in the housing market. This agrees with the intuition of Chinese society that the housing market is one of the dominant factors in inflation.

The other macro indicator, interest rate displays a negative impact on the increase. However, the results are only significant at five-percent level for the largest apartment type (three bedrooms). This agrees with intuition as consumers who choose larger apartments are more likely to obtain mortgage loans to finance the purchases. Future research can test this hypothesis to better understand the results.

The coefficients of the neighborhood variables are not as influential comparing to other explanatory variables. The shopping area indicator, "shopping", is only significant at one-percent level for the smallest apartment type. This is intuitive when we consider that consumers who choose larger apartments are more likely to also purchase cars, and thus care less about living in the shopping area.

The proxy for transportation accessibility, "metro," is only statistically significant at the 5 percent level for all three types of apartments, but with the opposite expected sign. One explanation for this opposite sign is that some metro stations in the urban area are above the ground, which might generate a noise level that is not utility-generating to some consumers and thus would have not be included in the model according to the assumption of the HPM. Therefore, future research can separate the stations into above-ground ones and underground ones, or look for better proxies such as bus systems to further inspect these explanations.[25]

In Shanghai, all neighborhoods are regulated to be standardized with facilities on grocery shopping, medical care and etc. In addition, there are more than ten central

[25] Table 5 gives the 95% confidence intervals of all coefficients on one-bedroom indices.

shopping areas within the urban region with more than one in each district. Because the convenience and accessibility of public transportation such as the metro lines and the local bus systems are also regulated to be part of the standardized neighborhood facilities in Shanghai, it is relatively easy to get to the city center and other shopping areas which are usually located at the intersection of multiple public transportation routes. These can also contribute to the insignificance of these neighborhood variables.

6. Robustness

To further support the main regression results, three robustness checks are conducted and included in this section. First, I will modify the threshold used for search-engine results in measuring school quality. People can argue that search results threshold set in the main analysis can be arbitrary. Testing different threshold levels can thus help evaluate the strength of the validity of this measure in the main regression. Recall that in the main regression, the threshold used is that a school shows up more than half of the time in the first 50 results. To check its robustness, I set the threshold to 60 percent, 70 percent, and 80 percent, respectively. The list of top-tier schools for each district didn't vary significantly, and the overall data set didn't change after I matched the new list with other variables.

Another different measure used for school quality in the robustness check is to have more tiers evaluating school quality. Instead of distinguishing elementary schools into top-tier and non-top-tier, I used a four-tier system.[26] Schools are categorized into best, good, okay, and never-mentioned using the same eWOM method. I used a slightly lower threshold, 20%, to identify schools in each tier because search results provided

[26] Top-tier schools use "重点"as key word. In the four-tier system, the key words are "重点", "好", "还可以"and nothing for those not-mentioned.

limited information. A sample for Changning district (cn) is included in Figure 3, where we mark "*" for the neighborhoods with a best school, "+" for good school neighborhoods, "-" for okay school neighborhoods, and nothing for the not-mentioned schools. Preliminary results show that there is a significant difference between the "best" schools and the other three tiers as a cohort, but they fail to support any significant difference among the remaining three tiers. This is consistent with the main results of this study, and it agrees with what parents think in China. Shanghai parents care more about whether a school is considered as top-tier or not when making purchasing decisions, and they don't measure school quality in a more continuous way.

My hypothesis is further supported by this third robustness check, an event study on Changning district (cn). The main data set in this study focus demonstrates the cross-spatial variation of school quality but lacks variation over time. Examining the variation over time is also interesting in understanding the causal relationship between school quality and residence prices. Jiang Wu Elementary School is the well-known top-tier school in this district. In 2009, the principal of Jiang Wu Elemenary School was appointed to Jiao Yu Xue Yuan Fu Xiao while he remained in his office at Jiang Wu Elementary School. Then in 2010, Jiao Yu Xue Yuan Fu Xiao officially became a new campus for Jiang Wu Elementary School. The changes in the attendance zones are illustrated in Figure 4, 5 and 6. Although the neighborhood that the main campus of Jiang Wu Elementary School covers is not in the housing-index data set, the new campus is. As shown in Figure 7, 2009 and 2010 are exactly when housing prices started to soar in the Zhongshan Gongyuan (zsgy) neighborhood. The rate of price change in this neighborhood exceeded that of Bei Xin Jing (bxj), which has been leading in Changning

district for the past decade. This event study suggests that future research can look into these special re-location events of top-tier schools in Shanghai to further inspect the effect that variation of education quality has on house prices.

7. Conclusion

In this research, I have sought to quantify the effects school quality has on real estate prices as well as the variation of this effect across different neighborhoods in urban Shanghai. By employing an adapted hedonic pricing model, integrating housing price indices that are controlled for structural characteristics, using eWOM search results to proxy for school quality, and controlling for non-school neighborhood characteristics such as shopping area and transportation accessibility, I was able to not only evaluate the magnitude of the capitalization of school quality, but also suggest evidence that school quality has a causal effect on housing prices.

Main results suggest that the effect of school quality on the price of residences in the neighborhood is statistically significant and serves as the dominant factor in consumers' decision-making processes within the Shanghai housing system. On the average of the six districts in this study, top-tier elementary schools increase the housing prices by about 42 percent, and the effect ranges widely from about 15 percent to 70 percent. In some districts, housing prices depend more on the school quality in the neighborhood than the distance to a center of business district. As for the neighborhood characteristics examined in this study, because of the standardized neighborhood system and mature transportation systems in Shanghai, they are not as influential.

