Neighborhood Quality and Housing Value:
Evidence from Urban Micro Data

Yi Wang

School of Economics and Management, Sichuan Agricultural University Chengdu Campus

No. 211, Hui Min Road, Wenjiang District, Chengdu 611130, Sichuan, China

E-mail: wang_yi@yahoo.cn

Ruiping Ran (Corresponding author)

School of Economics and Management, Sichuan Agricultural University Chengdu Campus

No. 211, Hui Min Road, Wenjiang District, Chengdu 611130, Sichuan, China

E-mail: ruipingran@yahoo.com.cn

Guoying Deng

Sichuan University, School of Economics

24 Yihuan Road, Chengdu 610065, Sichuan, China

E-mail: dengguoying@scu.edu.cn

Abstract

Using urban residential micro data from CHNS, this paper employs Hedonic Pricing Model to investigate the impact of Neighborhood Quality on housing value and its mechanism. We find that, Human Capitals measured by average schooling years and occupational diversity have significant positive effect while cultural capitals such as Ethnic Diversity have significant negative effect on housing value. Compared with the empirical results from developed counties, Social Capitals measured by length of tenure and own room rate have insignificant effect on housing value. In addition, having Kindergarten and Hospitals near the communities are positively correlated with the housing value while Internet cafe or transportation services are negative. We believe that these facts are closely related to the stage of economy transition and rapid urbanization in current China. The conclusions have important implications for the effective construction of harmonious commodities.

Keywords: Housing value, Neighborhood quality, Hedonic pricing model

1. Introduction

China has experienced rapid economic growth in the last two decades, and people's life standards have been significantly improved in the meantime. The public is increasingly concerned about the improvement of living conditions. The value of environment and resources can be "discovered" from the change of producer's income and consumer's utility, and the prices of various goods and services in the market system, which Freeman (2002), a well-known American environmental economist, defined as "indirect market value" of environmental changes. If "Indirect market value" is positive, it shows that people value this factor; on the contrary, if the "indirect market value" is significantly negative, this environmental factor would be harmful to residents. In this paper, we try to measure the indirect housing market value of community-level environmental factors, such as human capital, social and cultural capital, and also explain from a new perspective why the housing prices or rents will go up recently in China.

There are huge differences in housing conditions. Not only the structural characteristics of housing itself, but the surrounding community environment will result in significant differences in housing values (Jim and Chen, 2007). All of these factors can be modeled in the traditional "Hedonic Pricing Model" (Hereafter HPM). The

HPM has been used as a valuable tool in the valuation of housing because of its simplicity and usefulness. Housing in our country has been usually treated as homogeneous goods when building the housing price index. If the index increases, both the neighborhood quality improvement and demand or supply shock can be the reason, but we usually can't identify which one is the exact answer (Chen, 2005). So the HPM price index, which has considered the heterogeneous properties of housing, would be more helpful to determine how the supply and demand changes influence the housing prices, and provide the government more direct tools to develop the real estate market (Liu and Sun, 2006).

Now most of the studies about housing prices are from a macro perspective in China, but those studies using the aggregated macro data face more and more challenges, while the empirical studies using HPM method from the micro perspective are rare, especially that how the "quality of community" affect the value of housing is still blank. In this paper, using urban household micro-data and building neighborhood-level community environmental indicators, we try to examine in detail whether people value various neighborhood-level environments and also how it works.

2. Literature review

HPM is often used to estimate the market value of physical assets or non-market goods. The model is based on consumer preferences and utility theory, assuming that commodity prices depends on how much the utility all aspects of product attributes can give to the consumers. Rosen (1974) and others extend this approach to the analysis of real estate and other durable goods pricing. Then the literature divided housing properties into three categories, namely, location characteristics, structural characteristics and neighborhood characteristics (Mok, etc., 1995; Rosiers, et al., 2000; Jim and Chen, 2007, etc.). The reason why location characteristics and neighborhood characteristics affect property values can be explained by the Social interaction theory. Manski (1993) pointed out that the rising of non-cooperative game theory, family economics and the endogenous growth theory in recent decades has give birth to the theory of social interaction. In the real estate research, they are also known as "environmental spillovers", "neighborhood effect" or "peer effect", and Lu and Zhang (2007) unified them as "non-market interaction". But due to the definition of the concept and data inference problems, economists are still difficult to draw a consistent conclusion (Ioannidis, 2002). Fu (2005, 2008) summarized the literature and classifies three types of social-interaction-based social amenities: human capital, social capital, and cultural capital at residential neighborhood levels and by estimating hedonic housing models with social amenities, he found that different types of social capital have different effects on property values at different geographic levels

