Herd behavior in sales launch: an empirical study of the Chinese residential market

Qiaoyi Ye

A thesis in

The John Molson School of Business

Presented in Partial Fulfillment of the Requirements

For the Degree of Master of Science in Administration (Finance) at

Concordia University

Montreal, Quebec, Canada

May 2018

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Qiaoyi Ye

Entitled: Herd behavior in sales launch: an empirical study of the Chinese residential market

and submitted in particle fulfillment of the requirements for the degree of

Master of Science in Administration (Finance)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

____Chair

Dr. Mahesh C. Sharma

____Examiner

Dr. Ravi Mateti

____Examiner

Dr. Saif Ullah

____Supervisor

Dr. Tingyu Zhou

Approved by ____Graduate Program Director

____Dean of Faculty

Date: 14 May 2018

Abstract

Herd behavior in sales launch: an empirical study of the Chinese residential market

Qiaoyi Ye

This paper examines the herd behavior in sales launch decisions by developers in China's housing market. Using a sample of 6,930 development projects from 1997 to 2009, we find that the propensity of sales launch is positively related to the number of prior sales launches within a certain distance. This effect is more pronounced when the time interval between sales dates is shorter and when the distance is shorter. Furthermore, lead projects that are developed by reputable developers have a greater influence on later developers' decisions than non-reputable ones. These findings provide evidence of herd behavior in developers' sales launch decisions and are in line with informational and reputational herding theories.

Acknowledgments

First, I would like to express gratitude to my supervisor, Dr. Tingyu Zhou. I appreciate all the valuable advice and guidance she has given me not only for my thesis but also for my life in general. She has been a role model for me with her professional knowledge, rigorous attitude and passion for her career. I would also like to thank my committee members, Dr. Ravi Mateti and Dr. Saif Ullah, for their insightful comments and instructions.

I'm grateful to my friends, Xiaying Zhang, Yunhao Xie, Xiaguang Sun and Shufei Wu. I'm lucky to have met them in Montreal and I will always remember the good times we had together no matter where I am.

Last but not least, I'd like to give special thanks to my parents and Sonny. Thank you for your love and support.

Table of Contents

1. Introduction

Sales launch decisions are important in the real estate market. On the supply side, developers rushing to sell at the same time will cause a higher housing supply in the short term. Excess supply can drag down real estate prices. On the demand side, the “fallacy of a boom market” caused by excessive sales and marketing activities would affect market sentiment. When the investors become optimistic about the housing market, they tend to invest more in the market, pushing the housing price up. Therefore, the timing of a developer’s sales launch can influence housing price, stock volatility, and stability of housing market.

There are many factors that can influence a developer’s sales launch strategy. Some developers launch public sales only after the completion of construction. With the presale system, developers often sell before construction completion, and even before construction commencement. It gives developers more operating capital during development and better flexibility to respond to the market. However, it makes it harder for investors and policymakers to estimate the effect of sales launch. Understanding the timing of sales launch is of significant importance. Do developers make sales launch decisions solely based on the market situation and their own development plans? Or can they also be influenced by other developers?

This paper follows prior studies in strategic decisions and herd behavior. In the fields of economics and finance, herd behavior has been widely analyzed. The reasons that cause herd behavior are mainly information and reputation. For instance, Bikhchandani and Sharma (2000) find that investors with similar profit-maximizing goals and similar information tend to react similarly at the same time. Trueman (1994) finds that reputation-concerned analysts tend to release similar earnings forecasts as announced by other analysts before, even though it is not consistent with their own information. In real estate literature, herding is observed when developers make development decisions. Decoster and Strange (2012) use statistical herding to explain the causes of overbuilding. In their paper, developers learn from their predecessors. If inaccurate signals are spread, the following developers will still believe in the wrong signal and ignore their own information, thus causing overbuilding.

In this paper, we use a sample of residential development in China from 1997 to 2009. We examine the timing of developers' sales launch. The results from hazard proportional model show that there is an evident herd behavior among developers. First, given a certain time period and distance, the propensity of new sales launch is positively related to the number of prior sales launches. To be more specific, for a typical developer, an increase in the number of prior sales launches in the past 3 months within a 5-mile radius can increase the probability of his sales launch by 1.2%.

Next, we conduct a few robustness tests using different combinations of time intervals between sales launch dates and of distances between projects sites. Results show that the influence of predecessors is greater when the time interval between two public sales dates is shorter and when the distance between two project sites is shorter. Lastly, we take reputation into consideration and test whether more reputable developers have greater influence than less-reputable ones. Results show that if the predecessor is from the list of "Top 500 Real Estate Developers in China", its sales launch decisions can accelerate the follower's sales launch decisions by 6.1%. Results suggest similar conclusions for developers that are public firms.

Examining different subsamples, the results show that the herd behavior is more pronounced if the projects are located in more developed cities. The developers tend to herd more when there are with higher reputation concerns. Compared with projects developed in multiple phases, herd behavior is more significant within single-phase projects, possibly because they are faced with higher demand uncertainty.