This study focused on the cross-spatial variation of housing prices due to the effect of varying school quality in the neighborhoods. The robustness check indicates the

variation of school quality over time could be of interests as well. The example discussed in robustness also provides strong support in the "school district effect" and causal relationship. Future research can look for more data to further verify these findings.

Looking for more public school quality data can be another improvement to this research. Because the housing data obtained in this study covers only 51 neighborhoods in Shanghai, while only 30 neighborhoods remained after matching up with available data on school quality for over 8 years, some top-tier elementary schools are not studied as a result of lack of housing data in the corresponding neighborhoods. Since the 2005 education reform, each district commission in Shanghai is required to publicize all education information online. Future research can aim to gather more official data over a longer period of time to conduct research that is more robust.

In addition to public schools, some neighborhoods studied have private schools and, thus, it is hard to segregate the effect the private school might have on nearby residences. Therefore, another future avenue of research is to look at the effects of private schools. Consumers use a lot of heuristics and sometimes biases when making purchase decisions. There is evidence that in Shanghai, the housing prices in the neighborhood are greatly affected by a renowned private school, even if the private school has open recruitment or is a boarding school. Since the attendance zones of private schools vary to a great degree and are not as strict as public schools, it would create a downward bias in results.

Instead of the subject-test scores widely used for measuring school quality in the literature, this paper generates the top-tier school data set, which is a combination of government publicized data, search-engine results, and private connection information

due to the limitation of publicly available data. Although these measures might involve subjective judgment, various news and research (*Economics Information*, 2009; *Information Times*, 2011) supports most of them as they compared school reputation with annual major competition results.

Because of the nature of public resource allocation in the Shanghai education and housing systems, the focus on the Shanghai case provides a new perspective in eliminating both endogeneity and multicollinearity problems in the literature. It also provides strategic suggestions to consumers for their purchasing decisions, and meanwhile suggests urban planning policy implications (such as top-tier school re-location) to the government for shrinking price gaps among different neighborhoods in Shanghai. Another value this study can potentially add into the current literature is the data set generated for the Shanghai case. The neighborhood property prices, along with the corresponding education and geographical information are not directly available to the public. This large amount of data manually collected and integrated can serve as a reference data set available for future research.

8. References

Anderson, S. T., & West, S. (2006). Open space, residential property values, and spatial context. *Regional Science and Urban Economics, 36*(6), 773-773-789.

Bartholomew, K., & Ewing, R. (2011). Hedonic price effects of pedestrian- and transit-oriented development. *Journal of Planning Literature, 26*(1), 18-18-34.

Berends, M., & Penaloza, R. V. (2010). Increasing racial isolation and test score gaps in mathematics: A 30-year perspective. *Teachers College Record, 112*(4), 978-978-1007.

Bifulco, R., & Ladd, H. F. (2007). School choice, racial segregation, and test-score gaps: Evidence from NOrth carolina's charter school program. *Journal of Policy Analysis and Management, 26*(1), 31-31-56.

Bischoff, K. (2008). School district fragmentation and racial residential segregation: How do boundaries matter? *Urban Affairs Review, 44*(2), 182-182-217.

Black, S. E. (1999). Do better schools matter? parental evaluation of elementary education. *Quarterly Journal of Economics, 114*(2), 577-577-599.

Bogart, W. T., & Cromwell, B. A. (1997). How much more is a good school district worth? *National Tax Journal, 50*(2), 215-215-32.

Bogart, W. T., & Cromwell, B. A. (2000). How much more is a good school district worth? *Journal of Urban Economics, 47*(2), 280-280-305.

Brasington, D. M. (1999). Which measures of school quality does the housing market value? *The Journal of Real Estate Research, 18*(3), 395.

Brasington, D., & Haurin, D. R. (2006). Educational outcomes and house values: A test of the value added approach. *Journal of Regional Sicence, 46*(2), 245-245-68.

Brown, L. A., & Moore, E. G. (1970). The intra-urban migration process: A perspective. *Geografiska Annaler. Series B, Human Geography, 52*(1), 1-1-13.

Chen, J., & Hao Qianjin. (2007). The importance of location on shanghai housing price. *European Network of Housing Research (ENHR) International Conference - Sustainable Urban Areas,* Rotterdan, Holland.

Chen, J., & Hao, Q. (2008). The impacts of distance to CBD on housing prices in shanghai: A hedonic analysis. *Journal of Chinese Economic and Business Studies, 6*(3), 291-291-302.

Cheshire, P., & Sheppard, S. (2004). Capitalising the value of free schools: The impact of supply characteristics and uncertainty. *The Economic Journal, 114*, 397-397-424.

Cheung, M. Y., Luo, C., Sia, C. L., & Chen, H. (2009). Credibility of electronic word-of-mouth: informational and normative determinants of on-line consumer recommendations. *International Journal of Electronic Commerce, 13*(4), 9-9-38.

Chiodo, A. J., Hernandez-Murillo, R., & Owyang, M. T. (2010). Nonlinear effects of school quality on house prices. *Federal Reserve Bank of St. Louis Review, 92*(3), 185-185-204.

Cho, J., & Roy, S. (2004). Impact of search engines on page popularity. *Proceedings of the 13th International Conference on World Wide Web*, 20-20-29.