Here human capital among households reflect the knowledge, skills and information spillovers, and a spatial equilibrium mechanism lead to the equilibrium that well-educated people tend to attract each other, because living together can get higher level of utility. Rauch (1993) pointed out that the average human capital is a form of public goods, and rents in cities with more human capitals would be higher. The so-called social capital refers to social networks that can help people to obtain the labor market and housing market information resources. Bian (2004) summarizes the social network function as a transmission of information, communication, nurturing trust and reputation. DiPasquale and Glaeser (1999) found that an increase in the home ownership rate has positive effects at large geographic levels, which is consistent with the argument that homeowners have strong incentives to invest in social capital, and homeowners concerned more about their own community school leaders, more involved in local elections and to solve problems, and that the length of residence also affect the quality of the community. The so-called cultural capital refers to the sub-cultural environment such as customs, religious beliefs and cultural traditions that can help identify certain social groups. For example, Mieszkowski (1973) and Dubin and Sung (1990) show that the American racial and cultural differences have a major impact on housing price. Given the above, this paper will further study two external factors—neighborhood housing characteristics and location characteristics affect the housing value, compared with the internal housing features, and we also include the quality of local education, living, communication and transportation facilities conditions as the impact factors. According to Fu (2005, 2008) we classify three types of community-based social amenities: human capital, social capital, and cultural capital at community-neighborhood-based.

In recent years, there are growing literatures to study the problem of China's real estate market, but most of them focus on traditional supply or demand side of the housing market, housing prices and housing policy (Fu, 2000; Yuan and Fan, 2003), and little of them are micro-data based. Some exceptions are, For example, Jim and Chen (2007) which examined the feasibility of application of HPM in China, and used survey data about residents of Guangzhou to examine the impact of preferences and the spillover effects on housing prices. Ma and Li (2005) introduced theoretical basis, advantages and disadvantages of the HPM systems. Although Xu and Chen (2008) argue with Liu and Li (2005) whether there is social stratification in urban communities, but pay little attention to the neighborhood effect. Referring to Fu (2005, 2008), we try to examine the effect of a variety of

environmental characteristics on housing values, and give corresponding explanations.

3. Empirical evidence

3.1 Model

We divide the factors affecting housing value into two categories: internal structural characteristics(S) and external quality factors of community (N). Therefore, the basic model is set as:

$$V = f(S, N) \qquad (1)$$

According to our research purposes, we focus on the impact of the quality of the community on the housing value, with the structural features of housing held as control variables.

How to determine the form of Hedonic price function is a big problem, and literatures usually used Box-Cox transformation method (Mok, 1995), while in this paper we try to employ a more simple and useful method. Comparing with linear, logarithmic and semi-logarithmic functional form, we found that the semi-logarithmic model can well meet the requirements. Specifically, semi-logarithmic model residuals normality test skewness is 0.108, kurtosis 0.042, which is largerer than the critical value of chi-square P=0.927, so we can not reject normality of the residuals of the original assumptions. Therefore, the characteristic form of price function can be set as:

$$\ln V_{ic} = \alpha + \beta N_c + \gamma S_{ic} + \varepsilon_{ic} \qquad (2)$$

Here, i is individual households, c is community, $\alpha, \beta,$ and γ is the coefficient to be estimated. To resolve the random disturbance ε heteroskedasticity or spatial autocorrelation problem, we take heteroscedasticity robust standard deviation to deal with related issues. The median home price or rent is also commonly used to measure the housing value; we can similarly obtain the following regression model:

$$\ln V'_{ic} = \alpha + \beta N_c + \gamma S_{ic} + \varepsilon_{ic} \qquad (3)$$