The contribution of this paper is mainly in three aspects. First, there are few existing studies analyzing the herd behavior in sales launch decisions in the Chinese housing market except Lai et al. (2009). Second, instead of examining the developers' strategies in only one city (e.g. Tang and Wang, 2017) or several cities (e.g. Huang, 2014), it uses a large sample size with a time span of 13 years covering 49 cities, showing a more universal and reliable pattern nationwide. The Chinese housing market is also a good representative of emerging markets. Finally, as a strategic focus in China's economic development, the real estate industry is of crucial importance. Understanding the sales launch pattern of developers is very helpful for policymakers.

The remainder of the paper is organized as follows. Section 2 describes the existing literature on real estate development strategy, sales strategy, and herd behavior. Section 3 proposes the hypotheses. Section 4 introduces the methodology used in this paper. Section 5 shows the empirical results of the hypotheses and further robustness tests. Section 6 presents the conclusions.

2. Literature review

2.1 Strategic Decisions by Developers

The majority of prior studies on developers' strategic decisions focuses on development and sales. Unlike other financial markets, real estate market has poor liquidity (Shiller, 1994), more private information and information noises (Grenadier, 1999; Childs et al., 2002; Tang and Wang, 2017). On one hand, developers want to seize market opportunities when the housing demand increases; on the other hand, they do not want to overestimate market demand and end up with vacant apartments and outstanding loans. Thus, it is essential that developers make strategic development decisions and sales decisions.

One popular development strategy is to hold the development decisions until more information is observed in the housing market. Many scholars model the real estate development as a real option and call it "option to wait". For example, by examining why lots of valuable urban land was kept vacant in Los Angeles, Titman (1985) finds that under high uncertainty in real estate market, developers are better off holding their development decisions and waiting for additional information before any actions are taken. Quigg (1993) confirms the idea that "the option to wait has value" by analyzing a large sample of real estate transactions in Seattle and in his paper, the premia associated with option (time) is 6% of the land value. Somerville (2001) considers building permits in his paper and he concludes while builders may obtain their permits first, their decisions of exercising the permits change with new information, which is consistent with the “real option” theory. Bulan et al. (2009) also find significant evidence that uncertainty delays development strategy using Vancouver condominiums data. In their paper, the probability of investment drops 13 percent with one-standard deviation increase in return volatility.

One key factor that can affect the “option to wait” development strategy is competition. Williams (1993) concludes that since the supply of options is limited, the developers are imperfectly competitive. The imperfect competition among developers greatly decreases the value of options, leading to an earlier exercise of options. Grenadier (2002) argues that with competition, a typical firm cannot realize the full option premium by waiting to develop. This makes developers exercise their options sooner because the preemption of its competitor will diminish the value of options. Bulan et al. (2009) show that competition reduces the effect of volatility on development decisions. In their paper, return volatility leads to a decline in new construction, while the competition around makes the decline in construction less sensitive to volatility. Wang et al. (2016) find the same pattern by analyzing Chinese real estate market. Their empirical results show that competition will influence the effect of uncertainty on investment timing and accelerate the investment.

When it comes to sales strategies, many Asian countries and regions adopt the presale strategy. In the presale strategy, a developer can launch public sales before completion of construction, or even before commencement of construction. In this way, developers can secure buyers of uncompleted dwellings and homebuyers can get housing price security, both consumers and developers benefit from presale contracts (Edelstein et al., 2012). Besides, the equity from presales can be injected into development and reduce financing costs, solving developers’ financing constraint problem (Chan et al., 2008). Thus, the presale system has been widely applied in Asian countries and regions, especially in Hong Kong, Taiwan, Singapore, South Korea and Mainland China. It is becoming popular in North America and Europe as well.

Some scholars examine the patterns of presale strategies. For instance, Lai and Wang (1999) find the developers in Hong Kong usually presell their housing units one year before construction completion. Developers react promptly to a boom market by marketing the presale units immediately. Lai et al. (2004) consider a presale contract as a European option. The results in their paper show that a presale is superior to selling upon completion. When launching a presale is an option, developers should do so as soon as possible. Chan et al. (2008) find that in a market with nascent financing system, developers are more willing to take the advantage of presale and pursue aggressive strategies. Because in this way, the cash associated with downpayments can be invested in their projects, providing them with cost-saving efficiency.

2.2 Informational herding and reputation

Herding behavior is widely studied in economics and finance literature. The rationale behind it is that decision makers with similar information, facing similar decision choices and payoffs, tend to make similar decisions, therefore causing a behavioral convergence (Brown et al., 2006).

