Housing Choice Under School Admission Uncertainty: Evidence From China's School Lottery

Jiusheng Zhu*   Lifang Zhang†

Abstract

To relieve the strain on public school seats and mitigate competition for elite schools, Beijing introduced a school lottery that somewhat severed the link between housing and school assignments. Using housing transaction data from Beijing, we leverage the school lottery as a quasi-experiment to examine how school admission uncertainty affects parental decisions on housing choice. Our findings show that, while the school lottery reduced the school-quality premiums of Tier 1 school district houses (SDHs), it improved the premiums of Tier 2 SDHs. Besides, the lottery had heterogeneous effects in different school zonings. In zonings with more elite schools, premiums of SDHs either increased or remained unaffected. These indicate that households, in response to admission uncertainty introduced by the lottery, tended to choose SDHs with higher admission probabilities. This new pattern of housing choice mitigated the policy effect on narrowing price gaps between SDHs and non-SDHs, which was an unintended consequence from the policymakers' perspectives.

Keywords: Housing choice, School lottery, School choice, School admission uncertainty, Residence-based enrollment

*Monash University, Australia.
Email: zhujiusheng@mail.bnu.edu.cn.
ORCID : 0009-0003-0458-3506

†Corresponding author.
Beijing Normal University, China.
Email: zhanglifangceee@163.com.
ORCID: 0000-0001-5881-8735

[1] We are grateful to two anonymous referees, Juan Yang, Rigissa Megalokonomou, Xiaoyang Ye, and the editor for helpful discussions and suggestions.

1 Introduction

Public school choice is an integral part of housing choice, regardless of whether the school enrollment system offers multiple school choices or not. A large literature documents the relation between school choice and housing choice. These papers advance our understanding of how household's housing choice interacts with school choice. A strand of them develops theoretical models incorporating endogenous residential decision to examine how school choice affects housing choice (Avery and Pathak, 2021; Bayer et al., 2007; Ferreyra, 2007). Another strand provides empirical evidence on housing choice in response to school enrollment policies, including open enrollment (Brunner et al., 2012; Ely and Teske, 2015), school choice lottery (Bibler and Billings, 2020), and school desegregation (Baum-Snow and Lutz, 2011). However, because most of these studies primarily concentrate on school choice with multiple options, how households choose housings under residence-based enrollment coupled with uncertainty in admission is underexplored.

This paper provides novel evidence of how school admission uncertainty affects households' housing choices in the context of a residence-based enrollment system. Using housing transaction data from Beijing, we estimate the parental valuation on the probability of their children's enrollment in public elementary schools following a school lottery policy. Specifically, we leverage the school lottery policy as a quasi-experiment and estimate the policy effect on the school-quality premiums of school district houses (SDHs, henceforth). The policy was designed to alleviate the strain on public school seats and break the ties between housing and schools assignment, which introduces uncertainty to the admission process. Besides, we also build a theoretical model of housing choice incorporating school choice, which helps to illustrate households' choices under uncertainty.

Our findings indicate that the school lottery significantly reduced the school-quality premiums for Tier 1 SDHs but simultaneously increased those for Tier 2 SDHs. This occurred because the school lottery reduced admission probabilities for Tier 1 SDHs much more than it did for Tier 2 SDHs. Besides, the lottery had heterogeneous effects in different school zonings. In zonings with more elite schools, premiums for SDHs either increased or remained unaffected. These suggest that, in response to admission uncertainty introduced by the lottery, households tended to prefer SDHs with higher admission probabilities. This new pattern of housing choice restricted the policy effect on narrowing price gaps between

SDHs and non-SDHs, which was an unintended consequence from the policymakers' perspectives.

Previous studies largely focus on the school choices with multiple options in developed countries, which are not fully applicable to developing countries. This is because residence-based enrollment, an one-to-one assignment pattern, is the prevalent K-12 enrollment system in developing countries (Snyder et al., 2019). This pattern in developing countries leads to the equivalence of school choice to housing choice for households seeking quality education. While the effect of uncertainty on school choice in developing countries is inconclusive, papers in the context of multiple school choices would poorly shed light on our question.

Moreover, prior papers on China's residence-based enrollment focused on the capitalization effect of school quality on housing prices (Chan et al., 2020; Feng and Lu, 2013; Han et al., 2021b; Zheng et al., 2016). Few of them explored parental choice on school in response to enrollment reform. Han et al. (2021a) estimated the policy effect of restricting access to senior high schools on housing prices, but they did not take into account household's school or housing choice. Our study fills this gap by providing evidence on household's school and housing choice from China's residence-based enrollment.

The most relevant enrollment system to our paper is the school lottery, which introduces uncertainty to the admission process. A school lottery system is a random selection process used to determine which students will attend, and it is implemented when there are more students applying for admission than can be accommodated. There are several papers on how school lottery impacts households' school and housing choice. Andreyeva and Patrick (2017) and Bonilla-Mejía et al. (2020) took into account the school admission uncertainty in admission probability. However, as we mentioned above, their settings were based on school choice with multiple options instead of the traditional residence-based admission, limiting their application in developing countries.[1]

A growing literature investigated the policy effect of school lottery in China. Jin et al. (2023) and Chen and Li (2023) both examined parental school choice between public elite schools and private elite schools in response to uncertainty in private schools' lottery enrollment. However, public school enrollment was not subject to the lottery in their studies

[1]Andreyeva and Patrick (2017) and Bonilla-Mejía et al. (2020) leveraged school choice lottery's admission priority in the settings of school choice lottery and charter school, respectively.

and most households' school choices within public schools were not affected by uncertainty. Yang et al. (2024) and Tong et al. (2024) examined the school lottery policy in Beijing, which is the same policy studied in this paper. While they mainly estimated the effect on the premium of school district house and whether the policy contributed to education equality, they did not discuss anything related to school and housing choice. Compared to their work, we focus on the effect of uncertainty introduced by the lottery on parental choice of school and housing and evaluate the policy's effectiveness through this perspective. Furthermore, no study has previously estimated the effect of schools' admission probabilities on housing prices in China. We aim to estimate this effect by using the per capita school slot within a school district as a proxy for schools' admission probabilities.

This paper provides new evidence on the effect of uncertainty on household's decision-making on school and housing choice, particularly when options available are limited. We fill the gap of few empirical evidence on how school admission uncertainty influences housing choice under residence-based enrollment. Our findings also have policy implications for developing countries facing the dilemma of insufficient public school seats coupled with intensive competition among families for elite schools through housing choice.

The rest of the paper is structured as follows. Section 2 describes the background of school choice in China and the school lottery policy. Section 3 presents data sources and some important definitions. Section 4 is the theoretical analysis. Section 5 is the empirical analysis. Section 6 concludes.

2 Background

2.1 School choice in China

China's location-based assignment system shapes Chinese households' school choices. Since the implementation of "nearby enrollment policy", a one-to-one link between houses in each community and public schools was founded. Public schools designate an enrollment area known as *school district*, which is similar to the school catchment area in the US. For families living in a school district, housing ownership is a necessity for children to attend the schools in that school district. In this way, the requirement of housing ownership limits households' school choices, but creates a market with parental demand for good schools through housing choice. Therefore, school district house (henceforth, SDH), located in the school district of an elite elementary school, accordingly becomes a favored educational investment for families.

However, an irrational rise in SDH premiums has made quality public education increasingly unaffordable (Zheng et al., 2016). In many major cities, the school choice competition drove the prices of SDHs to new records, with the value of SDHs being more of an investment for speculation. "School district housing fever" has become a social phenomenon, especially in metropolitan areas like Beijing.[2] School-quality premiums of SDHs in Beijing kept rising for a long time (see Figure 1). According to estimates by Han et al. (2021b), the school-quality premiums for elite elementary schools in Beijing rose from 8.4% to 12.3% during 2013-2016, equivalent to the price of SDHs rising from 3,693 yuan per square meter to 5,649 yuan per square meter. This induced the policymakers to curb the rising trend of SDH premium.

Moreover, the strain of educational resources in megacities posed another challenge. Beijing, one of the most densely populated cities in China, witnesses a burgeoning population of more than 20 million, which put a strain on public resources, particularly on the access to elementary schools. According to Figure 2, the total number of elementary schools in Beijing kept decreasing, while the number of elementary school enrollments (the sum of Beijing and non-Beijing enrollments in the figure) was increasing. This imposed great pressure on the seats schools could provide and turned elite schools increasingly oversubscribed. In addition, the introduction of the "two-child policy" in 2014 would significantly increase new students and further overwhelm the school capacity in the near future. It was imperative for policymakers to figure out a solution to the dilemma.

2.2 The school lottery in Beijing

Given the two problems above, the Beijing Municipal Education Commission introduced the "multi-school zoning policy". It integrates several adjacent school districts into a school zoning and assigns students to schools within the zoning. The assignment features a random lottery. It was adopted in April 2018 for Dongcheng and Haidian, and in April 2020 for Xicheng (the three city districts are called "DHX" henceforth). The policy stated that housing purchased after the implementation time would be subject to the school lottery and be randomly assigned within the school zoning as long as the originally assigned school was

[2] "Why Anxious Parents Are Driving a Home Price Surge in China" - Bloomberg (2021)

oversubscribed. By severing the link between housing and the originally assigned schools, the policy created the uncertainty of school assignment, which was meant to reduce the SDH premiums and alleviate the strains on school seats in elite schools.

The enrollment regulation of the school lottery can be illustrated by Figure 3. The largest grey area in the figure represents the school zoning (henceforth, zoning) that are made up of four school districts denoted by the small grey areas. The school lottery works within every single zoning respectively. There are four schools in the zoning, with School A and B elite schools, School C and D non-elite schools. Besides, students are originally assigned to the school tied to their communities within the same school district. We assume that School A and School C are oversubscribed,[3] and new students from Community 1 and Community 2 would be assigned to another school in the zoning. According to the school lottery, new students from Community 1 would be assigned to School B with priority over School D, while those from Community 2 would be assigned to School D without other choice. This pattern implies that the school lottery was not designed to equalize access to elite schools. Instead, SDHs maintain an advantage due to their priorities of being assigned to another elite school, while non-SDHs remain tied with non-elite schools.

It is worth noting that the zoning areas were announced one day after the school lottery by the local government in the three sample city districts. And the zoning areas largely coincide with the jurisdiction area of the subdistrict office (subdistrict area, for short).[4] And the subdistrict area was initially designated by the local government based on the number of schools, population. And the local government did not distribute elite schools evenly across the subdistrict areas. Therefore, the number of elite schools in each zoning was determined solely by their location (see an example of zonings in Haidian in Figure A6).

Due to the compositional differences within the zonings, we categorize zonings into two groups: dense zonings and other zonings (see Figure 4 and Figure 5). Dense zonings are characterized by elite elementary schools constituting more than 60% of all elementary schools within the zoning.[5] To eliminate the risk that our results are derived by design, we replace 60% with 50% and 70% in the robustness tests to validate our results.

[3] Oversubscribed schools are those facing the dilemma of the number of applicants exceed the enrollment plan.

[4] The subdistrict office, also called “street committee”, is China’s primary-level organization in the urban area.

[5] The choice of the 60% threshold is derived by consulting professional property agents.

3 Data and definitions

3.1 Data source

This study utilizes Beijing's residential housing transactions from Lianjia, Beijing's largest real estate agent. The data is primarily sourced from transactions in Beijing's four major city districts: Dongcheng, Xicheng, Haidian, and Chaoyang, covering the period from January 2016 to January 2021. Every housing transaction of the dataset includes details such as transaction time, transaction price, property characteristics (e.g., area, layout, floor, elevator, year of construction, latitude and longitude coordinates), and neighborhood characteristics (e.g., greening rate, property fee, distance to Tiananmen) for each property transaction. The dataset comprises a total of 131,504 transactions spanning 3,612 neighborhoods within the four city districts.

To show the representativeness of our data, we compare two key indicators—transaction volume and house price growth rate—derived from our dataset with the corresponding official statistics. In Figure 6, we compare the annual transaction volume in our sample data with the officially published second-hand house transaction volume in the Beijing Real Estate Yearbook for the period 2017-2020. Notably, the sample volume of transaction data provided by Lianjia constitutes more than half of the official transaction volume. In Figure 7, we compare the monthly year-on-year house price growth rates in Beijing, calculated from both the sample data and official statistics. The curve of our sample generally aligns well with the official curves. Taken together, the sample data used in this paper is shown to be representative of Beijing's second-hand property market during the sample period.

3.2 Designation of the SDH in this paper

Before the commencement of each school year, elementary schools release their enrollment prospectus to the community, outlining the school's catchment area, referred to as the school district in this paper. We collected the annual enrollment prospectus of each elementary school in the four districts of the city. Subsequently, we matched all transaction properties in our dataset with their originally assigned elementary school based on the transaction date.

In order to explore the heterogeneous treatment effects of school lottery on the housing prices of SDH, we divide the SDH into three categories:

- **Tier 1 SDH (S1):** Housing is located within the school district of Tier 1 elite elementary schools. Tier 1 elite schools refer to the top 5 schools within their own city district.
- **Tier 2 SDH (S2):** Housing is located within the school district of Tier 2 elite elementary schools. Tier 2 elite schools refer to the top 6-20 schools within their own city district.
- **Other SDH (S3):** Housing is located within the school district of quasi-elite elementary schools. Quasi-elite elementary schools include branch schools that are associated to the elite schools.
- **Non-SDH (S4):** Housing is located within the school district of non-elite elementary schools.

The ranking of elementary schools used in this paper is based on a commonly accepted ranking of reputation among Beijing's local residents, also widely used in the literature (Han et al., 2021b; Huang et al., 2020; Zheng et al., 2016). To validate the school rankings, we collected the school-level average test scores from a citywide mathematics exam taken by all grade-6 students in Beijing schools in 2014, focusing on schools in the sample city districts. The first row of Table 1 shows that elite schools' average test scores in that exam were 12 points higher than those of the non-elite schools. Among elite schools, Tier 1 schools scored the highest (20 points higher than non-elite schools), followed by Tier 2 schools, which also scored 14 points higher than non-elite schools.