Omitted variables such as family background will usually lead to endogenous problem, so for getting consistent and unbiased estimates we combined two years' data (2004 and 2006) to compose panel data and used fixed effect methods. The Model is set as:

$$\ln V_{ict} = \alpha_i + \beta N_{ct} + \gamma S_{ict} + year06 + \varepsilon_{ict} \qquad (4)$$

Here, t is time indictor, $year06$ controls annual exogenous shocks, α_i controls some unobservable factors. We use "household code + line number" cluster of Huber/White heteroscedasticity standard deviation estimates and robust standard deviation method to obtain a consistent and efficient standard deviation.

3.2 Variable selection

We divide the variables of community quality further into human capital, social capital and cultural capital, and other services categories. Specifically, household human capital reflects the knowledge, skills and information spillovers, and its depth and breadth, and we proxy with "the average schooling years" or "the proportion of high school graduates and above" and "occupational diversity" (Fu, 2005; 2007; 2008). Occupational diversity index can be calculated as follows:

$$\text{Occupational diversity}_c = 1 - \sum_o \left(\frac{\text{Community worforce}_{oc}}{\sum_o \text{Community worforce}_{oc}} \right)^2 \qquad (5)$$

In Equation (5), c is the community, o is behalf of professional types (between 0 and 1, the greater the value the more diverse of the occupation). In general, the higher human capital the more obvious spillovers of information between community residents, and which are easier to capitalize into housing price. So the housing price will be much higher than others' (see the Table 1).

Social capital reflects the advantage of people's social networking. According to DiPasquale and Glaeser (1999) and Fu (2005, 2008), we use own room rate and the proportion of households living in a certain number of years to measure. The higher rate of room own, residents are more likely to participate in community activities, which would have a positive effect on housing value. However, the longer of living time, people are easier to be affected by traditional thinking around them, so the lower efficiency may be exit, otherwise shorter will lead to more innovation (Fu, 2005).

Since Cultural capital and national customs are closely related, we use national diversity to measure it, and the calculate method is similar with formula (5), just using national identify to be instead of Occupational diversity.

From our view, proportion of minority should be a good proxy variable of cultural capital. Because of language and cultural differences, multi-ethnic mixed relatively is not easy to communicate with each other, so we expected that cultural capital will have negative effects on housing values. Other facilities include: all types of schools and other educational and cultural facilities, parks, supermarkets, hospitals and other living facilities, Internet cafes and other communication services, bus and railway stations and other transportation facilities.

It is worth mentioning that the communities surround the bus or railway station is convenient to travel, which will increase the housing value, but on the other hand, the station's noise and security problems should have negative effect on the price or rent. Since the coexistence of two effects, we cannot determine what the exact effect of bus or railway station will have on housing price. In addition, the proximity zone (open cities, special economic zones) can also be characterized by the community surrounding environment. Traditionally, structural features can use housing age, housing construction materials or room numbers, size and other indicators to measure, but the kitchen equipment type is also a good variable (Ma and Li, 2003). Other control variables include: total household number or total resident number of communities, community poverty rate and so on. Based on common sense and the analysis above, different types of the quality of housing communities can have impact on housing values. The summary of all variables are listed in Table 1.

4. Data sources and descriptive statistics

4.1 Data sources

We obtained data, "China Health and Nutrition Survey" (CHNS)", mainly from Population Research center of United States University in North Carolina. They conduct this survey based on provinces (autonomous regions), city (county) and community (neighborhood, village) stratified random sampling method in China, and there are totally 4400 households in this dataset, and Data collection begins from 1989. In recent years, Many interested studies have being done using this data.