Information is one of the key reasons that causes herd behavior. Banerjee (1992) develops a sequential decision model to show the rationale behind herd behavior. In his theory, a typical agent tends to follow the decision made by most agents regardless of his own information. Bikhchandani and Sharma (2000) find that investors with similar profit-maximizing goals and similar information tend to react similarly at the same time. In an emerging financial market with weak reporting requirements and higher information acquisition costs, information cascade is more likely to arise. Chiang and Zheng (2010) find that herding activity exists in both advanced financial markets (except for the U.S.) and Asian markets, exists in both up and down markets, and is more profound in rising Asian markets. Information also brings "payoff externalities". Choi (1997) finds that the payoff externalities and information spillover together generate herding in the choice of technology. Jorgensen and Kirschenheiter (2003) also conclude that one's payoff can somehow depend on other decision makes' choices by analyzing risk disclosure regime. Zhang and Liu (2012) find the same pattern in microloan markets. One lender can estimate the creditworthiness by observing other lenders' decisions, therefore causing socially-correlated lending decisions.

Except for information, reputation also plays an important role in herd behavior. Scharfstein and Stein (1990) use group psychology to explain herd behavior among managers. In some cases, managers simply mimic other managers' investment decisions and ignore their own private information. It seems inefficient but it does make sense. With reputational concerns, managers don't want contrarian behavior to damage their reputation as sensible managers, so they choose what most managers choose to "share the blame". Trueman (1994) also finds that reputation-concerned analysts tend to release similar earnings forecasts as announced by other analysts before, even though it's not consistent with their private information. Kauffman and Li (2003) analyze the scenario under IT adoption. They conclude that IT managers imitate other managers' action because they believe by doing so, other people will have a positive impression on their capabilities. Therefore, they have an incentive to make decisions that are not in the interest of their firms, causing agency costs.

Herd behavior is observed in real estate markets too. Grenadier (1996) finds an abnormal construction boom during a recession (recession-induced construction boom). When the market starts to erode, developers simultaneously choose to start their buildings. He refers to the rapid succession of exercise strategies as "development cascades". The followers, even though they cannot benefit from the leading space without competition, can benefit from the information conveyed by the leader, and therefore have a better understanding of the value of their buildings with less costs. Chu and Sing (2007) find the explanation for short bursts and overbuilding under asymmetric duopoly. They conclude that when the relative price function is smaller between two developers, the preemptive threat is more critical. Decoster and Strange (2012) use statistical herding to explain overbuilding. In their paper, developers learn from their predecessors. If inaccurate signals are spread, the following developers will still believe in the wrong signal and ignore their own information, causing overbuilding. Povel et al. (2016) examine why hotels are built in booms using U.S. hotel industry data. They find that the decision to build a hotel is made under great uncertainty about future demand, so the builders rely heavily on information from other participants and peers in the industry. This finding is consistent with information-based herding explanations.

3. Hypothesis Development

Based on the existing literature, we expect to see a behavioral convergence in developers launching public sales when developers believe those developers who previously launched sales convey a positive signal about the real estate market (e.g. that there is a burst in housing demand). When the number of public sales launches within a certain time period and distance increases, an average developer tends to “follow” predecessors’ decisions and adjust his sales schedules too, trying to “catch up” with other developers. Thus, the first hypothesis is:

H1: Given a certain time period and distance, the propensity of a new sales launch is positively related to the number of prior sales launches.

Intuitively, an event that is more recent and within a shorter distance may have greater influence. Because the long-term housing demand and market situation probably do not significantly affect short-term changes in demand for a new project, the timing of decision-making is important. Besides, neighboring projects that share similar geographic and cultural characteristics also likely face similar market demand. So the second hypothesis is as follows:

H2: The positive relation in H1 is larger when the time period is shorter and the distance is shorter.

Though the incentives for managers to herd are similar, the value of the information spillover and the influence of actions may be different among firms. Shiller (1995) finds that herd behavior differs across groups. If the information comes from someone reliable, then people tend to accept and spread the information more quickly. Generally speaking, a more reputable firm has a greater influence on the market and gains more trust on the market. Therefore, we expect to see that a typical developer believes the information conveyed by a more reputable firm is more valuable and tends to herd more according to the sales actions from more reputational firms. Thus:

H3: The positive relation in H1 is larger when the prior projects are from reputable developers.

4. Methodology

We use survival analysis to test the above hypothesis on the timing of sales launches. Survival analysis examines the probability that an event happens, which in this paper, is the probability that a new real estate project launches public sales. Compared to OLS, survival analysis has several advantages. First, it can correctly incorporate information from both censored and uncensored data to estimate the parameters of the model. Hence it will not cause a sample selection bias. Second, it does not require an assumption that the observations are normally distributed, which is almost impossible for the outcome variable, time. Third, it can handle time-varying data that change values during the observation period. Lastly, it captures the sequential effect, which is perfect to explain the herd behavior in this paper (Cleves et al., 2004).