Clark, W. A. V., & Onaka, J. L. (1983). Life cycle and housing adjustment as explanations of residential mobility. *Urban Studies*, (1), 47-57.

Din, A., Bender, A., & Hoesli, M. (2001). Environmental variables and real estate prices. *Urban Studies, ,* 1989-1989-2000.

Downes, T. A., & Zabel, J. E. (2002). The impact of school characteristics on house prices: Chicago 1987-1991. *Journal of Urban Economics, 52*(1), 1-1-25.

Figlio, D. N., & Lucas, M. E. (2004). Whats in a grade? school report cards and the housing market. *American Economic Review, 11*, 591-604.

Fischel, W. A. (2001). *The homevoter hypothesis: How home values influence local government taxation, school finance, and land-use policies*. Cambridge, MA: Harvard University Press.

Gibbons, S., & Machin, S. (2003). Valuing english primary schools. *Journal of Urban Economics, 53*(2), 197.

Hanushek, E. A., & Rivkin, S. G. (1997). Understanding the 20th century explosion in U.S. school costs. *Journal of Human Resources, 32*, 35-35-68.

Haurin, D. R., & Brasington, D. (1996). School quality and real house prices: Inter- and intrametropolitan effects. *Journal of Housing Economics, 5*(4), 351-368.

Hayes, K. J., & Taylor, L. L. (1996). Neighborhood school characteristics: What signals quality to homebuyers. *Economic Review, Fourth Quarter*, 2-2-9.

He, D. (2009). *He dongchang on education (何东昌论教育)* People Education Press (人民教育出版社).

He, S., Liu, Y., Wu, F., & Webster, C. (2010). Social groups and housing differentiation in china's urban villages: An institutional interpretation. *Housing Studies, 25*(5), 671-671-691.

Hennig-Thurau, T., & Walsh, G. (2003). Electronic word-of-mouth: motives for and consequences of reading customer articulations on the internet. *International Journal of Electronic Commerce, 8*(2), 51-51-74.

Hilber, C. A. L., & Mayer, C. (2009). Why do households without children support local public schools? linking house price capitalization to school spending. *Journal of Urban Economics*, (1), 74-90.

Holme, J. J. (2002). Buying homes, buying schools: School choice and the social construction of school quality. *Harvard Educational Review*, (2), 177-177-206. Retrieved from http://www.metapress.com/content/u6272x676823788r/

Hu, P. (2008). Perfect the system, increase input, and promote a more equitable education. *Chinese Education & Society*, (1), 69-74.

Jiang, D. (2009, 贯彻教育方针 推进素质教育——读《何东昌论教育》. *Education Times (教育时报)*, pp. 2.

Kilpatrick, J. A., & Hefner, F.House price impacts of school district choice. *Growth and Change*, , 385.

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy*, *74*(2), 132-132-57.

Li, S., & Siu, Y. (2001). Residential mobility and urban restructuring under market transition: A study ofGuangzhou, china. *The Professional Geographer*, *53*(2), 219-219-229.

Li, Z., & Wu, F. (2006). Socio-spatial differentiation and residential inequalities in shanghai: A case study of three neighbourhoods. *Housing Studies*, *21*(5), 695-717.

Lin, T. (2009). Intraprovincial inequalities in basic education in mainland china. *Chinese Education and Society*, *42*(5), 3-3-10.

Morahan-Martin, J. (2004). How internet users find, evaluate, and use online health information: a cross-cultural review. *CyberPsychology & Behavior*, *7*(5), 497-497-510.

Oates, W. (1969). The effects of property taxes and local public spending on property values:An empirical study of tax capitalization and the tiebout hypothesis. *Journal of Political Economy*, , 957-957-15.

Park, D, Lee, J. & Han, I. (2007). The effect of on-line consumer reviews on consumer purchasing intention: the moderating role of involvement. *International Journal of Electronic Commerce*, *11*(4), 125-125-148.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, , 34-55.

Sayrs, L.Pooled time series analysis. In (pp. 10). Newbury Park, Ca: Sage.

Shanghai Existing Property Index Office (上海市二手房指数办公室). (2004 to 2011). *Shanghai existing property index monthly report (上海市二手房指数报告).* Shanghai: Shanghai Existing Property Index Office.

Shanghai Housing and Land Administration Bureau (SHLAB). (2003). *Shanghai real estate market 2003.* Beijing: Shanghai Social Science Academy Press.

Shanghai Local Chronicles Office (上海地方志办公室).*Shanghai local chronicles.* Retrieved from http://www.shtong.gov.cn

Shanghai Statistical Bureau (SSB). (2004 to 2011). *Shanghai statistical yearbook.* Shanghai: Shanghai Statistical Bureau.

Shen, T. (2011). 学区房热，家长称孩子读书是买房最大考虑因素. *SouFun,* Retrieved from http://zu.sz.soufun.com/rent/news/5412367.htm

Song, S., Yang, Z., & Guo, J. (2010). 小学教学质量对住房价格影响的研究. *Money China (财经界), 12*

Stadelmann, D. (2010). Which factors caitalize into house prices? A bayesian averaging approach. *Journal of Housing Economics, 19*(3), 180-180-204.

STATA. (2003). Cross-sectional time series. In (pp. 10). College Station, Texas: Stata Press.

The Trefis Team. (2011, Baidu girds for google battle in china. *Forbes,* Retrieved from http://www.forbes.com/sites/greatspeculations/2011/12/07/baidu-girds-for-google-battle-in-china/

Tiebout, C. (1956). A pure theory of local expenditures. *Journal of Political Economy, 64*(5), 416-416-24.