In this paper, the selected sample includes urban households from nine provinces (Liaoning, Hei Longjiang, Jiangsu, Shandong, Henan, Hubei, Hunan, Guangxi and Guizhou autonomous regions) in 2004 and 2006. To increase the sample size, and be consistent with the "Urbanization of China", in addition to urban residents, we also included neighborhood residents into the urban households in the county. The First reason we choose these two year's data is that the missing data in these two years are not too serious, and the most important, comparing with other years', these two-year's datasets are latest, so the community indicators to measure various types of capital in these two years would be more representative of China. At the same time, these two period data collection occurred after the housing reform in 1998, which can reflect more valuable information of the policy reform of China's housing market. To be much more convenient, we used household heads' education level, ethnic, occupational information, household income, housing conditions, community facilities and services around the integrated information to calculate the variables in this paper. For example, we take the ratio of graduation from high school or above level as the human capital depth's proxy variable for different communities; also we used the different head's careers to calculate community's occupational diversity indicators. Based on our questionnaire respondents of "Have your family removed since the last survey?" we calculated the "proportion of households living in a certain number of years" indirectly. Therefore, data for 2006 is "Ratio of living more than 2 years households", and 2004 data as "Ratio of living more than four years households." To get the variable "whether this is medical institutions in the communities", we selected "the most common to the medical institutions when neighborhood people need to see a doctor" to identify. Due to transportation costs are missed very seriously, we use "whether the train station or bus station near the community and whether there are open cities and special economics zones within two hours they can get there by car" to measure the location of housing characteristics. Many past studies have shown that the family kitchen equipment have a major impact on housing prices (Ma and Li, 2003). Because of the same missing data problem, we use "toilet type" as the structural characteristics' proxy variable. According to the official statistics of 2006 in China, if households' person income is under 785 Yuan (per year) they can be classified as the poor families, and using this data statistics we calculated our controlling variable "Poverty rate".

It is noteworthy noted that the housing values variable measurement in our paper are different with Ma and Li (2003) and Jim and Chen (2007). In these two papers, they use the listing price or transaction price as the housing value. But in CHNS database, there is only the living condition information of the surveyed households, in which there are three questions related to housing value: Firstly, to ask the rent of non-owned rental housing; Secondly, to ask the rent charged by the landlord; Thirdly, to ask the landlord's subjective judgment of housing prices. Taking into account of the missing data problem, we choose the rent of landlord and tenant as the housing valuation's proxy variable. Although there are big difference between Homeowners' subjective and objective

estimates of rent, the value of the property to make its own subjective estimate is based on similar homes in the surrounding environment or the actual rent, we believe that the subjective and objective rent have strong positive correlation, which is best way we can find to get the approximate value of the house. Sorting and merging the original data containing the above information and excluding missing values, we selected year 2006 (2004) including 874 (469) households in 66 (36) communities respectively to format the cross-sectional data and unbalanced panel data for our empirical research.

4.2 Descriptive statistics

Table 2 lists the main characteristics of the variable descriptive statistics of year 2006. From the statistics we can find that China's housing rent market is overall at a low level, but individuals vary considerably, with the maximum monthly rent be 8,000 Yuan and the smallest 6 Yuan. For the Community quality, the average schooling years is still low, and has not yet reached high school level, but individual differences are great. We also find the own room rates of China is 78%, which is significantly higher than U.S. average level 65%, and the ratio of tent to owned is about 3:10, while household living in the same community for more than two years occupies 95%. Ethnic diversity is low, and minorities accounted for only 12%, indicating that Han is the biggest nation in China, but a larger standard deviation existed, which is in line with China's nation characteristics named "big mixed, small communities". From the data statistics we also can find that in those cities and towns, their education and daily living facilities are convenient, especially for kindergarten and larger medical institutions. About 77% communities have a bus station nearby. There exists widespread use of indoor toilet, implying that the urban housing conditions are much better than before. Urban poverty rate of 6%-8% is in line with recent official statistics, but vary greatly across communities. In addition, our sample shows more than 50% communities are located in the "economic development zone", indicating that there is a growing trend of economic development zone (see the Table 2).