4.1 Survival function

In survival analysis, the outcome variable is the time until an event of interest happens. Let T be a non-negative random continuous time until the occurrence of an event. If an event happens during time t, it is called a "failure", otherwise it is said to "survive". Given the notions above, the survival function is defined as:

$$S(t) = Pr(t \leq T)$$

measuring the probability that an event does not happen (survives) during time t. Here in this paper, the survival function shows the probability that a typical real estate project does not launch public sales during time t. Accordingly, the failure function is defined as:

$$F(t) = Pr(T < t) = 1 - S(t) = \int_0^t f(t)dt$$

where $f(t)$ is its probability density function of $F(t)$. It measures the probability that an event happens (fails) during time t. In this paper, the failure function shows the probability that a real estate project launches public sales during time t.

4.2 Cox proportional hazards model

A survival function analyzes the probability an event happens, while a hazard model measures the potential that an event will happen based on the condition that it has survived up to the specific time t. The hazard model is defined as:

$$h(t) = \lim_{\Delta t \to 0} \frac{(t \leq T < t + \Delta t | T \geq t)}{\Delta t}$$

where $h(t)$ is the hazard rate of the event. It is a conditional probability that the event occurs between interval [t, $t + \Delta t$] given that it has not occurred before. In this paper, the hazard rate of a real estate project at time t means the probability that it launches public sales given that it has not launched public sales until time t.

In 1972, Dr. Cox introduced the proportional hazard model. In this model, the hazard rate is defined as:

$$h(t|X_i) = h_0(t)exp\{X_i\beta\}$$

where $h_0(t)$ is called the “baseline hazard function”, X_i is the vector of covariates and β is the conformable vector of parameters. In this model, the hazard rate of an event at time t is the product of the baseline hazard function and the exponential function of the linear combination of the covariates.

The benefit of the Cox proportional hazard model is that it is a semi-parametric model that imposes no assumption on the shape of the baseline hazard function. It also makes the interpretation of the parameters more straight-forward. The exponential of parameter β_i represents the expected change in the hazard ratio (the ratio of two hazard rates) given there is a one unit change in variable X_i, holding all other predictors constant. Thus the effect of the predictors is independent from time t and is proportional over time. So in this paper, by interpreting each parameter β_i, we can capture the effect it has on the propensity for a real estate project to launch public sales and the extent to which it affects the propensity.

5. Data and variable construction

In this paper, we focus on the Chinese real estate market. China has experienced a surging period that is still continuing today. Taking China as an example can provide a better understanding of developers' decision-making behavior in an emerging market. The data is obtained from the GTA CSMAR database. The initial dataset comprises construction and transaction information of 29,068 real estate projects in 70 major cities from August 1990 to November 2010. Variables include construction date, completion date, public sales date, project location, selling price, and project-level characteristics (property type, plot ratio, gross floor area, etc.).

The requirements for presales differ among countries and regions. According to the Urban Real Estate Administration Law of the People's Republic of China, as of January 1, 1995, a real estate developer is allowed to take an advanced sale (presale) only when the "funds put for construction of the houses for advanced sale have exceeded 25% of the total budgetary investment for the project". For a real estate developer in China, the cost of real estate development includes the cost of land requisition, expenses for pre-construction engineering, construction and installation, infrastructural projects and supplementary public utilities, and indirect project expenses. Usually, the cost of land requisition accounts for 20% of the total cost and the pre-construction engineering expenses account for 6% of the total cost.

Therefore, in this paper, we assume a real estate developer is legally allowed to launch public sales as long as the construction starts.

The outcome variable is the time a project takes from construction commencement to the launching of public sales. Projects with missing public sales dates are excluded. There are quite a few observations with completion dates but without commencement dates. To preserve number of observations, we fill up the missing values with the median estimated based on projects with the same completion year. We also exclude data from Hong Kong because of the social system differences with mainland China.

We exclude projects which are for office-use only and commercial-use only, while including residential projects and mixed-use development. Projects with a sales date before construction commencement date are excluded. We also exclude 15 projects with sales dates before 1997 (2 projects in 1993, 6 projects in 1995 and 7 projects in 1996) because the annual sample size is too small, which may cause an outlier problem.

After data cleaning, we are left with a full sample of 6,930 projects in 49 cities from January 1997 to December 2009. Figure 1 shows the number of projects and the average selling price of the projects in this paper on an annual basis.

[Insert Figure 1 here]

In 1997, China's real estate industry was in a recession because of the Asian financial crisis. To stimulate real estate economy, in 1998, the central government declared "The Notice of State Council on Further Deepening the Reform to Urban Housing System and Speeding up the Housing Construction". It marks the end of the welfare system and the beginning of the commoditization of residential real estate. Since 1998, the real estate sector has been identified as a strategic focus of China's economic development (Xu and Chen, 2012) and China entered an era of housing market boom (Chen and Wen, 2014). Figure 1 shows that most of the commencement of sales in this paper happened between 1999 and 2008, thus the data can capture the features of China's real estate market as a surging market. Though the annual housing price in the full sample experienced a decline from 1997 to 2000 and another modest decline from 2004 to 2007, the overall trend of the housing price is increasing.