Walden, M. L. (1990). Magnet schools and the differential impact of school quality on residential property values. *Journal of Real Estate Research, 5*(2), 221-221-230.

Wang, L. (2010). Impact of urban rapid transit on residential property values: An empirical analysis of the shanghai housing market. *The Chinese Economy, 43*(2), 33-33-52.

Wang, X., & Ge, J. (2010). 南京老城区学区房价格机制研究. *Co-Operative Economy & Science (合作经济与科技), 12*

Weimer, D. L., & Wolkoff, M. J. (2001). School performance and housing values: Using non-contiguous district and incorporation boundaries to identify school effects. *National, 54*(2)

Wen, J. (2011). 上海市长宁区学区房价格走势分析. *Economics Vision (Jing Ji Shi Jiao), 5*

Wu, W. (2002). Migrant housing in urban china: Choices and constraints. *Urban Affairs Review, 38*(1), 90-90-119.

Yaffee, R. (2003). A primer for panel data analysis. *Connect: Information Technology at NYU, Fall Edition*

Zahirovic-Herbert, V., & Turnbull, G. K. (2008). School quality, house prices and liquidity. *The Journal of Real Estate Finance and Economics, 37*(2), 113-113-130.

Zhang, J. Q., Craciun, G., & Shin, D. (2010). When does electronic word-of-mouth matter? A study of consumer product reviews. *Journal of Business Research, 63*(12), 1336-1336-1341.

Zhou, X. (2011). 中国父母一辈子都在为子女而活. Message posted to http://blog.people.com.cn/blog/c4/s335326,w1312060311588244

高房价已改变国人生活　即使啃老也只能蜗居. (2009) *Economics Information (经济参考报)*,

高龄学位房很抢手 家长提前买房银行放贷宽松. (2011) *Information Times(信息时报)*

Appendix

Housing System in Shanghai

Urban residential dwellings in China are relatively different from those in western countries. Before the open policy in 1978, housing in urban China was distributed by the employers, and thus seen as an employee's right but not a commodity. The 1998 housing reform in Shanghai lifted the welfare housing system. It also greatly improved living conditions by putting housing and neighborhood facility standardization into law.[27] After reform, the whole Shanghai housing market has been administratively divided into 106 residential zones, while the urban area of Shanghai is separated by three major rings: the inner ring, the middle ring and the outer ring. Although the 106 residential zones have varying areas, each zone has a neighborhood that is relatively homogeneous (Chen and Hao, 2006). The Shanghai housing market began to boom in 2000, but the disparities in the housing market between inside areas and outside areas of the inner ring also started to emerge. Each year between 2001 and 2005, the average price of commodity residential real estate achieved a double-digit growth rate over the previous year. In spring 2005, a transaction tax was imposed for sellers who had owned their own houses for less than two years. In the same year, the Central Bank raised interest rates on mortgages twice. The rooted socio-spatial inequalities were strengthened by institutional changes in the post-reform period. These reform policies have greatly developed the private real estate market in Shanghai but at the same time have revealed the social-spatial inequalities among different neighborhoods.

[27] These facilities cover healthcare, green area, education, entertainment, food market, bus system, etc.

There are seven types of housing in Shanghai, namely shanties, old lane houses, improved lane houses, old staff dwellings, new staff dwellings, apartments and single houses (Wu, 2002). Among them, apartments[28] have the best quality of the seven types in the inner city (Li and Wu, 2005). The land for housing within the middle ring reached and continues to reach saturation since the 1998 reform with housing market leaning towards existing property, while, of the newly built villas after 1949, 85.3 % were constructed in the suburban districts of Minhang, Songjiang, Baoshan and Pudong New Area (SHLAB, 2003, Shanghai Statistical Yearbook, 2004 to 2011). Researchers who have been concerned about spatial inequalities consider the social grouping before 1998 reform to be the major factor (Li and Wu, 2005). Other researchers have found the disparities to be correlated with other factors. He, Liu, Wu and Webster (2010) show a significant correlation between *hukou* and disparity in their empirical results. Li and Wu (2005) suggest in their study that only housing space is significantly correlated with social group disparity on a 1% level among different neighborhoods. Their results also suggest the absence of a definite relationship between socio-economic status and residence, which confirms the results of other urban studies research such as Li and Siu (2001). My study will not be biased by these inequalities because the Shanghai Existing Property Index data that I used in my study focused on the apartments with similar structural characteristics in the six districts that are all located inside the middle ring.

[28] Apartment here means of both apartments and single houses.

Education System in Shanghai

The Shanghai education system implements the nine-year compulsory education program proposed by the national government in 1984. According to the Shanghai Statistical Yearbooks from 2004 to 2011, more than 70% of school funding comes from municipal government at all levels every year, with about another 13% from tuition. The rest of the funding consists of enterprises' education sponsorships and private donations that usually go to private schools directly. There are studies that show the education expenditure inequalities within Shanghai (Lin, 2009; Hu, 2008; Li and Wu, 2006). The intra-municipal disparity of fiscal expenditure for education only lies between urban and suburban areas, but the absolute gap started to narrow since the 2005 reform (Li and Wu, 2006).

