Hanwen Zhang
Minzhi Meng
17 May 2019

Research Project

The Effect of Housing Price on Personal Consumption

Housing price has been the hottest discussed topic in China over the last twenty years ever since the Chinese government launched a series of market-oriented housing regulation reform in the year of 1999. After the reform, welfare housing system in China was abolished, thus unleashing a growing demand of private housing and boosting a rapid development of real estate industry in China (Liu and Xiong, 2017). This rapid transform and development of housing industry trigger great interest in researchers. Some researchers concern that the growing gap between housing price and income would depress consumer demand, and in turn constrains the development of economy. However, some researchers believe that any rise in house prices would serve as an increase in household wealth and thus boots up the consumption and market expectation, which is a positive market signal that would benefit the economy in the long run.

In this research, we examine if there is a positive relationship between housing price and non-housing consumption in five large cities in China. We look at annual data from the year of 2003 to the year of 2017 for five cities in China, including Beijing, Shanghai, Guangzhou, Shenzhen, and Hangzhou. These five cities have large population, booming market economy, and relative comprehensive data that we can look at. Follow the method of our references (Kuang, 2011), we estimate our model using pooled regression with OLS first. We then treat the data as panel to check for the random and fixed effect estimator, depending on the result of Hausman test. Our empirical results indicate that there is some evidence of a positive relationship between housing price and non-housing consumption.

Our theoretical model is based on the intertemporal choice model and microeconomic theory. Intertemporal choice model states that people make decisions about spending, investment, and so on by using discounted utility model in a way similar to that financial markets evaluate assets by using discounted cash flow. People would take into account possibilities in different periods to make current period's consumption and saving decisions (Berns, 2007). Thus, possibilities in past period should be included into model. Microeconomics theory states that people would make their consumption and saving decisions over different points of time by maximizing the utility they gain from each purchase and investment decision.

Therefore, if we want to examine the relationship between an individual's current consumption and the current housing price, we also need to include variables that would influence the individual's consumption and saving decision over different period of time. Referring to Kuang's analysis, we estimate that the non-housing consumption of a person in current period would be a function of the following: his or her non-housing consumption in last period; the price of housing area in current period; his or her disposable income in current period; his or her housing area in current period; his or her saving deposit in this period; and the rate of return on investment in this period. In the model, the dependent variable is the non-housing consumption in current period, the independent variable is the price of housing area in current period, and the control variables are the non-housing consumption in last period, the disposable income in current period, the housing area in current period, the saving deposit in current period, and the rate of return on investment in current period.

There are plenty of ongoing researches concerning the relationship between housing prices and consumption. It is generally believed that increasing housing price manifests its impact on consumption mainly in two aspects: on one hand, rising house price have a wealth effect, thus promote consumption and economic growth; on the other hand, rising house prices will increase the purchase cost of household and rents of tenants, thereby curbing housing consumption and non-housing consumption at the same time. The existing literature that utilize different data draw different conclusions regarding the relationship between housing price and consumption.

Weida Kuang in 2011 used household consumption data and housing market data from 1996 to 2008 for 35 large and medium-sized cities in China to estimate the relationship between housing price and consumption. Kuang's article shows that his research mainly focused on the impact of changing housing price on landlord household. This focus is due to the high homeownership rate of Chinese urban residents which is about 82.3% (Kuang 2011). Therefore, the data collected in these 35 cities mainly reflect landlord household's situation. Kuang believed that the domestic housing price has a negative effect on the national non-housing consumption in China. Guangcan Cui (2009) utilized panel method analysing data from 1995 to 2006 of 31 cities in China and found that housing price has a significant impact on total social consumption, in both short-term and long-term. And, the housing price's impact on wealth effect is obvious, but not evident on total social investment. However, some researchers, such as Skinner (1989) suggested that the influence of housing price on the consumption is trivial. Skinner used income dynamics panel data to examine the impact of rising house prices on consumption and savings. Skinner believed that housing price has smaller impact on consumption and after controlling for individual effect, housing prices have no significant effect on consumption. In addition, Engelhardt (1996) used data from 1984 to 1989 of landlord household to study the house prices and home owner saving behaviour. Engelhardt concluded that housing price does not have a negative effect on households' saving. We find that the model Kuang presented in his research includes most factors in interest, thus we decide to utilize his model in our research.