3.3 Admission probability

In order to capture the effect of admission uncertainty on parental housing choice, we need to measure the admission uncertainty using admission probability of each school. It is worth noting that there was no admission uncertainty before the policy, hence the probability is always 1 for all schools before the school lottery.

We calculated each school's admission probability after the school lottery, using the equation below:

$$\%\text{Admission} = \max\{1, \frac{N(\text{students per class}) \times N(\text{grade 1 classes in the school})}{N(\text{fam. w/ child and hukou in the school district})}\}$$

The number of grade 1 classes in each school is used as a proxy for the supply of school seat, while the number of families with local hukou and at least a child within the school district is used as a proxy for the demand of school seats.[6] The relevant statistics are reported in Table A2.

On the supply side of the school seats, we collected every school's class quantity from Dongcheng and Haidian in 2018 and that from Xicheng in 2020, consistent with each city district's policy timing. We did not collect every class's student quantities because the quantities could be very close because of a policy in Beijing stating that the number of students in every class must be no more than 40. And due to the fast growing student quantity in Beijing, the student quantity in each class could approach the upper limit of 40, which is supported by the statistics of the class's average student quantity from Beijing's four city districts shown in Table A2.

On the demand side, we need to obtain the number of families with hukou in each of the school districts and the proportion of families with hukou and at least a child in the corresponding school district.

To calculate the number of families in school districts, we collected data on the number of communities in each school district, the total number of communities in those city districts, and the total number of families with local hukou in the three city districts. The statistics above are from January 2017 to December 2017 for Haidian and Dongcheng, and from January 2019 to December 2019 for Xicheng. Then we calculated the average number of families with hukou per community in each city district by dividing the city district's total number of families by the city district's total number of communities. Next, we obtain the total number of families with hukou within each school district by multiplying the average number of families with hukou per community with the number of communities within the school district.

Additionally, to obtain the proportion of families with hukou and at least a child in the school district, we resort to the 2015 China Population Census dataset, which surveys 1% of the nationwide population at the county level. We collected data on the proportion of families, at the city-district level, that had local huukou and at least one child (aged 0Ĩ2) in each city district in 2015.[7] Then we multiplied the proportions with community-level

[6] We restricted the families to be the ones with local hukou because those without local hukou did not have access to local public schools.

[7] Before gathering the proportion, we excluded families without local hukou from the sample.

family quantity to form the denominator of the proxy for admission probability.

Taken together, the number of families with at least a child in the School District s from City District d is calculated by the following method:

$$N_{s,d}(\text{fam. w/ child \& hukou}) = P_d(\text{fam. w/ a child \& hukou}) \times \frac{N_d(\text{fam. w/ hukou})}{N_d(\text{communities})} \times N_{s,d}(\text{communities})$$

where $P_d(\text{family with a child})$ refers to the proportion of family with at least a child in the City District d.

Combining the supply side and demand side, we can calculate the admission probability in the School District s from City District d:

$$\%\text{Admission}_{s,d} = \max\{1, \frac{40 \times N_{s,d}(\text{grade 1 classes in the school})}{N_{s,d}(\text{fam. w/ child \& hukou})}\}$$

3.4 Summary statistics

Using the calculation process described above, we can derive the admission probabilities for different types of SDHs. As reported in Table 1, the admission probability for non-SDH remained at 1 after the policy. Meanwhile, Tier 1 SDH experienced the largest drop in admission probability (48.07%), whereas Tier 2 experienced a moderate decline (29.71%).

Except for the different type of schools' average test scores and admission probabilities, Table 1 also presents descriptive statistics of other variables for SDHs in the three sampe city districts.

The housing prices show that non-SDH are notably cheaper than SDH. The difference in housing prices among various types of SDH is even greater. Particularly, the housing prices between Tier 1 SDH and non-SDH differ by about 40,000 per square meter, reflecting a high degree of education premium in housing prices. Comparing variables related to living conditions, such as area, furnishing conditions, elevator, age, and greenery rate, non-SDHs have better living conditions. In contrast, living conditions for Tier 1 and Tier 2 SDH are very similar, with many being old and small. In terms of the sample size, SDHs account for about 45% of the whole sample, while Tier 1 SDHs constitute only about 6%. Together with the small number of grade 1 classes for Tier 1 SDH (91), these highlights the scarcity of high-quality education resources in Beijing.

4 Conceptual framework: A model of school choice

We model households' school choice in the context of purchasing SDHs for their children's enrollment in an elite school. It features households' utilities for housing characteristics, school qualities, and housing costs. We employ the model to illustrate households' decisions of school choice under the school lottery, in which households also take into account the admission probability.[8]

4.1 Model setup

Definition 1 (School quality). There are four types of schools for households to choose: Tier 1 elite schools, Tier 2 elite schools, quasi-elite schools, and non-elite schools, represented by s_1, s_2, s_3, s_4, respectively. And we set the school quality of s_1 relative to that of s_4 as y_1, thus y_2 for s_2, and y_3 for s_3. The school-quality ranking is $y_1 > y_2 > y_3$.

Definition 2 (Housing prices). The housing price, p, for SDHs are ranked based on the school quality. And we set the school-quality premium of Tier 1 SDH as p_1, equivalent to the relative price compared to that of non-SDH. Thus, the premium for Tier 2 SDH is p_2, and the premium for other SDH is p_3. The price ranking is $p_1 > p_2 > p_3$.

Following the utility function specification in Bayer et al. (2007), we model households' school choice in the form of residential decision. Let X_h represent the observable characteristics of housing choice h, including housing and community characteristics (e.g., size, age, floor, and plant area). Let p_h denote the price of housing choice h. Again, y_j^h represents the quality of school j that is originally tied to the housing h. Each household chooses its residence h to maximize its indirect utility function $V_i(h, j)$.

Household i's utility from housing choice h and attending school j is

$$V_i(h,j) = \alpha_i^X X_h + \underbrace{\alpha_i^y y_j^h}_{\substack{\text{utility from} \\ \text{school j}}} - \underbrace{\alpha_i^p p_j^h}_{\substack{\text{housing} \\ \text{cost of } h}} + \xi_h + \epsilon_i$$

The estimates of those variables are represented by α. The error term ξ_h captures the unobserved amenities of housing choice h, and ϵ_i is idiosyncratic preferences shocks over locations and schools.

[8] A reminder is that the model was designed to illustrate the school choice but not to simulate it, as we leave the examination work for the empirical analysis. Therefore, the parameters of the model are not estimated.

In addition, our model assumes sticky prices, meaning that prices do not react to shocks in the short term.

4.2 School lottery

Next, we consider the scenario under the school lottery that breaks the ties of housings with their originally assigned schools. And the new element added to the model is uncertainty, measured by the probability of admission to the originally assigned school.

Assumption (Admission probability). Due to the shortage of quality educational resources, s_1 is the fewest and also the most sought-after school, followed by s_2, and so on. The admission probabilities for originally assigned schools are $\gamma_1 = Pr(s_1|$ living in s_1 zone), $\gamma_2 = Pr(s_2|$ living in s_2 zone), $\gamma_3 = Pr(s_3|$ living in s_3 zone), and $\gamma_4 = Pr(s_4|$ living in s_4 zone). And we assume that the ranking of the admission probabilities is $\gamma_1 < \gamma_2 < \gamma_3 < \gamma_4$. Assuming when a s_1 school is oversubscribed, the probabilities of admission to another school are $\theta_1 = Pr(s_1^{'}|$ living in s_1 zone), $\theta_2 = Pr(s_2|$ living in s_1 zone), and $\theta_3 = Pr(s_3|$ living in s_1 zone).[9] Similarly, when a s_2 school is oversubscribed, the probabilities of admission to another school are $\theta_4 = Pr(s_2^{'}|$ living in s_2 zone), and $\theta_5 = Pr(s_3|$ living in s_2 zone).[10][11] The ranking of these probabilities is not necessary to be specified for the following analysis.

Under the school lottery, household i's expected utility from housing choice a, a Tier 1 SDH, would be

$$E[V_i(a, \underbrace{\hat{s}_1}_{\text{with uncertainty}})] = \alpha_i^X X_a + \gamma_1 \alpha_i^{y_1} y_1^a + \underbrace{\theta_1 \alpha_i^{y_1} y_1^a + \theta_2 \alpha_i^{y_2} y_2^a + \theta_3 \alpha_i^{y_3} y_3^a}_{\substack{\text{assigned to}\\ \text{other school}}} - \alpha_i^p p_1^a + \xi_a + \epsilon_i$$

Household i's expected utility from housing choice b, a Tier 2 SDH, would be

$$E[V_i(b, \underbrace{\hat{s}_2}_{\text{with uncertainty}})] = \alpha_i^X X_b + \gamma_2 \alpha_i^{y_2} y_2^b + \underbrace{\theta_4 \alpha_i^{y_2} y_2^b + \theta_5 \alpha_i^{y_3} y_3^b}_{\substack{\text{assigned to}\\ \text{other school}}} - \alpha_i^p p_2^b + \xi_b + \epsilon_i$$

Conclusion 1. Comparing housing choices a and b, achieving maximum utility depends on the trade-off between the additional costing of s_1 compared to s_2 and the ex-

[9] The $s_1^{'}$ refers to another Tier 1 school.

[10] The $s_2^{'}$ refers to another Tier 2 school.

[11] We exclude the probability $Pr(s_1|$ living in s_2 zone) because this scenario does not exist and hence should be zero.

pected utility from the school-quality of s_1. If s_1 compared to s_2 is overvalued for the expected utility from the school-quality of s_1, i and other households will prefer housing choice b until p_1 decreases to normal. The residential sorting in Tier 1 and Tier 2 SDHs awaits further empirical evidence.

4.3 Dense zoning

Furthermore, we consider the scenario of school choice within the dense zoning. Dense zoning features more elite schools and provides more seats for elite schools, which implies that the admission probabilities in dense zonings are higher than those in other zonings. Therefore, we add the condition of $\theta_j^D > \theta_j^O$, where θ_j^D denotes the probability of admission to a transferred school j in dense zonings and θ_j^O denotes that in other zonings.

Household i's expected utility from housing choice c , a Tier 1 SDH in the dense zoning, would be

$$\underbrace{E[V_i^D(c,\hat{s_1})]}_{\text{in the dense zoning}} = \alpha_i^X X_c + \gamma_1^D \alpha_i^{y_1} y_1^c + \theta_1^D \alpha_i^{y_1} y_1^c + \theta_2^D \alpha_i^{y_2} y_2^c + \theta_3^D \alpha_i^{y_3} y_3^c - \alpha_i^p p_1^c + \xi_c + \epsilon_i > \underbrace{E[V_i^O(a,\hat{s_1})]}_{\text{in other zonings}}$$

Household i's expected utility from housing choice d , a Tier 2 SDH in the dense zoning, would be

$$\underbrace{E[V_i^D(d,\hat{s_2})]}_{\text{in the dense zoning}} = \alpha_i^X X_d + \gamma_2^D \alpha_i^{y_2} y_2^d + \theta_4^D \alpha_i^{y_2} y_2^d + \theta_5^D \alpha_i^{y_3} y_3^d - \alpha_i^p p_2^d + \xi_d + \epsilon_i > \underbrace{E[V_i^O(b,\hat{s_2})]}_{\text{in other zonings}}$$

Conclusion 2. Based on the comparisons of expected utilities above, we can make a plausible prediction that the advantages of higher admission probability in dense zonings would intensify residential sorting behaviors for SDHs within those dense zonings. We will provide empirical evidence to validate this prediction in the following section.

5 Empirical analysis

5.1 Empirical strategy

Before examining the impact of school lottery on the school-quality premium, we need to measure the pre-policy school-quality premium for each type of SDH. This helps us understand how much premium of school quality exists for different SDHs and also validate the quality gap among different schools.

The school-quality premium of SDH is equivalent to the capitalization of school quality in housing prices, measured by the housing price increment of SDH relative to non-SDH. Specifically, we adopt the boundary fixed effect to compare a SDH with a non-SDH with the same amenities but within different school districts. This method eliminates the biases from unobserved amenities of housings (Bayer et al., 2007; Black, 1999). We control for the boundary fixed effect by matching every SDH with its nearest non-SDH within the distance of 500 meters. Meanwhile, we add covariates of housing characteristics and community characteristics to increase the precision of our estimates. These covariates are all outlined in the descriptive statistics (see Table 1).

The school-quality premiums are measured using the hedonic pricing model as follows:

$$\ln P_{itn} = \alpha_0 + \beta_1 SDH_{itn} + \sum_{n=1}^{N} \text{match}_n + \delta_t + \alpha_1 \text{X} + \epsilon_{itn} \tag{1}$$

where $\ln P_{itn}$ is the transaction price in logarithm of housing i, on transaction date t, located in community n. SDH_{itn} is a dummy variable of whether housing i is a SDH, whose coefficient β_1 measures the school-quality premium. $\sum_{n=1}^{N} \text{match}_n$ is the boundary fixed effects which match SDHs with their nearest non-SDH within the distance of 500 meters. X is a vector of housing characteristics and community characteristics. And δ_t is the year-month fixed effect. Lastly, ϵ_{itn} is an error term clustered by every community.

Next, we use the school lottery as a quasi-experiment and leverage difference-in-differences to identify the policy effect on housing prices, with SDH as the treatment group and non-SDH as the control group (see Table 2). In the model, we also control for the boundary fixed effect and the same covariates as those in Equation 1. We estimate the average treatment effects of school lottery using the following linear regression:

$$\ln P_{itn} = \alpha_0 + \beta_0 SDH_{itn} \times Post_t + \beta_1 SDH_{itn} + \sum_{n=1}^{N} \text{match}_n + \delta_t + \alpha_1 \text{X} + \epsilon_{itn} \tag{2}$$

where $SDH_{itn} \times Post_t$ is a dummy variable for whether the housing i is a SDH and the transaction date t is after the policy. β_0 captures the policy's treatment effect on school-quality premium of SDH. $\sum_{n=1}^{N} \text{match}_n$ is the boundary fixed effects which match SDHs with their nearest non-SDH within the distance of 500 meters. δ_t is the year-month fixed effect which absorbs the dummy for transaction date being after policy. X is a vector of housing characteristics and community characteristics. ϵ_{itn} is an error term, clustered by every community.

