The Effects of Transaction Taxes on the Housing Market in China: Evidence from Shanghai

Bozhen Liu*

November 6, 2022

Abstract

Many governments have encouraged low-income housing; however, affordability concerns for the middle class are often overlooked. In 2005 the Shanghai government began to promote the supply of middle-class houses (i.e., mid-sized houses at reasonable prices) with the introduction of relative lower transaction tax rates for housing units below 140 square meters (sqm). This paper uses novel data to assess the consequences of these transaction taxes on the housing market in Shanghai. I find that the combination of a 1.5% relative decrease in the deed tax and a 5.5% relative decrease in sales tax for housing units below 140 sqm increased the supply of units that were 100 - 140 sqm by 41.3% compared to units that were 140 - 180 sqm. I also find that the relative lower transaction taxes led to a 6.55% decrease in the average price of housing units below 140 sqm compared to those above, and I present evidence that suggests this price decrease resulted from shifts in the composition of units developed after the policy. Overall, my results show that transaction taxes in Shanghai had substantial consequences on the distribution of house sizes.

1 Introduction

Many governments have encouraged low-income housing; however, affordability concerns for the middle class are often overlooked. Since housing affordability is the relationship between housing costs and income, it not only concerns low-income households but also the middle class as it demands affordable access to

*Institution: The George Washington University; E-mail: bozliu@gwu.edu; Tel: (617)390-3993. I am extremely grateful for the extraordinary guidance from Prof. Stuart, as well as comments from Prof. Smith, Prof. Brooks and Prof. Regan. All errors are my own.

mid-sized houses. Middle-income housing affordability is socially and economically important because the middle class usually occupies the major share of population and is the predominance of a country's economic development. This paper studies the effect of a policy that promoted the construction of middle-income houses through introducing higher transaction tax rates for housing units larger than 140 sqm, or equivalently, relative lower transaction tax rates for units smaller than 140 sqm since tax rates on these units did not change, in Shanghai in May 2005.

The housing market's rapid development in China has significantly contributed to the country's economy since the commercialization of housing units in 1998. Barth et al. [2012] estimate that investment in residential housings accounted for about 25% of total fixed asset investments and one-sixth of the gross domestic product (GDP) growth in China. However, the dramatic rise in housing prices has negatively impacted people's livelihoods. Although low-income households are the most vulnerable to the housing price hike, middle-income households have also felt the increased financial burden. (Fang et al. [2015][1]). To sustain housing market growth while promoting home-ownership for middle- and low-income households, the government has carefully intervened in the market to encourage the supply of small and reasonably-priced houses. Unlike low-income housing, which usually relies on administrative regulation[2] or public subsidies to ensure the needs of households unable to afford private housing, middle-income housing in China is often encouraged through economic instruments. Among these instruments, differential taxes have been frequently adopted by the government to guide the market. Whether and how these instruments accomplish the goals has therefore become an important question for researchers and policy makers.

In May 2005, the Shanghai government classified housing units into ordinary and luxury properties and levied different transaction tax rates on these two groups. Specifically, housing units larger than 140 sqm were considered as luxury properties. After this policy change, the deed tax on the sales value for ordinary properties became 1.5% lower than that for luxury properties, and sales tax on the capital gain was 5.5% lower for ordinary properties than for luxury properties if the resale occurred two years after purchase. The purpose of this policy was to encourage the construction of ordinary properties.

Previous research has not been able to examine the effects of introducing the tax differentiation policy because of data paucity. I overcame this challenge through access to two comprehensive datasets from the Shanghai Housing Bureau, which contained the universe of housing transactions spanned the period of the

[1]Fang et al. [2015] find that low-income home-buyers generally endured price-to-income ratios of around eight, and that metric was sometimes over ten in developed cities.

[2]For example, a policy that requires developers to build certain amount of affordable houses within complexes.

policy introduction (i.e., from 2004 to 2007). Given that the Chinese housing market has been heavily regulated since 2005, to investigate the introduction of this policy could help avoid confounding with other policies and hence provide the clearest estimates of the policy effects.

Using supply data of new housing units from 2004 to 2007 in Shanghai, bunching estimations were first applied to identify whether the tax policy led to a discrete change in the supply of units at the 140 sqm threshold. The results find no evidence of bunching. This may due to the presence of optimization frictions that prevented developers from precisely adjusting the floor areas of housing units to meet the notch. Developers usually offer a few standardized structures within a hundred- or thousand-unit complex, and to provide a type of housing unit slightly below 140 sqm may not be the best use of resources given the maximum floor-area-ratio (FAR) and other site planning restrictions, and may not be the most popular choice among home-buyers.

Given that there was little local bunching behavior around the floor area threshold, I then investigated whether the tax differentiation policy had effects on housing units distributed in broader floor area ranges. Difference-in-differences (DID) strategies were employed to estimate the policy effects on the supply distribution of new housing units. Housing units in the (140, 180] sqm range were defined as the control group, and units in the (100, 140] sqm range were considered as the treatment group. The results show that, on average, there was almost a 41.3% increase in the supply of housing units in the treatment group after the policy compared to the supply of units in the control group. Moreover, the average treatment effects were not significant in the inner and middle annuli of Shanghai, whereas there was a 35.9% relative increase in the supply of treated units in the outer annulus. This finding may due to the increasing price elasticity of demand moving away from the city center. Home-buyers in the outer annulus are more price sensitive, hence the same degree of financial incentives could induce a larger response.

Furthermore, additional DID estimations were exploited to explore the policy effects on house prices. Standard economic theory predicts that taxes should be capitalized into housing prices (Oates [1969]; Yinger et al. [1988]). On the other hand, developers can adjust quality or the supply of housing units according to policy changes in the newly-built market, which can also affect the prices. Therefore, the overall policy effect on house prices is ambiguous that requires closer examination. Since the aforementioned supply data set did not contain price information, another sales data set that contained transaction information for new housing units from the same period was used to explore the policy effects on prices. DID estimations find that the policy had decreased the average price per sqm of housing units in the treatment group by

6.55% relative to the price of units in the control group. To investigate whether developers adjusted the characteristics of housing units in response to the tax differences, the decomposition method by Gelbach [2016] was applied to examine whether the decrease in treatment effects could be attributed to quality changes. The results find weak evidence due to the lack of appropriate quality variables. Moreover, I explored whether there were changes in the housing supply that affected prices. Housing units were assigned into groups according to the pre-policy price distribution and separate DID estimations were conducted for each group. The results show that in lower percentiles of the price distribution the average number of sales of the treated units increased significantly compared to the controlled units after the policy, which caused a relative decrease in the average price of the treated units. This evidence suggests that effects resulted from changes in the composition of housing units developed after the policy dominated the capitalization effects.

This study contributes to two main strands of the literature. First, this paper contributes to the growing research on market behavior in response to notched taxes. Many studies have estimated bunching at price or time notches (Besley et al. [2014]; Kopczuk and Munroe [2015]; Slemrod et al. [2017]; Best and Kleven [2018]; Tam [2018]). Since sales prices and timing of transaction are manipulable, most existing research has emphasized tax evasion and avoidance behavior in response to notched taxes (Kleven et al. [2011]; Rees-Jones [2018]; Slemrod [2019]Agarwal et al. [2020]; Fan et al. [2022]). This paper differs from the aforementioned literature for two main reasons. First, the notched taxes of the policy were based on the floor areas of housing units, which were subject to objective government measurements and hence manipulation was more difficult. Therefore, this paper examines real (supply-side) responses to the notched taxes rather than evasion behavior. Second, there were optimization frictions that prevented developers from precisely adjusting the floor area of housing units to meet the notch. As a result this paper explores the impacts of notched taxes on broader ranges rather than estimating its local effects around the notches.

Second, this paper sheds light on the limited literature on the impacts of taxation on the newly-built market. There have been several studies on tax policies in the housing market (e.g., Dachis et al. [2012]; Besley et al. [2014]). Unlike the majority of literature that has examined the influence of taxation in the resale market, this paper explores its effect in the newly-built market, which is equally important but remains understudied. In contrast to developed countries, where resale homes account for the majority of housing stocks, newly-built homes in China and other developing countries such as India, Vietnam, Thailand, etc., occupy a large share of the housing stock. A distinct feature of the newly-built market is that its supply-side response to policy changes is often more pronounced. For example, in the newly-built market, developers can change the structure of housing units or adjust their marketing strategies according policy changes, whereas in the resale market the housing structure and total supply are rather fixed. Therefore, it is critical to study how tax policies may affect the newly-built market differently.

Specifically, this paper seeks to determine how the tax differentiation policy affects the floor area distribution of new housing units and their prices. This paper is most close to Zhang et al. [2017], which examine the same policy but emphasize different contexts and times. Utilizing the price threshold, they find small bunching of transactions whose prices were below the tax notch. Moreover, their results show that the tax policy temporarily increased transaction volume among ordinary properties by 1.7% and decreased their prices by 7.3% compared to those of luxury properties. My paper differs from their research in two aspects. First, I examine market responses to the introduction of floor area thresholds, which could lead to more pronounced changes. Second, their paper studies the effects of policy adjustment, which may be confounded due to own persistent lagging effects as well as overlapping effects of other policies. Instead, my paper uses data that spanned the period of the policy introduction, which allows me to investigate the policy when it was first initiated and provides clearer estimations of the policy effects.

2 Institutional Background

Development of the Chinese housing market has been through three important phases (Wu et al. [2012]; Fang et al. [2015]). The first phase was from the founding of the People's Republic of China in 1949 until 1988. During this stage, the government controlled the ownership of land and was the monopoly supplier of housing. Housing units were developed by working units, which often were state-owned enterprises, and allocated to employees as a form of in-kind compensation. The second phase lasted between 1988 and 1998, when a series of housing reforms were implemented to privatize the market. The 1988 Constitutional Amendment upheld the government's land ownership but permitted leasing land use rights to private developers for up to 70 years. Since then, individuals were allowed to purchase housing units either from their working units at subsidized prices or from the private market. Finally, in 1998, the State Council issued the 23rd Decree, which terminated housing developments by working units and the decade-long housing allocation[3]. After that, individuals were required to purchase housing units in the private market.

The Chinese housing market has experienced a remarkable boom since the commercialization of hous-

[3]Housing developments and allocation in progress are permitted to be completed. The deadline could be determined by local governments with a little flexibility.

ing units in 1998. From 2003 to 2014, development of the total floor area increased by 100 billion square feet (National Bureau of Statistics of China, 2014; Chivakul et al. [2015]). The government encourages development of the housing market to boost GDP and employment. The ratio of real estate investment to total fixed asset investment increased from 13% in 1997 to 20% in 2009 (National Bureau of Statistics of China). As of 2014, the construction industry employed 29 million workers, representing 16% of urban employment (Glaeser et al. [2017]). In some years, land sales account for as much as 40% of local government revenue (Wu, 2015). The rapid development of the housing market has made it an indispensable engine for Chinese economic growth.

Research has documented double-digit growth of housing prices in China (Ahuja et al. [2010]; Fang et al. [2015]; Chen and Wen [2017]). The dramatic rise in housing prices naturally raises concerns of affordability. To sustain the housing market growth while promoting home-ownership for middle- and low-income households, the Chinese government has carefully intervened in the market. In 2005 the government began the cyclical housing market interventions to mitigate the prolonged price hikes, such as to increase interest rates and taxes. Since then, various forms of tightening and loosening policies have been employed to calm and boost the market, respectively, under different economic circumstances (Zhou [2016]; Deng et al. [2012][4]).

Unlike low-income housing, which usually relies on administrative regulation or public subsidies to ensure the needs of households unable to afford private housing, middle-income housing in China is often encouraged through economic instruments. The tax differentiation policy is among the most typically used instruments for the government to guide the market. It provides financial incentives for purchases of smaller housing units and imposes heavier taxes on purchases of larger units. As this policy could help the government to collect tax revenue by differentiating purchases while promoting middle-income housing, it has played an important role in guiding the housing structure in China.

In May 2005 the Chinese central government launched an initiative to classify housing units into ordinary and luxury properties and levied different transaction tax rates on their respective transactions, such that luxury properties were taxed at higher rates. There were three necessary requirements for housing units to be qualified as ordinary properties: (1) the FAR of the complex was greater than or equal to one; (2) the sales price per sqm of the housing unit was less than the maximum price set by municipal governments[5];

[4] Deng et al. [2012] document notable mean reversion in the annual growth of housing prices, which could be explained by the frequent policy adjustments.

[5] Taking into account the local price level, the price cap could be set differently for regions within cities (e.g., the government

and (3) the floor area of the housing unit was less than or equal to 120 sqm. The central government allowed for a maximum of 20% upward adjustment of the floor area requirement by municipal governments. In late May 2005, the Shanghai government announced the localized requirements, which set the floor area threshold at 140 sqm and the policy went into effect in June 2005[6]. The purpose of this policy was to encourage construction of ordinary properties. The overall analysis of this paper is centered on the floor area requirement.

