How does location affect the outcome of housing-induced poverty in China?

Dengli Yang

Abstract

As Chinese economy booms, Chinese households are experiencing hiking housing prices and expenditures. As the result, the housing affordability issue is hotly discussed in China, especially in large and developed cities like Beijing and Shanghai where housing prices are increasing rapidly. This paper utilizes imputed rent instead of reported expenditures to determine the severity of housing-induced poverty in China. From the results, large cities in China are only experiencing comparatively moderate housing unaffordability problem. Hebei, Guangdong, and Guangxi have much higher percentages of housing-induced poor. Moreover, households in the Northeast and East also have higher likelihoods of encountering housing-induced poverty.

Introduction

Housing prices in China are reaching new heights. over the last ten years, housing prices in Beijing and Shanghai have been tenfold. Because of this increase in price, Chinese have deep burden on housing expenditures and difficulties in paying for rent, especially in large cities. Chinese are believed to be overspending on housing. This paper aims to examine the severity of Chinese housing affordability problem by using rent-to-income ratio, residual income, and housing-induced poverty. Housing-induced poverty exists when housing costs are so high that households cannot afford their minimum and adequate non-housing needs.

The first goal is to estimate housing expenditures for all households since information about homeowners is missing. To accomplish the goal this paper comes up with a hedonic regression to impute rent for all households, employing 2011 China

Household Financial Survey(CHFS) data. After that, I will calculate rent-to-income ratio, residual income, and proportion of housing-induced poor for every province. Whether a household is in housing-induced poverty will be determined based on the residual income. This paper calculates both the overall and provincial proportions of Chinese in housing-induced poverty. To be more precise, households are in housing-induced poverty if households are not in absolute poverty but have insufficient residual income. Lastly, this paper examines the households' characteristics that affect the outcome of housing-induced poverty.

Although other researchers have mentioned housing-induced poverty when they are evaluating Chinese housing affordability, this paper investigate into housing-induced poverty in China. Compared to previous papers, this paper will give more insights about the severity of housing-induced poverty in different provinces or regions. Moreover, the characteristics that affect the probability of experiencing housing-induced poverty will be analyzed. From the estimations, we will have a better idea of Chinese housing affordability problem. By understanding more about housing-induced poverty in China, Chinese policy will be more effective and efficient in addressing and solving housing unaffordability.

Literature Review

Before analyzing the housing-induced poverty in China, I will review housing affordability and its measures. Housing affordability has been gathering worldwide attention. Economists have been studying affordability problems in different countries, including Australia (Gan and Hill, 2009), China (Zax, 2000; Rosen, 2000; Mak et al,

2007; and Yi et al, 2016), and UK (Whitehead, 1991; and Stone, 2006). China's housing affordability problems seem to be apparent, yet research is still needed to further explain, understand, and ultimately solve the problem. The idea that houses are merit goods is explained and elaborated by Whitehead (1991). Hancock (1993) also gives a detailed explanation on the properties of housing.

Housing affordability measures the ability of households to purchase or rent houses at reasonable prices. If housing prices or rent become unreasonable, the households will have affordability problems. The first measure of housing affordability is the price-to-income ratio. The income and price here can be either the mean or the median income and price. This ratio conveys how much a house is worth in terms of household income. Hancock (1993) elaborates and defines housing affordability in a more theoretical way (for a thorough theoretical definition see Hancock, 1993). Generally, the reasonable range for price-to-income ratio is between 3-6, which is evident for most developed countries (Yi et al., 2016). Since Yi et al. observe the ratio in China is significantly above 6, they conclude the houses in China are unaffordable. The high ratio of price-to-income in China is also mentioned in Rosen et al. (2000) and Zax (2000).

Although this ratio is convenient to compute because we only need the price and the income, it's not flawless. The first drawback is that it does not contain information about the quality of the house. The increase of the price-to-income ratio can be attributed to improvement in quality such as amenities, size, and other conditions. Linneman et al. (1992) point out that the rising ratio is due to improvement in amenities such as more bathrooms and better environmental conditioning. Secondly, the median prices and

incomes are not very representative of all households. Gan and Hill (2009) mention that the median of prices and incomes misrepresent the low-income households in Australia which constitute a large section of the Australian population. Using aggregate prices and income, the affordability problem is not significant for some people, but for lower-income households the affordability problem is persistent. The last drawback is that standardizing housing is a troublesome approach in this ratio because houses vary in many ways. Chen et al. (2010) mention the difficulty of standardizing house because houses in Shanghai are segmented. Since houses in Shanghai are differentiated in many aspects, it's very difficult to standardize houses such that the standardization can accurately represent the aggregate houses. Whitehead (1993) also mentions the difficulty of accurately identifying the average housing needed.

The price-to-income ratio may stay within a reasonable ratio for some households, yet they might be living in houses with low and unacceptable quality. Because of the disadvantage of price-to-income ratio, Lerman and Reeder (1987) come up with the idea about "the affordability of adequate housing" which incorporates the quality of the house into the price-to-income ratio. In their paper they conclude that the traditional price-to-income ratio overestimates the severity of affordability problem because without controlling the quality constraint, some households not experiencing affordability problem are classified as troublesome households. If some households have a higher ratio simply because they want a luxury house or a "taste for penthouse", they don't have affordability issues. On the other hand, households which have a ratio within a reasonable range but live in low-condition houses might need housing assistance. Moreover, after controlling the quality of the houses, the location should be controlled as well for

research because housing prices vary tremendously in different locations. Fisher et al. (2009) conduct research on housing in Boston relying on the assumption that housing affordability depends on location.