However, DID estimation omits the effects of school admission probabilities which serve as the mechanism. Therefore, we resort to 2SLS to examine the effects of schools' admission probabilities on housing prices. The first stage of 2SLS is as follows:

$$\%Admission_{itn} = \alpha_0 + \beta_0 SDH_{itn} \times Post_t + \beta_1 SDH_{itn} + \sum_{n=1}^{N} \text{match}_n + \delta_t + \alpha_1 \text{X} + \epsilon_{itn} \tag{3}$$

where $\%Admission_{itn}$ is the housing i's probability of getting admission from the originally assigned school at time t. $SDH_{itn} \times Post_t$ serves as the instrument variable, and the other variables are as described above. The second-stage is below:

$$\ln P_{itn} = \alpha_0 + \beta_0 \cdot \%\widehat{Admission}_{itn} + \beta_1 SDH_{itn} + \sum_{n=1}^{N} \text{match}_n + \delta_t + \alpha_1 \text{X} + \epsilon_{itn} \tag{4}$$

where $\%\widehat{Admission}$ represents the predicted values from estimating Equation (3).

It is worth noting that the school lottery is implemented at different timing in DHX (see Table 2). Difference-in-differences with variation in treatment timing may pose a threat to the robustness of the result due to heterogeneous treatment effects (de Chaisemartin and D'Haultfœuille, 2020; Goodman-Bacon, 2021; Sun and Abraham, 2021). In the subsequent robustness test, we validate our DID estimation results by using two alternative robust DID models.

To estimate the dynamic effects and test the parallel trend assumption, we use the event-study as follows:

$$\ln P_{itn} = \alpha_0 + \sum_{k \geq -m}^{M} \beta_0 SDH_{itn} \times I_t(t = \tau_i + k) + \beta_1 SDH_{itn} + \sum_{n=1}^{N} match_n + \delta_t + \alpha_1 \text{X} + \epsilon_{itn} \tag{5}$$

where $I_t(t = \tau_i + k)$ refers to the k_{th} month after school lottery is implemented.

5.2 Empirical results

5.2.1 School-quality premiums of SDHs

Table 3 reports the results of school-quality premium in housing prices before the school lottery was announced. Columns (1) and (2) do not include any covariates and fixed effects.

Columns (3) and (4) include covariates and year-month fixed effects, while columns (5) and (6) additionally include boundary fixed effects by matching SDHs and non-SDHs within 500 meters. By adding covariates and fixed effects, the estimates of school quality premiums gradually decrease but remain statistically significant.

After including boundary fixed effects, the overall school-quality premium is shown to be 7.7% in column (5), while the premium of Tier 1, Tier 2, and quasi-elite schools are 16%, 7.85% and 5.3%, respectively in column (6). The premiums imply that SDHs are generally 7.7% more expensive that non-SDHs, with the the price gap reaching 16% for Tier 1 SDHs compared to non-SDHs. The notable variation in the school-quality premiums of different SDHs show the disparity of school-quality among different schools and highlight educational resource inequality in Beijing's major city districts. Besides, the school-quality premiums we estimated above are comparable to those found in the literature within the context of Beijing (Chan et al., 2020; Chen and Li, 2023; Han et al., 2021b; Huang et al., 2020; Zheng et al., 2016), giving credit to our results.

5.2.2 Policy impacts on school-quality premiums

In Table 4, column (1) displays the impact on SDH premiums, including samples of all SDHs and non-SDHs. Columns (2) and (3) depict the impact on the premiums of Tier 1 and Tier 2 SDHs, respectively, with samples comprising non-SDHs and Tier 1 or Tier 2 SDHs. All regressions from the three columns include covariates, month-year fixed effects, and boundary fixed effects.

Using DID estimation from Equation (2), Panel A reports the baseline results of impact of school lottery on SDHs premiums. The school lottery slightly increased SDH premiums by 0.29 percentage points, but statistically insignificantly. The most obvious policy effect is on Tier 1 SDHs, with their premiums significantly falling by 3.91 percentage points (p-value < 0.01), equivalent to a 24% reduction of premiums. In contrast, the policy somehow increased the premiums of Tier 2 SDHs by 1.33 percentage points (p-value < 0.05).

However, the DID estimation results do not account for the mechanism of schools' admission probabilities, which should play a key role in explaining the policy's effect. Next, we use 2SLS to estimate the effects of schools' admission probabilities on SDHs premiums, based on regressions of Equations (3) and (4). The second-stage and the first-stage results are reported in Panel B and Panel C, respectively.

In Panel C, the first-stage results show that the school lottery reduced admission probabilities in all SDHs by an average of 34.21 percentage points compared to non-SDHs. [12] Specifically, Tier 1 SDHs witnessed a dramatic decline in admission probability of 48.03 percentage points, while a smaller fall in admission probability of 29.19 for Tier 2 SDHs. These resulted in admission probabilities of 51.93% for Tier 1 SDHs and 70.29% for Tier 2 SDHs. The disparity in admission probability may explain the great difference in policy effect between Tier 1 and Tier 2 SDHs, as shown in Panel A.

The second-stage results in Panel B show that a 10-percentage-point increase in admission probability would increase the premiums of Tier 1 SDHs by 0.814 percentage points (p-value < 0.01), but would somehow decrease the premiums of Tier 2 SDHs by 0.456 percentage points (p-value < 0.05).

To explain why premiums of Tier 2 SDHs increase as admission probabilities decline, we must consider the relative admission probability between Tier 1 and Tier 2 SDHs. Although the admission probability for Tier 2 SDHs decreased, the more pronounced decline in Tier 1 SDHs makes Tier 2 SDHs more appealing to parents due to less relative uncertainty.

In Panel D and E, we use relative admission probability as the new IV and restrict the sample with only Tier 1 and Tier 2 SDHs. The first-stage results show that the admission probability of Tier 2 SDHs increased 18.84 percentage point compared to Tier 1 SDHs. The relative increase correspondingly increases the premiums by 27.81 percentage points for both Tier 1 and Tier 2 SDHs. These results suggest that the relative admission probability for schools is an essential part of parental housing choice decisions.

Taken together, the school lottery suppressed the premiums of Tier 1 SDHs while raising those of Tier 2 SDHs, narrowing the gap between the two kinds of SDHs. This pattern may be attributed to households' strategy of treating Tier 2 SDHs as a safe substitute for Tier 1 SDHs due to the gap in admission probabilities. We will explore this pattern further in the next subsection.

[12] According to Table 1, the admission probabilities of non-SDHs remained at 1 after the policy. The decline of 34.21, relative to non-SDHs, implies the average admission probability of all SDHs to be (1-0.3421)=0.6579. The same idea applies to the following analysis.

5.2.3 Heterogeneity in different zonings

As we mentioned in subsection 3.2, zonings are comprised by several school districts and elite schools distribute unevenly in zonings, giving rise to the dense zonings where elite schools account for at least 60% of all elementary schools. Under the school lottery, students may be assigned to other schools within the zonings when their originally assigned school is oversubscribed. This setting might benefit the dense zonings because SDHs in dense zonings are less likely to be assigned to a non-elite school compared to zonings with fewer elite schools. To test this hypothesis, we examine the heterogeneity of policy effects in the dense zonings and other zonings.

In Table 5, Panel A and Panel B present the results of DID and 2SLS estimations of policy effects on Tier 1 and Tier 2 SDHs in dense zongings and other zonings, respectively. The results from DID and 2SLS estimations are similar to the pattern shown in Table 4 that the policy decreased the premiums of Tier 1 SDHs but increased the premiums of Tier 2 SDHs. However, a comparison of the two panels reveals that SDHs' premiums in dense zonings increased relative to those in other zonings after the policy. Specifically, Tier 2 SDHs in dense zonings saw a significant increase in the premiums, while the premiums of Tier 1 SDHs almost remained largely unchanged. In contrast, Tier 1 SDHs in other zonings experienced a significant decrease in premiums, whereas the premiums of Tier 2 SDHs remained largely unchanged. This disparity aligns with the previous hypothesis.

Combined with the results from Table 4, these patterns suggest that households prefer to buy SDHs with lower admission uncertainty and greater access to alternative elite schools. This strategy improved the premiums of Tier 2 SDHs while reducing those of Tier 1 SDHs, as demand shifted from Tier 1 SDHs to Tier 2 SDHs.

5.2.4 Policy's dynamic impacts

We then examine the school lottery's dynamic impacts on school-quality premiums of SDHs using the event-study method based on the unit of month. Figure 8 depicts the dynamic effects on the premiums of all SDHs before and after the month the school lottery was announced.

We found that the pre-treatment trend is centered around zero, which means that the SDH premiums were stable before the policy was announced. This finding validates the parallel trend assumption required for DID.

Moreover, the premiums decreased consistently until the 18th month following the policy's implementation, after which they gradually began to increase. The rising trend may result from households adopting a new strategy of buying SDHs with lower admission uncertainty.

5.3 Further discussion: Spillover of the school lottery

During our sample period, three city districts (DHX) of the four major city districts in Beijing implemented the school lottery except for Chaoyang. We next use housing samples in Chaoyang to examine if Chaoyang was affected by the spillover from the three treated city districts. Table A2 shows the comparison of several key indexes ranging from economic development to education resource, indicating that Chaoyang is very similar to the other three city districts in these aspects. This implies that Chaoyang's SDHs might be an ideal alternative for households to avoid the risk of admission uncertainty.

In Table 6, we use the same regressions of Equation (2) but uses Chaoyang's SDHs as the treatment group and its non-SDH as the control group. Panel A shows the results when setting the policy timing to be the same as Dongcheng and Haidian, set at April 2018. Panel B shows the results of policy timing same as Xicheng, set at April 2020. The results in both panels reveal that the SDH premiums increased after the simulated timing, contrary to the pattern observed in DXH. This indicates a spillover effect from the three treated city districts, where demand shifted from DHX to Chaoyang, as the significant increase in SDH premiums would not have occurred without this spillover.

Furthermore, Figure A1 depicts the price trend of SDHs and non-SDHs in Chaoyang. There are obvious surges in SDHs' prices after the two timings of policy announcement (right after the black dash line), compared to non-SDHs. In addition, the housing transaction of SDH in Chaoyang also sheds light on the spillover. Figure A2 depicts the monthly transaction volume of SDH in Chaoyang. After the school lottery was announced in Xicheng, there was a sudden increase in the transaction volume of SDHs. Taken together, the evidence above justifies the existence of spillover on Chaoyang.

5.4 Robustness check

5.4.1 Using a more continuous school ranking

In the previous results, we use a discrete and predetermined school ranking to estimate the policy effect on SDHs' premiums, which could result in arbitrary results. Next, we attempt to use a more continuous school ranking as an alternative for the robustness check.

Based on each school's average test score from the citywide mathematics exam, we divided all elite schools into 10 deciles and separately ran the regression of Equation (2). The results are illustrated in Figure 9. The deciles and their corresponding average test scores are shown in the X-axis. According to the descriptive statistics from Table A3, the average test scores are 86.67 for Tier 1 schools and 80.90 for Tier 2 schools. Thus, Tier 1 schools should cluster in the 100th decile, while Tier 2 schools mostly cluster in the 60th to 80th deciles. And the 90th decile should be mixed with both Tier 1 and Tier 2 schools.

Consistent with the previous pattern, the policy significantly decreased the premiums of Tier 1 SDHs while increasing the premiums of Tier 2 SDHs. Besides, the estimate of the 90th decile is close to zero, probably because it includes a combination of the two school tiers.

5.4.2 Event-study with robust DID models

As we mentioned earlier, the policy timing vary in our sample (see Table 2). This reminds us a caveat that the DID estimation could be compromised by the variation in with treatment timing (Callaway and Sant'Anna, 2020; de Chaisemartin and D'Haultfœuille, 2020; Sun and Abraham, 2021). We use two alternative robust DID models proposed by Callaway and Sant'Anna (2020) and Sun and Abraham (2021) and employ them along with our standard DID model in the event-study.

Figure 10 presents the comparison of three models. The robust DID models are marked by red and blue, respectively, while the standard DID model is in black. There are no obvious pre-trends in all three models. Besides, the three models' trends are almost parallel to each other, suggesting that our estimations using standard DID are robust even with various treatment timing.

5.4.3 Justifying using non-SDHs as the control group

One crucial assumption for DID estimation is that the control group is not affected by the school lottery. Given that our DID estimations use the non-SDHs in DHX as the control group, the school lottery may affect the housing prices of non-SDHs.

To examine this possibility, we use Chaoyang's non-SDHs as the control group for DHX's non-SDHs. This is because Chaoyang's non-SDHs are not alternatives to DHX's SDHs that are subject to the policy. Thus, there wouldn't be any spillover from the policy on Chaoyang's non-SDHs. In Figure A3, we compare the time trends of average housing prices among non-SDHs from DHX and Chaoyang. Except for the COVID lockdown period in early 2020, the three trends largely align with each other after the two timings of policy announcement.

Next, we use DHX's non-SDHs as the treatment group and Chaoyang's non-SDHs as the control group. Based on the Equation (3), Table A4 shows the DID estimation results, with $District \times Post$ showing the policy effect on the three treated city districts. In both of the two timings of policy announcement, there are no significant effects on the housing prices of DHX's non-SDHs. This suggests that DHX's non-SDHs were largely unaffected by the school lottery.

5.4.4 Changing the baseline proportion of elite schools in the dense zonings

We next relax the restriction of setting the proportion of elite schools at 60% in the dense zonings. In order to test the robustness of the baseline proportion, we simulate two scenarios by increasing it to 70% and also decreasing it to 50%.

Figure A4 compares the results of three sets of proportions in the dense zonings and other zonings. Across three panels in the dense zonings, the premiums of Tier 1 SDHs all insignificantly decreased after the school lottery, while the premiums of Tier 2 SDHs increased after the policy. In contrast, the premiums of Tier 1 SDHs significantly decreased for all three sets of proportions, while those of Tier 2 SDHs were almost unaffected by the policy for all three sets of proportions. This pattern suggests that our results are robust to different proportions of elite schools in the dense zonings.

5.4.5 Impact of the COVID-19 pandemic

After the COVID-19 pandemic broke out in early 2020, lockdown in Beijing suspended the housing market from late January to mid-March. Even though the lockdown was canceled afterwards, the threat posed by the pandemic was still present. Thus, the long-lasting uncertainties caused by the pandemic may have discouraged or postponed household's willingness to purchase SDHs because of their substantial costs. This mechanism could challenge our results.