5. Results

5.1 Cross-sectional data results

Table 3 is the regression results of model (2). For testing the robustness, based on the first column, in the second column, we use "average schooling years" to be instead of "the proportion of high school graduates and above", the "ethnic diversity" to be instead of "the proportion of ethnic minorities", in the third column, we add "bus station" in the regression. As expected, all variables, for example, the depth and breadth of human capital; average schooling years and occupational diversity indicators are all very significant, indicating that the spillover effect of human capital can significantly enhance the community housing prices. Specifically, the average years of education for each additional year, prices rose about 0.049%, much lower than U.S. (0.1% -0.2%) (Rauch, 1993). Ethnic diversity has a significant negative effect on housing price, which indicated the influence of national culture is more broad and deep, or may be due to ethnic discrimination. However, by comparing the Han and Minorities households' average housing prices, we can find housing price of the former is much higher than the latter, which number is 132.26 (= 419.13-286.87) Yuan, so we ruled out the speculation about whether there is ethnic discrimination, and get that only the national identity can influence the quality of community and thus affect the housing price. Own room rate is significantly positive, which is consistent with the fact of the developed countries (DiPasquale and Glaeser, 1999; Fu, 2005), but "the proportion of households living in a certain number of years" is negative, although which is not significant. Possible explanation for this, comparing with developed courtiers, China's economic market mechanisms is in the process of transition and rapid urbanization currently, for many reasons including the housing complex sources; urban mobile population and short history of the community, these factors will certainly affect the community's cultural and security situation, thereby affecting its housing prices. In addition, other community's quality indicators, including kindergarten and hospital have more significant positive effect, which is similar to some existing researches' findings (such as Jim and Chen, 2007). In contrast, Internet cafes, bus and train stations have negative effect on house prices (see the Table 3).

5.2 Robustness test

Some studies used house' median rent to measure community's housing value (such as Fu, 2005). The same way, in model (3), we change the dependent variable into the median rent, regression results, seen in Table 4. In the first column, in addition to coefficient slightly changes, the results are basically consistent with the findings in Table 3. In reality, "similar goals" or similar people are often easier to live together, but these hidden preferences are not easy to detect or measure, which is likely to be affected by some special individual factors. If we just only use observable data, we can not resolve omitted variable problem (Manski, 1993). Also, people are often based on their education level; household income and occupations to choose where to live in the city-specific

region (Liu and Li, 2005; Xu and Shen, 2008), Therefore, these variables are likely to be structure endogenous (Rosiers, 2000). Furthermore, some proxy variables of community quality may have measurement errors, which can easily lead to bias OLS estimates of coefficients and inconsistence. In order to deal with this problem, we combined the data of year 2004 and year 2006 to compose an unbalanced panel data, using model (4) of the fixed effects approach to do further empirical test. In Table 4, the second column is the results of mixed cross-sectional regressions (Pooled OLS), and the third column is the results of fixed effects regression. It is clear that, regardless of which column, comparing with table 3, there are no big changes, indicating that in addition to the individual characteristics of the omitted variable, other types of endogenous problem is not serious in our regression model. We believe that the use of panel data fixed effects methods can be largely to eliminate the sample selection problem. Studies have shown that although there currently partitioned between rich and poor living phenomenon, but because of class consciousness and community culture are far from overlapping, Community layered living phenomenon has not yet appeared (Xu and Shen, 2008), which can also be obtained from the part of social capital to explain housing prices. Compared with the results of cross-sectional data, the average years of education about the spillover effects dropped to 0.07, but the "proportion of households living in a certain number of years" has become very significant, indicating that social capital can affect house prices, but it is not robust (see the Table 4).

6. Conclusions and policy recommendations

In this paper, Using 2004 and 2006 CHNS urban household survey micro data, by building community-level social capital and other environmental services, we use HPM to explore how the community's quality to impact the local housing values. Cross-sectional data regression results show that human capital and cultural capital of communities significantly affected prices or rents, and social capital remains to be further studied. To address possible endogenous problems, we constructed a panel data model. After that we obtained more consistent results, and Community layered living phenomenon is not apparent in China.

We give the explanation is that compared with developed countries, China's urban households own room rate is much more higher, but China is a country with rapid economic transition and rapid urbanization, some social problems result in the community quality can not well capitalized to housing price. But after the open up reform, China's educational undertakings have being repeatedly to pick up human capital spillover effects, which are more obvious to raise the housing price of some good communities. As the big difference of national language, culture, cultural capital reduced the community housing prices. These conclusions are consistent with the developed countries. Further, despite a higher rate of own room rate and relatively long time living, but the community culture and class culture   have no been overlapped, so the domestic urban residents living in a sub-layer effect (Sorting Effect) is not yet obvious.