Because of the nine-year compulsory education policy,[29] students enroll in the public schools unless they are admitted by a private school. According to the Shanghai Statistical Yearbook 2004 to 2011, private schools hold no more than 10% of the entire municipal student body every year in Shanghai. A student can only enroll in the public school with an attendance zone that covers the neighborhood where the student's *hukou* is located.[30] No two elementary schools have overlapped attendance zones and each neighborhood is covered by only one elementary school. Therefore, the only way to get into a specific public primary school turns out to be living in the corresponding attendance zone.

[29] Nine-year compulsory education in China refers to elementary (1st to 5th or 6th grade) and middle school (6th or 7th to 9th grade) education.

30 The location of a *hukou* is defined by where the head of household lives Immigrants to Shanghai can obtain Shanghai *hukou* only if they are special elites introduced by the government or their spouses hold the citizenship.

Tables and Figures

Table 1 – A Sample Report for the Shanghai Existing Property Index

Area	snll (上南六里)					
Structure	Single Bedroom		Double Bedroom		Triple Bedroom	
Date	Price	Index	Price	Index	Price	Index
2008/2/1	3142	1	3106	1	3182	1
2005/2/2	3171	1.0092	3152	1.0148	3159	0.9928
2012/2/3	3228	1.0274	3182	1.0245	3170	0.9962

Table 2 – Index Size Standards

Apartment Structure	Size (m^2)		
	Mean	Min	Max
oneBedroom	35	32	37
twoBedroom	66	62	70
threeBedroom	88	85	90

Note:
(i) The report dates for different areas are usually collected every four weeks but on different dates.

Table 3 – Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
oneBedroom	2450	3.459	0.955	1.750	5.822
twoBedroom	2450	3.403	0.929	1.681	5.655
threeBedroom	2450	3.366	0.922	1.682	5.590
topTier	2450	0.171	0.376	0	1
shopping	2450	0.2	0.400	0	1
neighbor	2450	15.496	8.667	1	30
dateYM	2450	576.793	23.739	533	618
interest	2420	2.874	1.164	1.09	5.91
metro	2450	0.300	0.459	0	1
cpiadjusted	2450	1.105	0.0618	1.016	1.251
district1	2450	0.234	0.424	0	1
district2	2450	0.0992	0.299	0	1
district3	2450	0.134	0.341	0	1
district4	2450	0.231	0.422	0	1
district5	2450	0.167	0.373	0	1
district6	2450	0.134	0.341	0	1

Notes:
topTier: 1 = has a top-tier school, 0 = has a non-top-tier school
shopping: 1 = neighborhood is within the shopping area
metro: 1 = metro station available within the neighborhood
interest: 10 observations fewer than other variables due to missing data on Jan-2006
cpiadjusted: adjusted monthly CPI index
district1 = cn, district2 = hp, district3 = ja, district4 = pd, district5 = xh, district6 = yp

Table 4 – Correlation Between Independent Variables

	topTier	shopping	neighbor	districtDum	dateYM	interest	cpiadjusted
topTier	1						
shopping	-0.0016	1					
neighbor	-0.1507	0.3181	1				
districtDum	0.1566	0.0295	0.0486	1			
dateYM	0.0014	0.0005	0.0011	0.0071	1		
interest	-0.0019	-0.0005	0.0029	0.0015	0.2792	1	
cpiadjusted	0.0006	0.0003	0.0018	0.0064	0.9567	0.4122	1

Notes:
(i) dateYM consists of year and month and represents the record dates of the observations
(ii) cpiadjusted is the adjusted monthly CPI that has a base 100 of the value from the same month in 2003

Table 5 – Model Specifications in OLS Regressions

Model	I	II	III	IV	V
topTier	0.337***	0.333***	0.351***	0.313***	0.266***
	(0.0508)	(0.0507)	(0.0509)	(0.0205)	(0.0200)
shopping	-	0.181***	0.125***	0.130***	0.124***
		(0.0477)	(0.0504)	(0.0203)	(0.0194)
metro	-	-	-0.148***	-0.142***	-0.0999***
			(0.0442)	(0.0178)	(0.0175)
interest	-	-	-	-0.150***	-0.148***
				(0.00724)	(0.00681)
cpiadjusted	-	-	-	14.88***	14.87***
				(0.137)	(0.128)
constant	3.402***	3.366***	3.419***	-12.59***	-12.58***
	(0.0210)	(0.0229)	(0.0278)	(0.144)	(0.135)
Observations	2450	2450	2450	2420	2420
Adjusted R^2	0.017	0.023	0.027	0.843	0.862

Notes:
(i) Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1
(ii) Absorb(District) in Model V with 6 categories: F(5, 2409) = 64.747 (0.000)

Table 6 – Between Effects and GLS Specifications

Model	I	II
topTier	0.269***	0.298***
	(0.0922)	(0.0537)
shopping	0.123	0.136**
	(0.0879)	(0.0545)
Metro	-0.100	-0.0901**
	(0.0794)	(0.0413)
Interest	(omitted)	-0.00629
		(0.00445)
cpiadjusted	(omitted)	6.945***
		(0.186)
district2	0.279**	0.263***
	(0.130)	(0.0872)
district3	0.139	0.146**
	(0.116)	(0.0616)
district4	0.175*	0.230***
	(0.0966)	(0.0602)
district5	-0.150	-0.121**
	(.107)	(0.0515)
district6	0.204*	0.208***
	(0.115)	(0.0595)
Constant	-3.342***	-4.436***
	(0.0738)	(0.208)
Observations	2420	2420
R^2-within	0.011	-
R^2-between	0.642	-
R^2-overall	0.048	-