Linneman et al. (1992) then raise a very valuable question which asks what can the change of ratio reflect. Why the ratio should be increased, decreased, or stay constant in the first place? Moreover, since most households purchase houses with the help of mortgages, financial costs such as interest rates should also be incorporated to understand the affordability problem. They argue that lower interest rates bridge the gap between what a household can afford and how much a house costs. By having lower interest rates, houses are more affordable because less interest accrues over time.

There is another frequently employed ratio called the expenditure-to-income ratio. This ratio dates to the19th century where economists such as Engel started collecting household budget surveys. Since then economists have researched the details of household budgets to obtain information on how households spent their money (for details about the history of budget studies see Hulchanski, 1995). Commonly, a house is affordable for a household if the expenditure-to-income ratio is below 30% (Joice,2014). Roughly speaking the adage or the rule of thumb “a week’s salary for a month’s rent” is equivalent. This ratio is more straightforward and gives researchers insight about how much households are spending on housing.

The residual income approach is differentiated from the expenditure-to-income ratio because the total income is not as insightful as the residual income which conveys the remaining income after paying housing costs. Even though the expenditure-to-income ratio falls below 30%, some households might not have enough money leftover because

they don't have enough money initially. In this case, they are still considered as people in need. By aiding their housing expenditures, they will have enough money to afford the minimum and adequate requirement for non-housing goods. Mulliner et al. (2013) assess the sustainability of housing affordability by using this residual approach. Whitehead (1991) also employs residual income to conduct his research.

What is the minimum and adequate residual income for a household? Stone (2006) goes on to explain the minimum amount needed for an adequate living. He uses the data from the Bureau of Labor Statistics (BLS) Lower Budgets, which stopped publishing after 1981. After 1981, Stone adjusts the budget to the inflation for his research. However, this budget is outdated because the components of a typical household's basket of goods has changed and some goods become futile over the decades. Therefore, the Bureau of Labor Statistics (BLS) Lower Budgets adjusted for inflation is not an appropriate measure for the minimum residual income needed. Whether this lower budget underestimates or overestimates the minimum amount needed is uncertain. In contrast to Stone, Kutty (2005) simply assumes a third of income will be attributed to housing expenditures and the rest is for non-housing goods. Her proportion of residual income to total income comes from the previous studies of household budgets.

We look at the residual income because there is a tradeoff between housing expenditures and residual income. The squeezed effect or the tradeoff between housing and non-housing expenditures is explained by what Stone called "shelter poverty" (1993). Stone argues that households must face tradeoffs between housing and non-housing goods. When housing expenditures are too high, eventually and inevitably the non-housing expenditures will shrink to the level that adequate consumption for non-housing

goods is infringed. Bramley (1990) alludes to the idea of this tradeoff and defines housing affordability as “live on without failing below some poverty standard”. Kutty (2005) names “shelter poverty” as “housing-induced poverty” to stress the induced and squeezing effect of housing expenditures. Most researchers agree that if a household, after paying for housing, cannot afford enough non-housing goods, then it is in housing-induced poverty. By classifying households with the definition of housing-induced poverty or shelter poverty, outreach can be more efficient and effective. Simply, we can help poor households, but this strategy is not the most efficient one. Resources can be reallocated from households below the poverty line and not experiencing housing-induced poverty to households above the poverty line and experiencing housing-induced poverty.

Stone and Kutty both use observed expenditures to calculate the number of households in housing-induced poverty. For renters, the observed expenditures are rent and utilities. For homeowners, the observed expenditures are mortgage payments, utilities, taxes, and insurance. They make comparison between renters and owners in their papers and discover that greater proportion of renters are experiencing housing induced poverty.

Hulchanski (1995) argues that there is no best way to measure housing affordability. In fact, the ratio measures of any kind may result in inefficiency due to the limitations of ratio measures. Simply put, the public or the private cannot just help households with high ratios and neglect households with low ratios. Some households in need of assistance are neglected because their ratios are within a reasonable range while some households who are living finely receive attention because their affordability is

infringed. Moreover, the ratios should be distinguished among different sizes or types of households because the variations in households depend not only on the income of the households, but also the size of the households. Thalmann (1991) concludes the ratio depends on household size, and (2009) constructs a model to determine the exact assistance needed for different household sizes. Usually more assistance and greater attention occur if the households have a larger size. Stone (2006) also probes the affordability problem of the UK by distinguishing different households by the number of members in the households. By undertaking the size of household, Stone determines the amount of assistance needed.

Housing affordability impacts other issues, such as health. Pollack et al (2010) test their hypothesis that high housing expenditures will squeeze non-housing expenditures and thus may cause harm to the health and well-being of household members. Leventhal and Newman (2010) attempt to answer the relationship between housing affordability and children's well-being. In their paper, they conclude the affordability problem has no effect on children's outcomes like health, performance, school achievement, and economic attainment. However, the authors explain this result by displaying the positive impacts of housing unaffordability: often the schools are better, crimes rates are lower, and amenities are improved near expensive houses. These positive impacts of housing unaffordability offset the negative impacts of housing unaffordability.

Solutions for the affordability problem are possible. Chen et al. (2010) contend that if market can't adjust itself in response to the housing affordability problem, then market should provide smaller houses. Smaller dwelling units will be less expensive and therefore more people can afford them. What's more, mortgages should always be

available to households when people plan to purchase houses because the help from financial institutions is imperative for most people. Lastly, one possible solution Stone (1993) offered is to lower the requirement for affordable housing so that lower-quality, yet cheaper houses will be available to poor households. If housing prices cannot drop below current prices, some poorer households are willing to sacrifice the quality of a house in exchange for a decrease in price.