The policy was initiated in May 2005 and subsequently adjusted several times according to various market conditions. One benefits of this paper's focus on the policy initiation is that it may show the largest effects on the housing market because it represented the largest change in tax rates. Another key advantage is that it could help avoid confounding with other policies given the frequent policy interventions afterwards. Although there were other policies happened during the examined period, such as the introduction of 90/70 rule in June 2006 that required 70% of housing units in new complexes have to be under 90 sqm, the increase of the minimum down payment requirement from 20% to 30% for units above 90 sqm, and the increase of holding period for sales tax relief from two years to five years, later robustness tests demonstrate that their impacts were minimal to the result. Moreover, despite the policy was introduced nationally, only the effects in Shanghai were studied due to data limitations. Since China's housing market is relatively new, micro-level data availability was rare until 2004 as Shanghai became the first city to digitize housing transactions.

Figure 1 shows how a typical housing complex looks in China. The government sets a maximum FAR for any buildable land and the developer who acquires the land will then decide on the housing structure to be built. In order to maximize profits, developers typically build up to the FAR limit with high-rise building complexes that comprise limited types of standardized housing structures.

Table 1 demonstrates the tax liabilities for the two types of properties. There were two different tax liabilities between ordinary and luxury properties. First, the deed tax for ordinary properties was 1.5% of the sales price versus 3% for luxury properties. Second, sales tax (which was only paid at resale) was zero for ordinary properties versus 5.5% for capital gains on luxury properties if the resale occurred after two years from the time of purchase. However, if the resale occurred within two years, sales tax would be 5.5% of the total resale price for both types of properties. Despite the fact that home-buyers only need to pay the deed tax when buying new housing units, sales tax will affect their cost of resale and therefore should be

sets different caps for the three annuli in Shanghai). The cap would be adjusted periodically to account for price appreciation and inflation.

[6]See Appendix 1 for the detailed localized requirements.

Figure 1: A Typical Housing Complex in China

Source: `https://ganzhou.newhouse.fang.com/2014-08-26/13619917.htm`

factored into their decision-making process if they are rational.

Table 1: Tax Liabilities for Ordinary and Luxury Properties

Type		**Ordinary Properties**	**Luxury Properties**
Deed Tax		1.5% of the sales price	3% of the sales price
Sales Tax	$<$ two years	5.5% of the resale price	
	$\geq$ two years	Zero	5.5% of the capital gain

The deed tax for ordinary properties was 1.5% of the sales price versus 3% for luxury properties. Sales tax (which was only paid at resale) was zero for ordinary properties versus 5.5% for capital gains on luxury properties if the resale occurred after two years from the time of purchase. However, if the resale occurred within two years, the sales tax would be 5.5% of the total resale price for both types of properties.

Theoretically, the skyrocketing housing price should make sales tax a critical factor in individuals' home-buying decisions as it directly affects the after-tax resale value. Figure 2 shows the yearly supply of total floor area and average sales price per sqm of new housing units in Shanghai between 2002 and 2010. The supply of total floor area had been declined since 2004 due to the scarcity of residential land. Its pattern reflects major economic cycles during this period, where in the economic downturn (i.e., years 2007 and 2009) more land was leased to developers to boost GDP from construction and hence brought larger supply of total floor area. In addition, the average sales price increased more than three-fold during this period. The economy's heavy reliance on the housing market and frequent government interventions have emboldened individuals to believe that the market is "too big to fail" (Fang et al. [2015]). The perception of continuous price appreciation and the lack of suitable investment vehicles (Fang et al. [2015]; Chen and Wen [2017]) has induced strong incentives for individuals to invest in the housing market. Therefore, the expectation of housing units' after-tax resale value could play an important role in home-buyers' purchasing decisions.

Figure 2: Sales of New Homes 2002 - 2010 in Shanghai

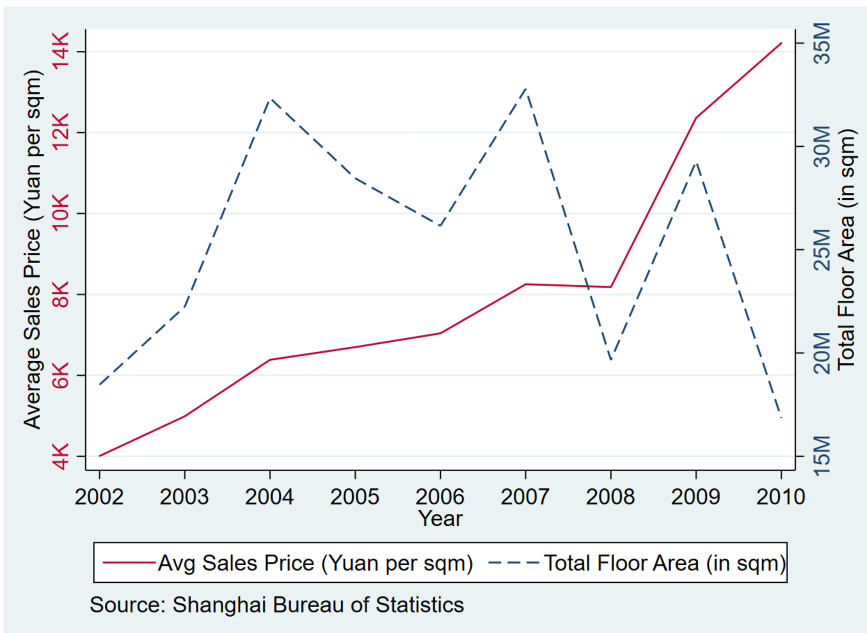


Figure 3 presents few scenarios to demonstrate the differences in tax liabilities between ordinary and luxury properties. The calculation of tax liabilities used the tax rates shown in Table 1 and the house prices shown in Figure 2. Figure 4 graphically illustrates the policy-induced variation in taxes. The figure considers two housing units that were just below or above 140 sqm, but otherwise identical. The deed tax was calculated by multiplying the average sales price per sqm of the year of purchase by the floor area and corresponding tax rates. Similarly, sales tax was calculated by multiplying the resale price by the floor area and corresponding tax rates. Since there was only price information for new housing units available, the resale price was obtained by discounting the average sales price of new units of the year in which the resale occurred by a 3% annual depreciation rate[7]. Finally, sales tax was discounted back to

[7] I assume the housing unit lasts for 30 years, thus the depreciation rate is about 3%.

the 2005 currency equivalent using a 3.5% risk-free rate[8]. The net differences in tax liabilities between ordinary and luxury properties were shown in the 2005 currency and translated to percentage based on the total purchase price. These differences ranged from ¥14,065 to ¥48,964, or 1.5% to 5.2% of the purchase price, in different circumstances. As housing prices continued to grow over time, the main differences in tax liabilities arose from sales tax, which was based on the price appreciation. However, since it is not possible to disentangle effects of the two taxes given that they were implemented simultaneously and that sales taxes were capitalized and dependent on expectations of future housing prices, this paper does not intend to study their respective marginal effects, and instead analyzes the combination of the two taxes.

Figure 3: Scenario Analysis

Scenarios	140- Sqm (Ordinary)		140+ Sqm (Luxury)		Net Difference	Net Difference
	Deed Tax	Sales Tax	Deed Tax	Sales Tax	(in 2005 Yuan)	(in Percentage)
I - A home buyer bought the housing unit in 2005 and sold it in 2006.	¥14,065.80	¥50,796.32	¥28,131.60	¥50,796.32	¥14,065.80	1.5%
II - A home buyer bought the housing unit in 2005 and sold it in 2007 (after two years).	¥14,065.80	¥0.00	¥28,131.60	¥7,618.04	¥21,683.84	2.3%
III - A home buyer bought the housing unit in 2005 and sold it in 2008.	¥14,065.80	¥0.00	¥28,131.60	¥5,192.29	¥19,258.10	2.1%
IV - A home buyer bought the housing unit in 2005 and sold it in 2009.	¥14,065.80	¥0.00	¥28,131.60	¥28,063.83	¥42,129.63	4.5%
V - A home buyer bought the housing unit in 2005 and sold it in 2010.	¥14,065.80	¥0.00	¥28,131.60	¥34,898.11	¥48,963.91	5.2%

The calculation of tax liabilities relied on the tax rates shown in Table 1 and the housing prices shown in Figure 2. The deed tax was calculated by multiplying the average sales price per sqm of the year of purchase by the floor area and corresponding tax rates. Sales tax was calculated by multiplying the resale price by the floor area and corresponding tax rates. Since there was only price information for new housing units available, the resale price was obtained by discounting the average sales price of new units of the year in which the resale occurred by a 3% annual depreciation rate. Finally, sales tax was discounted back to the 2005 currency equivalent using a 3.5% risk-free rate. The net differences in tax liabilities between ordinary and luxury properties were shown in the 2005 currency and translated to percentage based on the total purchase price.

The effects of this policy could emerge relatively quickly because of the presale mechanism of new housing units. In China, the majority of sales of new housing units are made through presale, whereby developers list and sell the right to claim units that are to be built in the future (usually within few years)[9]. Therefore, developers are able to adjust their strategies rather quickly in response to policy changes.

[8] According to tradingeconomics.com, the average 10-year China government bond yield between 2005 and 2010 was approximately 3.5%.

[9] According to Win.d and Huatai Securities, the ratio of presale home to complete home in China were about 65%, 70%, and 75% in 2005, 2006, and 2007, respectively

Figure 4: Differences in Taxes

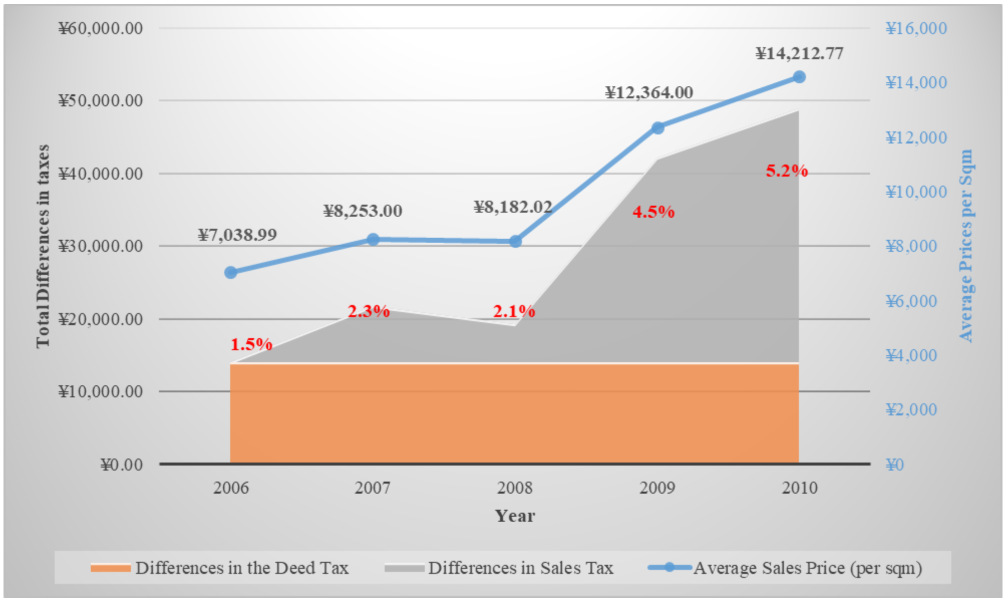


Figure 4 visually presents the estimations in Figure 3. The orange area represents the differences in the deed tax between ordinary and luxury properties. The gray area are the differences in sales tax between the two type of properties. The blue line denotes the yearly average price per sqm of new houses in Shanghai.

3 Housing Supply

3.1 Supply Data

Since the study is interested in understanding how the tax differentiation policy affected the floor area distribution of new housing units, it emphasizes new housing supply. I use two privileged datasets from dataln[10], which is a leading data service company specializing in the real estate market in China. It was allowed, with permission, to collect supply and sales data of housing units from the Bureau of Housing Administration in Shanghai. The study has been granted access to the raw supply and sales data of housing units in Shanghai from April 2004 to June 2007[11]. The supply dataset contained information on new housing unit reported to the Bureau of Housing Administration for listing approval[12]. The majority of this supply was "future supply" which developers reported as housing units to-be-built in the future. However, the dataset was not able to distinguish already-built from to-be-built housing units. Since the projected sales

[10] `https://www.dataln.com/views/main/login.jsp.html`

[11] The sales data began from June 2004.

[12] According to the Law of the People's Republic of China on Urban Real Estate Administration, developers can apply for permission to list presale housing units after funds put for construction have exceeded 25% of the total budgetary investment for the complex.

price is set after the permission is granted, this dataset did not contain price information. Similarly, the sales dataset used in this and the following sections collected information on the sales of new housing units, which also included presales, but the proportion of such "futures" would be smaller than those in the supply dataset. This is because the sales data included transactions of housing units that had been on the market for a while and sold in the current period. The supply dataset had 675,781 observations while the sales dataset contained 471,336 observations. The number of sales is smaller because not all the supply will be listed on the market at once after the permission is granted, instead developers may list new housing units in phases according to different market environments.