To exclude this possibility, we calculated the year-over-year growth rates of housing transaction for SDHs and non-SDHs in 2020. Besides, we also collected China's reported COVID cases in 2020 as a reference of the COVID pandemic intensity. Figure A5 includes four city districts' year-on-year growth rates of both SDH and non-SDH transactions, as well as the reported COVID cases. The figure shows that all city districts' housing transactions returned to normal after COVID subsided in April 2020, with growth rates remaining consistently above zero for the rest of the year. This implies that the pandemic's negative impact on the housing market was a short-term disruption rather than a long-lasting recession.

To further address the threat posed by COVID on parental housing choice, we exclude the Xicheng sample and examine the results using only the samples from Dongcheng and Haidian. The results are shown in Table A6. Comparing the two panels, the pattern does not vary much except that SDHs' premiums are relative higher in Panel A than those in Panel B. This may be attributed to the strategy that families resorted to SDHs from Xicheng after the policy took effect in Dongcheng and Haidian but before the policy was announced in Xicheng.

6 Conclusion

Using a large micro dataset on residential housing transaction from Beijing, this paper examines how school admission uncertainty affects parental choice of housing. We leverage the school lottery policy as a quasi-experiment and employ difference-in-differences to estimate the policy effect on school-quality premiums of school district houses (SDHs). Our findings show that, while the school lottery reduced the school-quality premiums of Tier 1 school district houses, it improved the premiums of Tier 2 SDHs. This is due to the higher admission uncertainty of Tier 1 SDHs than that of Tier 2 SDHs. Additionally, the school lottery had heterogeneous effects on the premiums of SDH in different school zonings. While the premiums of Tier 1 SDHs in dense zonings were unaffected by the lottery, those of Tier 2 SDHs in dense zonings were significantly improved. In contrast, the pattern shown in other zonings were reversed. This can be attributed to dense zoning offering more alternatives elite schools for SDHs not assigned to their original elite school.

The results indicate households resorted to SDH with higher admission probability in response to uncertainty from the school lottery. However, the new pattern of housing choice mitigated the policy effect on narrowing price gaps between SDHs and non-SDHs, which was an unintended consequence from policymakers' perspectives.

The findings from our study contribute to the growing body of literature by providing fresh evidence on the impacts of school district policies on school and housing choice. In particular, our results emphasize the effects of school admission uncertainty on housing prices in the context of residence-based enrollment. These insights are crucial for understanding the unintended and far-reaching consequences of school district reforms, offering policymakers a more comprehensive view of how such policies can shape not only the educational landscape but also the housing market and broader economic outcomes. Besides, our findings also have policy implications for developing countries facing the dilemma of insufficient public school seats coupled with intensive competition among families for elite schools through housing choice.

References

Andreyeva, E. and Patrick, C. (2017). Paying for priority in school choice: Capitalization effects of charter school admission zones. *Journal of Urban Economics*, 100:19–32.

Avery, C. and Pathak, P. A. (2021). The Distributional Consequences of Public School Choice. *American Economic Review*, 111(1):129–152.

Baum-Snow, N. and Lutz, B. F. (2011). School Desegregation, School Choice, and Changes in Residential Location Patterns by Race. *American Economic Review*, 101(7):3019–3046.

Bayer, P., Ferreira, F., and McMillan, R. (2007). A Unified Framework for Measuring Preferences for Schools and Neighborhoods. *Journal of Political Economy*, 115(4):588–638.

Bibler, A. and Billings, S. B. (2020). Win or Lose: Residential Sorting After a School Choice Lottery. *The Review of Economics and Statistics*, 102(3):457–472.

Black, S. E. (1999). Do Better Schools Matter? Parental Valuation of Elementary Education. *The Quarterly Journal of Economics*, 114(2):577–599.

Bonilla-Mejía, L., Lopez, E., and McMillen, D. (2020). House prices and school choice: Evidence from Chicago's magnet schools' proximity lottery. *Journal of Regional Science*, 60(1):33–55.

Brunner, E. J., Cho, S.-W., and Reback, R. (2012). Mobility, housing markets, and schools: Estimating the effects of inter-district choice programs. *Journal of Public Economics*, 96(7):604–614.

Callaway, B. and Sant'Anna, P. H. (2020). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, page S0304407620303948.

Chan, J., Fang, X., Wang, Z., Zai, X., and Zhang, Q. (2020). Valuing primary schools in urban China. *Journal of Urban Economics*, 115:103183.

Chen, J. and Li, R. (2023). Pay for elite private schools or pay for higher housing prices? Evidence from an exogenous policy shock. *Journal of Housing Economics*, 60:101934.

de Chaisemartin, C. and D'Haultfœuille, X. (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review*, 110(9):2964–2996.

Ely, T. L. and Teske, P. (2015). Implications of Public School Choice for Residential Location Decisions. *Urban Affairs Review*, 51(2):175–204.

Feng, H. and Lu, M. (2013). School quality and housing prices: Empirical evidence from a natural experiment in Shanghai, China. *Journal of Housing Economics*, 22(4):291–307.

Ferreyra, M. M. (2007). Estimating the Effects of Private School Vouchers in Multidistrict Economies. *American Economic Review*, 97(3):789–817.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, page S0304407621001445.

Han, J., Cui, L., and Yu, H. (2021a). Pricing the value of the chance to gain admission to an elite senior high school in Beijing: The effect of the LDHSE policy on resale housing prices. *Cities*, 115:103238.

Han, X., Shen, Y., and Zhao, B. (2021b). Winning at the starting line: The primary school premium and housing prices in Beijing. *China Economic Quarterly International*, 1(1):29–42.

Huang, B., He, X., Xu, L., and Zhu, Y. (2020). Elite school designation and housing prices-quasi-experimental evidence from Beijing, China☆. *Journal of Housing Economics*, 50:101730.

Jin, Z., Wang, X., and Huang, B. (2023). The enrolment reform of schools and housing price: Empirical evidence from Shanghai, China. *International Review of Economics & Finance*, 84:262–273.

Snyder, T. D., de Brey, C., and Dillow, S. A. (2019). Digest of Education Statistics 2017, 53rd Edition. NCES 2018-070. *National Center for Education Statistics*.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.

Tong, J., Zhang, C., Yue, T., Bai, Y., and Shao, L. (2024). The distributional effects of introducing a lottery system in school assignment rule: Evidence from an experiment in Beijing, China. *International Journal of Educational Development*, 107:103030.

Yang, S., Luo, R., Zhang, P., Zou, L., Qu, S., Zhang, C., Zhou, L., and Hu, S. (2024). The effect of a multi-school zoning policy on the capitalization of educational resources: A quasi-natural experiment in Beijing. *Cities*, 152:105153.

Zheng, S., Hu, W., and Wang, R. (2016). How Much Is a Good School Worth in Beijing? Identifying Price Premium with Paired Resale and Rental Data. *The Journal of Real Estate Finance and Economics*, 53(2):184–199.

Figures and Tables

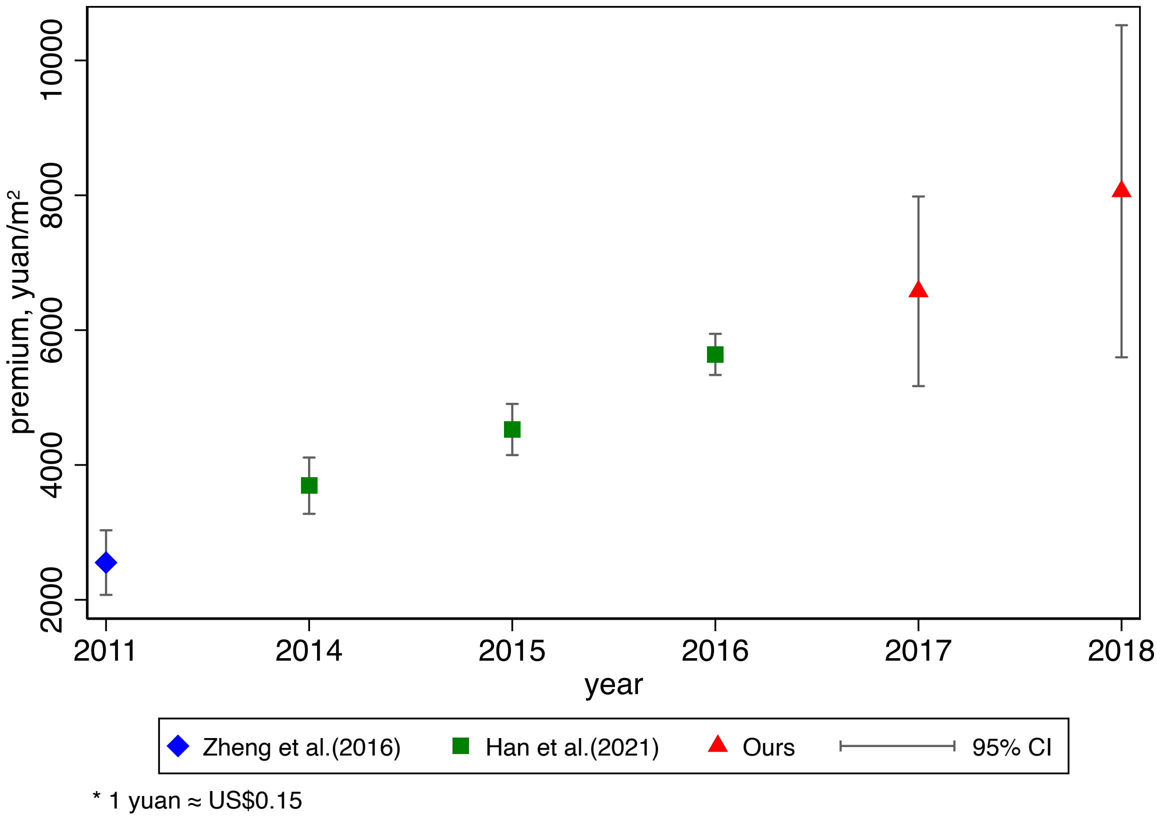
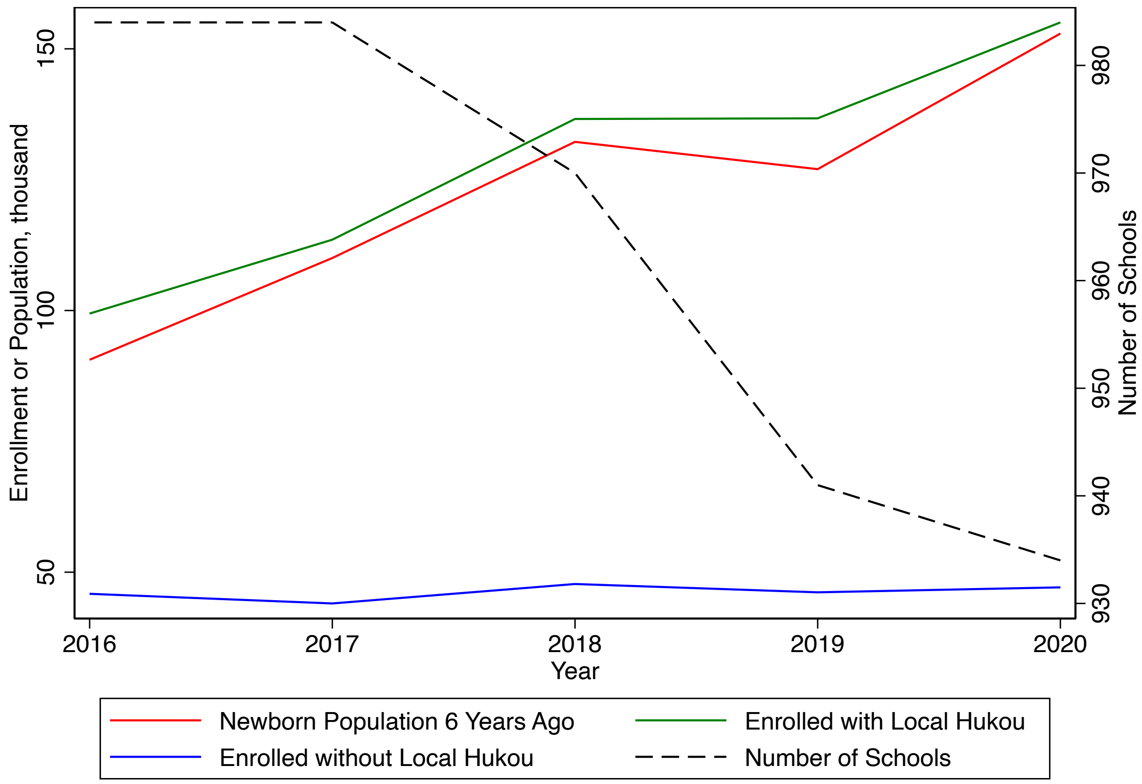


Figure 1: An increasing trend of school-quality premium in Beijing.

Notes: The purple square refers to the school-quality premium estimated by Zheng et al. (2016) in 2011; the green squares refer to the school-quality premium estimated by Han et al. (2021b); the red triangles refer to the school-quality estimated by us.



Source: *Beijing Statistical Yearbook*

Figure 2: School capacity increasingly overwhelmed in Beijing.

Notes: (1) The red curve represents the newborn population in Beijing 6 years ago. (2) The green curve represents the number of newly enrolled elementary school students with Beijing's local Hukou. (3) The blue curve represents the number of newly enrolled elementary school students without Beijing's local Hukou. (4) The black dashed curve represents the number of elementary schools in Beijing.