Notes:
(i) Standard errors in parentheses:
*** p<0.01, ** p<0.05, * p<0.1
(ii) Model I: between effect regression
(iii) Model II: GLS regression after biases corrected

Table 7 – GLS Regression Results

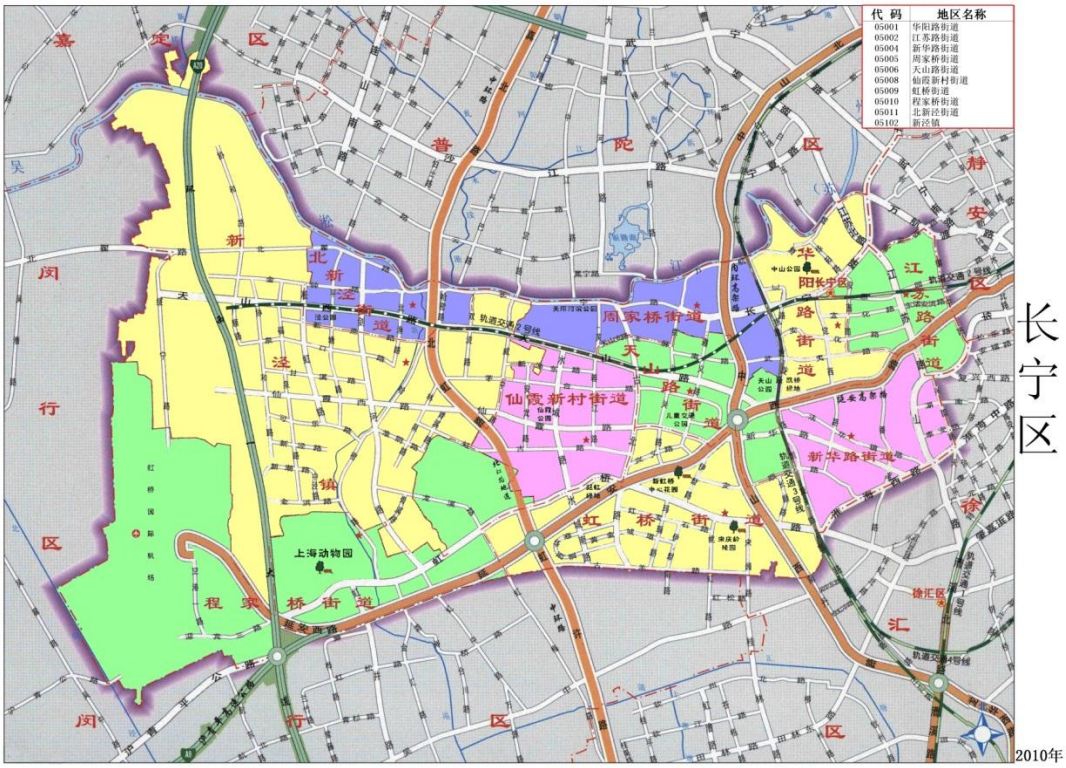
Variables	oneBedroom	twoBedroom	threeBedroom
topTier	0.281***	0.210***	0.291***
shopping	0.149***	0.0828	0.0628
metro	-0.0891**	-0.0936**	-0.0795**
interest	-0.00473	-0.00566	-0.0110**
cpiadjusted	6.907***	7.323***	8.568***
district2=hp	0.258***	0.327***	0.296***
district3=ja	0.142**	0.247***	-0.000807
district4=pd	0.225***	0.288***	0.330***
district5=xh	-0.126**	-0.0432	-0.0836**
district6=yp	0.204***	0.324***	0.406***
constant	-4.394***	-4.925***	-6.312***

Notes:
(i) observations: 2420
(ii) number of neighborhoods: 30
(iii) omitted: district1 = cn

Table 8 – Reformatted Results for Top-tier Effects

	One Bedroom	Two Bedrooms	Three Bedrooms	
district1=cn	28.1%	21.0%	29.1%	26.1%
district2=hp	53.9%	53.7%	58.7%	55.4%
district3=ja	42.3%	45.7%	29.0%	39.0%
district4=pd	50.6%	49.8%	62.1%	54.2%
district5=xh	15.5%	16.7%	20.7%	17.6%
district6=yp	48.5%	53.4%	69.7%	57.2%
	39.8%	40.1%	44.9%	41.6%