In this section, the supply dataset was used to examine the policy effects on the floor area distribution of new housing units. Compared to sales data, supply data reflects more timely responses to the policy. Table 2 displays several summary statistics for the supply dataset. Since the data collection was in the remote past when digitization was newly applied, many variables were missing. Here, only the statistics of available variables of interest are reported. After data cleaning there were 675,781 new housing units with a mean floor area of 115.7 sqm. As Shanghai is a monocentric city, the administrative defined inner annulus is the closest annulus to the city center, whereas the middle and outer annuli are further away from the city center. Figure 5 presents the map of Shanghai with the three annuli identified. The average floor area of housing unit was the largest in the inner annulus. In China, the wealthy generally live closer to the city center because they value superior social amenities, which are often located near the center, and disfavor commuting to work. Therefore, the average floor area of housing unit is larger for those closer to the city center. The average floor area also decreased from 126.0 to 109.5 sqm after the policy. The average number of bedrooms was 2.3, which was similar to that of housing units in different annuli. Because of the one child policy, a typical household usually consists of two parents and one child, and hence two- and three-bedroom housing units are most common. FAR and greening rates are complex-level attributes. The average FAR was 1.8 and it declined monotonically from the inner to outer annuli, which was consistent with the literature on the FAR gradient for the monocentric city model (e.g. McMillen [2006]; Barr and Cohen [2014]). The greening rate is measured as the ratio between green space area and total land area of the complex and is often marketed as a desired amenity of complexes. The average greening rate was 40.4, which was also similar to that of complexes in different annuli.

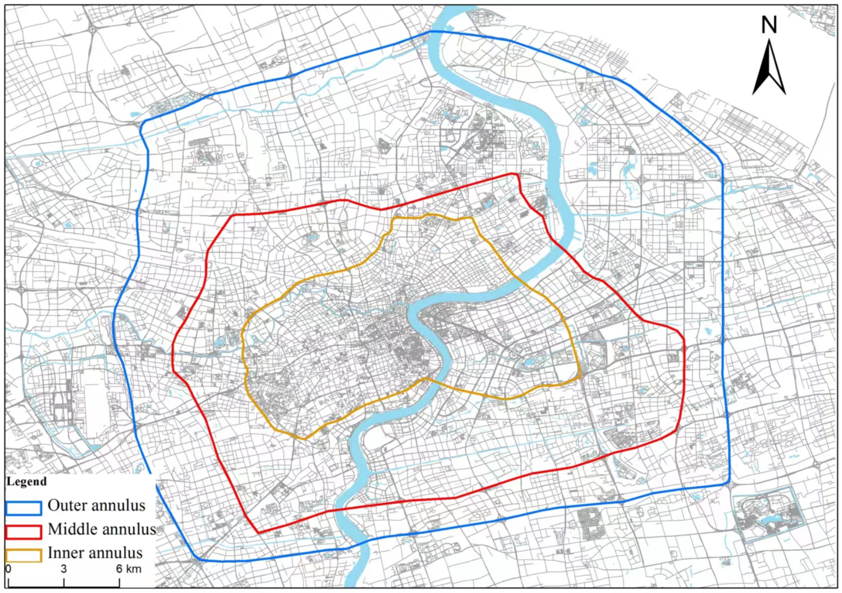
Figure 6 plots the floor area distribution of new housing units before and after the tax differentiation policy was implemented with size bins of five sqm. The red vertical line refers to the floor area threshold

Table 2: Summary Statistics of the Supply Dataset

Variable	Obs	Mean	Std. dev.	Min	Max
Panel A: Full Sample					
Floor area	675781	115.7	56.16	5.7	5004.8
Number of bedrooms	510069	2.3	0.70	1.0	5.0
Floor-to-area ratio	618573	1.8	0.83	0.1	40.0
Greening rate	621264	40.4	8.69	0.2	90.0
Panel B: Inner Annulus					
Floor area	71863	133.9	61.19	25.4	1239.6
Number of bedrooms	52767	2.4	0.81	1.0	5.0
Floor-to-area ratio	67295	3.0	1.18	0.2	36.0
Greening rate	67238	40.7	10.08	0.2	78.0
Panel C: Middle Annulus					
Floor area	195557	113.6	54.00	16.1	5004.8
Number of bedrooms	143730	2.3	0.70	1.0	5.0
Floor-to-area ratio	184213	1.9	0.61	0.2	40.0
Greening rate	185033	41.7	8.23	0.3	79.0
Panel D: Outer Annulus					
Floor area	408361	113.4	55.65	5.7	3434.0
Number of bedrooms	313572	2.3	0.68	1.0	5.0
Floor-to-area ratio	367065	1.5	0.60	0.1	8.9
Greening rate	368993	39.6	8.56	0.3	90.0
Panel E: Before the Policy					
Floor area	250239	126.0	61.58	22.9	3434.0
Number of bedrooms	115351	2.4	0.72	1.0	5.0
Floor-to-area ratio	229532	1.8	0.96	0.2	36.0
Greening rate	230408	41.4	8.92	0.3	90.0
Panel F: After the Policy					
Floor area	425542	109.5	51.74	5.7	5004.8
Number of bedrooms	394718	2.3	0.69	1.0	5.0
Floor-to-area ratio	389041	1.7	0.74	0.1	40.0
Greening rate	390856	39.7	8.49	0.2	90.0

The summary statistics was calculated from the full sample from April 2004 to June 2007. The calculation was based on residential housing units.

Figure 5: A Map of Shanghai



that differentiated between the two types of properties. The floor area of housing units before the policy appears to be normally distributed; however, after the policy was implemented, the distribution seems to become left skewed. The bunching of housing units in 50 - 60 and 70 - 80 sqm bins in 2005 - 2007 was because of the affordable housing program, which required developers to build housing units of 50 - 60 and 70 - 80 sqm for two- and three-household families, respectively. Although the program was launched around the same period as the tax differentiation policy under examination in this paper, it is still possible that this program will not confound the analysis because the target population was fairly different between the two policies. The government set strict criteria for affordable housing application and resale to prevent arbitrage. Home-buyers who could afford private housing units will not and cannot participate in the program. Also, as the floor area threshold of the tax differentiation policy was separated from the size requirement of affordable housing program, it is plausible to assume that the impacts of the program on the study sample was limited.

Figure 6: Floor Area Distribution of New Housing Units

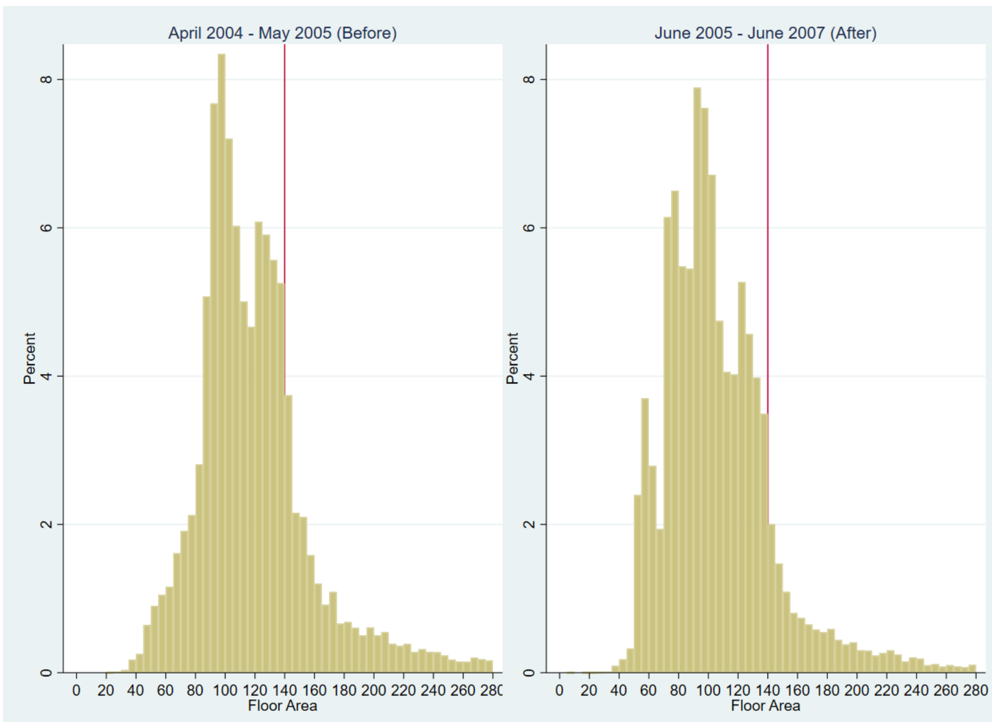


Figure 6 plots the floor area distribution of new housing units before and after the tax differentiation policy with size bins of five sqm. The red vertical line refers to the floor area threshold that differentiates between the two types of properties. The bunching of housing units in 50 - 60 and 70 - 80 sqm bins in 2005 - 2007 was because of the affordable housing program, which required developers to build housing units of 50 - 60 and 70 - 80 sqm for two- and three-household families, respectively.

4 Bunching Estimation

Following the conventional literature on notch analysis (Chetty et al. [2011]; Kleven and Waseem [2013]) and the comprehensive guide to bunching estimation in Best and Kleven [2018], similar procedures were applied to examine the policy effects on the floor area distribution of new housing units around the size threshold. The baseline specification takes the following form:

$$c_i = \sum_{j=0}^{q} \beta_i (z_i)^j + \sum_{k=\bar{h}_v^-}^{\bar{h}_v^+} \gamma_k I\{i = k\} + \mu_i \tag{1}$$

where c_i is the number of transactions in floor area bin i, z_i is the distance between bin i and the threshold $\bar{v}$, and q is the order of the polynomial. The second term excludes a region (v^-, v^+) around the notch that could be distorted by bunching behavior, and μ_i is a residual that reflects mis-specification of the density equation.

$$\hat{B} = \sum_{i=v^-}^{\bar{v}} (c_i - \hat{c}_i) \tag{2}$$

$$\hat{M} = \sum_{i=\bar{v}}^{v^+} (\hat{c}_i - c_i) \tag{3}$$

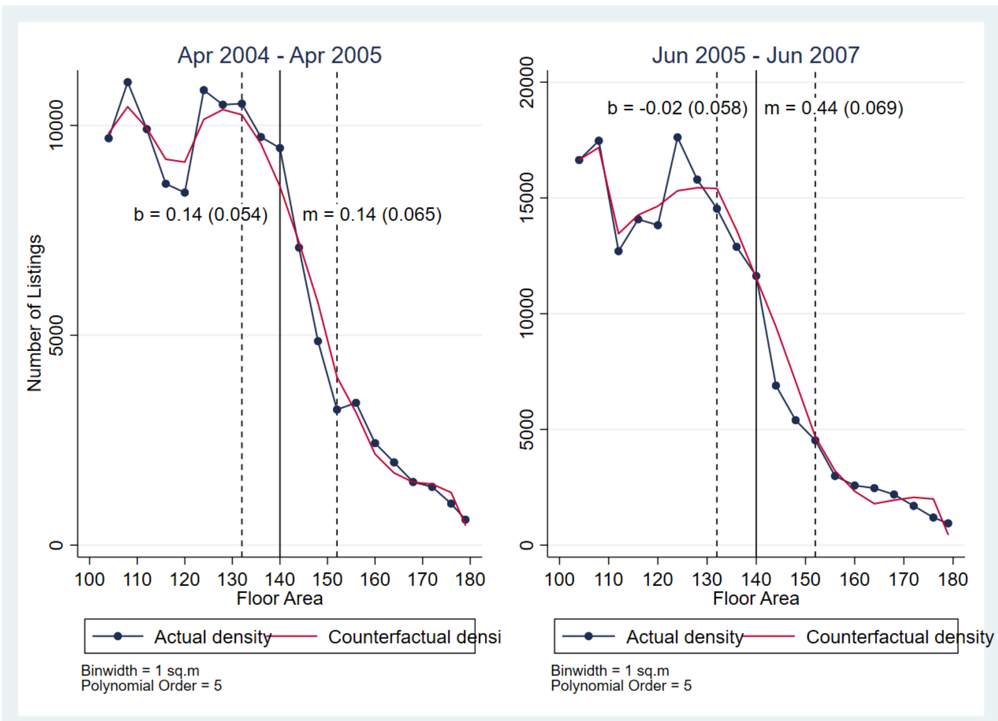
Equation (1) estimates the counterfactual floor area distribution of new housing supply in the absence of the policy. $\hat{B}$ estimates the excess bunching by taking the difference between the observed and counterfactual counts and $\hat{M}$ estimates the missing mass. The estimation aims to examine the changes in $\hat{B}$ and $\hat{M}$ after the policy implementation.

Figure 7 illustrates the results of the bunching estimation. The estimation used bins of one sqm and a polynomial of order five to fit the observed distribution[13]. Both the actual and the counterfactual densities were shown by aggregating bins up to four-sqm wide. The vertical dashed lines denote the upper and lower bounds of excluded region around the notch, and they were set at the point where the observed density changed signs of slopes. b and m were B and M scaled by the average counterfactual frequency in the excluded range, with its standard error displayed in parentheses. All standard errors were calculated by bootstrapping the procedure 200 times.