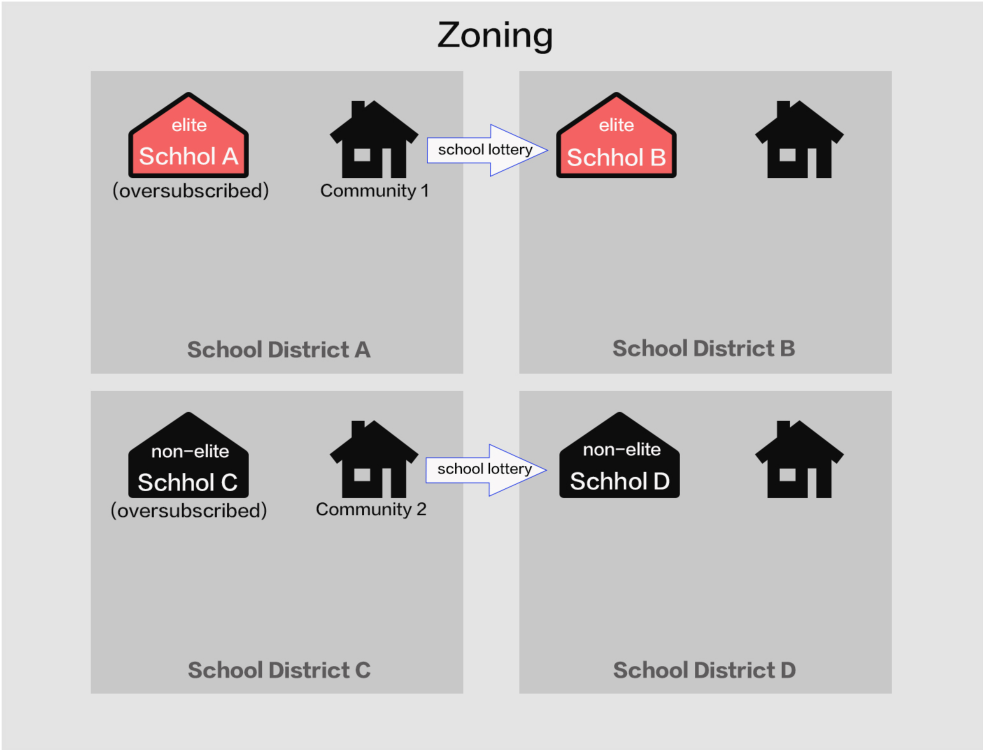


Figure 3: An illustration for the school lottery in a zoning.

Notes: (1) Schools in red color are elite schools, while schools in black color are non-elite schools. (2) "Oversubscribed" means that the number of applicants exceeds the enrollment plan. (3) The arrow of the "school lottery" stands for some students who are assigned to the neighboring school under the school lottery. (4) The four school districts constitute a zoning.

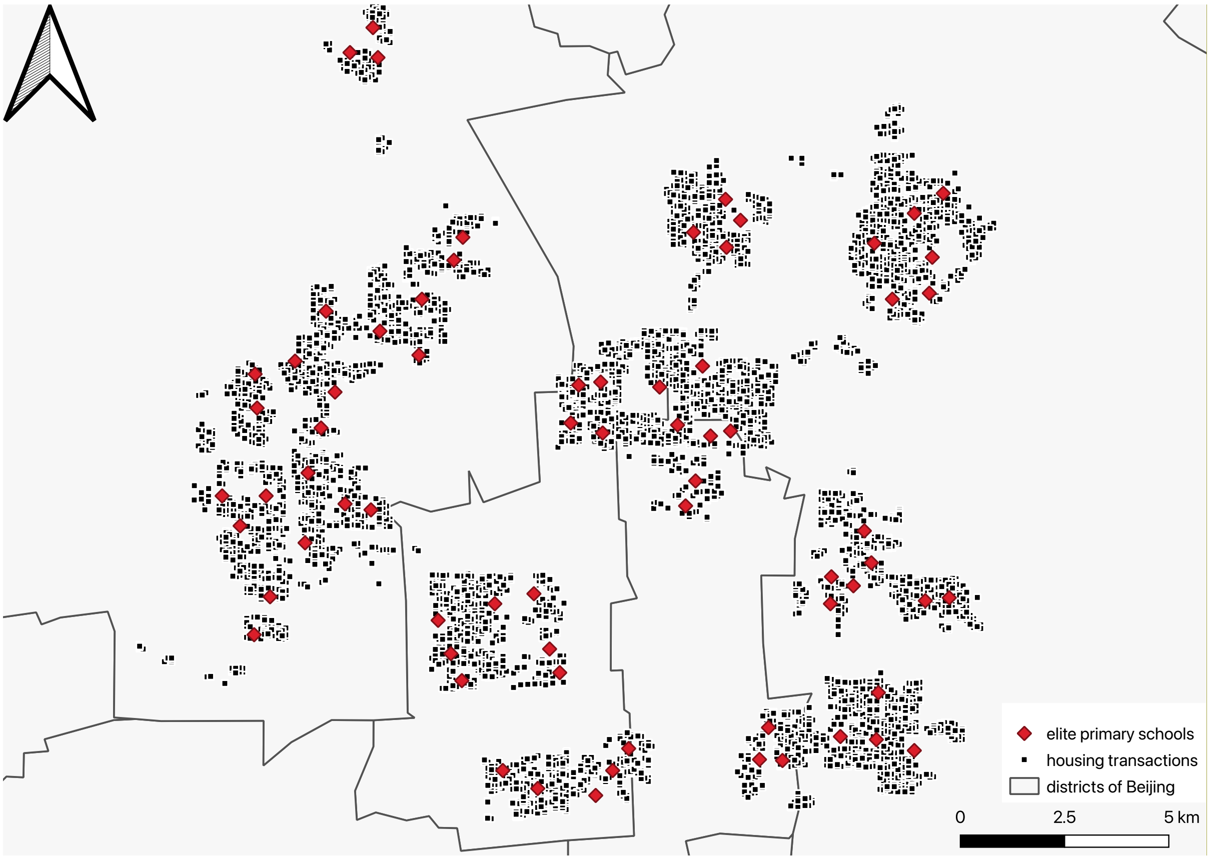


Figure 4: Dense zonings.

Notes: This figure shows the distribution of elite schools, represented by the red dots, in dense zonings outlined by the black dots.

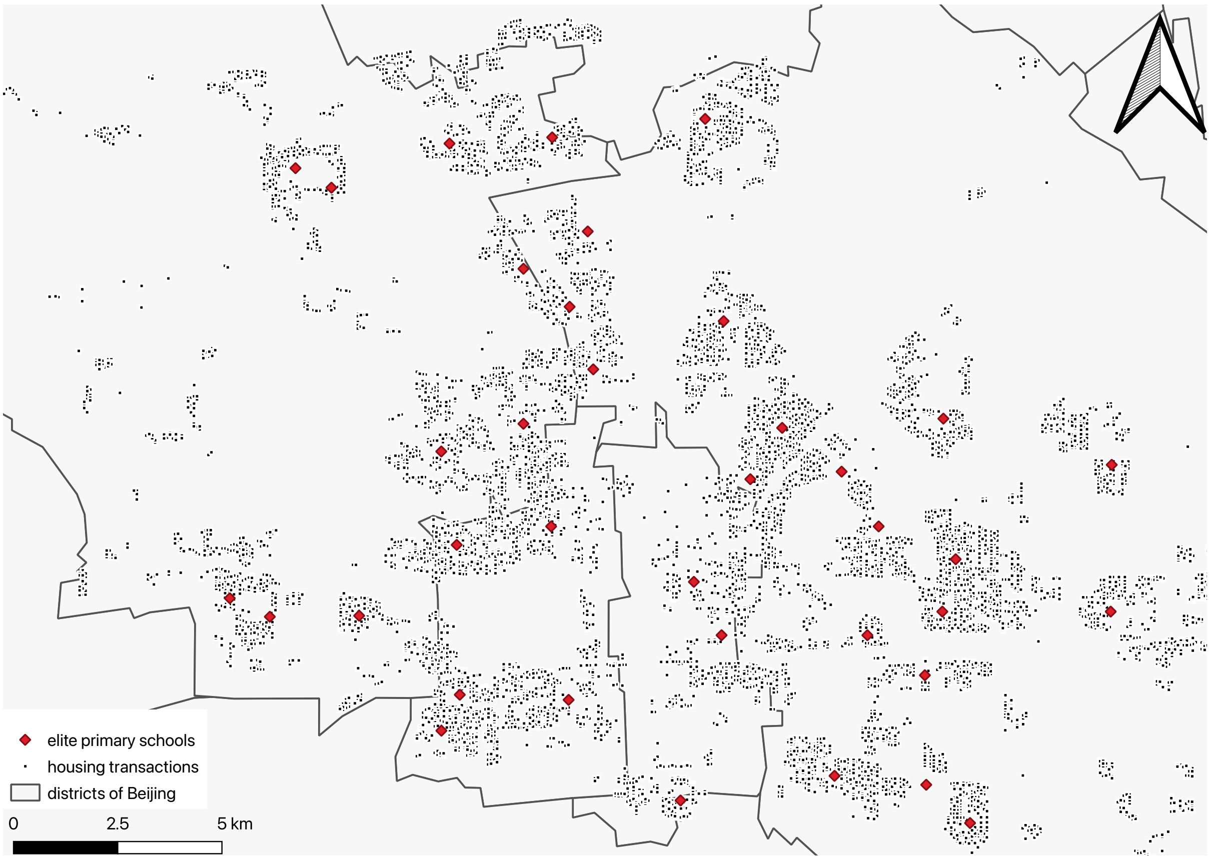


Figure 5: The other zonings.

Notes: This figure shows the distribution of elite schools, represented by the red dots, in the other zonings outlined by the black dots. In contrast with Figure 4, the distribution of elite schools are more sparse within the other zonings than dense zonings.

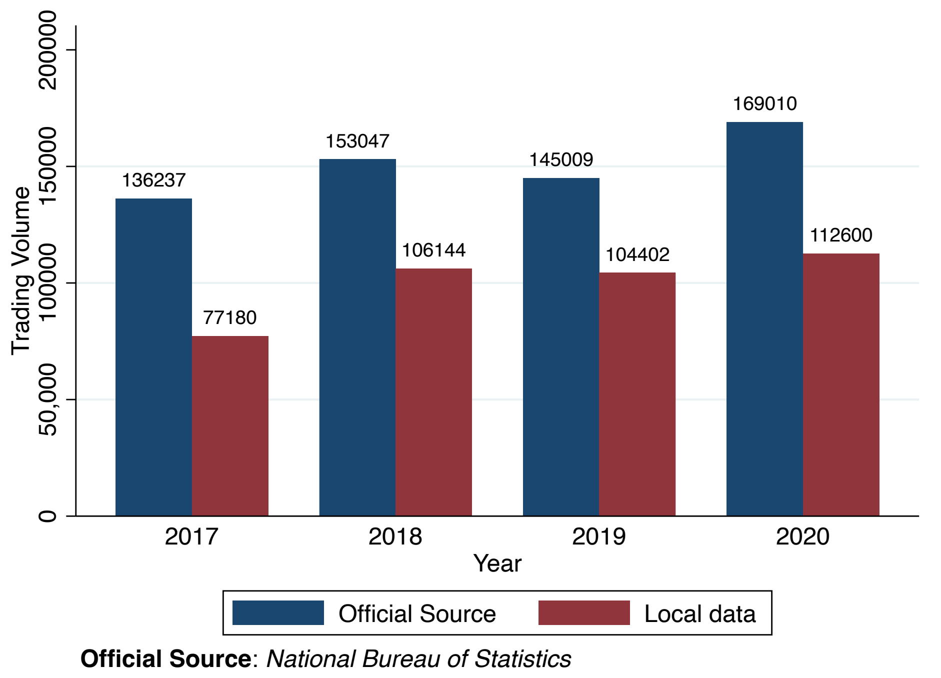


Figure 6: Transaction volume of second-hand houses in Beijing.

Notes: The figure compares the trading volume between official data from National Bureau of Statistics and the local data we used in this paper. The trading volume from local data is close to that of the official data, suggesting the local data is representative.

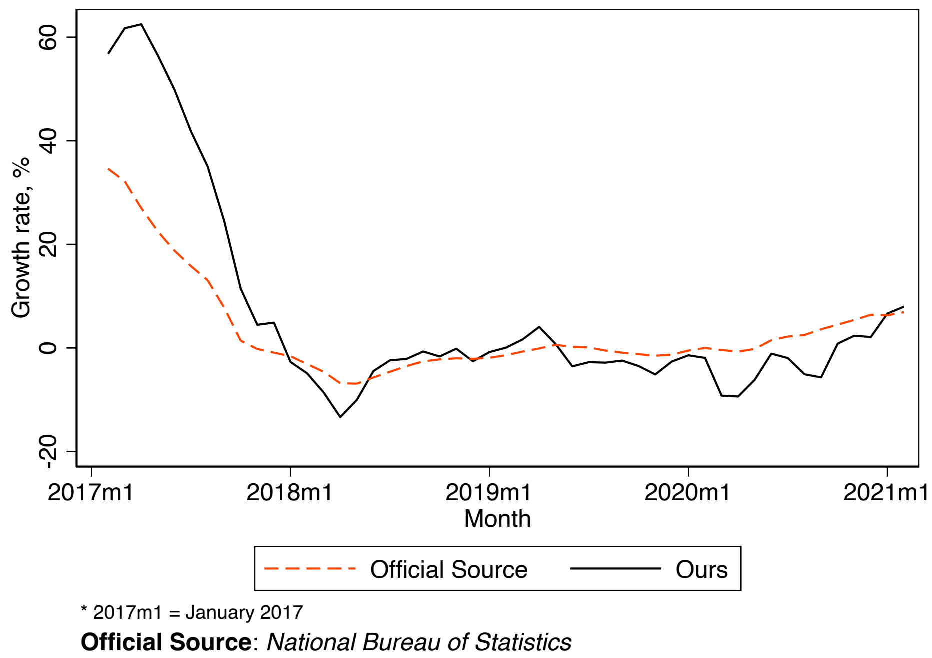


Figure 7: Year-over-year growth rate of housing prices.

Notes: This figure compares year-over-year growth rate of housing prices between the official data and the local data. The two curves align with each other, indicating the representativeness of the local data.