At the policy level, we believe that since cultural capital will result in relatively low housing prices, the government should improve their living conditions, and to plan a special urban areas inhabited by ethnic minorities and further strengthen the "small communities" trend. In addition to all forms of education continue to support the development, appropriate control of the urbanization process, strengthening the urban flow of personnel management, which will be conducive to harmonious urban community building. From R^2 of the regression model, the survey has at least 32% of housing prices is due to community factors such as quality improvement, and the remaining 50% of the price are due to the supply and demand and other factors. Part of Housing prices rise reflects the improvement of surrounding environment, which is in line with characteristics of pricing principles and can promote local economic prosperity, but if the price is much higher than people's purchasing power, we think the government should pay more attention to industrial structure and housing price bubble.

References

DiPasquale, D., and W. Wheaton. (1995). *Urban Economics and Real Estate Markets*. New Jersey: Prentice Hall, (Chapter 5).

DiPasquale, D., and E. Glaeser. (1999). Incentives and Social Capital: Are Homeowners Better Citizens? *Journal of Urban Economics*, 45, 354-384. http://dx.doi.org/10.1006/juec.1998.2098.

Dubin, R., and C. Sung. (1990). Specification of Hedonic Regressions: Non-Nested Tests on Measures of Neighborhood Quality. *Journal of Urban Economics*, 27, 97-110. http://dx.doi.org/10.1016/0094-1190 (90)90027-K.

Fu, S. (2005). What has been capitalized into Property Values: Human Capital, Social Capital, or Cultural Capital? [Online] Available: http://riem.swufe.edu.cn/new/techupload/paper/20069720324058107.pdf.

Fu, S. (2007). Sexual Orientation and Neighborhood Quality: Do Same-Sex Couples Make Better Communities? [Online] Available: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=951727 (October 31, 2007).

Fu, S. (2007). Smart Café Cities: Testing Human Capital Externalities in the Boston Metropolitan Area. *Journal of Urban Economics*, 61, 86-111. http://dx.doi.org/10.1016/j.jue.2006.06.002

Ioannides, Y. (2002). Residential Neighborhood Effects. *Regional Science and Urban Economics*, 32, 145-165. http://dx.doi.org/10.1016/S0166-0462(01)00082-5

Jim C.Y., and W. Y. Chen. (2007). Consumption Preferences and Environmental Externalities: A Hedonic Analysis of the Housing Market in Guangzhou. *Geoforum*, 38, 414-431. http://dx.doi.org/10.1016/j.geoforum.2006.10.002.

King, T., and P. Mieszkowski. (1973). Racial Discrimination, Segregation, and the Price of Housing. *Journal of Political Economy*, 81, 590-606. [Online] Available: http://www.jstor.org/pss/1831027

Manski, C. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies*, 60, 531–542. [Online] Available: http://www.jstor.org/pss/2298123

Mok, H., P.P.K. Chan, and Y. Cho. (1995). A Hedonic Price Model for Private Properties in Hong Kong. *The Journal of Real Estate Finance and Economics*, 10, 37-48. http://dx.doi.org/10.1007/BF01099610

Rauch, J. (1993). Productivity Gains from Geographic Concentration of Human Capital: Evidence from Cities. *Journal of Urban Economics*, 34, 380-400. http://dx.doi.org/10.1006/juec.1993.1042

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Perfect Competition. *Journal of Political Economy*, 82, 34-55. [Online] Available: http://www.jstor.org/pss/1830899

Rosiers, F., M. Theriault, and P. Villeneuve (2000). Sorting out Access and Neighbourhood Factors in Hedonic Price Modeling. *The Journal of Property Investment and Finance*, Vol.18, pp.291-315. http://dx.doi.org/10.1108/14635780010338245

Table 1. Expectation direction of different variables to housing value

human capital	Average schooling years and the proportion of high school graduates and above	+
	Occupational diversity	+
social capital	Own room rate	+
	proportion of households living in a certain number of years	?
cultural capital	ethnic diversity or minority proportion	-
Other services	Schools, hospitals, etc.	+
	Bus and railway stations, etc.	?