The results find little bunching below and a weak missing mass above the notch. The bunching mass

[13] Housing units were grouped into bins of one-sqm wide and decimal numbers were rounded up to the following integer, e.g., a housing unit with 140.3 sqm was grouped into the bin of 141 sqm.

Figure 7: Bunching Estimations Before and After the Policy Implementation



The estimation used bins of one sqm and a polynomial of order five. Both the actual and the counterfactual densities were shown by aggregating bins up to four-sqm wide. The vertical dashed lines denote the upper and lower bounds of excluded region around the notch, and they were set at the point where the observed density changed signs of slopes. b and m were B and M scaled by the average counterfactual frequency in the excluded range, with its standard error shown in parentheses.

before the policy was .14, whereas it decreased to -.02 after the implementation. It implies that the mass of supply in the excluded range below the floor area threshold decreased, which contradicts with the assumption that there should be bunching of new housing supply below the notch given the tax incentives for smaller units. The missing mass above the threshold increased from .14 to .44 after the policy, which suggests a decrease of supply in that excluded range and supports the hypothesis. However, these estimations were not statistically significant. This puzzling finding could due to optimization frictions such that adjustment or information costs in shaping behavioral responses to tax incentives is high (Kleven [2016]). Optimization frictions have been documented in various studies. For instance, Kleven and Waseem [2013] find little adjustment of labor supply in response to tax incentives among the majority of workers due to optimization frictions. Similarly, Liu et al. [2021] show that 90% of firms do not adjust turnover according to value-added tax notches. In this case, one reason that there was little bunching could due to frictions in construction, which prevent developers from precisely adjusting the floor areas of housing units to meet the notch. Developers usually offer a small selection of standardized structures within a hundred- or thousand-unit complex. Therefore offering a type of housing units sized at slightly below 140 sqm may not be the optimal use of resources given the maximum FAR and other site planning restrictions, and may not be the most popular choice among home-buyers.

Despite there being little impacts on the floor area distribution of new housing supply around the threshold, there may still be extensive effects such that behavioral responses were present throughout the distribution instead of being concentrated around the cutoff. In the following section, DID strategies were employed to study such extensive responses.

4.1 DID Estimation

In this section, DID strategies were applied to assess the effects of tax differentiation policy on the floor area distribution of new housing units. According to the floor area threshold, housing units above the cutoff could be defined as one group and units below could be defined as the other group. To avoid confounding with other policies[14], only housing units ranged from 100 to 180 sqm were examined. However, it is possible that the policy did not have an impact on housing units in such broad ranges, therefore the first challenge is to identify the range of potential effects to provide a basis for defining the control and treatment groups.

[14] Potential confounding policies include the affordable housing program, which set specific floor area requirements (60 - 70 and 70 - 80 for two- and three-household families, respectively) for housing units, as well as the 90/70 policy, which required 70% of new units in each complex to be under 90 sqm.

Housing units in (100, 180] sqm were assigned to eight groups with equal-width, which $Group_1$ consisted of units in (100, 110] sqm, $Group_2$ consisted of units in (110, 120] sqm and so on. The group of housing units in (100, 110] sqm was used as the reference group ($Group_1$). Figure 8 plots the quarterly normalized log supply of new housing units in each group compared to that of the reference group. The quarterly normalized log supply was defined as the quarterly log supply of housing units in the specific group minus its average quarterly log supply before the policy (i.e., Q2 2004 - Q1 2005). This figure illustrates that compared to the supply of housing units in the reference group, supply in (110, 120] and (120, 130] sqm were relatively similar over the examined period. On the other hand, supply in other groups decreased relative to that in the reference group after the policy was implemented. These findings align with the hypothesis that the supply of housing units above 140 sqm should decrease relative to the supply of units below the threshold after the policy. However, supply of housing units in (130, 140] sqm also decreased compared to the reference group. This could also due to the optimization frictions aforementioned in the bunching analysis. These figures provide a basis for defining housing units in (100, 130] sqm as one group and units in (130, 180] sqm as the other group. However, although the supply of housing units in (130, 140] sqm also decreased, these units should be grouped with those in (100, 130] sqm because of the definition of ordinary properties. Therefore, to highlight the effects of introducing the relative lower transaction tax rates for ordinary properties, it is plausible to define housing units in (140, 180] sqm as the control and units in (100, 140] sqm as the treatment groups[15].

To further validate the selection of control and treatment groups, pseudo-DID estimations were employed to numerically identify where the impacts occurred taking the following form:

$$n_{it} = \alpha + \sum_{j=2}^{j=8} \beta_j Group_j + \sigma Post + \sum_{j=2}^{j=8} \delta_j Group_j \times Post + \varepsilon_{it} \tag{4}$$

where n_{it} is the log count of housing units whose floor areas are in bin i at time t. $Group_j$ is an indicator dummy variable that equals 1 if bin i is in group j and 0 otherwise. *Post* is a dummy variable that equals 1 if n_{it} is after May 2005 and 0 otherwise. According to Bertrand et al. [2004] and Cameron and Miller [2015], the error term ε_{it} was clustered by floor area bins to account for serial correlation. Columns (1) to (4) in Table 3 represent four specifications with different fixed effects. The interested coefficients are $\delta_j s$ which indicate the average policy effects on the supply of new housing units in different groups compared to those

[15] In Appendix B I provide identical DID analysis using housing units in (130, 160] and (100, 130] sqm as the control and treatment groups, respectively, and similar conclusions were attained.

Figure 8: Quarterly Normalized Log Supply of New Housing Units

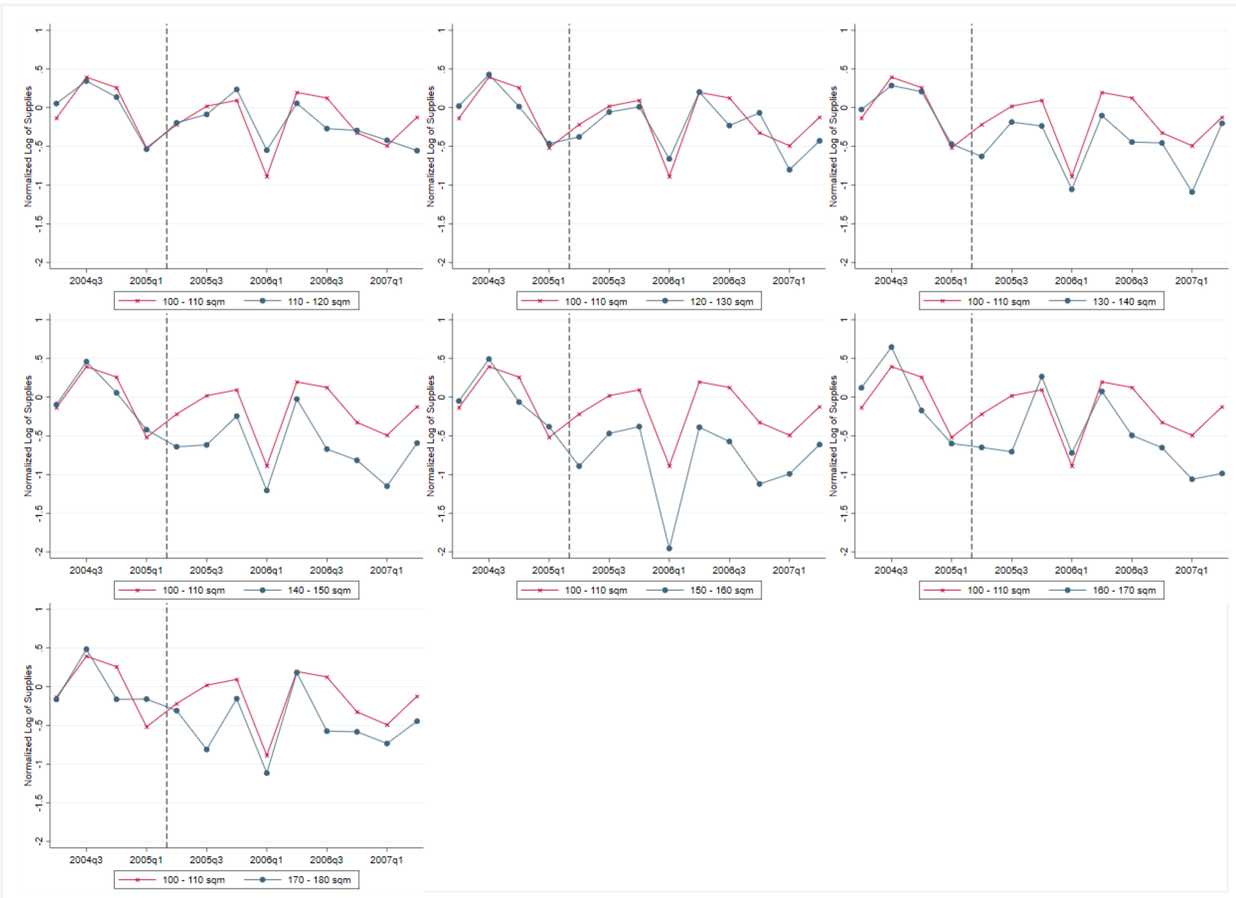


Figure 8 plots the quarterly normalized log supply of new housing units in each group compared to that of the reference group ((100, 110] sqm). The quarterly normalized log supply was defined as the quarterly log supply of new housing units in the specific group minus its average quarterly log supply before the policy (i.e., Q2 2004 - Q1 2005).

of the reference group.

Table 3 displays the results. Compared to the reference group, the average treatment effects on the supply of housing units in (110, 120] and (120, 130] sqm were not significant. Moreover, the treatment effects for groups above 140 sqm were negative and significant relative to those for the reference group. The supply of housing units in (130, 140] sqm also decreased about 21.0% compared to the supply in the reference group. These estimations are consistent with the findings shown in Figure 8, and hence provide supports for the above selection of the control and treatment groups.

Table 3: Policy Effects on the Supply of Housing Units for Different Groups

	(1)	(2)	(3)	(4)
(110, 120]×Post	0.0455	0.0455	0.0455	0.0455
	(0.116)	(0.0854)	(0.0853)	(0.0858)
(120, 130]×Post	0.0523	0.0523	0.0523	0.0523
	(0.124)	(0.0912)	(0.0911)	(0.0916)
(130, 140]×Post	-0.236**	-0.236***	-0.236***	-0.236***
	(0.120)	(0.0769)	(0.0768)	(0.0773)
(140, 150]×Post	-0.366***	-0.366**	-0.366**	-0.366**
	(0.137)	(0.153)	(0.152)	(0.153)
(150, 160]×Post	-0.687***	-0.687***	-0.687***	-0.687***
	(0.136)	(0.121)	(0.121)	(0.122)
(160, 170]×Post	-0.219	-0.219**	-0.219**	-0.219**
	(0.149)	(0.105)	(0.105)	(0.106)
(170, 180]×Post	-0.251*	-0.251**	-0.251**	-0.251**
	(0.142)	(0.119)	(0.119)	(0.120)
Bin FE			Yes	Yes
Month FE				Yes
Standard Error	Robust	Cluster	Cluster	Cluster
Observations	3120	3120	3120	3120
Adj. R-sq	0.491	0.491	0.500	0.651

Standard errors in parentheses

In all regressions, the dependent variable is lcount (log of count).

The reference group is (100, 110] sqm.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Overall, the findings above provide evidence that using DID methods could be more reasonable to assess the policy effects than using bunching estimations. The non-positive treatment effect for the (130, 140] sqm group and negative effects for groups above 140 sqm are consistent with the results of bunching estimations, i.e., that there was little bunching below and some missing mass above the threshold, respectively. However, as the results also show that the policy did not only affect the supply of housing units around the tax notch, using DID estimations could be more plausible to capture the policy effects on housing units distributed in broader floor area ranges.

By defining housing units in (140, 180] sqm as the control group and units in (100, 140] sqm as the treatment group, Figure 9 shows the quarterly normalized log supply of new housing units for these two groups. The graph displays certain significant patterns. First, supply was usually inactive in the first quarter due to the holiday season and rebounded quickly after that period. The Lunar New Year is the biggest festival of the year in China and typically begins from late January to early February. While the celebration takes only fifteen days, companies are busy conducting annual reviews and preparing plans for the next lunar year around this period; therefore, it is a drag on economic activity. Most land supply is released to developers by the government after this period, thus the supply of new housing units in the first quarter is often limited. There was no notable rebound in supply in the second quarter of 2005 due to the increase in regulations during that period. Since the second quarter of 2005, the government has initiated a series of policies to cool the overheated housing market, which had caused a continuous decline in the housing supply. Overall, the graph illustrates a parallel trend between the control and treatment groups before the policy that diverged after the policy was introduced, such that the quarterly normalized log supply of the treated units increased relative to that of the controlled units.