Data and Methodology

The data comes from Chinese Household Financial Data (CHFS). Although the most recent data is from 2013, 2011 data is used in this paper since the 2013 data is unavailable to me. The 2011 data contains 8438 households from 25 provinces: Beijing, Tianjin, Hebei, Shanxi, Liaoning, Jilin, Heilongjiang, Shanghai, Jiangsu, Zhejiang, Anhui, Jiangxi, Shandong, Henan, Hubei, Hunan, Guangdong, Guangxi, Chongqing, Sichuan, Guizhou, Yunan, Shanxi, Gansu, and Qinghai. Renters only originate in 22 provinces: Guizhou, Chongqing, and Shanxi have no renters in the data. Another data source is China Household Income Project (CHIP). This data contains more information on the conditions of dwelling units, but the income information is limited. As the result, CHFS data is chosen despite its inadequate information on the conditions of dwelling units. 52 households have negative reported income, which is less than 1% of total households. The main reasons for negative reported income are due to the negative profit from investing in stock market and operating business.

Renters make up 1329 of the households. These households can be either renting a whole or a partial dwelling unit.[1] Out of 1329 households, 960 households are renting

1. Partial dwelling unit: like a bedroom.

whole dwelling units. The remaining 369 households are renting partial dwelling units. The rest of the households are living in privately-owned dwelling units. When estimating the coefficients to impute rent, only the households renting the whole dwelling units will be used.

Since most households are living in dwelling units that they own, it's imperative to calculate their imputed rent expenditures because they are not living for "free". Previously researchers are not imputing the rent for homeowners. Rather, they only use observed and reported expenditures for homeowners. For homeowners not paying any rent, Kutty and Stone consider them as living for free. The opportunity cost of living in a self-owned dwelling unit should also be taken into consideration.

Hedonic regression is employed to calculate the imputed rent. In order to impute the rent for all households, including renters, the following regression determines the estimated coefficients for intercept, area, and minute-to-downtown.

$$y = \beta_0 + \beta_1 area + \beta_2 mindt + \varepsilon 2 \quad 2$$

Note that the area here is the useable area of the household. It's differentiated from the construction area. The reason for using useable area over construction area is that renting households didn't report the construction area for their dwelling units. Although self-owned households reported the construction area and construction area is more related to rent, the coefficient for the construction area cannot be estimated due to lack of information. To check for the interaction effect between area and minute-to-downtown, a separate regression needs to add an interaction term and check its statistical significance.

2. $Y_{t:}$ Annual rent expenditures for household t
   $Area_t$: the useable area of a dwelling unit (measured in square meters) for household t
   $Mindt_{t:}$ Minute-to-downtown (measured in minutes) for household t

$$y = \beta_0 + \beta_1 area + \beta_2 mindt + \beta_3 area * mindt + \varepsilon$$

To consider the provincial effect on housing rent, this following regression has all the province dummy variables[3]. The logic is to test for any provincial effect. If the coefficients are different for different provinces, then each province has different provincial effect on rent.

$$y = \beta_0 + \beta_1 area + \beta_2 mindt + \sum \beta_i x_i + \varepsilon \quad (3)$$

The next regression breaks area and minute-to-downtown into categories since the impact of a variable on rent could be different for different values. The objective is to find a breaking point such that the coefficients are different and statistically significant. If there exists a first breaking point, other potential breaking points require further investigation as well. Failing to find a breakpoint will leave the regression the same as before. Repeating the same process would discover reliable breaking point or points for minute-to-downtown as well.

$$y = \beta_0 + \beta_1 area_1 + \beta_2 area_2 + \beta_3 mindt_1 + \beta_4 mindt_2 + \sum \beta_i x_i + \varepsilon$$

After obtaining estimated coefficients, imputed rent for all households can be calculated. This paper estimates rent for 8104 households out of 8438 households. The households from the provinces Chongqing, Guizhou, and Shanxi are deleted due to the absence of renters.

When rent is imputed for all households, Household income also needs to be adjusted in order to determine the rent-to-income ratio. For those households living in self-owned dwelling units, income isn't merely the reported income. Their adjusted

3. x_i: the province dummy. 1 if household in this province, 0 if not.

income should be their opportunity cost, the imputed rent, plus reported income. For renters, income stays the same since they are paying for their dwelling units. By adjusting the income to opportunity cost, the real income level is determined.

Calculating adjusted rent-to-income ratio uses imputed rent and adjusted total income. Theoretically, the adjusted ratio is the same as unadjusted ratio for renters since their total income is the same. If imputed rent is estimated correctly, it should approximately equal to reported rent for renters. Similarly, adjusted residual income is the approximately same for renters if rent is imputed correctly. For self-owned households, adjusted residual income is the adjusted income minus the imputed rent.

Households are in absolute poverty if adjusted total income is below a certain level. Households are in housing-induced poverty if there are not in absolute poverty, but after paying for housing, have insufficient residual income for non-housing goods. Residual income is insufficient if it is below a proportion of the certain level. Researches uses a proportion based on previous budget studies. From the findings in budget studies, they implement the average proportion for households as the reasonable proportion. By dividing the number of housing-induced households by the total number of households, the percentage of housing-induced poor is calculated. This percentage is calculated for different provinces to observe any provincial difference.

Most researchers are using official poverty line to determine the proportion of households in absolute poverty. In her paper, Kutty (2005) employs the official American poverty line and assumes that the minimum and adequate residual income is two-thirds of the poverty line. This paper employs a different line because the official Chinese poverty line is unsatisfying when it is compared to international standard. The official Chinese

poverty line in 2011 was 2300 yuan. It is much less than the poverty line set by World Bank. In this paper, I will use World Bank's poverty standard, which is $1.25 per day. By adjusting for currency, the poverty standard in this paper is 2965.63 yuan.[4] It's approximately 30% higher than the 2011 official Chinese poverty line. Households are in near-poverty if income is less than 150% of the poverty line. Using 2965.53 as poverty line, the near-poverty line is 4448.45 yuan. Households in near-poverty are inclined to be in housing-induced poverty, after paying for housing costs.