Note: "+"denotes positive effect, "-" denotes negative effect, "?"is not able to determine.

Table 2. Descriptive statistics

Variables	Mean	Stand error	Minimum	Maximum
Rent	402.94	476.39	6	8000
Average schooling years	8.66	2.80	2.79	14
proportion of high school graduates and above	0.48	0.26	0.10	1
Occupational diversity	0.91	0.10	0.58	1
proportion of households living in a certain number of years	0.95	0.06	0.79	1
Own room rate	0.78	0.21	0.05	1
ethnic diversity	0.08	0.16	0	0.67
minority proportion	0.12	0.26	0	1
Internet cafe	0.84	0.37	0	1
kindergarten	0.36	0.48	0	1
Middle and primary schools	0.41	0.49	0	1
supermarkets	0.20	0.40	0	1
park	0.38	0.49	0	1
Medical institution	0.54	0.50	0	1
Bus station	0.77	0.42	0	1
Railway station	0.29	0.45	0	1
Economic development zone	0.58	0.49	0	1
Transportation costs	7.62	3.87	1	12
Toilet type	1.76	1.52	0	8
Poverty rate	0.06	0.08	0	0.55
Total household of community	1926.72	1331.87	140	4935

Table 3. Cross-sectional data results

Independent variables	Regression One	Regression Two	Regression Three
Average schooling years	0.048**		0.050***
	(0.0188)		(0.0192)
Occupational diversity	1.801***	1.221***	1.878***
	(0.399)	(0.404)	(0.409)
proportion of households living in a certain number of years	-0.607	-0.068	-0.619
	(0.646)	(0.655)	(0.644)
Own room rate	0.992***	1.026***	1.002***
	(0.284)	(0.287)	(0.282)
Ethnic diversity	-1.342***		-1.353***
	(0.279)		(0.279)
Internet cafe	-0.498***	-0.373***	-0.505***
	(0.101)	(0.105)	(0.102)
Kindergarten	0.195**	0.219**	0.193**
	(0.095)	(0.103)	(0.095)
Medical institution	0.311***	0.259***	0.321***
	(0.076)	(0.079)	(0.077)
Railway station	-0.249***	-0.339***	-0.234***
	(0.066)	(0.072)	(0.068)
proportion of high school graduates and above		0.406**	
		(0.184)	
Minority proportion		-0.469**	
		(0.188)	
Bus station			-0.067
			(0.094)
Observation	874	874	874
R^2	0.336	0.319	0.337

Note: Dependent Variable is "log (room rent)"; Heteroscedasticity robust standard deviation in brackets. "*", "**"and"***"denotes significant at 90%, 95% and 99% level. All regressions control the fixed of province, toilet, poverty rate and Households number of communities.

Table 4. Robustness test results

Independent variables	Regression One	Regression Two	Regression Three
Average schooling years	0.056***	0.032***	0.0267*
	(0.015)	(0.015)	(0.016)
Occupational diversity	1.744***	0.855***	0.858***
	(0.261)	(0.225)	(0.231)
proportion of households living in a certain number of years	-0.740	-0.987***	-1.093***
	(0.460)	(0.249)	(0.277)
Own room rate	1.335***	0.756***	0.746***
	(0.249)	(0.175)	(0.184)
Ethnic diversity	-1.467***	-0.783***	-0.759***
	(0.164)	(0.187)	(0.195)
Internet cafe	-0.584***	-0.309***	-0.321***
	(0.081)	(0.074)	(0.073)
Kindergarten	0.264***	0.200***	0.208***
	(0.068)	(0.062)	(0.063)
Medical institution	0.477***	0.102*	0.113**
	(0.056)	(0.055)	(0.056)
Railway station	-0.237***	-0.378***	-0.373***
	(0.047)	(0.053)	(0.056)
Observation	874	1343	1343
R^2	0.487	0.336	0.337

Note: Dependent variable of First column is "log (median rent)", and "long (room rent)" in the last two column.