Following Kuang's method, we base our theoretical model on intertemporal choice theory and we make the following assumptions to simplify our theoretical model: (1) At time 0, consumer j in society has to spend some of his or her disposable income y_{j0} to cover the residential expense (which is calculated as certain percentage θ of housing price P_0 multiply by housing area h_0). The rest of income will be divided into daily consumption for good x_{j0}, deposits S_{j0} and rational investment a_{j0}. (2) At time 1, j will use up his or her wealth (disposable income y_{j1} plus return from last period's deposits and investment) on the daily goods consumption x_{j1} and the residential expense θ*P_1 h_1. (3) Rates of return of deposits and investment at time 0 are r_0 and R_0 respectively. (4) The depreciation rate is η. Then, the utility maximization function is:

$$Max\sum_{t=0}^{1}U(X_{jt},h_{jt}) = Max\left(lnX_{j0} + lnh_{j0} + \frac{1}{1+\eta}lnx_{j1} + \frac{1}{1+\eta}lnh_{j1}\right)$$

s.t. $x_{j0} + \theta p_{j0}h_{j0} + S_{j0} + a_{j0} = y_{j0}$

$$x_{j1} + \theta p_{j1}h_{j1} = y_{j1} + (1+r_0)S_{j0} + (1+R_0)a_{j0}$$

After calculating the first order condition and reorganizing the function, we find that: under our assumptions, the increase in housing price would lead to an increase in consumption of daily necessities; increasing disposable income can stimulate increase of daily consumption as well as housing expenses; and the more increase in deposit and investment, the less in the consumption of daily necessities. We expect the variables in our model to have the same directions. The increase in housing price lead to an increase in consumption could be explained by the income effect. House serves as household's main asset. Thus, an increasing price in housing make household wealthier than before and consume more. The increase in deposit and investment leads to a decrease in the consumption. This could be explained by substitute effect. As more wealth would allocate in deposit and investment, the individual would have to lower his or her consumption.

Following our theoretical model, we constructed the functional form of the model such that:

$$con_{jt} = \beta_0 + \beta_1 con_{j,t-1} + \beta_2 p_{j,t} + \beta_3 inc_{jt} + \beta_4 ha_{jt} + \beta_5 ha_{jt} + \beta_6 dep_{jt} + \beta_7 roi_{jt} + \varepsilon_j$$

where the dependent variable is consumption (con_{jt}); independent variable is housing price ($p_{j,t}$); control variables include: lagged consumption ($con_{j,t-1}$), income (inc_{jt}), housing area (ha_{jt}), and deposit (dep_{jt}). In addition, j ranges from 1 to 5, standing for five different cities in China (Beijing, Shanghai, Guangzhou, Shenzhen, and Hangzhou); t ranges from the year 2004 to the year 2017.

Initially, we considered taking a log form of each of the variables to examine the percentage change relationship between independent variables and independent variable. However, we decided not to include the log form. Because we find that variable roi has five negative values, which could not be taken into logarithm form. At first, we tried to transform all the values of variable roi into positive numbers by adding roi_{jt} with one plus the absolute value of the lowest value of roi of each city. However, we found this transformation might be problematic for further analysis. What's more, after conduction this transformation, we found the variables in our model would have different level of integration. This would also be a problem for our further analysis of the model. Therefore, we decide to drop the log and use the original form for our data. In this way, we could utilize all the values of roi and we found all of variables are stationary after first differencing.

Table 1 in the following describes all the variables in our econometric model, Table 2 summarizes the observation number, frequency, observation unit, time period, unite of measure for each data series.