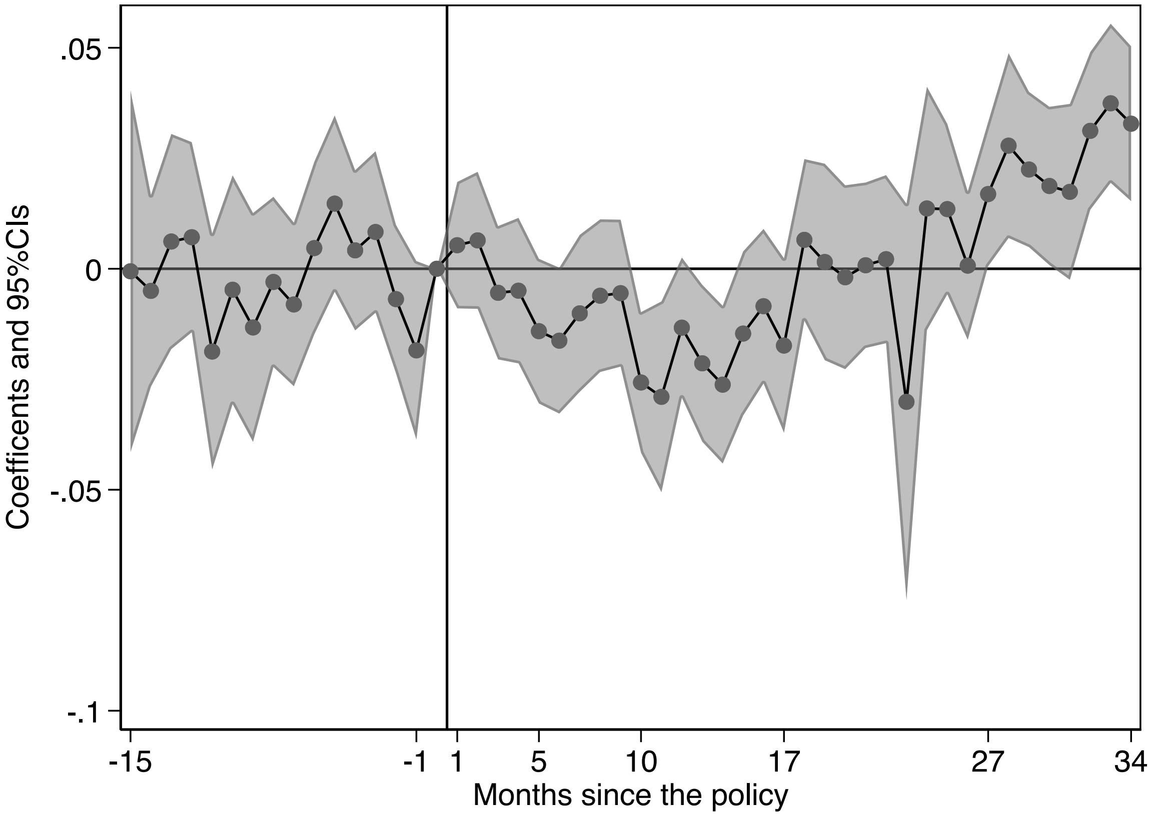


Figure 8: Event study of the policy.

Notes: This figure depicts the event study of the policy's dynamic effects for the three city districts. The trend fluctuated around the zero line prior to the policy, validating the parallel trend assumption for DID estimation. After the policy was implemented, the trend declined for 17 months before recovering. This suggests that the school lottery suppressed the premiums of SDHs initially, but the effect diminished over the long term.

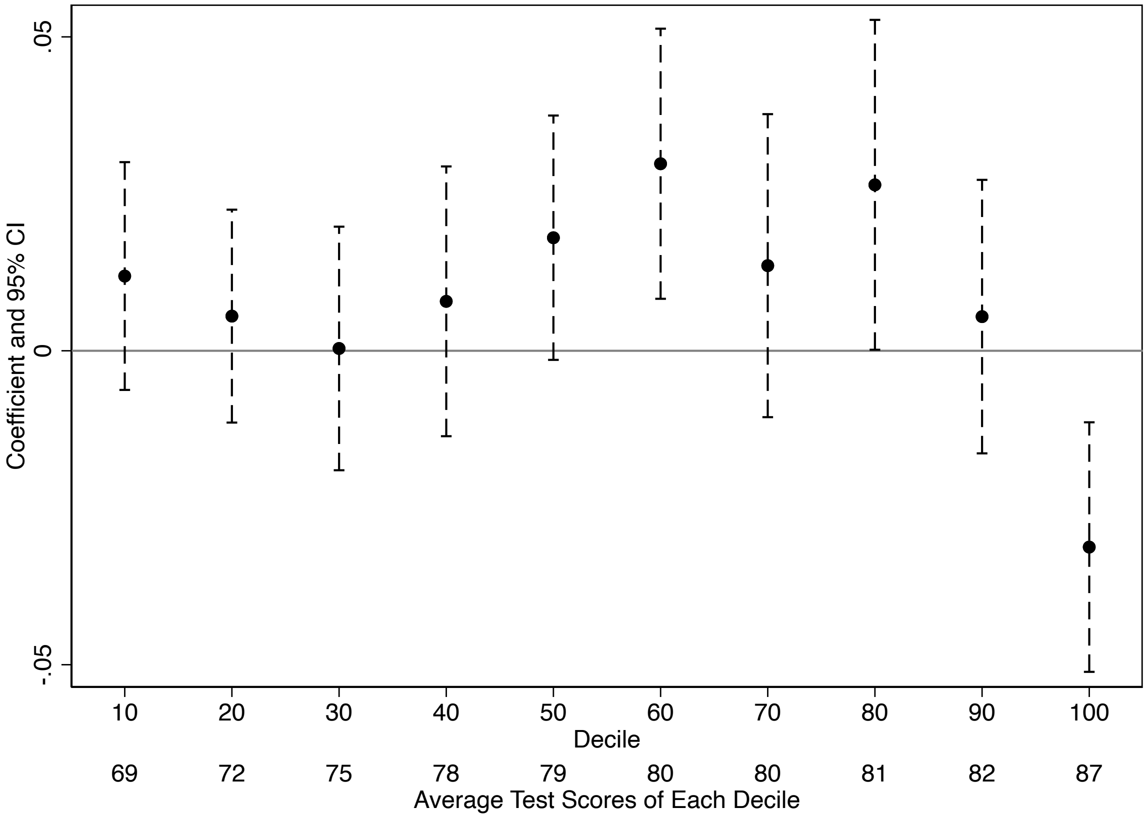


Figure 9: Effects of school lottery by school-quality decile.

Notes: This figure shows the policy effects and their 95% confidence intervals across different school quality represented by the school-quality deciles. Under the school-quality deciles are the corresponding average test scores for each decile.

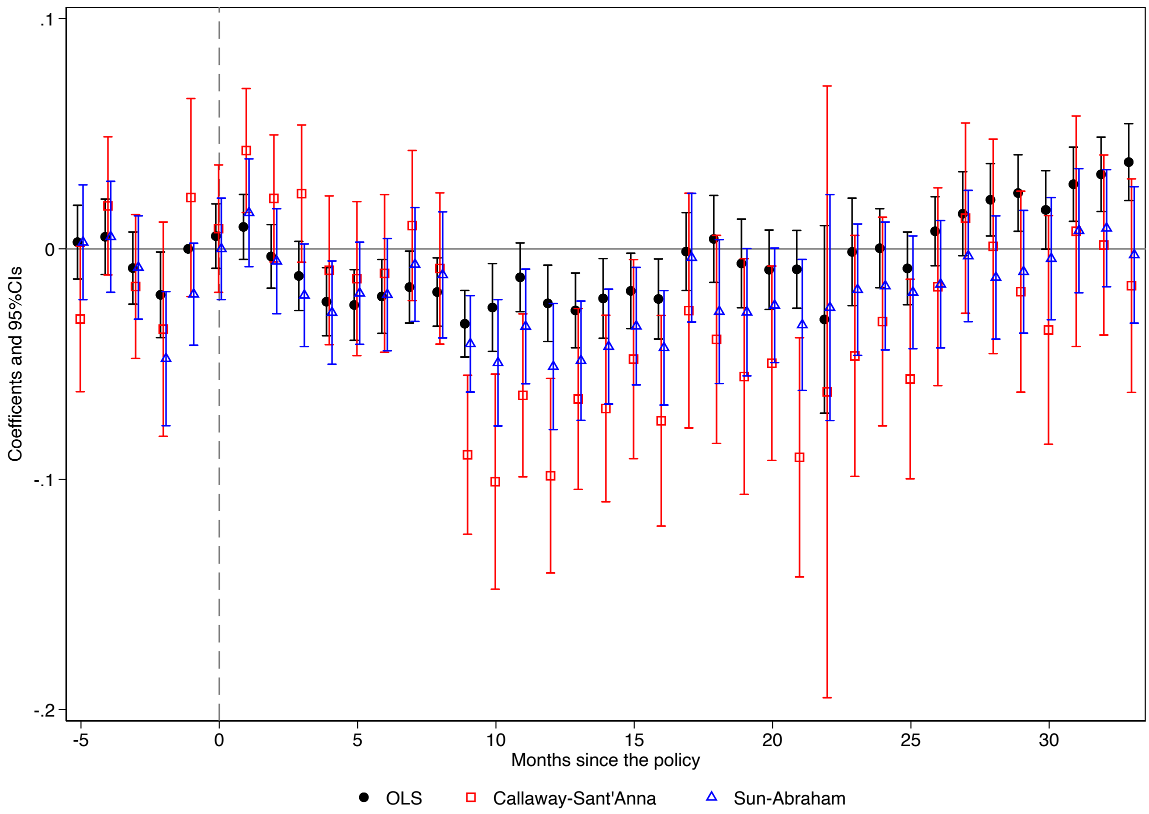


Figure 10: Event study with alternative robust DID models.

Notes: This figure depicts event study using three DID models: OLS, Callaway-Sant'Anna Robust DID (Callaway and Sant'Anna, 2020), and Sun-Abraham Robust DID (Sun and Abraham, 2021). Comparing the three trends reveals no significant disparity, suggesting the robustness of using the OLS model for the staggered DID setting.

Table 1: Descriptive statistics of samples from Dongcheng, Haidian, and Xicheng

	Non-SDH	SDH	Tier 1 SDH	Tier 2 SDH	Other SDH
	(1)	(2)	(3)	(4)	(5)
Panel A. School statistics					
School's characteristics					
Average test score (max=100)	66.21	78.58	86.67	80.90	73.03
Total no. of grade 1 classes	712	685	91	310	284
School district's characteristics					
Total no. of communities	327	378	59	168	151
% Admission probability	100	64.95	51.93	70.29	67.41
Panel B. Housing statistics					
Housing price(Yuan/m^2)	83,166	103,284	121,890	102,901	95,304
Architectural features					
Area(m^2)	74.1	71.0	74.3	70.4	70.7
Well-decorated (Yes=1)	0.45	0.36	0.34	0.34	0.39
Elevator (Yes=1)	0.52	0.49	0.60	0.50	0.51
Entrance guard (Yes=1)	0.77	0.74	0.70	0.75	0.71
Building age (Year)	24.60	28.55	25.79	29.68	28.12
Community characteristics					
Dist. to assigned school (KM)	1.59	1.51	1.51	1.51	1.56
Management fee (Yuan/m^2)	1.46	1.42	1.59	1.38	1.41
Green area (%)	28.44	26.47	25.34	26.93	26.16
Distance to CBD (KM)	7.06	6.88	6.12	6.02	5.99
Hospital nearby (Yes = 1)	0.51	0.56	0.37	0.59	0.59
Supermarket nearby (Yes = 1)	0.88	0.91	0.90	0.92	0.90
Observations	33,762	28,537	3,763	14,962	9,792

Notes: The table reports the mean of each variable. The average test score is the originally assigned school's average test score from a citywide mathematics exam, which was taken by all grade-6 students in Beijing in 2014. The class number in grade 1 is the total number of classes in the originally assigned school's grade 1. Community number is the total number of communities in the school district. The admission probability is measured by Class/Community family, the ratio between the number of originally assigned school's grade 1 class and the number of families in the communities in the associated school district.

Table 2: Treatment and control group designation in staggered DID

Group Designation	Housing Category	Policy Timing
Treatment Group	Dongcheng SDHs, Haidian SDHs	April 18, 2018
	Xicheng SDHs	April 20, 2020
Control Group	Non-SDHs in DHX	Null

Notes: This table presents the assignment of treatment group and control group: the SDHs from DHX are the treatment group, while non-SDHs in DHX are the control group. Aditionally, the SDHs from Dongcheng and Haidian are subject to policy timing of April 18, 2018, while those from Xicheng subject to policy timing of April 20, 2020.