A different proportion of income on non-housing goods is also implemented. Since the conventional ratio two-thirds is based on previous households' budgets studies, the Chinese ratio should be determined by using Chinese household budget studies. Based on the numbers from *China Statistics Yearbook* (2011), it's calculated that, on average, 25.7% of the income was distributed to the housing expenditures. This proportion is derived using the expenditures and income sections.[5] As the result, the housing-induced poverty line is 2203.46 yuan.

Using the international poverty line and average proportion of housing expenditures for Chinese, this paper computes the proportion of households in housing-induced poverty for 22 provinces. In summary, a household is in absolute poverty if adjusted income is below 2965.63yuan. A household is in housing induced poverty if adjusted income is at least 2965.63 yuan and residual income is below 2203.46 yuan.

When the outcome of housing-induced poverty is determined for every household, a regression is utilized to estimate the probability of housing-induced poverty for Chinese

4. Exchange rate in 2011 on average was 1 dollar to 6.5 yuan.

5. The proportion is calculated based on the average housing expenditures, mean income, and distribution of Chinese population.

households. The dependent variable is the 1 if a household is determined to be in housing-induced poverty and 0 if not. Independent variables are number of people, number of children, number of elderly, and regional dummy variables.

Result

Choosing the regression

From table 1, the estimated coefficients for the first regression are 6913.4, 78.75, and -2.84, respectively. The anticipated sign for area is positive since the larger the dwelling unit, the higher the rent is. In this regression, the coefficient for area is indeed positive. The 90% confidence interval is between 62.32 and 95.18 and reassures the positive sign. The expected sign for minute-to-downtown is negative, since if a household is further away from downtown, rent is cheaper. The coefficient for minute-to-downtown is marginally small at -2.84: Holding area constant, rent decreases by only 2.84 yuan for every extra minute. Even though the sign corresponds to the prediction, this value implies that minute-to-downtown has a relatively small impact on rent in this regression. When this coefficient is tested at a 90% confidence interval, the troubling interval -35.46 and 29.79 indicates the impact of minute-to-downtown on rent is ambiguous. Since R-squared is only 0.057, this regression cannot accurately estimate the rent for most households.

Nonlinear variables such as $Area^2$ and $Mindt^2$ are also considered. However, since these nonlinear variables yield lower t statistics values, compared to linear variables, they are not in regressions. The second regression adds an interaction variable to observe the

4. Exchange rate in 2011 on average was 1 dollar to 6.5 yuan.

5. The proportion is calculated based on the average housing expenditures, mean income, and distribution of Chinese population.

relationship between area and minute-to-downtown. The sign of interaction variable is positive, which suggests that holding area constant, as minute-to-downtown increases, the rent decreases to a less extent. In other words, an additional minute-to-downtown has a greater increase in rent for larger dwelling units. However, the coefficient is not reliable because it's not statistically significant. Also, R-squared only increases by 0.002, from 0.057 to 0.059, which is essentially nothing.

As the result of previous regressions which yield trivial R-squared, locational variables are added. 21 dummy variables are in the third regression. Even though 25 provinces total are available in the whole sample, the subsample for renting households only has 22 provinces left because three provinces (Chongqing, Guizhou, and Shanxi) don't have single observations available. Moreover, to eliminate the dummy variable trap, which resulted in singularity, the province with the least observations, Guangxi, is also dropped. Although many provinces such as Shanxi, Jiangsu, and Qinghai only contain few renting households, all provinces are retained since this paper aims to evaluate the holistic China.

Table 1: Four Regressions to Impute Rent

	(1) Without province dummy variables	(2)With interaction variable	(3) Without categories	(4) With province dummy variables
Variable				
Intercept	6913.4	7802.68	-696.22	-163.93
Area	78.75***	63.83***	86.38***	
Mindt	-2.84	-39.61	-67.85***	
Area*Mindt		0.68		
Area1				146.65***
Area2				93.45***
Mindt1				-234.05**

Mindt2				-76.97***
Beijing (114)[6]			17909.8	15455.43
Tianjin (26)			6478.74	4160.74
Hebei (30)			-106.49	-2056.46
Shanxi (3)			-1352.54	-3038.8
Liaoning (22)			1292.73	-1561
Jilin (10)			1338.02	-904.69
Heilongjiang (64)			2865.17	602.5
Shanghai (59)			10644.4	7798.7
Jiangsu (4)			7379.69	6473.64
Zhejiang (35)			7342.96	5057.28
Anhui (17)			5478.48	2682.16
Jiangxi (13)			5124.62	3144.81
Shandong (19)			817.86	-912.34
Henan (25)			6.38	-2330.92
Hubei (24)			-1229.73	-4297.52
Hunan (54)			-432.55	-2879.03
Guangdong (65)			31476.85	28503.61
Sichuan (18)			-2515.93	-4673.16
Yunnan (18)			1446.54	-473.97
Gansu (10)			203.45	-2172.1
Qinghai (5)			3894.91	632.24
Guangxi (2) omitted				
R-squared	0.057	0.059	0.481	0.492

*p<0.1 **p<0.05 ***<0.01

Area1: area smaller than 90 square meters.
Area 2: area equal to or larger than 90 but less than 180 square meters. Area larger than 180 is omitted
Mindt1: minute-to-downtown smaller than 20 minutes.
Mindt 2: minute-to-downtown equal to or larger than 20 but less than 150 minutes. Minute larger than 150 is omitted.