We define the outcome variable as the number of months a project takes from construction commencement to the launching of public sales, denoted by *TIME_TO_SALE*. The test variables and key control variables are defined as follows:

Herd behavior. To test Hypothesis 1, we examine whether the propensity for a new project to launch public sales is influenced by the number of nearby projects that launched public sales before it. In other words, when it is a time period where many real estate managers are deciding to sell their projects, we examine whether the manager of a typical project will "follow" them. The herd behavior is measured by the variable *SALES_COUNT* which defines the number of projects that launch public sales before project i in the previous 3 months and within a 5-mile radius.

Reputation. To test hypothesis 3, we distinguish a general real estate developer from a reputable real estate developer. Based on the variable *SALES_COUNT*, we create the variable *TOP500_COUNT* which defines the number of developers that are on the list of "Top 500 Real Estate Developers in China" according to the China Real Estate Association. We also create the variable *PUBLIC_COUNT* which defines the number of developers that are public firms. To be more specific, *TOP500_COUNT* represents the number of projects from the top 500 real estate developers that launch public sales before project i in the previous 3 months and within a 5-mile radius; *PUBLIC_COUNT* represents the number of projects from public firms that launch public sales before project i in the previous 3 months and within a 5-mile radius. In this paper, we assume that a public real estate developer is more reputable and has greater influence than a top 500 real estate developer.

Information transparency. CBNweekly classifies Chinese cities into different tiers based on their commercial resources clustering, activities of citizens, the varieties of people's lifestyles, and future development potential. A higher-tier city means it is more developed, is more attractive to business, and has a more active and competitive market. In a higher-tier city, due to the higher levels of market competition, there is also higher information asymmetry and transparency. We create a dummy variable *TIER_1_CITY*, which is coded as “1” if the project is located in either a first-tier city or a new first-tier city, “0” otherwise.

Appreciation. Following Bulan et al. (2009), we use the change in housing price to control for expected price appreciation. Instead of using the change rate of one-quarter-ahead expected housing price and current housing price, we calculate the appreciation rate by dividing the average selling price of a commercial house in the current year by the average selling price of a commercial house in the previous year, and then subtracting one. The housing price data is obtained from National Bureau of Statistics of China. It is on a nationwide-annual basis and is adjusted for inflation using Consumer Price Index (2000=100). The variable is denoted as *APPRECIATION* and it is expressed as a percentage.

Price level. Following Wang et al. (2016), we use the housing price as a control for market situation. This is necessary because the data in this paper has a long time span (13 years) covering 49 cities. The variable is calculated as the median of the average selling price for each project on a citywide-annual basis, denoted as *PRICE_LEVEL*. The price is adjusted for inflation using Consumer Price Index (2000=100).

Other controls. We include several hedonic variables to control for different characteristics among different projects. *DOMESTIC* controls for the sales areas and is equal to “1” if the project is sold domestically only, “0” otherwise. *DECORATION* controls for the decoration condition of the project and is equal to “1” if the apartments in the project are decorated, ”0” otherwise. *MIXED_BUILDING* controls for the building type of the project and is equal to “1” if the project contains more than one type of building (e.g. low-rise building and high-rise building), “0” otherwise. *PHASED* controls for the phasing strategy of a project and is equal to “1” if the project is phased, “0” otherwise. *VILLA* controls for the property type of the project and is equal to “1” if the project is a villa or townhouse (luxury properties), “0” otherwise. *PLOT_RATIO* is calculated as the gross floor area (GFA) divided by project plot area. *GREEN_RATIO* represents the ratio of public green space to the plot area of the project. *GFA* represents the gross floor area of a project.

Given the above variables, the baseline model in this paper can be written as:

$$h(t|X_i) = h_0(t)exp\{\beta_1 SALES\_COUNT_i + \beta_2 TIER\_1\_CITY_i + \beta_3 APPRECIATION_i$$
$$+\beta_4 PRICE\_LEVEL_i + \beta_5 DOMESTIC_i + \beta_6 DECORATION_i + \beta_7 MIXED\_BUILDING_i$$
$$+\beta_8 PHASED_i + \beta_{10} VILLA_i + \beta_{11} PLOT\_RATIO_i + \beta_{12} GREEN\_RATIO_i + \beta_{13} GFA_i\}$$

Figure 2 shows the number of projects and average selling price on a citywide basis (in order of tier 1 cities, new tier 1 cities and other cities).

[Insert Figure 2 here]

Figure 2 shows that tier 1 cities (Beijing, Shanghai, Shenzhen, Guangzhou) account for 51% of the full sample size. Because tier 1 cities are the most developed cities are in mainland China, the average housing price of tier 1 cities (7282 yuan/sq.m) is significantly higher than that of new tier 1 cities' (5500 yuan/sq.m) and other cities' (3890 yuan/sq.m). This simply suggests a positive relationship between housing prices and the development level of the cities.