Table 3: School-quality premiums of SDHs before the school lottery

	Dependent variable: ln(housing price)					
	(1)	(2)	(3)	(4)	(5)	(6)
SDH	0.2137***		0.0930***		0.0770***	
	(0.0227)		(0.0115)		(0.0108)	
Tier 1 SDH		0.3757***		0.1655***		0.1600***
		(0.274)		(0.0173)		(0.0244)
Tier 2 SDH		0.2112***		0.0989***		0.0785***
		(0.0239)		(0.0112)		(0.0123)
Other SDH		0.1500***		0.0596***		0.0530***
		(0.0334)		(0.0129)		(0.0149)
Boundary FE	No	No	No	No	Yes	Yes
Year-month FE	No	No	Yes	Yes	Yes	Yes
Covariates	No	No	Yes	Yes	Yes	Yes
Match distance	No	No	No	No	500m	500m
Observations	21,630	21,630	21,484	21,484	20,396	20,396
R-squared	0.753	0.757	0.843	0.844	0.841	0.842

Notes: (1) "Match" stands for matching the SDH with the nearest non-SDH. While "unlimited" stands for not limiting the matching distance, "500m" is setting the matching distance as 500 meters. (2) Standard errors are clustered at the community level and are reported in parentheses, * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 4: Policy effects on the school-quality premiums

	SDH	Tier 1 SDH	Tier 2 SDH
	(1)	(2)	(3)
	Dependent variable: ln(housing price)		
Panel A. DID estimation (Reduced form)			
$SDH \times Post$	0.0029	-0.0391***	0.0133**
	(0.0047)	(0.0096)	(0.0057)
R-squared	0.8326	0.8689	0.8425
Sample	$\sum_{i=1}^{4} S_i$	$S_1 + S_4$	$S_2 + S_4$
Observations	57,969	35,833	46,366
Panel B. 2SLS estimation			
%Admission	-0.0085	0.0814***	-0.0456**
	(0.0131)	(0.0205)	(0.0197)
Sample	$\sum_{i=1}^{4} S_i$	$S_1 + S_4$	$S_2 + S_4$
Observations	57,969	35,833	46,366
Panel C. First stage for 2SLS			
	Dependent variable: *%Admission*		
$SDH \times Post$	-0.3421***	-0.4803***	-0.2919***
	(0.0121)	(0.0108)	(0.089)
F-statistics for IV	893	1978	1188
Sample	$\sum_{i=1}^{4} S_i$	$S_1 + S_4$	$S_2 + S_4$
Observations	57,969	35,833	46,366
Panel D. Alternative 2SLS			
	Dependent variable: ln(housing price)		
Relative %Admission		0.2781**	0.2781**
		(0.1212)	(0.1212)
Sample		$S_1 + S_2$	$S_1 + S_2$
Observations		18,725	18,725
Panel E. Alternative first stage			
	Dependent variable: *Relative %Admission*		
$Tier1 \times Post$		-0.1884***	
		(0.0095)	
$Tier2 \times Post$			0.1884***
			(0.0095)
F-statistics for IV		364	364
Sample		$S_1 + S_2$	$S_1 + S_2$
Observations		18,725	18,725
Covariates	Yes	Yes	Yes
Boundary FE	Yes	Yes	Yes
Year-month FE	Yes	Yes	Yes

Notes: (1) The table shows the policy effects using three estimation methods: DID, OLS, and 2SLS. The first three panels share the same dependent variable, while the last panel uses admission probability as the dependent variable. (2) Standard errors are clustered at the community level and reported in parentheses, * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 5: Policy effects on school-quality premiums in different zonings

	Tier 1 SDH			Tier 2 SDH		
	DID	2SLS	1st Stage	DID	2SLS	1st Stage
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	ln(housing price)		*%Admission*	ln(housing price)		*%Admission*
Panel A. Dense Zonings						
$SDH \times Post$	-0.0004		-0.4110***	0.0174***		-0.2500***
	(0.0073)		(0.0343)	(0.0055)		(0.0167)
%Admission		0.0010			-0.0696**	
		(0.0197)			(0.0349)	
R-squared	0.8645			0.783		
F-statistic for IV			249			254
Observations	7,465	7,465	7,465	13,235	13,235	13,235
Panel B. Other Zonings						
$SDH \times Post$	-0.0581***		-0.5399***	0.0019		-0.3189***
	(0.0087)		(0.0370)	(0.0043)		(0.0170)
%Admission		0.1076***			-0.0060	
		(0.0212)			(0.0202)	
R-squared	0.8526			0.8380		
F-statistic for IV			260			366
Observations	29,077	29,077	29,077	33,424	33,424	33,424
Sample	$S_1 + S_4$	$S_1 + S_4$	$S_1 + S_4$	$S_2 + S_4$	$S_2 + S_4$	$S_2 + S_4$
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Boundary FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-month FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: 1. Columns (1) and (4) show the results from DID estimation. Columns (2) and (5) present the results from the second-stage estimations of 2SLS, while columns (3) and (6) show the first-stage estimations of 2SLS. Additionally, the fist-stage estimation uses admission probabilty as the dependent variable. 2. Standard errors are clustered at the community level and reported in parentheses, * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 6: Spillover from the three treated city districts on Chaoyang

	Dependent variable: ln(housing price)		
	SDH	Tier 1 SDH	Tier 2 SDH
	(1)	(2)	(3)
Panel A. Policy timing: April 2018			
$SDH \times Post$	0.0218**	0.0532**	0.0259**
	(0.0094)	(0.0235)	(0.0106)
Sample	$\sum_{i=1}^{4} S_i$	$S_1 + S_4$	$S_2 + S_4$
Covariates	Yes	Yes	Yes
Boundary FE	Yes	Yes	Yes
Year-month FE	Yes	Yes	Yes
Observations	66,649	58,933	63,128
R-squared	0.731	0.739	0.736
Panel B. Policy timing: April 2020			
$SDH \times Post$	0.0345***	0.0521***	0.0384***
	(0.0083)	(0.0079)	(0.0120)
Sample	$\sum_{i=1}^{4} S_i$	$S_1 + S_4$	$S_2 + S_4$
Covariates	Yes	Yes	Yes
Boundary FE	Yes	Yes	Yes
Year-month FE	Yes	Yes	Yes
Observations	66,649	58,933	63,128
R-squared	0.732	0.740	0.737

Notes: (1) Panel A presents the results of spillover using the policy timing of Haidian district and Dongcheng district. Panel B shows the results of spillover using the policy timing of Xicheng district. (2) Standard errors are clustered at the community level and reported in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix

A.1 Further discussion: School distance variance

It is possible that by purchasing out of Tier 1 SDHs, families are also mitigating variance in possible distance to assigned school.

Table A5 shows the correlation between the distance to the assigned school and the housing prices. The commuting distance to the school is statistically significantly valued by families, especially by those buying Tier 1 SDHs. Therefore, families who would have purchased Tier 1 SDHs without the school lottery value commuting distance and attempt to reduce uncertainty in the distance to their assigned school by purchasing out of Tier 1 SDHs.

Tables and graphs

Table A1: The summary of admission probability for different SDHs

	Non-SDH	Tier 1 SDH	Tier 2 SDH	Other SDH
	(1)	(2)	(3)	(4)
Admission probability				
Mean	1	0.5193	0.7029	0.6741
S.D.	0	0.0322	0.0408	0.0645
Max	1	0.6948	0.8919	0.9399
Min	1	0.3366	0.5178	0.4905

Notes: The table reports the average values, standard deviations, maximum, and minimum values of admission probability for different types of SDHs.

Table A2: Comparisons among the main city districts in Beijing

	Dongcheng	Xicheng	Haidian	Chaoyang
Economic conditions				
GDP (million yuan)	24.26	42.44	64.80	60.94
Per capita GDP (thousand yuan)	295	360	193	169
Per capita disposable income (yuan)	75,547	81,678	78,178	70,746
Demographic conditions				
Population (thousand)	822	1,179	3,358	3,605
No. of fam. w/ hukou (hundred)	516	696	1089	1602
No. of communities	204	221	280	339
No. of families per community	253	314	399	432
% fam. w/ hukou & a child	36.95%	34.01%	39.23%	43.80%
Educational conditions				
No. of Tier 1 elementary schools	4	6	5	3
No. of Tier 2 elementary schools	12	15	16	13
No. of elementary schools	61	58	84	83
No. of grade 1 students	11,025	14,328	30,484	28,505
No. of grade 1 classes	280	358	759	702
Grade 1 student / Class	39.4	40.0	40.2	40.6

Notes: The table reports the summary statistics from four city districts in terms of economy, demography, and education.

Table A3: Average test scores of different school tiers

	N	Mean	SE	Min	Max
Tier 1 School	18	86.67	1.44	81.67	91.71
Tier 2 school	56	80.9	1.7	72.82	87
Other Elite School	41	73.03	2.19	65.03	81.02
Non-elite School	171	66.21	1.56	59.74	72.14

Notes: This table summarizes the average school-level test scores across different types of schools.

Table A4: Robustness check: Non-SDH from DHX vs Non-SDH from Chaoyang

Policy timing	April 2018	April 2020
	(1)	(2)
$District \times Post$	0.0051	0.0089
	(0.0053)	(0.0083)
Treatment city district	Dongcheng + Haidian	Xicheng
Sample	S_4	S_4
Covariates	Yes	Yes
Boundary FE	No	No
Year-month FE	Yes	Yes
Observations	90,814	66,386
R-squared	0.7047	0.7031

Notes: Standard errors are clustered at the community level and reported in parentheses, * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A5: Housing price and the distance from home to school, sample from DHX

	Non-SDH	SDH	Tier 1 SDH
	(1)	(2)	(3)
Dist. to school	-0.006	-0.015*	-0.034**
	(0.006)	(0.008)	(0.016)
Sample	S_4	$S_1+S_2+S_3$	S_1
Covariates	Yes	Yes	Yes
Year-month FE	Yes	Yes	Yes
Observations	33,762	28,537	3,763
R-squared	0.856	0.843	0.872

Notes: Standard errors are clustered at the community level and reported in parentheses, * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A6: Policy effects on the school-quality premiums in Dongcheng and Haidian

	Dependent variable: ln(housing price)		
	SDH	Tier 1 SDH	Tier 2 SDH
	(1)	(2)	(3)
Panel A. Full sample			
$SDH \times Post$	0.0095*	-0.0198***	0.0098*
	(0.0058)	(0.0073)	(0.0051)
Sample	$\sum_{i=1}^{4} S_i$	S_1+S_4	S_2+S_4
Covariates	Yes	Yes	Yes
Boundary FE	Yes	Yes	Yes
Year-month FE	Yes	Yes	Yes
Observations	40,462	25,876	32,894
R-squared	0.7959	0.8402	0.8097
Panel B. Discarding sample after 2020			
$SDH \times Post$	-0.0020	-0.0286***	0.0068
	(0.0037)	(0.0071)	(0.0047)
Sample	$\sum_{i=1}^{4} S_i$	S_1+S_4	S_2+S_4
Covariates	Yes	Yes	Yes
Boundary FE	Yes	Yes	Yes
Year-month FE	Yes	Yes	Yes
Observations	26,681	16,933	21,655
R-squared	0.8045	0.8505	0.8208

Notes: Standard errors are clustered at the community level and reported in parentheses, * p < 0.1, ** p < 0.05, *** p < 0.01.

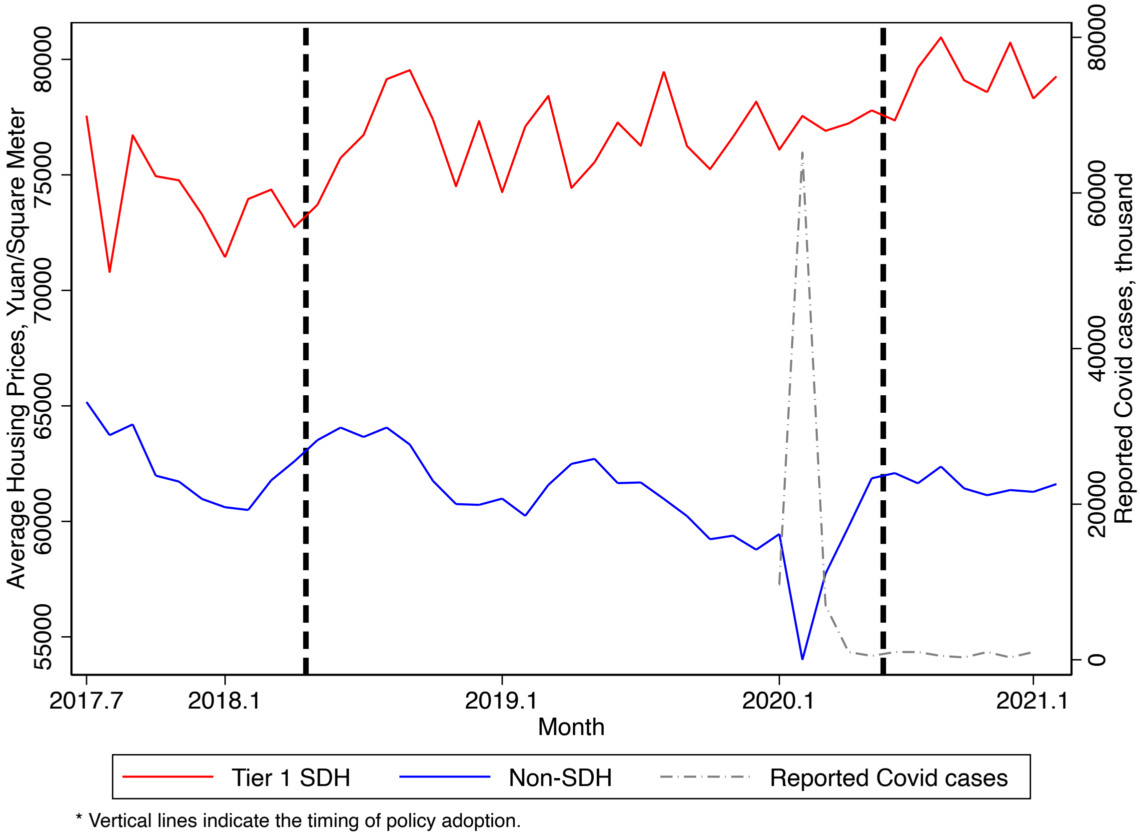


Figure A1: Monthly Housing Prices of Tier 1 SDH and non-SDH in Chaoyang

Notes: This figure illustrates the trends of monthly housing prices from Tier 1 SDH and non-SDH in Chaoyang. The two vertical dash lines are the timing of two policy adoption. The red line represented by the Tier 1 SDH shows two surges right after the policy adopted, suggesting spillover from the city districts affected by the polciy.