With categorical variables, not only the coefficients for smaller area and larger area are different, they are both statistically significant at 0.01 level. By distinguishing smaller and larger area with a breakpoint of 90, the coefficients for both variables are still statistically significant. From the regression, rent increases more if a dwelling unit is smaller. The coefficients for both sizes increase from 86.38 to 146.65 for smaller are

6. Number of households in the province is indicated in the parentheses.

and 93.45 for larger area. This means that if a dwelling unit is closer to downtown, travel time has a larger negative effect on rent. The significance level for minute-to-downtown decreases from 0.01 to 0.05, which is still statistically significant enough to interpret the result. The difference between less minutes and more minutes is significant. The reason for choosing 90 and 20 as the breakpoints is because other break points make coefficients statistically insignificant. Further dividing area and minute-to-downtown into more categories doesn't yield better coefficients.

Distributions for renting and all households are very distinct. The mean and median area for renters are 65.15 and 59.9, respectively. Only 20% of the renters have an area above 90. The mean and median minute-to-downtown are 23.73 and 15, respectively. About half of the renting households have minute-to-downtown above 20. For all households, the mean and median area are 101.15 and 80 respectively. The mean and median minute-to-downtown are 38.02 and 30 respectively. From the comparison, their distributions vary notably. However, out of these four regressions, the last one is chosen to impute rent for all households.

Comparing imputed rent and reported rent

The imputed rent for renters is greater than the reported rent for renters by 10.69%. Therefore, the imputed rent approach overestimates the rent for renters. However, there is smaller difference between reported rent for renters and imputed rent for all households. This suggests that if the rent for renters and all households has the same distribution, then this model accurately estimates the imputed rent for all households.

For renters, the range for imputed rent is smaller than that for reported rent This appearance suggests that the rent imputed has a much more concentrated distribution. The mean only shifts to the right by 10.68%, but the range is shrunk from 99976 to 41197.74. Since the change in range has little effect on the mean of rent, it will not make the following computations on housing-induced poverty incorrect.

Table 2: Summary Statistics on Rent

Renters	Observations	Mean	Median	Min	Max
Reported Rent	637	11977.04	6000	24	100000
Imputed Rent	637	13258.24	10770.63	51.54	41249.28
All households					
Imputed Rent	6889	12222.64	8662.97	7.53	108294.4

The correlation coefficient between estimated rent expenditures and household income is also checked. Without restricting rent and income, the correlation between estimated rent and household income is 0.185. Restricting the estimated rent to be nonnegative, the correlation coefficient is 0.184. Restricting both estimated rent and household income to be nonnegative, the correlation is 0.183. A correlation of 0.183 suggests a moderate correlation between estimated rent and income. There is no significant difference among these correlation coefficients; however, households with negative estimated rent and reported income are discarded to avoid confusion. Neglecting households with negative imputed rent and income is justified since only a very small fraction of households has negative reported income and estimated rent. Deleting them doesn't make a notable distinction.

The main reasons for negative estimated rent are relatively smaller dwelling units and greater time to downtown. The average area for households with nonnegative

estimated rent is 104.79 whereas the average area for households with negative estimated rent is 57.19. The average minute-to-downtown for households with nonnegative estimated rent is 33.18, whereas the average minute-to-downtown for households with negative estimated rent is 106.03. Since area and minute-to-downtown have a positive and negative impact on estimated rent, respectively, it's not surprising that households with negative estimated rent have significantly smaller dwelling units and are much further from downtown. Subsequent results only focus on nonnegative imputed rent and estimated income for convenience.

Adjusting income and residual income

Adjusted income is equal to the reported income for renters and is equal to imputed rent plus reported rent for homeowners. Renters don't have adjusted income since they are paying for rent. From the table, renters have approximately 50% greater unadjusted income compared to all households. This large difference is mostly attributed to a higher income for renting households. After adjusting income, renters still make 25.19% more than all households. However, the medians for renters and all households are much closer. According to the median, renters only make 16.75% more than all households. After adjusting the income, renters have a lower median. After all, the income distribution for renters is more skewed toward the right, regardless of the modification.

Table 3: Summary Statistics for Income and Residual Income

All households	Observations	Mean	Median	Min	Max
Observed Income	6889	57685.45	30664.26	0	3000000
Adjusted Income	6889	68351.21	41089.87	48.945	3056984
Residual Income	6889	45462.8	20553.93	-68001.7	2988723

Adjusted Residual Income	6889	56128.57	29172.73	0	3000000
Renters					
Observed Income	637	85568.09	35800	-50000	3000000
Residual Income	637	72919.4	26778.72	-79000	2960000
Adjusted Residual Income	637	72267.85	29937.59	40.43	2977284

The difference between observed income and adjusted residual income is precisely the average imputed rent for renters. The same difference is slightly smaller for all households. By restricting both estimated rent and adjusted income to be positive, the correlation between them is 0.2422. Although it's still a moderate correlation, it is a much higher correlation than that between estimated rent and adjusted income before adjusting the income.

Calculating percentage of housing-induced poverty

Without any constraint, the rent-to-income ratio isn't reliable. the mean rent-to-income ratio for all households is 2.8672 and the median is 0.2526. Clearly, this mean is unrealistic because the maximum average rent-to-income ratio is one since households can't spend more than they earn on housing. This ratio might be greater than one for some households living in their parents' or friends' dwelling units, but these households are discarded because they have significantly higher ratios which disturb the mean.[7] The average ratio is significantly increased by the outliers: Compared to the mean, the median is relatively low. By strictly constraining the ratio between 0 and 1, the mean is 0.2867 and the median is 0.2153. In this case the mean and median are relatively close to each other compared to other cases. The issue here is that only 5992 households have ratios between 0 and 1 which suggests that more than a

7. For instance, one of the household made 1 yuan in 2011 and the reported rent was 4400.

thousand households have greater imputed rent. By strictly restricting the ratio between 0 and 1, the measurement error is 14.4% which is a relatively high measurement error. Many homeowners have less income when compared to their estimated rent.