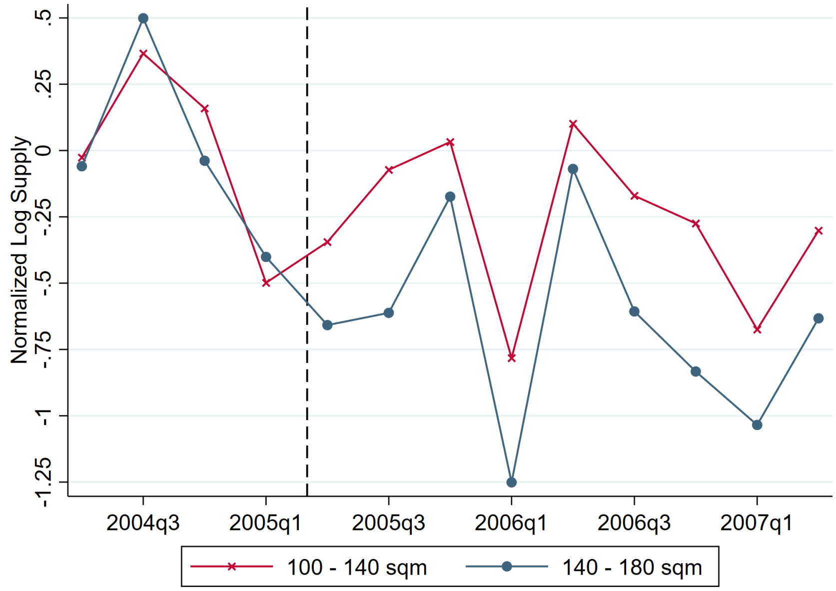
Similar to Equation (4), the baseline specification of the DID estimation takes the following form:

$$n_{it} = \alpha_0 + \alpha_1 Post + \alpha_2 Treat_i + \alpha_3 Post \times Treat_i + \varepsilon_{it} \tag{5}$$

where $Treat_i$ is a dummy variable that equals 1 if the floor area of bin i is in the treatment group and 0 otherwise.

Table 4 reports the DID results. Columns (1) - (3) represent estimations with different fixed effects for the full sample. The DID estimator of the model was .346 and significant at the 1% level after controlling for the bin and month fixed effects. This implies that, on average, there was almost a 41.3% increase in the supply of housing units in the treatment group after the policy compared to the control group. Moreover, given there are certain differences in attributes among annuli (e.g., the number of supply of housing units, FAR, households income level), understanding the policy effect in these different regions could be important

Figure 9: Quarterly Normalized Log Supply of New Housing Units



The quarterly normalized log supply was defined as the log supply in that quarter minus its average log supply before the policy (i.e., Q2 2004 - Q1 2005).

to policy makers. Columns (4) - (6) estimate the average treatment effects on the supply of housing units in three annuli. Estimations show that the magnitude of treatment effects increased monotonically from the inner to outer annuli. Compared to the supply of housing units in the control group, the policy had caused 10.2% and 15.6% increases in the supply of treated units in the inner and mid annuli, respectively. However, these treatment effects were not statistically significant. On the other hand, the DID estimator of the outer annulus was .307 and significant at the 1% level, implying a 35.9% increase in the supply of housing units in the treatment group after the policy, relative to that in the control group. One explanation for the increasing treatment effects from the inner to the outer annuli is the difference in price elasticity of demand in these areas. Since the wealthy often live near the city center, price elasticity of demand is lower in areas closer to the city center. Therefore, the same degree of financial incentives could induce larger responses moving away from the center. Zhou [2016] confirms that the housing market in downtown areas overreact less to tightening policies compared to the suburban market. They also argue that since demand in the downtown market may be less financially constrained, tightening policies could be less relevant.

Table 4: Policy Effects on New Housing Supply

	(1) All	(2) All	(3) All	(4) Inner	(5) Middle	(6) Outer
Treat×Post	0.346***	0.346***	0.346***	0.0975	0.145	0.307***
	(0.0707)	(0.0712)	(0.0712)	(0.0945)	(0.103)	(0.102)
Bin FE	Yes		Yes	Yes	Yes	Yes
Month FE		Yes	Yes	Yes	Yes	Yes
Observations	3120	3120	3120	3120	3120	3120
Adj. R-sq	0.498	0.570	0.649	0.254	0.504	0.580

Standard errors in parentheses

In all regressions, the dependent variable is lcount (log of count).

Clustered standard errors

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.2 Robustness Check

In this section, several robustness checks were conducted to validate the above results. Since there were two other criteria that could cause housing units to become non-ordinary and thus not exempt from higher transaction taxes despite the floor area requirement, these potentially confounding observations were dropped to examine the intended policy effect in a friction-less environment. Specifically, observations that did not qualify for the FAR or price requirements for ordinary properties were dropped. Recalling that the maximum FAR of a complex is determined exogenously by the government, if it is less than one, any housing units within that complex would be automatically considered to be luxury properties regardless of their floor area and prices. Therefore, the supply of housing units in those complexes could be immune to the floor area requirement, as developers would not have incentives to build smaller units there. To reduce such noises, in this section the analysis was restricted to complexes whose FARs were greater than or equal to one after the policy was implemented.

Neighborhoods whose overall housing prices did not qualify for the ordinary status requirement were also excluded. Because housing prices are endogenously determined by developers who may have taken the policy into account, to simply exclude housing units whose price exceeded the requirement may otherwise lead to selection biases. Rather, to exclude areas where the overall price exceeded the requirement could be more reasonable. Specifically, using the sales dataset with price information, I identified and retained neighborhoods where 75% of their housing sales before the policy (i.e., June 2004 - May 2005) were below the price cap[16]. If prices of the majority of housing units in a neighborhood exceed the requirement, there must be some external factors associated with such high prices, for instance, superior location. Therefore, the policy may provide little incentive for developers in these neighborhoods to provide housing units below 140 sqm because their prices would most likely exceed the price cap. The price requirement is a weaker constrain on housing units to claim the tax benefit. The ordinary or luxury status is not assigned permanently, each time a resale occurs the status will be reevaluated. On the other hand, once the FAR of the complex or the floor area of the housing unit does not qualify for the requirement, the unit would be considered as luxury property permanently, and thus higher deed tax and sales tax would apply. However, if the price of a housing unit exceeds the requirement in its first sale the higher sales tax may not necessarily apply. Because if its resale price does not exceed the price cap[17], the housing unit could still be considered as ordinary property given the FAR and the floor area requirements are met[18].

After excluding the above-mentioned observations, two sub-samples were formed. One is called the higher-expected-impact sample, which contained the rest of full sample after the sample exclusion. Another sample is the lower-expected-impact sample, which only included observations in those excluded neighborhoods. Observations in the higher-expected-impact sample were expected to be affected the most by the policy, whereas those in the lower-expected-impact sample were assumed to have smaller impacts. Table 5 shows the estimations for the two samples. The DID estimator of the higher-expected-impact sample was .502 and significant at the 1% level. It implies that, on average, there was almost a 65.2% increase in the supply of housing units in the treatment group after the policy compared to that in the control group. The estimator of the lower-expected-impact sample was .201 and also significant at the 10% level. The result implies that the policy caused an average 22.6% increase in the supply of treated units relative to that of the controlled units. The magnitudes and significance level of the policy effect for the higher-expected-impact sample were notably higher than that for the lower-expected-impact sample, which is consistent with the hypothesis.

Next, event studies of the treatment effect with a 95% confidence interval (CI) for the full and the higher-expected-impact samples were depicted. Figure 10 illustrates the results. In spite of the limited data availability before the policy, the figure shows that compared to the supply of housing units in the

[16] The government set three different price caps for housing units in the three annuli: ￥17,500, ￥10,000, and ￥7,500 per sqm in the inner, mid, and outer annuli, respectively

[17] Recall that the government periodically adjusts the price cap to reflect inflation and price appreciation.

[18] Even if the actual resale price exceeds the price cap, there are studies documenting price manipulations in the resale market to dodge extra tax liability (Agarwal et al. [2020], Dai and Xu [2022]).

Table 5: Policy Effects on the Supply of Higer- & Lower-Expected-Impact Samples

	Higher-Expected-Impact Sample			Lower-Expected-Impact Sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Treat×Post	0.502***	0.502***	0.502***	0.201*	0.201*	0.201*
	(0.0813)	(0.0818)	(0.0818)	(0.103)	(0.104)	(0.104)
Bin FE	Yes		Yes	Yes		Yes
Month FE		Yes	Yes		Yes	Yes
Observations	3120	3120	3120	3120	3120	3120
Adj. R-sq	0.520	0.560	0.650	0.205	0.288	0.319

Standard errors in parentheses

In all regressions, the dependent variable is lcount (log of count).

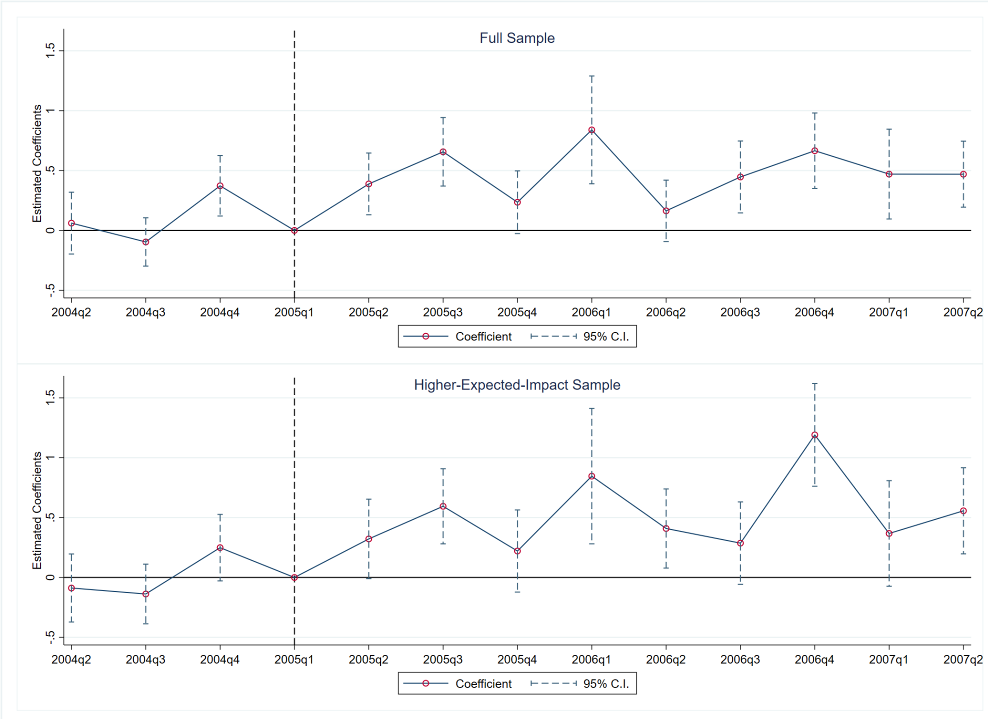
Clustered standard errors

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

control group, the relative increase in the supply of treatment group occurred shortly after the policy was implemented and essentially lasted throughout the examined period. The treatment effects were slightly insignificant in some of the quarters, which was possibly because of the high volatility in the Chinese housing market. Since the housing market was dominated by a relatively small number of large developers at the moment, their behavior may cause volatility in the overall market statistics. The treatment effects were more volatile for the higher-expected-impact sample than the full sample as fewer observations with higher impacts were included. Although the treatment effects were not strictly significant in each period, the overall timing and significance of the effects provide suggestive evidence for the validity of the DID estimations.

Finally, the impacts of possible confounding policies were explored. Since the housing market is heavily regulated in China, there is often policy overlap that makes analysis challenging. During the analyzed period, there were few policies that may potentially affect the estimation. In June 2006, the government announced a policy that increased down payment requirements for housing units above 90 sqm from 20% to 30% and required that at least 70% of the units in each complex be under or equal to 90 sqm. Since the new policy cutoff was away from the threshold of the examined policy, it may not significantly confound the above estimations. However, as developers were required to build a greater amount of smaller housing units, the supply of units above 90 sqm may not fall proportionally, which would require further investigation. Another potential confounding policy was the increase of holding period for sales tax relief from two years to five years. As shown in Table 1, resales of housing units within two years of purchases were given

Figure 10: Event Study with 95% CI for the Full & the Higher-Expected-Impact Samples



certain benefits to their sales taxes. This minimum holding period was increased to five years in June 2006. However, the impact of this policy is expected to be minimal because houses are illiquid assets that do not trade within a short period of time[19], especially in a market dominated with pre-sale houses.