Table 1

Variable	Definition	Description
con_{jt}	consumption	Per capita expenditure of urban households
$con_{j,t-1}$	lagged consumption	Per capita expenditure of urban households last year
$p_{j,t}$	housing price	Residential commodity house purchase price
inc_{jt}	income	Per capita disposable income of urban households
ha_{jt}	housing area	Per capita housing sales area
dep_{jt}	deposit	Per capita savings deposit of urban households
roi_{jt}	return on investment	The national rate of return on investment in the stock market

Table 2

Variable	Frequency	Obs. unit	Time Period (year)	No. of Obs. (per group)	Unit of Measure
con_{jt}	annual	cities	2004-2017	14	Yuan
$con_{j,t-1}$	annual	cities	2003-2016	14	Yuan
$p_{j,t}$	annual	cities	2004-2017	14	Yuan per sq.m
$inc_{\square t}$	annual	cities	2004-2017	14	Yuan
ha_{jt}	annual	cities	2004-2017	14	sq.m
dep_{jt}	annual	cities	2004-2017	14	Yuan
roi_{jt}	annual	country	2004-2017	14	percentage points

*The measuring units of variables con_{jt}, $con_{j,t-1}$, $p_{j,t}$, inc_{jt}, dep_{jt} are values of nominal Yuan (Chinese currency).

In order to eliminate the effects of inflation, we transform the values of con_{jt}, $con_{j,t-1}$, $p_{j,t}$,

inc_{jt}, dep_{jt} from nominal number into real number by dividing each of the value by the corresponding level of inflation that year for the specific city, setting base year as the year of 2003. In addition, we take the first difference of each variable to get stationary data series. We found all the variables have same order of integration. After first differencing, all data series are integrated of order zero.

As for the data source, we collected data of variables $\mathrm{con}_{\mathrm{jt}}$, $\mathrm{p}_{\mathrm{j,t}}$, $\mathrm{inc}_{\mathrm{jt}}$, $\mathrm{ha}_{\mathrm{jt}}$, $\mathrm{dep}_{\mathrm{jt}}$ from National Bureau of Statistics of China. Here are the links of website where we find our data.

consumption($\mathrm{con}_{\mathrm{jt}}$):http://data.stats.gov.cn/easyquery.htm?cn=E0103&zb=A0A00&reg=110000&sj=2017

price ($\mathrm{p}_{\mathrm{j,t}}$): http://data.stats.gov.cn/easyquery.htm?cn=E0105&zb=A03&reg=110000&sj=2017

income ($\mathrm{inc}_{\mathrm{jt}}$):

http://data.stats.gov.cn/easyquery.htm?cn=E0103&zb=A0A00&reg=110000&sj=2017

housing area ($\mathrm{ha}_{\mathrm{jt}}$):

http://data.stats.gov.cn/easyquery.htm?cn=E0105&zb=A03&reg=110000&sj=2017

deposit ($\mathrm{dep}_{\mathrm{jt}}$):

http://data.stats.gov.cn/easyquery.htm?cn=E0105&zb=A04&reg=110000&sj=2017

The rate of inflation we use to conduct the deflation for the nominal variable: http://data.stats.gov.cn/easyquery.htm?cn=E0103&zb=A0901&reg=110000&sj=2017

For financial variable $\mathrm{roi}_{\mathrm{jt}}$, we obtain the disclosed data from CSMAR Database (http://us.gtadata.com.proxy.library.nyu.edu/SingleTable/DataBaseInfo?nodeid=16685&tbid=2606&data_key=%u4E0A%u8BC1%u7EFC%u5408%u6307%u6570%u4E00%u5E74%u5E02%u573A%u56DE%u62A5%u7387%28%25%29).

To be noted that all the data information of variables $\mathrm{con}_{\mathrm{jt}}$, $\mathrm{p}_{\mathrm{j,t}}$, $\mathrm{inc}_{\mathrm{jt}}$, $\mathrm{ha}_{\mathrm{jt}}$, and $\mathrm{dep}_{\mathrm{jt}}$

could be found in the Chinese version webpage of National Bureau of Statistics of China. However, when changed to English version webpage, we could only found date information about con_{jt}, inc_{jt}, dep_{jt} but not variables housing, $p_{j,t}$ and ha_{jt}.