After adjusting the income and strictly restricting the ratio between 0 and 1, the mean and median are 0.3007 and 0.2148, respectively.[7] 7. After adjusting the income, the mean is 0.5663 and the median is 0.2268. This is the mean and median without restriction. After restricting the ratio between 0 and 10, the mean and median are 0.3614 and 0.2256. Only 417 households have higher estimated rent than adjusted income. From among with 7306 households, 6889 households maintain ratios between 0 and 1, which corresponds to a measure error of 4.7%. Adjusting the income reduces the measurement error.

Table 4: Adjusted Ratio, Adjusted Residual Income, and Percentage

Overall	Households	Adjusted Ratio	Adjusted Residual Income	Housing-induced poverty	Percentage
	6889	0.315	56128.56	491	7.13
Province					
Beijing	374	0.2966	97983.79	10	2.67
Tianjin	189	0.248	56724.22	15	7.94
Hebei	326	0.3847	25009.49	44	13.5
Shanxi	176	0.2981	34957.88	7	3.98
Liaoning	292	0.225	41106.18	19	6.51
Jilin	236	0.2814	32516.82	21	8.9
Heilongjiang	349	0.2808	41579.88	30	8.6
Shanghai	458	0.2811	84601.83	22	4.8
Jiangsu	397	0.3732	60888.26	23	5.79
Zhejiang	502	0.299	83261.19	33	6.57
Anhui	339	0.3323	44659.41	19	5.6
Jiangxi	167	0.3363	39725.37	10	5.99
Shandong	403	0.2848	44937.54	35	8.68
Henan	562	0.3092	32204.91	35	6.23

Hubei	443	0.265	30439.67	18	4.06
Hunan	371	0.2469	40328.32	23	6.2
Guangdong	664	0.5138	117339.1	88	13.25
Guangxi	70	0.4365	24890.1	10	14.29
Sichuan	179	0.2466	34734.62	7	3.91
Yunnan	159	0.2973	23832.35	13	8.18
Gansu	171	0.2213	29106.68	7	4.09
Qinghai	62	0.1653	81010.12	2	3.23

Province: 22 provinces total.
Housing-induced poverty: number of housing-induced households in the province.
Percentage: percentage of housing-induced households in the province.

However, there is a discrepancy between the distributions for renting households and all households. For renting households, the mean is 0.1943 and the median is 0.1309 after restricting the ratio between 0 and 1. The measurement error is 10%, whereas 10% of the renting households have a ratio below 0 or above 1. Renters' mean and median are significantly lower than the mean (0.3007) and median (0.2148) for all households with the same restriction. This is somewhat reasonable since renters have higher income and live in much smaller dwelling units. The average area for all households is almost double the average area of renting households and the average minute-to-downtown is marginally greater for all households. In short, renting households are renting dwelling units that have significantly smaller sizes and are relatively further away from downtown.

Surprisingly, Qinghai has the lowest rent-to-income ratio and a very high adjusted residual income, behind only the most developed provinces like Guangdong, Beijing, Shanghai, and Zhejiang. However, since Qinghai only contains 62 observations, errors are possible. Guangdong has an adjusted ratio of 0.5138, the highest in 22 provinces; it also has the highest adjusted residual income. Guangxi also has a ratio above 0.4, which

is the second highest. Guangxi's situation is reasonable since it has the second lowest residual income. Overall the average adjusted ratio is 0.315, considerably larger than 0.257, the average ratio calculated from *Chinese Statistics Yearbook* (2011).

The total number of households in absolute poverty is 587, or .8.52% of total households, without adjusting the income. After adjusting the income, only 0.68% of all households are in absolute poverty. This difference is mostly attributed to the increase in adjusted income, which considers the opportunity cost. Before adjusting income, 982 households, or 14.25% households, are in housing-induced poverty. This is because many homeowners have income less than their imputed rent. After adjusting the income, 491 households or 7.13% of the total households, are in housing-induced poverty. The percentages of absolute poverty and housing-induced poverty decline sharply. Comparing to all households, 11.62% renters are in absolute poverty and 2.20% renters are in housing-induced poverty. Renters suffers more from absolute poverty and have a relatively moderate percentage in housing-induced poverty.

Using Kutty and Stone's measure to calculate the proportion of housing-induced poor, imputed rent is replaced by the observed expenditures such as rent and utilities for renters and mortgage payments and utilities for owners. Under their measure, the mean and median expenditure-to-income ratios are 0.1222 and 0.049, respectively. This result is much lower than the measure this paper employed since most homeowners only have utilities as their housing expenditures. Even though approximately 8% households have housing mortgages, most of the mortgage payments cannot be calculated due to missing information. Therefore, without comparing with different measures, Stone and Kutty's measure underestimates the expenditures for Chinese households.

Following their measure, 2.09% households are in housing-induced poverty. This percentage is close to the percentage of renters in housing-induced poverty, but it's much lower than the percentage for all households. This conventional measure, which only uses observed housing expenditures, computes a much lower percentage compared to my measure. Since imputed rent for renters overestimates the actual rent for all households by 2.09%, a greater proportion of households are in housing-induced poverty under my approach.

Table 4 also illustrates the percentage of households in housing-induced poverty for 22 provinces. Beijing has the lowest percentage of housing-induced poor and Guangxi has the highest. Guangxi, Hebei, and Guangdong have considerably higher percentages, which suggest households in these three provinces suffer much more than households in other provinces. To my surprise, Beijing has the lowest percentage of housing-induced poverty. However, Beijing is widely regarded as a city with unaffordable housing.