Because the duration model prohibits multiple events by the same developer within the same time interval, in this paper each project is treated as an independent event. It is assumed that decisions to launch public sales are made by different managers, especially for projects from the same developer and for different phases of the same project. To mitigate the effect of outliers, we winsorize all the continuous variables at the 1% and 99% levels.

Table 1 summarizes the definitions of variables outlined above. The explanatory variables are classified into three categories: (i) herd parameter testing the herd behavior, (ii) market characteristics including city level and housing price, (iii) hedonic variables for the projects.

[Insert Table 1 here]

6. Empirical results

6.1 Summary statistics

Table 2 presents the summary statistics. In the full sample, the average time that a project takes from the commencement of construction to the launch of public sales is 12 months, with a median time of 8 months. We further plot a more visualized survival probability curve in Figure 3. It shows that half of the projects take less than 8 months to launch public sales, and 75% of them take less than 14 months. The shortest time to sale is only 1 month and the longest time to sale is 169 months. The slope of the survival probability curve indicates a significant cluster of public sales decisions not long after the beginning of construction, while it gets more dispersive as time goes on.

[Insert Table 2 here]

[Insert Figure 3 here]

For a typical project, the average number of projects that launch public sales 3 months prior to it and within a 5-mile radius is nearly 10, while the average number of projects developed by top 500 developers drops to 0.6 and the average number of projects developed by public developers further drops to 0.5. In the full sample, 80% of the projects are located in either tier 1 or new tier 1 city, indicating a hotter property market in more developed cities.

Table 3 reports the correlation matrix among all the variables used in this paper. The results show a moderate positive linear correlation between *TIER_1_CITY* and *PRICE_LEVEL*. This is because *PRICE_LEVEL* is created on a citywide-annual basis, and since *TIER_1_CITY* differentiates the development level of cities, a more developed city is intuitively associated with a higher housing price. In addition, there is also a strong linear correlation between the three herd parameters (*SALES_COUNT*, *TOP500_COUNT* and *PUBLIC_COUNT*) but these variables will be tested individually in different regressions. Other absolute values of pairwise correlation coefficients are all under 0.3, suggesting there is little concern on multicollinearity.

[Insert Table 3 here]

6.2 Baseline specification: Hypothesis 1

Table 4 shows the empirical results for hypothesis 1. Recall that in the proportional hazard model, the effect of an estimated coefficient on the hazard rate is exponential. Therefore for each estimated coefficient β, the effect on hazard rate is e^{β}, holding all else constant. A positive β means an increase in the variable has a positive impact on the hazard rate, which means it accelerates the probability that an event happens; a negative β means an increase in the variable delays the probability that an event happens.

[Insert Table 4 here]

We first test the impact of herd parameter in column (1). Next, we add market characteristics in column (2), then all hedonic variables together in column (3). Column (1)-(3) are based on our baseline model. In all the three regressions, *SALES_COUNT* has a significant positive effect on the time a project takes from construction commencement to launching public sales. The results in baseline model show that the likelihood that a project launches public sales increases by 1.2% with one more project launching public sales prior to it within 3 months and within a 5-mile radius. This result is consistent with the idea that competition accelerates development decisions (Bulan et al., 2009; Wang et al., 2016). Furthermore, the significant positive coefficient of *TIER_1_CITY* shows that in a more developed city (tier 1 city or new tier 1 city), developers choose to sell their projects 28% faster than those in less developed cities.

This result is in line with the "informational herding" theory. When the city is more developed, the competition there is fiercer, thus increasing the transparency of information. For a typical developer, the "option to wait" cannot provide much value given that the developer already benefits from the information spillover and informational externalities. On the contrary, for a typical developer in less developed cities, where the competition is less fierce and the information in market is less transparent, the “option to wait” provides more value when dealing with market uncertainty.

Lastly, the significant positive coefficient of *APPRECIATION* indicates that a one percentage point increase in housing price in the previous year accelerates the decisions to launch public sales by 79%. Rising prices can provide capital gains for developers to overcome liquidity constraints, so the expected housing price appreciation might lead to a higher hazard rate (Bulan et al., 2009).

6.3 Herding under different time period and distance: Hypothesis 2

To test hypothesis 2, we modify the construction of herd parameter *SALES_COUNT* to see if herd behavior is more pronounced when the distance between a typical project and the project that launched public sales prior to it is shorter and when the time period between the two public sales dates is shorter. Following Bulan et al. (2009), we reconstruct *SALES_COUNT* with different distances and time periods. Recall that in the baseline model, *SALES_COUNT* is defined as the number of projects that launch public sales before project i in the previous 3 months and within a 5-mile radius. We create three new *SALES_COUNT* parameters with 3 months and 10-mile radius, 6 months and 5-mile radius, and 6 months and 10-mile radius. The results are shown in Table 5.