Figure 1- Administrative Map for Neighborhoods in Changning (cn) District

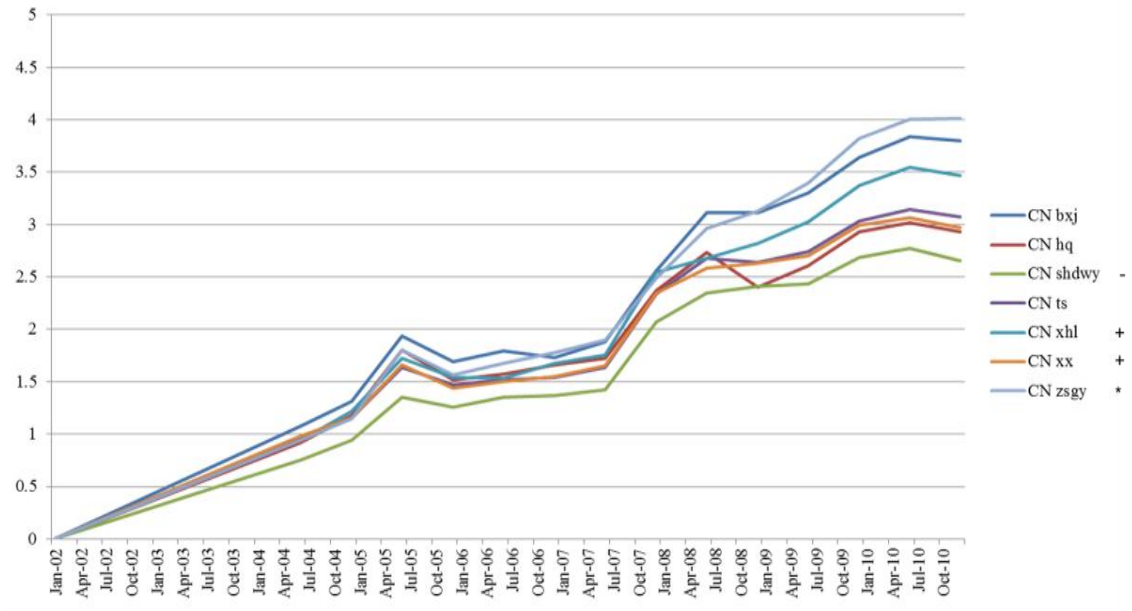


The colored regions represent different sub-district under the district Changning (cn) in 2010

Figure 2 - Shanghai Existing Property Index Control Points within Changning (cn) District

This map provides the location of the control points picked for housing price index in the district Changning (cn)

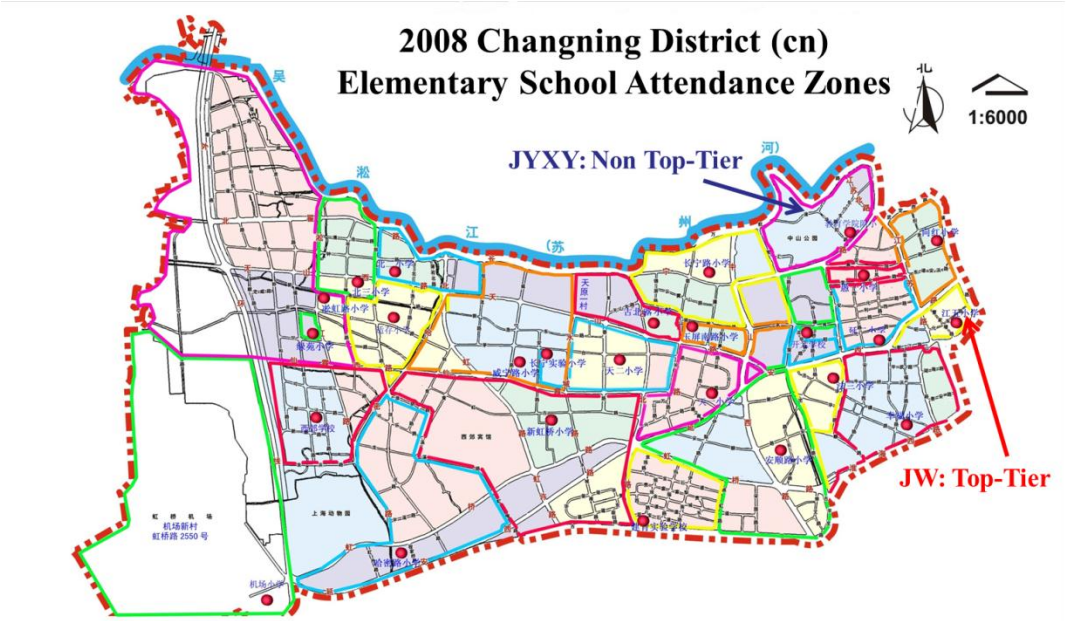
Figure 3 - Housing Price over Time for Changning (cn) with Four Tiers



Area "zsgy" started to have top-tier school around 2008

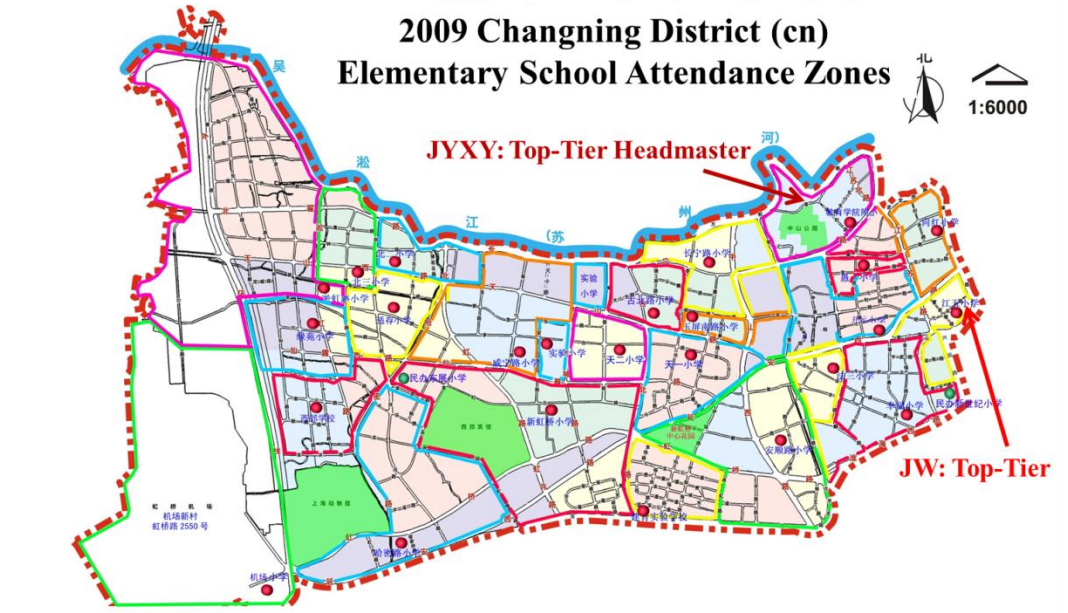
No top-tier primary school is under any of the neighborhoods given above except for the neighborhood zsgy (labeled in blue), in which the relocation of top-tier school happened around 2009.
Shopping area is zsgy.