Households with insufficient income always suffer from housing-induced poverty. For households not in absolute poverty, 7.20% of them are in housing-induced poverty. For households in absolute poverty, 100% of them are in housing-induced poverty. However, out of 538 households in housing-induced poverty, only 47 households are in absolute poverty. Most households are in housing-induced poverty not because they are poor initially, but because they, after paying for housing, have inadequate residual income. For households in near-poverty, 90.48% of them are in housing-induced poverty. Although households in near-poverty are not in absolute poverty, they are very likely to experience housing-induced poverty as well.

Table 5: Estimating the Probability of Housing-induced Poverty

	(1) With province dummies	(2) Without province dummies	(3) Without regional dummies	(4) With regional dummies
Observations	6889	6889	6889	6889
Intercept	0.06**	0.11***	0.05	0.07***
Number of people	-0.01***	-0.01***	-0.01***	-0.01***
No children	0.02***	0.02***	0.02***	0.02***
two or more children	0.03***	0.03***	0.03***	0.04***
one child omitted				
No elderly		-0.01	0.01	0.01
One elderly		-0.02	0.02	0.01
two or more omitted				
Northeast				0.02**
East				0.02***
West				0.01
Middle omitted				
Beijing	-0.02		-0.02	
Tianjin	0.03		0.03	
Hebei	0.1***		0.1***	
Shanxi	0.01		0.01	
Liaoning	0.03		0.03	
Jilin	0.06		0.06	
Heilongjiang	0.05		0.05	
Shanghai	0.01		0.01	
Jiangsu	0.03		0.03	
Zhejiang	0.03		0.03	
Anhui	0.03		0.03	
Jiangxi	0.04		0.04	
Shandong	0.05		0.05	
Henan	0.03		0.03	
Hubei	0.01		0.01	
Hunan	0.03		0.03	
Guangdong	0.11***		0.11***	
Guangxi	0.12***		0.12***	
Sichuan	0.01		0.01	
Yunnan	0.05		0.05	
Gansu	0.02		0.02	
Qinghai omitted				

*p<0.1 **p<0.05 ***<0.01

Northeast: Liaoning, Heilongjiang, and Jilin
East: Beijing, Tianjin, Hebei, Shanghai, Jiangsu, Zhejiang, Shandong, and Guangdong
Middle: Shanxi, Anhui, Jiangxi, Hunan, Hubei, and Henan
West: Guangxi, Sichuan, Yunnan, Gansu, and Qinghai

Intercepts are always statistically significant except for the third regression which doesn't have constraints on location. The results for number of elderly are mixed. The coefficients for elderly are negative in the first regression and are positive in second and third regressions. Overall the number of elderly has no statistically significant impact on the likelihood of housing-induced poverty. Although the above regressions divide number of elderly into three subgroups, separating number of elderly into less or more subgroups doesn't yield a better result.

Number of children has a positive impact on outcome for households if they have no children or more than one child, compared to households with one child. Households with no children are more likely to be in housing-induced poverty since they don't have the financial ability and stability to afford a child. Therefor for households with no children, they are poorer than households with one child. Households with one child are not impoverished by having a child. Moreover, households with more than one child are more likely to be in housing-induced poverty since they can't simultaneously cover the costs for all the children and afford adequate non-housing goods.

In these four regressions, number of people has a negative impact on the probability, but it's relatively small at -.01. It's surprising that a greater number of people in a household decreases the likelihood. One possible reason is that when number of people increases, number of working adults increases as well. Thus, a household with more income sources will less likely be in housing-induced poverty. Holding number of people constant, a household is more likely to be in housing-induced poverty if it has no

child because of the same reason: lack of financial ability. A household with more than one child also has a higher probability because it will have less working adults, the income sources.

In the regression with provincial dummies, only three provinces (Hebei, Guangdong, and Guangxi) have statistically significant coefficients. Belonging to these three provinces strongly affects the probability: it's at least 10% more likely to be in housing-induced poverty if households come from these three provinces. They also have the highest coefficients in the provinces. The high percentages are reflected in the previous tables, whereas they have the highest percentages of housing-induced poor and highest rent-to income ratios. Despite being statistically insignificant for most provinces, they have positive coefficients. Beijing is the only province with negative coefficients, which corresponds to the lowest percentage of housing-induced households.

Households in the Northeast and East have higher chances of falling into housing-induced poverty, compared to households in the middle. The coefficients for them are statistically significant at 0.05 and 0.01 level respectively. Although households in the West also have higher chances of falling into housing-induced poverty, it's not as high as households in Northeast and East, and not statistically significant.

Conclusion

Conventional measure fails to accurately estimate the percentage of housing-induced poor due to missing information. However, imputed-rent measure tends to overestimate the severity of housing-induced poverty since imputed rent is higher than the actual rent and homeowners are usually less likely to suffer from housing

expenditures. When taken the opportunity cost into account, homeowners have greater probability of falling into housing-induced poverty and having "housing affordability" problem.

Although in some provinces and regions, housing-induced poverty is not very rampant, it is not neglectable. Three provinces Hebei, Guangdong, and Guangxi require special attention since they have significantly higher percentages of housing-induced poor, compared to other provinces. Another takeaway is that the public policy should focus more on provinces that are less populated and less developed. For large and developed cities like Shanghai and Beijing, they are only experiencing a relatively moderate housing unaffordability problem (in terms of housing-induced poverty). Regionally, Households in the Northeast and in the East have higher probability of suffering housing-induced poverty. Government should provide more assistance in these two regions, compared to the West and the Middle.

Households with no children and more than one child also deserve more attention. Since Chinese government is relaxing the One Child Policy and encouraging households to have more children, government should not let housing costs deter Chinese households from having children, especially for households with no children. Policy should aim to help households with no children so that China will have larger younger population and avoid aging population.