Since these confounding policies were introduced in June 2006, to identify the effects of these policies, two time indicators for the periods from June 2005 to May 2006 and from June 2006 to June 2007 were created and the treatment effects for these two periods were separately estimated. Equation (5) was modified to reflect the two periods as the following:

$$n_{it} = \alpha_0 + \alpha_1 Period_1 + \alpha_2 Period_2 + \alpha_3 Treat_i + \alpha_4 Period_1 \times Treat_i + \alpha_5 Period_2 \times Treat_i + \varepsilon_{it} \tag{6}$$

where instead of using *Post* as an indicator to represent the after policy period, now $Period_1$ equals 1 if the supply of housing units was between June 2005 and May 2006, and $Period_2$ equals 1 if the housing supply was between June 2006 and June 2007 and 0 otherwise.

Table 6 displays estimations of Equation (6) for the full and higher-expected-impact samples. The results indicate that the treatment effects for the first period, when only the tax differentiation policy was in

[19]Statistics show that the average holding period of houses in China are 6 - 8 years.

effect, were still positive and significant. Although the magnitude of the effects was smaller than it was in Table 4 for the full sample and Table 5 for the higher-expected-impact sample, where the two periods were combined, it nevertheless validates the previous conclusion that the tax differentiation policy increased the supply of housing units in the treatment group relative to the control group. The treatment effects for the second period were also positive and significant for the two samples. The magnitude of the relative increase in the supply of housing unis below 140 sqm rose in response to the new policies. If the increase of minimum holding period for sales tax relief had an impact, it should make ordinary properties less attractive than it was before the new policy because now it became harder to claim the tax advantage of buying a ordinary property. Therefore, the treatment effect should be smaller in this period. One potential explanation of the increase in treatment effects in the second period is the fact that to meet the 70% requirement and avoid paying higher down payments, some developers built two housing units of less than 90 sqm next to each other. Home-buyers who demand larger housing units could then buy two small units adjacent to each other and break down the wall between them to form a single unit. Therefore, it is reasonable to expect the supply of smaller housing units to increase even more relative to that of larger units after the new policy. Another possible reason is that optimization frictions may have less impacts over a longer time horizon as developers could adjust more easily. Overall these two forces dominated effects of the increased holding period and showed larger treatment effects after the new policies.

Table 6: Policy Effects on New Housing Supply in Two Periods

	Full Sample			Higher-Expected-Impact Sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Treat×Period I	0.300***	0.300***	0.300***	0.452***	0.452***	0.452***
	(0.0748)	(0.0752)	(0.0752)	(0.0957)	(0.0963)	(0.0963)
Treat×Period II	0.397***	0.397***	0.397***	0.556***	0.556***	0.556***
	(0.0917)	(0.0923)	(0.0923)	(0.1000)	(0.101)	(0.101)
Bin FE	Yes		Yes	Yes		Yes
Month FE		Yes	Yes		Yes	Yes
Observations	3120	3120	3120	3120	3120	3120
Adj. R-sq	0.501	0.570	0.649	0.522	0.561	0.650

Standard errors in parentheses

In all regressions, the dependent variable is lcount (log of count).

Clustered standard errors

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5 Housing Prices

5.1 Sales Data

The preceding discussion demonstrates that, given the tax incentives, the policy had induced the market to provide more housing units in (100, 140] sqm than those in (140, 180] sqm. In this section the study explores the policy effects on the prices of new housing units and seeks to understand how the capitalization of taxes combined with changes in the composition of new housing units, will affect the overall prices. Since the supply dataset did not contain price information, a sales dataset which included individual transactions of new housing units in Shanghai from July 2004 to June 2007 was used for analysis.

Table 7 summarizes the variables of interest after data cleaning. There were 471,336 observations in the dataset with an average unit price (i.e., average price per sqm) of ¥7869.0. The price pattern was consistent with the standard urban model of housing prices decreasing as the distance from the city center increases. The inner annulus had the highest average price per sqm of ¥15331.5, followed by the middle annulus of ¥8815.2, and the outer annulus had the lowest average price of ¥5443.8. The average price per sqm remained at about the same level after the policy. The statistics for the rest of variables were similar to those of the supply data. The average floor area was still the largest in the inner annulus and it decreased after the policy. The average FAR also declined monotonically from the inner to outer annuli, whereas greening rates were similar for complexes in different annuli.

Similar to Figure 9, Figure 11 illustrates the quarterly normalized log count of sales. Before the policy, the number of sales in the two groups moved accordingly, whereas after the policy the trends diverged such that the number of sales in (100, 140] sqm consistently and notably surpassed those in (140, 180] sqm. The number of sales dropped dramatically for both groups during the second quarter of 2005, when the housing market in China was experiencing a series of tightening regulations to calm the overheated market. The magnitude of decrease in the number of sales was larger than the number of supply shown in Figure 9[20] because developers can hoard housing units during a heavily regulated period to avoid unfavorable sales. Also the increase in interest rates could more directly affect the sales rather than the supply, thus caused a more apparent decline in the number of sales. Research has also confirmed that the number of housing transactions usually drops in the beginning of tightening cycles. Zhou [2016] documents that the housing market usually overreacts to tightening policies when they are announced, when housing prices and

[20] Recall that the supply in Figure 9 was not the actual listing on the market, rather it represented the new stock of housing supply.

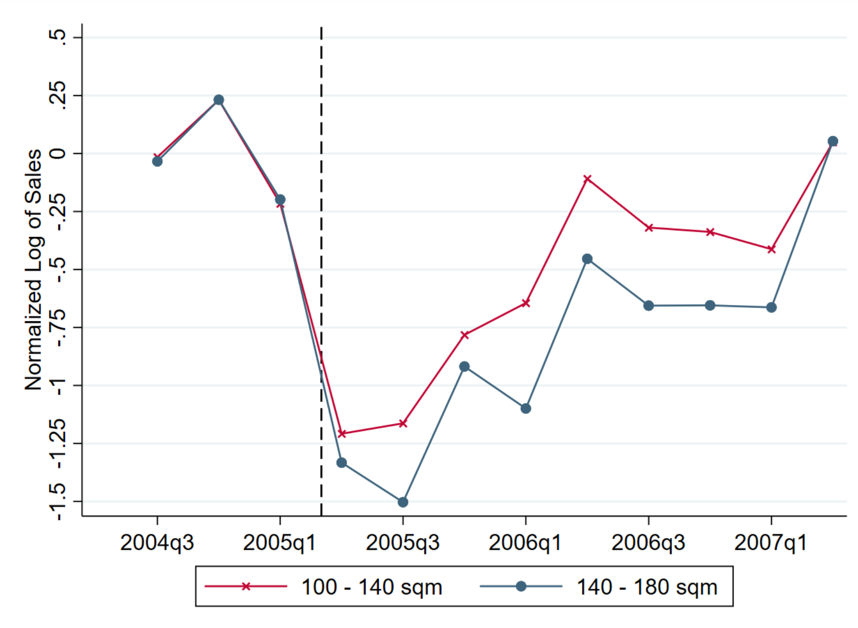
Table 7: Summary Statistics of the Sales Dataset

Variable	Obs	Mean	Std. dev.	Min	Max
Panel A: Full Sample					
Floor area	471336	122.0	59.33	6.0	5005.0
Price per sqm	471336	7869.0	5066.56	50.0	142240.0
Floor-to-area ratio	437936	1.8	0.87	0.1	40.0
Greening rate	436748	41.2	8.87	0.2	90.0
Panel B: Inner Annulus					
Floor area	60815	134.7	59.79	25.0	1102.0
Price per sqm	60815	15331.5	7220.58	750.0	142240.0
Floor-to-area ratio	57169	3.0	1.18	0.2	36.0
Greening rate	57050	40.8	10.46	0.2	73.0
Panel C: Middle Annulus					
Floor area	160696	115.6	55.29	31.0	5005.0
Price per sqm	160696	8815.2	3505.65	170.0	69600.0
Floor-to-area ratio	152386	1.9	0.63	0.2	40.0
Greening rate	153446	41.8	7.95	0.3	79.0
Panel D: Outer Annulus					
Floor area	249825	123.1	61.14	6.0	1422.5
Price per sq.m	249825	5443.8	2796.51	50.0	129330.0
Floor-to-area ratio	228381	1.4	0.60	0.1	8.9
Greening rate	226252	40.9	9.01	0.3	90.0
Panel E: Before the Policy					
Floor area	165251	127.6	58.84	6.0	1843.9
Price per sq.m	165251	7862.4	4511.72	400.0	129330.0
Floor-to-area ratio	154092	1.8	0.94	0.2	36.0
Greening rate	153405	41.6	9.25	0.3	90.0
Panel F: After the Policy					
Floor area	306085	119.0	59.37	7.0	5005.0
Price per sq.m	306085	7872.6	5342.22	50.0	142240.0
Floor-to-area ratio	283844	1.8	0.83	0.1	40.0
Greening rate	283343	41.0	8.65	0.2	90.0

The summary statistics was calculated from the full sample from July 2004 to June 2007. The calculation was based on residential housing units only.

transaction volumes usually drop and then quickly rebound.

Figure 11: Quarterly Normalized Log Count of Sales



The quarterly normalized log number of sales was defined as the log numbers of sales in that quarter minus the average log number of sales before the policy (i.e., Q3 2004 - Q1 2005).

As mentioned above, although the sales dataset also contained presale housing units, the proportion of such "futures" would be smaller than those in the supply dataset. Therefore, the number of sales may not yield an accurate and timely reflection of the impacts of the policy, and it is reasonable to expect smaller treatment effects for the number of sales than for the supply of housing units. Table 8 displays the results of identical DID estimations of Equation (5) but using the sales data. The results similarly suggest that the policy had increased the number of sales in the treatment group by about 21.2% relative to those in the control group. The smaller treatment effects for the number of sales than for the supply is consistent with the hypothesis that there were more lagging policy effects in the number of sales than in the supply of housing units.

5.2 DID Estimation

Standard economic theory holds that taxes should be capitalized into housing prices (Oates [1969]; Yinger et al. [1988]). On the other hand, developers can adjust quality or the supply of housing units according to policy changes in the newly-built market, which can also affect the prices. In this section, the overall policy

Table 8: Policy Effects on the Number of Sales

	(1)	(2)	(3)	(4)
Treat×Post	0.192***	0.192***	0.192***	0.192***
	(0.0449)	(0.0449)	(0.0451)	(0.0451)
Bin FE		Yes		Yes
Month FE			Yes	Yes
Observations	2880	2880	2880	2880
Adj. R-sq	0.543	0.642	0.780	0.887

Standard errors in parentheses

In all regressions, the dependent variable is lcount (log of count).

Clustered standard errors

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

effect on new house prices was examined and possible causes were investigated to explain the results.

Firstly, the overall policy effect on housing prices was assessed by estimating the following specification:

$$lunitPr_{it} = \alpha_0 + \alpha_1 Post + \alpha_2 Treat_i + \alpha_3 Post \times Treat_i + \varepsilon_{it} \tag{7}$$

where $lunitPr_{it}$ is the log price per sqm of housing units i at transaction date t. Estimations of Equation (7) may not identify the effect of taxes on prices, all else equal, because housing units that were sold after the policy were mostly built after its implementation. Therefore, there could be selection into both the control and treatment groups, which has been verified by the results in the following analysis. However, because the purpose of this study is to understand the policy effects on the prices of new housing units, it should take into account selections into the control and treatment groups. Since each housing unit only appeared once in the dataset, individual unit-level fixed effects cannot be used. Instead, higher level fixed effects could be employed to control for the time-invariant unobservables. One possibility is to use complex fixed effects because housing units within the same complex share similar characteristics, such as FAR, greening rates, location, and etc. However, it is still possible that a complex does not contain housing units both in the control and treatment groups before and after the policy. To ensure that the use of complex fixed effects is plausible, I compared DID estimations of the full sample with those of a restricted sample, where the restricted sample limited complexes to those contained the control and treatment groups both before and after the policy. Table 9 displays the comparison results. Column (1) represents the estimation of the treatment effect for the full model without complex fixed effects. Columns (2) and (3) represent estimations for the restricted model without and with complex fixed effects, respectively. Comparing results in columns (1) and (2) shows that the treatment effects for the full and restricted samples were substantially different in terms of both magnitude and significance level. In addition, observations and the adjusted R-squares of the restricted sample were much smaller than those of the full sample. Therefore, the results indicate that using the restricted sample with complex fixed effects would be inappropriate.

Table 9: Sample Comparison by Complex Fixed Effect

	Full Sample	Restricted Sample	
	(1)	(2)	(3)
Treat×Post	-0.129***	0.0166	-0.0330**
	(0.0361)	(0.0433)	(0.0143)
Project FE			Yes
Observations	253470	55042	55042
Adj. R-sq	0.0488	0.0156	0.898

Standard errors in parentheses

In all regressions, the dependent variable is lunitPr (log of price per sqm).

Clustered standard errors

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

To tackle this challenge, an even higher level fixed effect was investigated. The neighborhood is the smallest administrative division that is used to segment housing areas in China. In a neighborhood, there is shared topography, environment, housing, transportation, amenities, and other natural and social resources. Table 10 compares the full with the restricted sample, where the restricted sample only retained neighborhoods that contained the control and treatment groups both before and after the policy. The results in columns (1) and (2) show that DID estimators of the full and restricted sample were comparable in terms of both magnitude and significance level. Additionally, only a small portion of observations would be lost if the restricted sample were to be used, and the adjusted R-square increased notably when controlling for the neighborhood fixed effect, which demonstrates that this fixed effect could explain some of the variations in house prices. Therefore, the following analysis is built on the restricted sample with the neighborhood fixed effect.