The data utilized is cross-sectional panel data, thus we will first conduct the panel unit root test to decide whether it is stationary. If not, we will take the differencing method to achieve stationary data so that to use linear regression. Considering the small data size, we will conduct OLS pooled regression, Fixed Effect regression and Random effect respective. We will choose the suitable model according to the results from these regression analyses. We believe that the Random Effect/Fixed model will be more reasonable than others, because there should be some autocorrelation relationship among disturbances for different cities. Such autocorrelation could be explained by individual effects. The Fixed/Random effect model will be unbiased, consistent and efficient for our panel data.

Before the regression analysis, we conducted the unit root test for each variable and decide the level of integration of them. According to the result, at the 95% level of significance, all the variables are all I(1), except variable *roi* which is stationary of I(0). *roi* is stationary because that the data values of the five cities in each year are the same. Considering the non-stationarity of variables $\mathrm{con}_{\mathrm{jt}}$, $\mathrm{p}_{\mathrm{j,t}}$, $\mathrm{inc}_{\mathrm{jt}}$, $\mathrm{ha}_{\mathrm{jt}}$, $\mathrm{dep}_{\mathrm{jt}}$, we first difference these variables. There are two good consequences brought by the first differencing: first, linear regression becomes possible; second, the differenced variables are much more uncorrelated with each other (Details are given in the following Table 3 and Table 4).

Table 3

Correlation	con	lag con	inc	p	roi	dep	ha
con	1						
lag con	0.9406	1					

inc	0.9266	0.9024	1				
p	0.6583	0.632	0.6864	1			
roi	-0.1169	-0.1761	-0.1853	-0.1035	1		
dep	0.7397	0.7192	0.7053	0.328	-0.2017	1	
ha	-0.2621	-0.2796	-0.2314	-0.3801	0.1345	-0.2126	1

**Note: variables are all deflated data.*

Table 4

Correlation	d_con	d_lag_con	d_inc	d_p	d_roi	d_dep	d_ha
d_con	1						
d_lag_con	-0.461	1					
d_inc	0.2232	0.1264	1				
d_p	0.3042	-0.1691	0.19	1			
d_roi	0.2132	-0.2437	0.1183	0.1826	1		
d_dep	-0.2038	0.0984	-0.03	-0.1621	-0.3924	1	
d_ha	-0.0063	-0.0695	0.002	0.1438	0.3819	-0.0309	1

**Note: variables are all deflated data; d_(variable name) stands for data after first difference*

After we transformed the data by first differencing, we conducted the fixed effect, random effect, and pooled regression of our data, respectively. The results of regressions are presented in Table 5.1, 5.2, and 5.3 in the following pages. The result of fixed effect regression shows that our model does not fit the regression, given the p-value of F-statistics is over 0.18. Moreover, we notice that the p-value of the coefficients of variable *roi* after first differenced is 0.955 and the coefficient of variable *ha* after first differenced is over -300. This would affect the accuracy of our estimation. The same circumstance occurs in the random effect regression and pooled regression as well. Therefore, we decided to drop these two independent variables and then conducted these three regressions again.

The three regression results improve notably after dropping variables *roi* and *ha*. In the fixed effect regression, the coefficient of variable housing price *h* is statistically different from zero at 95% level of significance (with p-value equals to 0.012). The coefficient value of housing price *ha* equals to 0.2600951, indicating that on average, with one yuan per square meter increase in the change of housing price *ha*, the change of consumption will increase 0.26 yuan. For lagged variable consumption con_{t-1}, the coefficient value -0.469415 indicates that on average, with one yuan increase in the change of lagged consumption, the change of consumption will decrease 0.47 yuan. For variable income *inc*, the coefficient value 0.1559365 indicates that on average, with one yuan increase in the change of income, the change of consumption will increase 0.16 yuan. For variable deposit *dep*, the coefficient value -0.0768772 indicates that on average, with one yuan increase in the change of deposit, the change of consumption will decrease of 0.08 yuan.