One limitation of imputed-rent approach is the number of variables in the regression. With only two dependent variables, the imputed rent is questionable. R-squared is at most 0.492, which implies more than half of the households aren't captured. Although there are few more choices for dependent variables, they are not chosen for

various reasons. Initially, the regression has one dummy variable indicating rural or urban location of the household, it is dropped due to insufficient number of households in some provinces. Number of renting households is also an issue in this paper. It's barely adequate to analyze Chinese renters with just 637 households from 22 provinces.

There are no independent variables indicating the quality of dwelling units. Furnishing is one possible independent variable that is discarded. Although renting households reported their furnishing level of their dwelling units, the survey didn't inquire the same information for homeowners. On the other hand, owners reported the number of rooms in dwelling units while renters didn't. Thus number of rooms also can't be included in the regression. Compared to my regression to impute the rent, Zax (2000) has more than 10 dependent variables which estimate the rent more accurately.

Using the regressions from renters and then imputing the rent for all households is questionable because renting households can't represent all households. They have different distributions for area and minute-to-downtown. The coefficients for categorical variables from the regression might not be correct for all households.

From the table for renting households, we can see that many provinces only have few households. Provinces like Beijing, Shanghai, and Guangdong have considerably larger number of renting households and therefore the results are more believable. Shanxi, Jiangsu, and Guangxi only have few observations available. The imputed rent for households in these three provinces are therefore very inaccurate.

Endogeneity issue exists when I am estimating the probability of housing-induced poverty. Income of a household might simultaneously affect probability on the left and

number of people and number of children on the right. With higher income, a household tends to have lower probability, more people, and more children. This problem is unaddressed in this paper and requires further discussion.

References

Beer, A., Kearins, B. & Pieters, H. 2007, "Housing affordability and planning in Australia: The challenge of policy under neol-liberalism", *Housing studies,* vol. 22, no. 1, pp. 11-24.

Bramley, G. 1990, “Access, affordability and housing need”.

Chen, J., Hao, Q., Stephens, M. 2010, "Assessing Housing Affordability in Post-reform China: A Case Study of Shanghai", *Housing Studies,* vol. 25, no. 6, pp. 877-901.

Fisher, L., Pollakowski, H. & Zabel, J. 2009, "Amenity-Based Housing Affordability Indexes",*REAL ESTATE ECONOMICS,* vol. 37, no. 4, pp. 705-746.

Gan, Q. & Hill, R.J. 2009, "Measuring housing affordability: Looking beyond the median",*Journal of Housing Economics,* vol. 18, no. 2, pp. 115-125.

Hancock, K.E. 1993, "'Can Pay? Won't Pay?' or Economic Principles of 'Affordability'", *Urban Studies,* vol. 30, no. 1, pp. 127-145.

Hulchanski, J.D. 1995, "The Concept of Housing Affordability: Six Contemporary Uses of the Housing Expenditure-to-Income Ratio", *Housing Studies,* vol. 10, no. 4, pp. 471-491.

Joice, P. 2014, "Measuring Housing Affordability", *Cityscape*, vol. 16, no. 1, pp. 299.

Kutty, N.K. 2005, "A new measure of housing affordability: Estimates and analytical results",*Housing Policy Debate,* vol. 16, no. 1, pp. 113-142.

Lerman, D.L. & Reeder, W.J. 1987, "The Affordability of Adequate Housing", *Areuea Journal,* vol. 15, no. 4, pp. 389.

Leventhal, T. & Newman, S. 2010, "Housing and child development", *Children and Youth Services Review,* vol. 32, no. 9, pp. 1165-1174.

Linneman, P.D. & Megbolugbe, I.F. 1992, "Housing Affordability: Myth or Reality?", *Urban Studies,* vol. 29, no. 3-4, pp. 369-392.

Mak, S.W.K., Choy, L.H.T. & Ho, W.K.O. 2007, "Privatization, housing conditions and affordability in the People's Republic of China", *Habitat International,* vol. 31, no. 2, pp. 177-192.

Mulliner, E., Smallbone, K. & Maliene, V. 2013, "An assessment of sustainable housing affordability using a multiple criteria decision making method", *Omega-International Journal OF Management Science,* vol. 41, no. 2, pp. 270-279.

Pollack, C.E., Griffin, B.A. & Lynch, J. 2010, "Housing Affordability and Health Among Homeowners and Renters", *American Journal of Preventive Medicine,* vol. 39, no. 6, pp. 515-521.

Rosen, K.T. & Ross, M.C. 2000, "Increasing Home Ownership in Urban China: Notes on the Problem of Affordability", *Housing Studies,* vol. 15, no. 1, pp. 77-88.

Stone, M.E. 1993, "Shelter poverty: new ideas on housing affordability" in , pp. xi+423-xi+423.

Stone, M.E. 2006, "What is housing affordability? The case for the residual income approach", *Housing Policy Debate,* vol. 17, no. 1, pp. 151-184.

Stone, M.E. 2006, "A Housing Affordability Standard for the UK", *Housing Studies,* vol. 21, no. 4, pp. 453-476.

Thalmann, P. 1999, "Identifying Households Which Need Housing Assistance", *Urban Studies,* vol. 36, no. 11, pp. 1933-1947.

Thalmann, P. 2003, "'House poor' or simply 'poor'?", *Journal of Housing Economics,* vol. 12, no. 4, pp. 291-317.

Whitehead, C.M.E. 1991, "From Need to Affordability: An Analysis of UK Housing Objectives", *Urban Studies,* vol. 28, no. 6, pp. 871-887.

Yi, D., Huang, Y. & Fan, G. 2016, "Social Capital and Housing Affordability: Evidence from China", *Emerging Markets Finance and Trade*, vol. 52, no. 8, pp. 1728-1743

Zax, Jeffrey. 2000, "Housing Reform in Urban China".

China Statistics Yearbook. 2011.