Columns (1) - (4) in Table 11 display the estimation results of Equation (7). The DID estimator of column (4) was -.0677 and significant at the 1% level, which indicates that the policy had decreased the average price per sqm of housing units in the treatment group by about 6.55% compared to the price of units in the control group. Most of the literature on tax capitalization in housing prices is developed on the

Table 10: Sample Comparison by Neighborhood Fixed Effect

	Full Sample	Restricted Sample	
	(1)	(2)	(3)
Treat×Post	-0.129***	-0.118***	-0.0761***
	(0.0361)	(0.0354)	(0.0157)
Neighborhood FE			Yes
Observations	253470	246139	246139
Adj. R-sq	0.0488	0.0451	0.770

Standard errors in parentheses

In all regressions, the dependent variable is lunitPr (log of price per sqm).

Clustered standard errors

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

resale market or conditional on the housing supply and quality remaining unchanged (Dachis et al. [2012], Besley et al. [2014]). By contrast, this study explores the response of the newly-built market, in which these characteristics may change. Also, because the dataset lacked information pertaining to the quality of housing units, the estimation could not control for variations in quality. Therefore, the results may reflect composition effects such that developers adjusted quality or changed the supply of housing units, which may cause the relative prices of the treated units to fall. The smaller treatment effect shown in column (2) than that in column (1) when the neighborhood fixed effect was included suggests that developers increasingly built treated units in less expensive neighborhoods after the policy, which is consistent with the composition hypothesis.

Columns (5) - (8) in Table 11 provide estimations for the higher-expected-impact sample. The higher-expected-impact sample was similarly defined as in the above section except a stricter criterion was applied here. Since FAR and greening rates are approximate measures of housing quality and the sales data could reflect lagging effects, narrowing down the higher-expected-impact sample may help to identify the policy effects. Therefore, the higher-expected-impact sample retained neighborhoods where 90% of housing sales before the policy were below the price cap[21]. Estimations in columns (5) - (8) show larger treatment effects of the higher-expected-impact sample relative to those of the full sample in columns (1) - (4). For the higher-expected-impact sample, the estimation in column (8) suggests that the policy had decreased the average price per sqm of housing units in the treatment group by about 8.64% compared to the price of units in the control group and the result was significant at the 1% level.

[21]This chosen percentile was 75% in the above section.

Table 11: Policy Effects on Prices of New Housing Units

	Full Sample				Higher-Expected-Impact Sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treat×Post	-0.118***	-0.0761***	-0.0689***	-0.0677***	-0.185***	-0.0998***	-0.0938***	-0.0904***
	(0.0354)	(0.0157)	(0.0150)	(0.0139)	(0.0535)	(0.0248)	(0.0241)	(0.0215)
Neighborhood FE		Yes	Yes	Yes		Yes	Yes	Yes
Bin FE			Yes	Yes			Yes	Yes
Month FE				Yes				Yes
Observations	246139	246139	246139	246139	113403	113403	113403	113403
Adj. R-sq	0.0451	0.770	0.774	0.791	0.0444	0.735	0.740	0.758

Standard errors in parentheses

Clustered standard errors

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

To find evidence of the composition hypothesis, which asserts that the quality of housing units in the treatment group may be downgraded relative to that of units in the control group after the policy, the decomposition method by Gelbach [2016] was applied to explore the degree to which the observed quality variables (i.e., FAR, greening rates) could be attributed to the changes in treatment effects. Specifically, the base and full models were defined as below:

Base model:

$$lunitPr_{it} = \alpha_0 + \alpha_1 Post + \alpha_2 Treat_i + \alpha_3 Post \times Treat_i + \varepsilon_{it} \tag{8}$$

Full model:

$$lunitPr_{it} = \alpha_0 + \alpha_1 Post + \alpha_2 Treat_i + \alpha_3 Post \times Treat_i + \alpha_4 FAR_{it} + \alpha_5 green_{it} + \varepsilon_{it} \tag{9}$$

Using a pair of regressions such that the base specification only has X_1 variables whereas the full specification contains the additional quality variables X_2, the decomposition method allows one to examine how much of the changes in X_1 coefficient can be attributed to variables in X_2 as moving from the base to the full specifications. In this case, X_1 contained the treatment effect indicator $Post \times Treat$. X_2 included the FAR and greening rates, which were the only observed quality variables despite they are complex-level attributes.

Table 12 shows estimations of the decomposition method for the full and the higher-expected-impact samples. The DID estimators of the base and the full models for the full sample were -.0679 and -.0698, respectively, and both were significant at the 1% level. The difference in treatment effects between the two models was .0018 but not statistically significant. Although the signs and significance level of estimators in column (3) did not provide evidence for the composition hypothesis, it is possible that there were effects among observations that were more likely to be affected by the policy. Columns (4) - (6) in Table 12 report effects for the higher-expected-impact samples. The difference in treatment effects between the base and full models was -.0079, where decreases in the FAR and greening rates had contributed about 28% and 72% to the negative treatment effects, respectively. Even though these estimations were still not statistically significant due to the use of the FAR and greening rates as quality variables, they at least provide weak evidence for the composition hypothesis that a portion of the price decrease could be attributed to quality adjustment[22]. Kopczuk and Munroe [2015] similarly find that there are weaker bunching effects for newly-built properties that sell when already finished, thus suggesting the presence of real or supply-side quality adjustments. They expect that negotiating a purchase before construction is finished allows for significant response in terms of the quality of the finishes, appliances, and other amenities, which allows for price reductions driven by adjustments in property characteristics.

Table 12: Decomposition of the Treatment Effects

	Full Sample			Higher-Expected-Impact Sample		
	(1) Base	(2) Full	(3) Explained	(4) Base	(5) Full	(6) Explained
Treat×Post	-0.0679***	-0.0698***	0.0018	-0.0969***	-0.0890***	-0.0079
	(0.0137)	(0.0137)	(0.0019)	(0.0208)	(0.0207)	(0.0049)
FAR		Yes	0.0001		Yes	-0.0022
			(0.0003)			(0.0025)
green		Yes	0.0017		Yes	-0.0057
			(0.0019)			(0.0047)
Neighborhood FE	Yes	Yes	Yes	Yes	Yes	Yes
Bin FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

Clustered standard errors

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The composition hypothesis could expect that changes in the supply of housing units may also lead to relative price decreases in the treatment group compared to the control group. To investigate this hypothesis, separate DID estimations were conducted for housing units in different price groups. Specifically, housing units were assigned into different price groups according to the pre-policy price distribution, such that the

[22]Zheng et al. [2012] also show that higher greening rates are associated with higher housing prices.

first group included housing units whose prices were below the 25th percentile of sales price before the policy, the second group included units priced between the 25th and 50th percentiles of pre-policy sales prices, and so on. Then, the number of home sales within each price group were aggregated by quarters separately for the control and treatment groups. For each price group, Figure 12 plots the quarterly normalized log counts of home sales for the treatment and control groups. The figure shows that the number of home sales in the treatment group increased more notably in lower price groups compared to those in the control group. To see this more explicitly, Figure 13 similarly plots the normalized log counts through aggregating housing units by those above and below the pre-policy medium price. This figure more clearly illustrates that the number of sales of the treated housing units at lower price percentiles had increased compared to the controlled units after the policy, and hence caused a relative decrease in the average price of units in the treatment group.

Figure 12: Quarterly Normalized Log Count of Sales in Different Price Quantiles

The quarterly normalized log counts of sales was defined as the log numbers of sales in that quarter minus the average log number of sales before the policy (i.e., Q3 2004 - Q1 2005).

Table 13 numerically presents the findings in Figure 12 and Figure 13. It shows that the treatment effect is the largest for the group of housing units whose prices were below the 25th percentile of the price

Figure 13: Quarterly Normalized Log Count of Sales Under and Above the Medium Price

The quarterly normalized log counts of sales was defined as the log numbers of sales in that quarter minus the average log number of sales before the policy (i.e., Q3 2004 - Q1 2005).

distribution. It estimates that, on average, the policy had caused 56.7% increase in the number of sales of housing units in the treatment group compared to those in the control group. For housing units in the second and fourth quantiles of the price distribution, the number of sales of the treated units increased moderately relative to the controlled units, by 18.8% and 22.1%, respectively. The estimation in column (5) indicates that for housing units below the medium of pre-policy price, the policy had caused a 46.5% increase in the number of sales of the treated units compared to the controlled units, whereas the result in column (6) shows that the treatment effect is unapparent and not statistically significant for units above the medium of pre-policy price. Therefore, these evidence suggest there are composition effects such that the supply of the treated housing units increased at lower percentiles of the price distribution compared to the controlled units, which led to relative price decreases of the treatment group.

5.3 Robustness Check

Finally, two robustness checks were conducted to verify the results in this section. The first test was to perform an event study of the tax differentiation policy to examine its impacts throughout the studied period. Figure 14 plots the event study for both the full and higher-expected-impact samples. The figure shows that the difference in prices of housing units in the control and treatment groups became significant shortly

Table 13: Policy Effects on the Number of Home Sales in Different Price Groups

	(1) Under 25th	(2) 25th - 50th	(3) 50th - 75th	(4) Above 75th	(5) Below 50th	(6) Above 50th
Treat×Post	0.449***	0.172**	-0.0680	0.200***	0.382***	0.0562
	(0.0720)	(0.0848)	(0.0801)	(0.0678)	(0.0666)	(0.0498)
Bin FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2880	2880	2880	2880	2880	2880
Adj. R-sq	0.824	0.768	0.724	0.672	0.854	0.790

Standard errors in parentheses

In all regressions, the dependent variable is lcount (log of count).

Clustered standard errors

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

after the policy was implemented. That difference in prices persisted almost throughout the post period. Overall, the timing and significance of the treatment effect provide solid evidence for the credibility of DID estimations.

Similar to the above section, the last but not least robustness test was to identify potential confounding effects of overlapping policies. An adjustment of Equation (7) to the following form would be able to assess the overlapping effects:

$$lunitPr_{it} = \alpha_0 + \alpha_1 Period_1 + \alpha_2 Period_2 + \alpha_3 Treat_i + \alpha_4 Period_1 \times Treat_i + \alpha_5 Period_2 \times Treat_i + \varepsilon_{it} \quad (10)$$

where $Period_1$ that equals 1 if the sale was between June 2005 and May 2006, and $Period_2$ that equals 1 if the the sale was between June 2006 and June 2007 and 0 otherwise.

Table 14 demonstrates results for the full and the higher-expected-impact samples. The positive and significant coefficients of $Treat \times Period_1$ maintained the conclusions that the tax differentiation policy caused the average price of the treated housing units to decrease relative to that of the controlled units. In addition, the treatment effects for the second period decreased notably as the confounding policy was introduced. Since the new policy further encouraged the supply of smaller housing units as shown in section 4, it could induce more profound composition effects such that the supply of lower-priced units in the treatment group would increase more. Therefore, the average price of treated housing units may increase further compared to price of units in the control group.

Figure 14: Event Study with 95% C.I.