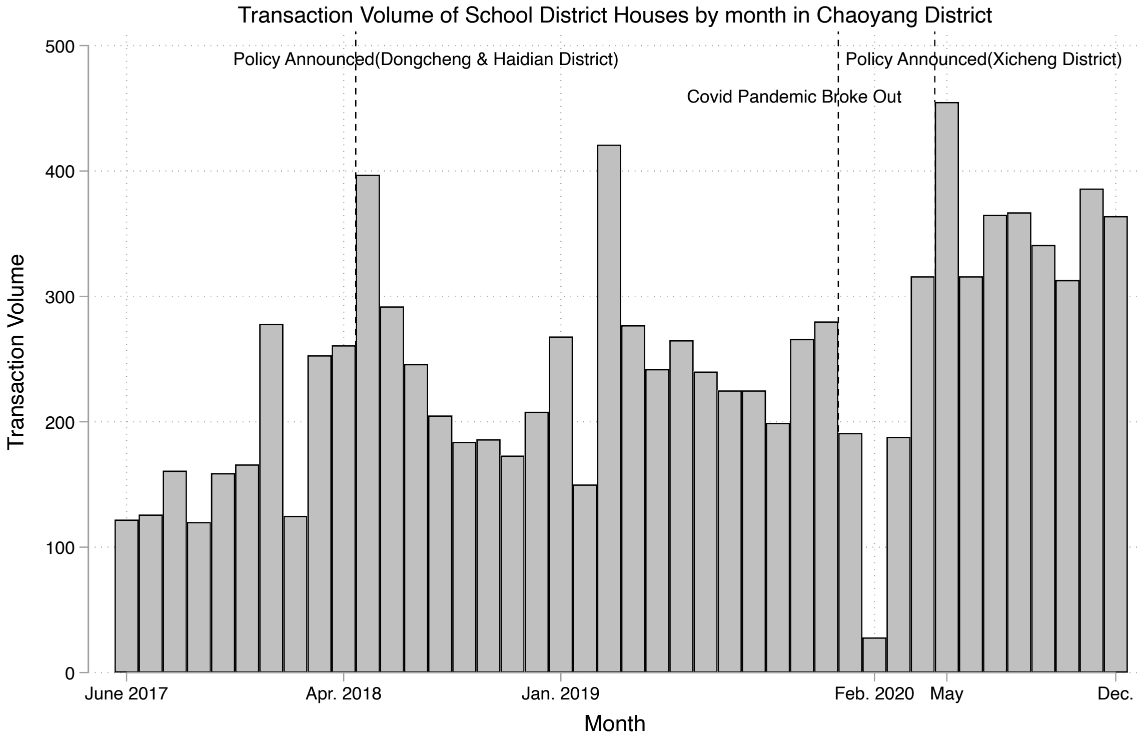


Figure A2: Monthly Transaction Volume of School District Homes in Chaoyang District.

Notes: This figure illustrates the monthly transaction volume of SDHs from Chaoyang. There were surges right after the policy was adopted in the adjacent city district, suggesting the presence of spillover from the treated districts.

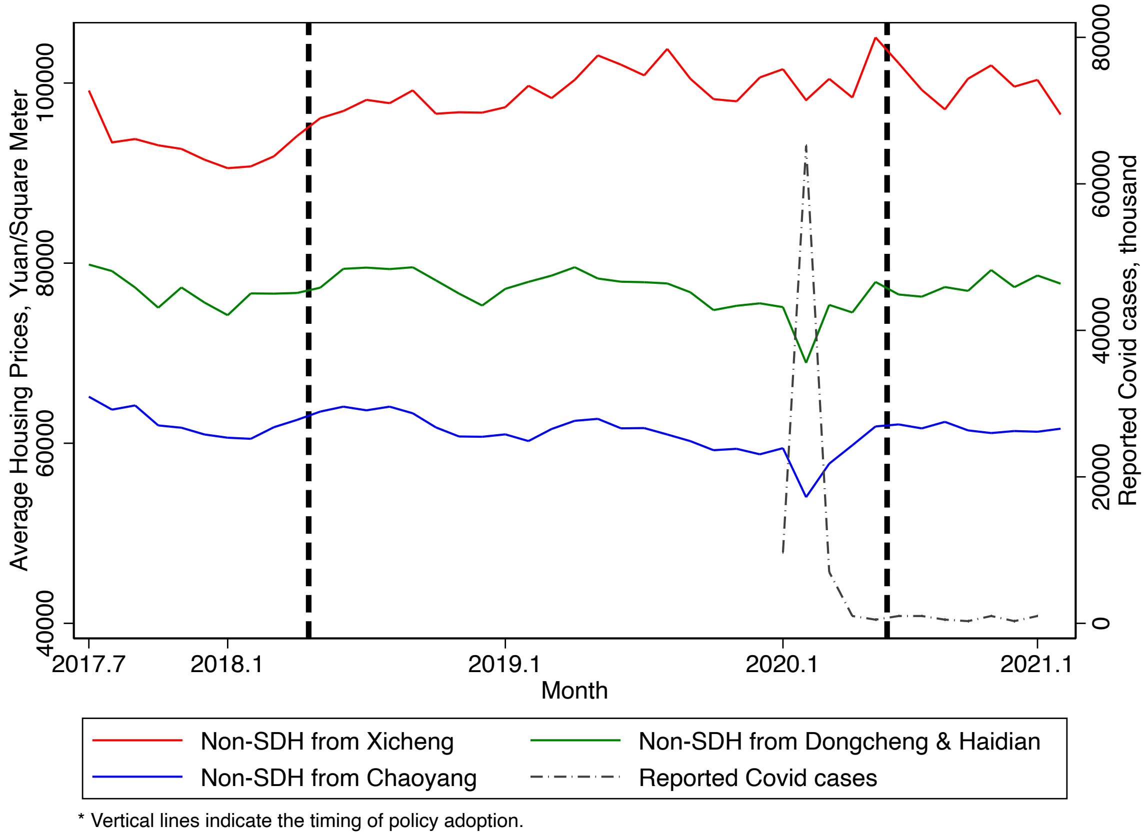


Figure A3: Monthly Housing Prices of non-SDH in different city districts.

Notes: The figure shows the trends of non-SDHs from three treated city districts. There were no great fluctuation around the timging of the policy adoption, validating their status as the control group for DID.

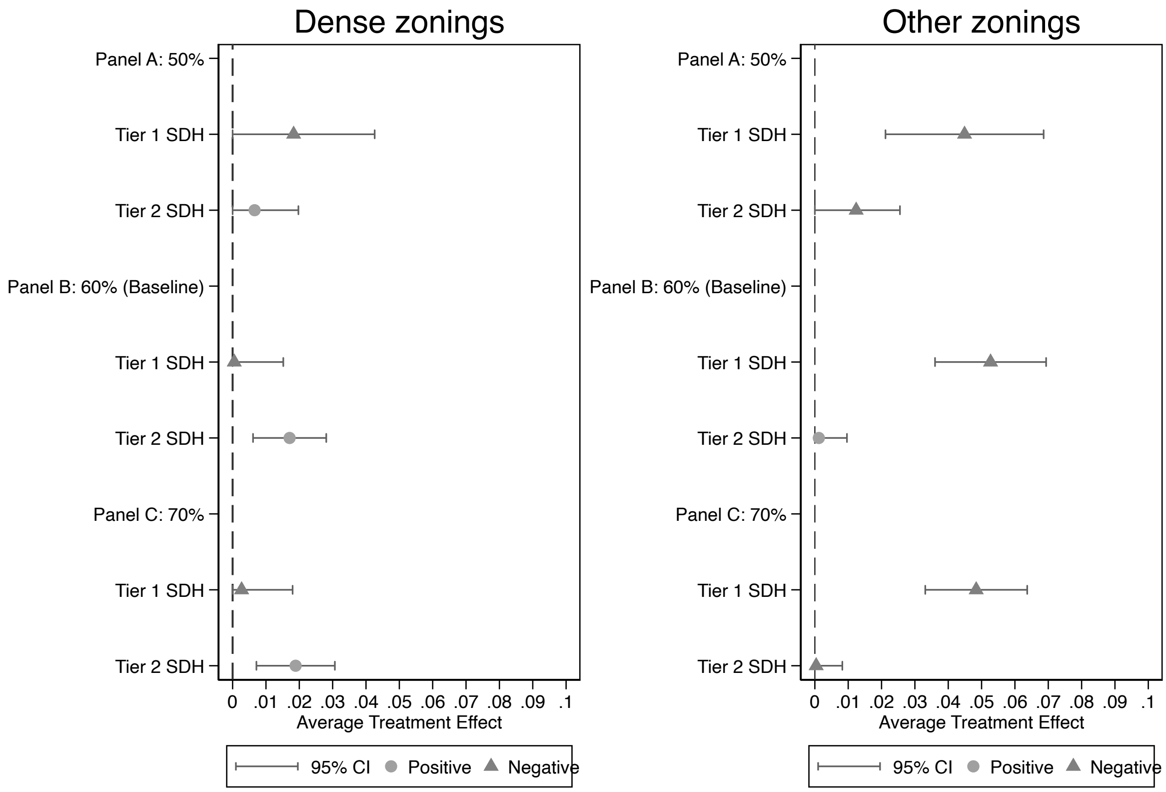
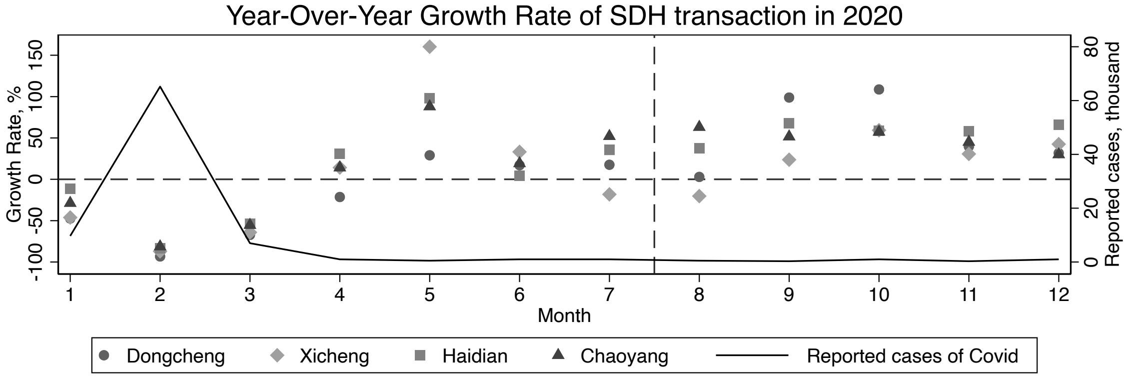


Figure A4: Regression results of dense and other zonings with different ratios.

Notes: (1) The left figure depicts the policy's treatment effects on SDHs with different proportions of elite schools in the dense zonings, while the right one depicts those in other zonings. (2) Panel A, B, and C set the proportion of elite schools in the dense zonings as 50%, 60%, and 70%, respectively.



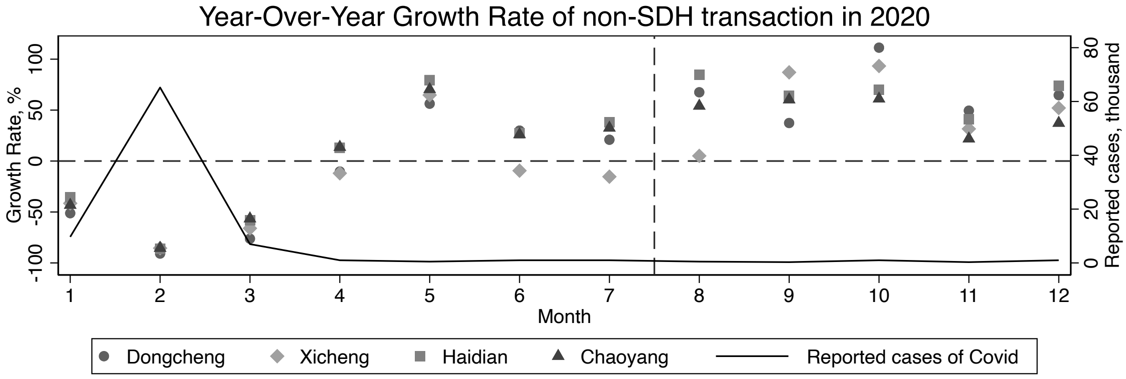


Figure A5: Year-over-year change in SDH transactions in 2020.

Notes: The graph illustrates how year-over-year growth rate of housing transaction responded to COVID pandemic. After reported cases of COVID faded away, the growth rate in all city districts turned positive and housing transaction returned to normal.

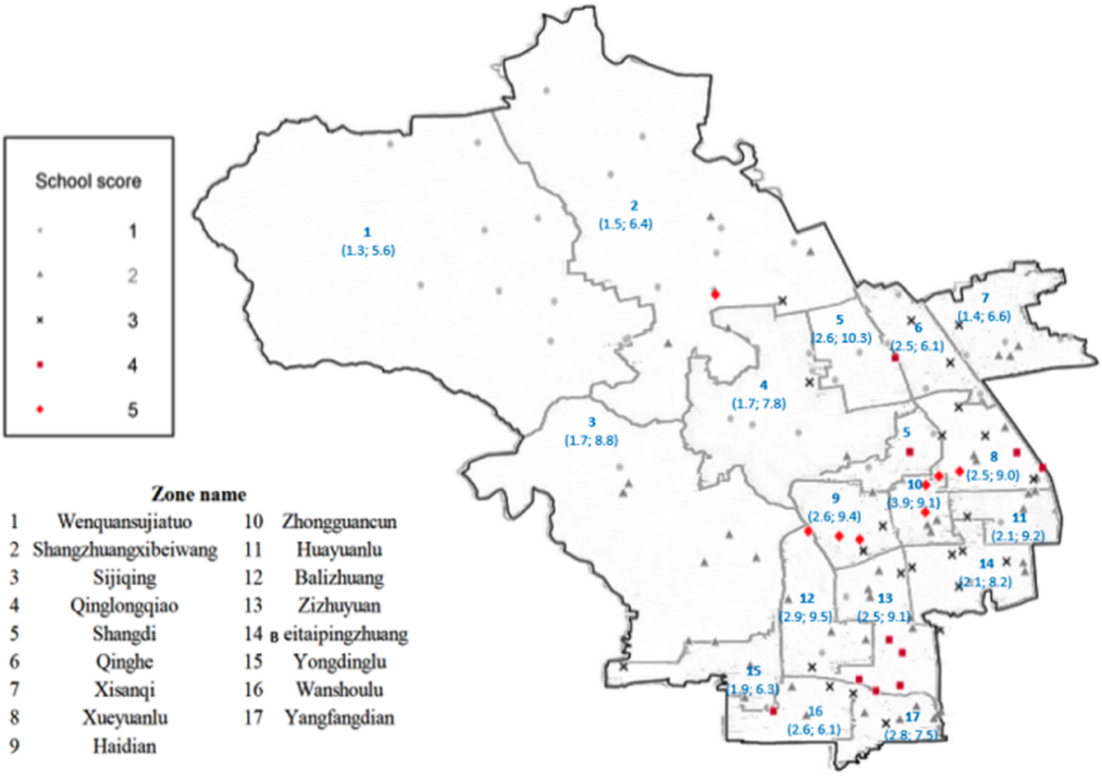


Figure A6: Haidian zoning boundary map

Notes: The graph illustrates the distribution of elementary schools across school zonings in Haidian District.