Comparing the results of three regression methods, we find that the coefficient of each variable has the same sign: negative for lagged consumption, positive for income, negative for deposits, and positive for housing price. The magnitude of all the coefficient values seems similar to each other in the three regressions as well. However, given the p-value 0.1875 of the Fixed effect regression's F-test, we fail to reject the null hypothesis that the fixed effect model does not fit our model. Whereas, it should be noted that the p-values of variable deposit's coefficient differ significantly in the random effect and OLS regression. The p-value of variable deposit in the random effect regression equals to zero, which is much lower than that in OLS regression (p-value=0.201). Therefore, the regression results of the random Effect regression seem more acceptable.

Following the regression analysis, we conducted tests for heteroskedasticity and autocorrelation. As shown in table 6, the p-value for lrtest is only 0.0333, thus we reject the null hypothesis that there exists heteroskedasticity in the panel time series data. The p-value of Wooldridge test for autocorrelation is 0.9004, thus we fail to reject the null hypothesis that there is no first-order autocorrelation in our panel data. Therefore, at 95% significance level, there are no heteroskedasticity and autocorrelation. Furthermore, we adopt three approaches (Kao, Pedroni and Westerlund) to testing the cointegration among variables. According to the p-values from each test, we reject the null hypothesis that there is no cointegration among variables.

Based on the results of the random effect model, we can conclude that the change of housing price does have a positive effect on the change of consumption for citizens in five main cities of China. The conclusion also in agreement with our expectation. However, the coefficient value of variable housing price is around 0.26 yuan. This implies that the positive effect of housing price on consumption is limited. This result agrees with what Kuang found in his research, but Kuang's results show more significant impact compared to ours result. This may due to the small size of our data set (70 observations in total). The small data size also makes it hard to conduct the Hausman test for our model. This problem can be solved by extending the horizontal size of the panel data, such as collecting data from 30 cities instead of just 5 main cities in China.

To summarize, the increase of housing price has a positive effect on the consumption of citizens in China's five main cities. However, this positive effect is trivial, which in agrees with Skinner's (1989) finding but differs from Kuang's. Considering the fact that our model is mainly based on Kuang's model, we supposed to have a result similar to Kuang's. The main reason of this difference may be that our data size is smaller than Kuang's panel data set which across 35 main cities in China. We also find that: the increase of consumption in last period has a negative effect on current period consumption; the increase of the deposit will discourage the current period consumption; the increase of disposable income will stimulate citizen's current peirod consumption. These findings agree with common sense and the theory of consumers. The conclusion proves our hypothesis that the increasing housing price will stimulate the consumption. To make the results more applicable and supportable, we need more work in further research. It will be reasonable to extend our panel data both horizontally and vertically, in other words, extend both the number of cities and the number of years observed. In this way, we would be able to conduct the Hausman Test. Besides, we could look for other variable instead of the ROI of Shanghai Index to get a better representation of the return on investment rate for our econometric model. Last but not least, more control variables could be added into our model to reduce the weight of constant term in the final result, so that to get a higher R-squared value.

Table 5.1				
Fixed Effect				
Variables	Coefficient	Standard Error (non-robust)	P-value	Observation
d_lag_con	-0.474277	0.1061237	0.011	65
d_inc	0.1539612	0.1601029	0.391	65
d_dep	-0.0783025	0.0410923	0.129	65
d_p	0.2750593	0.0637494	0.012	65
d_roi	-0.1796842	2.987247	0.955	65
d_ha	-300.9105	230.3672	0.262	65
constant	13.70337	4.763767	0.045	65
		R-square: within=0.3978		
		F test that all u_i=0: F(4, 54) = 1.60 , Prob > F = 0.1880		
d_lag_con	-0.469415	0.1058175	0	65
d_inc	0.1559365	0.1305834	0.237	65

d_dep	-0.0768772	0.046755	0.106	65
d_p	0.2600951	0.1193038	0.033	65
constant	13.92606	3.812505	0.001	65
		R-square: within=0.3978		
		F test that all u_i=0: F(4, 56) = 1.60 , Prob > F = 0.1875		