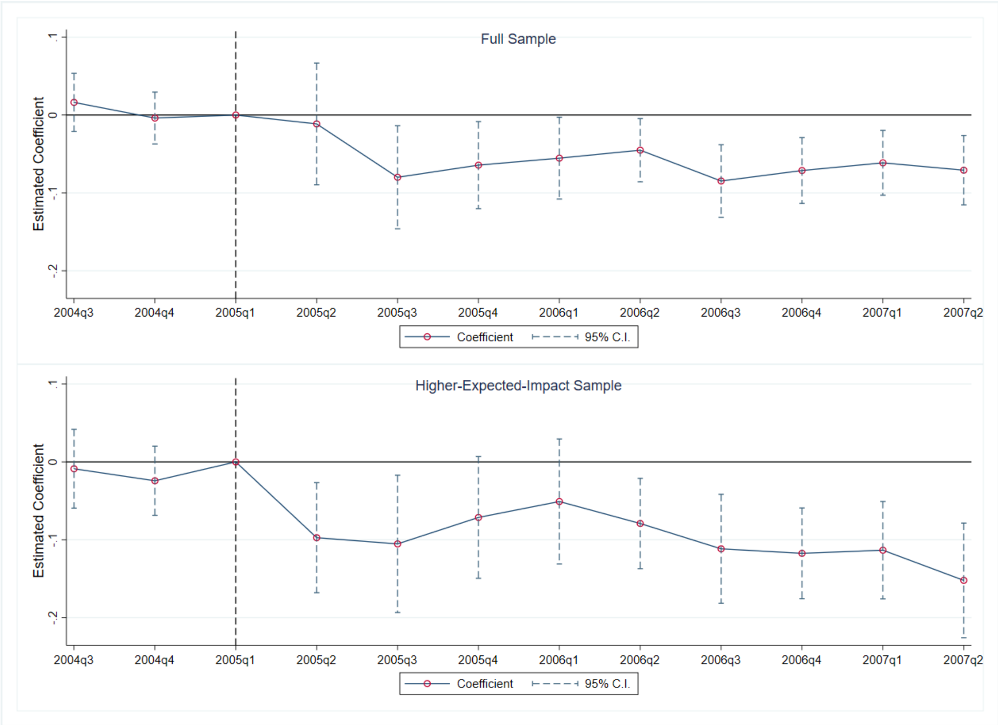


Table 14: Policy Effects on New Home Prices in Two Periods

	Full Sample			Higher-Expected-Impact Sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Treat×Period I	-0.0606***	-0.0533***	-0.0573***	-0.0676**	-0.0584**	-0.0591**
	(0.0191)	(0.0182)	(0.0176)	(0.0286)	(0.0276)	(0.0251)
Treat×Period II	-0.0836***	-0.0772***	-0.0745***	-0.117***	-0.114***	-0.109***
	(0.0169)	(0.0163)	(0.0153)	(0.0271)	(0.0262)	(0.0248)
Neighborhood FE	Yes	Yes	Yes	Yes	Yes	Yes
Bin FE		Yes	Yes		Yes	Yes
Month FE			Yes			Yes
Observations	246139	246139	246139	113403	113403	113403
Adj. R-sq	0.774	0.779	0.791	0.741	0.746	0.759

Standard errors in parentheses

Clustered standard errors

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6 Conclusion

The paper seek to understand the effects of a critical instrument in promoting middle-income housing in China namely the tax differentiation policy, which introduced relative lower transaction tax rates for housing units below 140 sqm in Shanghai in May 2005. Specifically, utilizing two comprehensive datasets from the Shanghai Housing Bureau that contained the universe of housing transactions spanned the period of policy introduction, it answered the question that how the tax differentiation policy affected the floor area distribution of new housing units and their prices in Shanghai from 2004 to 2007. Due to optimization frictions, estimations found little bunching below the floor area threshold and weak missing mass above the cutoff in response to the notched taxes. Using housing units in (100, 140] sqm as the treatment group and those in (140, 180] sqm as the control group, DID estimations showed that, on average, the combination of a 1.5% relative decrease in the deed tax and a 5.5% relative decrease in sales tax on housing units in the treatment group increased their supply by 41.3%, compared to the supply of units in the control group. The results also found that the policy had decreased the average price per sqm of treated units by 6.55% relative to the controlled units. Applying the decomposition method found weak evidence of quality adjustments to the price decrease given the lack of appropriate quality variables. Further DID estimations by housing units in different price distribution showed that the supply of treated units in lower percentiles of the pre-policy price distribution increased after the policy compared to that of the controlled units, which led to a relative decrease in the average prices of the treatment group. These evidence suggested that the price decrease in the treatment group relative to the control group resulted from shifts in the composition of housing units developed after the policy.

This paper contributes to the research on market behavior in response to notched taxes by exploring the impacts of notched taxes on broader distribution rather than estimating its local effects around the notches. It also sheds light on the limited literature on the impacts of taxation on the newly-built housing market. It explores how tax policies may affect the newly-built market differently given flexible house supply and quality adjustment.

Given that the housing market is heavily regulated in China, to assess the policy effects at its initiation could help avoid confounding with other polices, and thus provide more credible results. The study was able to gain access to two administrative datasets which allowed it to examine the effects of tax differentiation policy around the introduction, which is the biggest advantage of the paper. However, the data available during this period contained little information on property characteristics. As a result, an important direction for future work is to understand how the tax policy affected the quality of new housing units and the results on house prices. Another possible route to complete the research is to explore the resale market with information of repeated sales to study the capitalization of taxes conditioning on property quality being unchanged and a fixed stock of house supply.

References

S. Agarwal, K. Li, J. Wu, and J. Yan. Tax evasion, capital gains taxes, and the housing market. *Journal of Public Economics*, 188:104222, 2020.

A. Ahuja, L. Cheung, G. Han, N. Porter, and W. Zhang. Are house prices rising too fast in china? (10/274), 2010.

J. Barr and J. P. Cohen. The floor area ratio gradient: New york city, 1890–2009. *Regional Science and Urban Economics*, 48:110–119, 2014.

J. R. Barth, M. Lea, and T. Li. China's housing market: Is a bubble about to burst, 2012. URL `https://ssrn.com/abstract=2191087`.

M. Bertrand, E. Duflo, and S. Mullainathan. How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, 119:249–275, 2004.

T. Besley, N. Meads, and P. Surico. The incidence of transaction taxes: Evidence from a stamp duty holiday. *Journal of Public Economics*, 119:61–70, 2014.

M. C. Best and H. J. Kleven. Housing market responses to transaction taxes: Evidence from notches and stimulus in the u.k. *The Review of Economic Studies*, 85:157–193, 2018.

A. C. Cameron and D. L. Miller. A practitioner's guide to cluster–robust inference. *Journal of Human Resources*, 50:317–372, 2015.

K. Chen and Y. Wen. The great housing boom of china. *American Economic Journal: Macroeconomics*, 9: 73–114, 2017.

R. Chetty, J. N. Friedman, T. Olsen, and L. Pistaferri. Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from danish tax records. *The Quarterly Journal of Economics*, 126: 749–804, 2011.

M. Chivakul, W. R. Lam, X. Liu, W. S. Maliszewski, and A. Schipke. Understanding residential real estate in china. (15/84), 2015.

B. Dachis, G. Duranton, and M. A. Turner. The effects of land transfer taxes on real estate markets: evidence from a natural experiment in toronto. *Journal of Economic Geography*, 12:327–354, 2012.

Y. Dai and Y. Xu. Cheating under regulation: Evidence from “yin–and–yang” contracts on beijing’s housing market. *Sustainability*, 14:13346, 2022.

Y. Deng, J. Gyourko, and J. Wu. Land and house price measurement in china. (18403), 2012. URL `http://www.nber.org/papers/w18403`.

Z. Fan, X. Wang, and H. Zhang. Understanding misreporting: New evidence from responses to housing transaction tax, 2022.

H. Fang, Q. Gu, W. Xiong, and L.-A. Zhou. Demystifying the chinese housing boom. (21112), 2015. URL `http://www.nber.org/papers/w21112`.

J. B. Gelbach. When do covariates matter? and which ones, and how much? *Journal of Labor Economics*, 34:509–543, 2016.

E. Glaeser, W. Huang, Y. Ma, and A. Shleifer. A real estate boom with chinese characteristics. *Journal of Economic Perspectives*, 31:93–116, 2017.

H. J. Kleven. Bunching. *Annual Review of Economics*, 8:435–464, 2016.

H. J. Kleven and M. Waseem. Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from pakistan. *The Quarterly Journal of Economics*, 128:669–723, 2013.

H. J. Kleven, M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez. Unwilling or unable to cheat? evidence from a tax audit experiment in denmark. *Econometrica*, 79:651–692, 2011.

W. Kopczuk and D. Munroe. Mansion tax: The effect of transfer taxes on the residential real estate market. *American Economic Journal: Economic Policy*, 7:214–257, 2015.

L. Liu, B. Lockwood, M. Almunia, and E. H. F. Tam. Vat notches, voluntary registration, and bunching: Theory and u.k. evidence. *The Review of Economics and Statistics*, 103:151–164, 2021.

D. P. McMillen. *Testing for Monocentricity*, chapter 8. Wiley Blackwell, 2006.

W. E. Oates. The effects of property taxes and local public spending on property values: An empirical study of tax capitalization and the tiebout hypothesis. *Journal of Political Economy*, 77:957–971, 1969.

A. Rees-Jones. Quantifying loss–averse tax manipulation. *Review of Economic Studies*, 85:1251–1278, 2018.

J. Slemrod. Tax compliance and enforcement. *Journal of Economic Literature*, 57:904—-954, 2019.

J. Slemrod, C. Weber, and H. Shan. The behavioral response to housing transfer taxes: Evidence from a notched change in d.c. policy. *Journal of Urban Economics*, 100:137–153, 2017.

H. F. Tam. Behavioural response to time notche in transaction tax: Evidence from stamp duty in hong kong and singapore, 2018.

J. Wu, J. Gyourko, and Y. Deng. Evaluating conditions in major chinese housing markets. *Regional Science and Urban Economics*, 42:531–543, 2012.

J. Yinger, A. Borsch-Supan, H. S. Bloom, and H. F. Ladd. Academic Press, 1988.

C. Zhang, G. Bao, and G. Feng. Responses of housing market to changes in transaction taxes: Evidence from a natural experiment in shanghai, china, 2017.

S. Zheng, J. Wu, M. E. Kahn, and Y. Deng. The nascent market for "green" real estate in beijing. *European Economic Review*, 56:974–984, 2012.

Z. Zhou. Overreaction to policy changes in the housing market: Evidence from shanghai. *Regional Science and Urban Economics*, 58:26–41, 2016.

Table A1: The Localized Requirements for Ordinary Properties in Shanghai

Date	FAR	Floor Area	Prices
June 2005	≥ 1.0	≤ 140 sqm	Inner Annulus: $\leq$ ¥$17,500$ per sqm; Middle Annulus: $\leq$ ¥$10,000$ per sqm; Outer Annulus: $\leq$ ¥$7,000$ per sqm
November 2008	≥ 1.0	≤ 140 Sqm	Inner Annulus: ≤ 2.45 million CNY per housing unit; Middle Annulus: ≤ 1.40 million CNY per housing unit; Outer Annulus: ≤ 0.98 million CNY per housing unit
March 2012	≥ 1.0	≤ 140 Sqm	Inner Annulus: ≤ 3.30 million CNY per housing unit; Millde Annulus: ≤ 2.00 million CNY per housing unit; Outer Annulus: ≤ 1.60 million CNY per housing unit
November 2014	≥ 1.0	≤ 140 Sqm	Inner Annulus: ≤ 4.50 million CNY per housing unit; Middle Annulus: ≤ 3.10 million CNY per housing unit; Outer Annulus: ≤ 2.30 million CNY per housing unit

A Appendix A

B Appendix B

Table B1 compares the DID results between two different selections of the control and treatment groups. Columns (1) - (4) present the results for the based selection throughout the paper; columns (5) - (8) shows the results for an alternative selection, which housing units in (130, 160] sqm were defined as the control group, and those in (100, 130] sqm were defined as the treatment group. Coefficients of the treatment effect were larger for the alternative sample relative to those for the based sample. The estimation in column (5) shows that, on average, there was almost a 58.7% increase in the supply of housing units in the treatment group after the policy compared to the control group. Similar to the based sample the treatment effects also increased monotonically from the inner to outer annuli for the alternative sample.

Table B2 and Figure B1 related the results of two robustness tests for the based and alternative samples. Both results in Table B17 and Figure 15 provide evidence for a timely and consistent policy effect when the alternative sample was chosen.

Although the magnitude of policy effects between these two selections of the control and treatment groups differed in some degrees, the significance and consistency of the effect remained. The findings for the two samples similarly concluded that the policy had induced an increase in the supply of smaller housing units relative to larger units.

Table B1: Policy Effects on New Housing Supply

	(100, 140] vs (140, 180]				(100, 130] vs (130, 160]			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	Inner	Middle	Outer	All	Inner	Middle	Outer
Treat×Post	0.346^{***}	0.0975	0.145	0.307^{***}	0.462^{***}	0.0849	0.199	0.640^{***}
	(0.0712)	(0.0945)	(0.103)	(0.102)	(0.0786)	(0.104)	(0.121)	(0.0944)
Bin FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3120	3120	3120	3120	2340	2340	2340	2340
Adj. R-sq	0.649	0.254	0.504	0.580	0.598	0.214	0.445	0.543

Standard errors in parentheses

In all regressions, the dependent variable is lcount (log of count).

Clustered standard errors

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B2: Policy Effects on New House Supply: Two Periods

	(100, 140] vs (140, 180]			(100, 130] vs (130, 160]		
	(1)	(2)	(3)	(4)	(5)	(6)
Treat×Period I	0.300^{***}	0.300^{***}	0.300^{***}	0.477^{***}	0.477^{***}	0.477^{***}
	(0.0748)	(0.0752)	(0.0752)	(0.0754)	(0.0760)	(0.0760)
Treat×Period II	0.397^{***}	0.397^{***}	0.397^{***}	0.447^{***}	0.447^{***}	0.447^{***}
	(0.0917)	(0.0923)	(0.0923)	(0.103)	(0.104)	(0.104)
Bin FE	Yes		Yes	Yes		Yes
Month FE		Yes	Yes		Yes	Yes
Observations	3120	3120	3120	2340	2340	2340
Adj. R-sq	0.501	0.570	0.649	0.371	0.440	0.598

Standard errors in parentheses

In all regressions, the dependent variable is lcount (log of count).

Clustered standard errors

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure B1: Event Study with 95% C.I.