Figure 4 - 2008 Elementary School Attendance Zones for Changning (cn) District



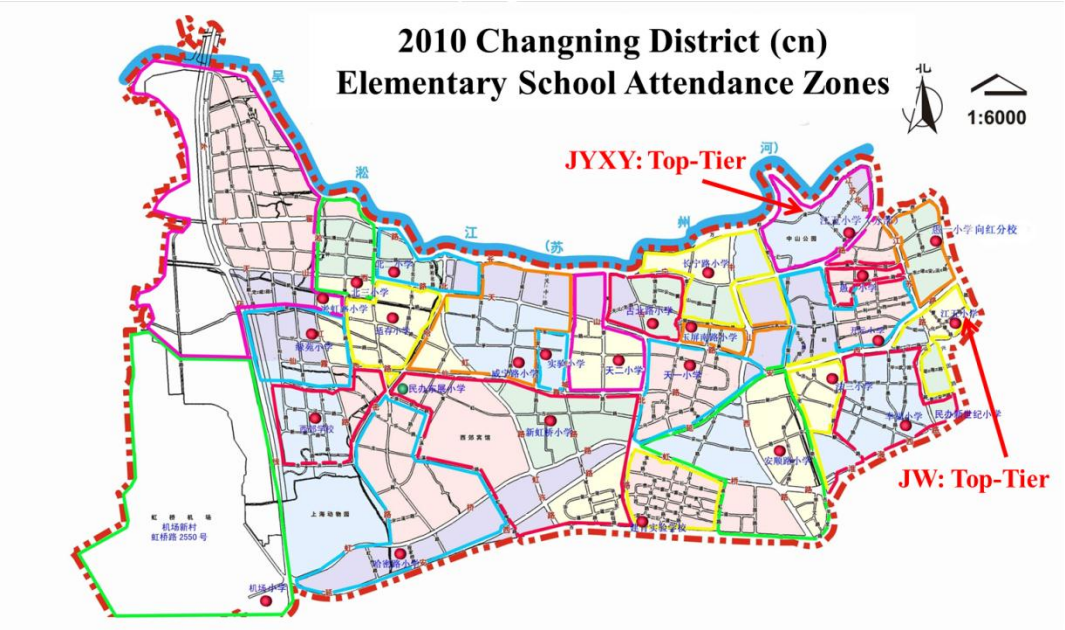
This map provides the school attendance zone information for district Changning (cn) in 2008

Figure 5 - 2009 Elementary School Attendance Zones for Changning (cn) District



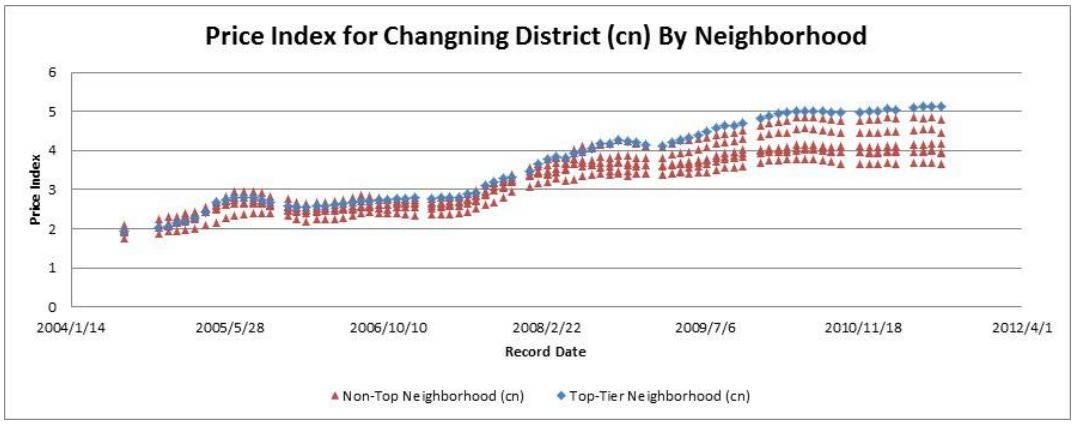
This map provides the school attendance zone information for district Changning (cn) in 2009

Figure 6 - 2010 Elementary School Attendance Zones for Changning (cn) District



This map provides the school attendance zone information for district Changning (cn) in 2010

Figure 7 - Housing Price over Time for Changning (cn) with Top-tier Labeled



No top-tier primary school is under any of the neighborhoods given above except for the neighborhood zsgy (labeled in blue), in which the relocation of top-tier school happened around 2009.
Shopping area is zsgy.