**Note: variables are all deflated data; d_(variable name) stands for data after first difference*

Table 5.2				
Random Effect				
Variables	Coefficient	Standard Error (non-robust)	P-value	Observation
d_lag_con	-0.4467488	0.0861263	0	65
d_inc	0.273091	0.1149055	0.017	65
d_dep	-0.0510167	0.0274995	0.064	65
d_p	0.1764748	0.0485668	0	65
d_roi	0.8181655	2.338216	0.726	65
d_ha	-278.9957	158.6091	0.079	65
constant	11.3556	2.98074	0	65
		Prob > chi2 = 0.0000		
		F test that all u_i=0: F(4, 54) = 1.60 , Prob > F = 0.1880		
d_lag_con	-0.4493181	0.0842298	0	65
d_inc	0.2785733	0.1110222	0.012	65
d_dep	-0.0551942	0.0138407	0	65
d_p	0.1672431	0.05087	0.001	65
constant	11.74392	2.859913	0	65
		Prob > chi2 = 0.0000		

**Note: variables are all deflated data; d_(variable name) stands for data after first difference.*

Table 5.3				
OLS (pooled)				
Variables	Coefficient	Standard Error (non-robust)	P-value	Observation
d_lag_con	-0.4467488	0.1547436	0.005	65
d_inc	0.273091	0.1552861	0.084	65
d_dep	-0.0510167	0.0487386	0.3	65

d_p	0.1764748	0.0837582	0.039	65
d_roi	0.8181655	4.577149	0.859	65
d_ha	-278.9957	381.1548	0.467	65
constant	11.3556	3.846994	0.005	65
		Prob > F = 0.0131		
		R-squared = 0.3430		
d_lag_con	-0.469415	-0.4493181	0.1531197	0.005
d_inc	0.1559365	0.2785733	0.1429916	0.056

(continued)

d_dep	-0.0768772	-0.0551942	0.0426893	0.201
d_p	0.2600951	0.1672431	0.0815053	0.045
constant	13.92606	11.74392	3.846974	0.003
		Prob > F = 0.0056		
		R-squared = 0.3381		

**Note: variables are all deflated data; d_(variable name) stands for data after first difference.*

Table 6			
Test	Distribution	Test statistics	P-value
lrtest for heteroscedasticity	chi2(4)	10.46	0.0333
Wooldridge test for autocorrelation	F(1,4)	0.018	0.9004
Kao test for cointegration			
Modified Dickey-Fuller t		-8.8532	0
Dickey-Fuller t		-10.5306	0
Augmented Dickey-Fuller t		-3.2361	0
Unadjusted modified Dickey-Fuller t		-8.7924	0
Unadjusted Dickey-Fuller t		-10.5291	0
Pedroni test for cointegration			
Modified Phillips-Perron t		1.5698	0.0582
Phillips-Perron t		-4.7451	0
Augmented Dickey-Fuller t		-5.5251	0

Westerlund test for cointegration		-0.9517	0

Work Cited

Berns, Gregory S., David Laibson, and George Loewenstein. 2007. Intertemporal choice—toward

an integrative framework. Trends in Cognitive Sciences 11(11): 482-488.

Cui, G.(2009). An Empirical Study of the Interaction between Real Estate Price and Macroeconomics: Based on Panel Data analysis of 31 Provinces in China. Economic Theory and Economic Management, Vol.1.

Engelhardt, Gary V.“House Prices and Home Owner Saving Behavior” Regional Science and Urban Economics,1996,Vol.26,pp.313-336

Kuang, W. (2011). Housing price and urban household consumption in China. WorldEconomy, 10,21–34 (in Chinese).

Liu, C. and Xiong, W. (2017). *China's Real Estate Market*. [ebook] p.1. Available at: http://wxiong.mycpanel.princeton.edu/papers/HousingChapter.pdf [Accessed 4 May 2019].

Skinner, Jonathan(1996). “Housing Wealth and Aggregate Saving” Regional Science and Urban Economics,1989,Vol.19,pp.305 − 324