[Insert Table 5 here]

In column (2), the effect of *SALES_COUNT* is still positive and significant, but due to the greater distance between projects, the extent of the effect is reduced. The likelihood that a project launches public sales drops from 1.2% to 0.7% with one more project launching public sales within a 10-mile radius. Similarly, the likelihood that a project launches public sales drops from 1.2% to 0.7% with one more project launching public sales 6 months prior to it, as shown in column (3). Not surprisingly, the coefficient in column (4) shows the least herding effect. When the time period is extended to 6 months and the distance is extended to 10 miles, the increase of one project taking public sales actions can only lead to a 0.4%

higher probability that other projects will also start to sell. These results strongly support hypothesis 2.

6.4 Reputation: Hypothesis 3

Because people tend to accept and spread the information more quickly if it comes from someone trustful (Shiller, 1995), we test hypothesis 3 by using herd parameters representing different levels of reputation (*TOP500_COUNT*, *PUBLIC_COUNT*). In the full sample, there are 458 projects that are developed by developers from the “Top 500” list, and 397 projects that are developed by public developers. Intuitively, a public developer has a higher reputation and greater influence than a “Top 500” developer because of the higher standard they need to meet to be public. The results are shown in Table 6.

[Insert Table 6 here]

In column (2), compared with baseline model, *TOP500_COUNT* has a significantly greater effect on the propensity that a project will launch public sales. The action of public sales is accelerated by 6.1% with one more “Top 500” project launching public sales in the past 3 months within 5 miles. When it comes to public developers, the effect is even greater. In column (3), an increase in *PUBLIC_SALES* leads to a 6.5% increase in the probability of public sales actions. These results strongly support hypothesis 3 that herd behavior is more pronounced when the lead developers are reputable.

6.5 Robustness tests

6.5.1 Herd behavior across cities

In the baseline model, we differentiate the level of cities by adding a dummy variable *TIER_1_CITY*, which captures the difference in the value of “option to wait” in developed cities versus less developed cities. To further test the herd behavior across cities, we create a sub-sample containing projects located only in “tier 1 cities” (Beijing, Shanghai, Shenzhen, Guangzhou). We also create another sub-sample containing projects located in Beijing because it is the capital city and because housing prices in Beijing have always represented the highest housing price level in China. Most importantly, neighborhood areas in Beijing are clearly separated by the "ring roads" which depict the urbanization process. The results are shown in Table 7.

[Insert Table 7 here]

The coefficient of *SALES_COUNT* in the first column suggests that in tier 1 cities, the propensity of launching public sales for a project is increased by 1.3% with one more project launching public sales, which is higher than in the baseline model. In the second column, we add a categorical variable *RING* to control for the geographical position of the project in Beijing. Beijing is one of the few cities that possess multiple ring roads, with a lower ring road that is closer to the city center. In this paper, we define *RING* = i as the project is located between the $i^{th}$and $(i+1)^{th}$ ring roads. The coefficient of *SALES_COUNT* still shows a strong herd behavior and the herd effect is greater than that of tier 1 cities. In Beijing, the propensity of a project taking sales action is increased by 1.7% when there is one more project taking sales action 3 months prior to it within 5 miles. The changes in the coefficients of *SALES_COUNT* suggest that in a more competitive market, the herd behavior of launching public sales is more pronounced.

For coefficients of variable *RING*, the projects beyond the 6^{th} ring road (*RING* = 6) are the benchmark. The coefficients show that the value of "option to wait" is significantly higher for projects located within the 6^{th}ring road than beyond the 6^{th} ring road, but there is no evident pattern across the ring roads. Wang et al. (2016) find that higher land prices accelerate development, while in this paper, the result is opposite. Projects within the 6^{th} ring road tend to delay public sales compared with projects beyond the 6^{th} ring road. The underlying reason may be that because Beijing is the capital of China, stricter land-use regulation causes development decisions to take longer for projects closer to city center. Besides, due to high housing prices, few people can afford a new house close to city center, thus making the suburban housing market more popular.

6.5.2 Herding with different reputation concerns

In reputational herding theory, decision makers tend to mimic others' decisions because being consistent with the majority makes them feel more sensible. However, the extent of reputational herding varies among decision makers with different reputation concerns. Scharfstein and Stein (1990) find that less skilled managers have higher reputation concerns and benefit more from herding. Brown et al. (2006) also conclude that less reputable managers exhibit a greater tendency to herd. Therefore, we divide the full sample into two separate sub-samples containing projects developed by public developers and projects developed by non-public developers. We then test the effect of *PUBLIC_COUNT* on time to launch public sales in each sub-sample to see whether the effect is different between two groups. The results are shown in Table 8.

[Insert Table 8 here]

The coefficient in the second column shows that for projects developed by non-public developers, the propensity for them to launch public sales increases by 1.7% when there is one more project where a public developer has launched public sales 3 months prior and within 5 miles. However, for projects developed by public developers, the effect of *PUBLIC_COUNT* is insignificant. This finding is consistent with the idea that the herd behavior is more pronounced for less reputable decision makers than reputable decision makers.

6.5.3 Single-phased development versus development with multiple phases

Except for delaying development and preselling projects, development by phases (phasing strategy) is also commonly used by real estate developers to deal with market uncertainty. When faced with market demand uncertainty, developers tend to lower the price in earlier units to make sure there is sufficient demand, then sequentially increase the price in later phases when the market demand is clearer (Lai et al., 2004). Tang and Wang (2017) find a significant acceleration effect of competition on projects that are single-phase, while the effect is not significant for phased projects as a whole. To test if the herding effect is different between projects adopting a phasing strategy and projects without a phasing strategy, we create two sub-samples containing phased and single-phase projects separately. The results are shown in Table 9.

[Insert Table 9 here]

In the multi-phased projects group, the coefficient of *SALES_COUNT* does not have a significant effect on the developers' selling decisions, whereas in the single-phase projects group, an increase in *SALES_COUNT* leads to a 1.3% rise in the propensity that a project takes a selling action. The results are in line with Tang and Wang's findings. For phased projects, the demand uncertainty has been largely reduced by selling the first units, so the information spillover does not have much benefit. However, for single-phase projects, the demand uncertainty still exists, so developers have higher incentives to follow others' decisions, and therefore the herding effect is more pronounced.

7. Conclusion

This paper explores herd behavior in developers' public launch decisions in China's housing market. The empirical results strongly support the hypotheses and provide new evidence for existing theories. First, a sales launch decision can be accelerated by the number of previous sales launches. The effect is more pronounced when the time interval between two sales dates is shorter and when the distance between two project sites is shorter. In addition, developers have a higher tendency to herd when the lead developers are reputable. These findings are in line with the theory of informational herding and reputational herding, they also show that competition can accelerate developers' sales launch decisions.

The herding patterns of different groups are examined in further robustness tests. Results show that herd behavior is more evident in more competitive markets, among less reputable developers, and in projects without a phasing strategy.

This paper also has some implications from a policy perspective. In China, real estate sector has been a strategic focus in economic development since 1998 and the government usually implements short-term policies to deal with temporary fluctuations in the housing market. The findings in this paper can help policymakers have a better understanding of the herd behavior in sales launch decisions among developers. This enables policymakers to implement more efficient policy measures to help maintain a healthy and stable housing market.

One limitation of this paper is that the data is more concentrated in higher-tier cities due to the nature of the existing data set. In addition, the effect of factors on developers' decisions to launch public sales are examined individually. In the future, it would be interesting to test the combined effect of factors and herd behavior in lower-tier cities. It is also worth investigating the effect of policies on public launch decisions.

References

Banerjee, A. V. (1992). A simple model of herd behavior. *The quarterly journal of economics*, *107*(3), 797-817.

Bikhchandani, S., & Sharma, S. (2000). Herd behavior in financial markets. *IMF Staff papers*, *47*(3), 279-310.

Brown, N., Gordon, L., & Wermers, R. (2006). Herd behavior in voluntary disclosure decisions: An examination of capital expenditure forecasts.

Bulan, L., Mayer, C., & Somerville, C. T. (2009). Irreversible investment, real options, and competition: Evidence from real estate development. *Journal of Urban Economics*, *65*(3), 237-251.

Chan, S. H., Fang, F., Yang, J., 2014. Presales, leverage decisions, and risk shifting. Journal of Real Estate Research 36, 475-509.

Chan, S., Fang, F., & Yang, J. (2008). Presales, financing constraints, and developers' production decisions. *Journal of Real Estate Research*, *30*(3), 345-376.

Chau, K. W., Wong, S. K., & Yiu, C. Y. (2007). Housing quality in the forward contracts market. *The Journal of Real Estate Finance and Economics*, *34*(3), 313-325.

Chen, K., & Wen, Y. (2014). The great housing boom of China. Working paper, *Federal Reserve Bank of St. Louis*.

Chiang, T. C., & Zheng, D. (2010). An empirical analysis of herd behavior in global stock markets. *Journal of Banking & Finance*, *34*(8), 1911-1921.

Childs, P. D., Ott, S. H., & Riddiough, T. J. (2002). Optimal valuation of claims on noisy real assets: Theory and an application. *Real estate economics*, *30*(3), 415-443.

Choi, J. P. (1997). Herd behavior, the" penguin effect," and the suppression of informational diffusion: an analysis of informational externalities and payoff interdependency. *The Rand Journal of Economics*, 407-425.

Chu, Y., & Sing, T. F. (2007). Optimal timing of real estate investment under an asymmetric duopoly. *The Journal of Real Estate Finance and Economics*, *34*(3), 327-345.

Cleves, M., Gould, W. W., Gutiérrez, R. G., & Marchenko, Y. (2004). Introduction to survival data analysis with Stata, revised edition.

Cox, D. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological), 34*(2), 187-220.

Cunningham, C., 2006. House price uncertainty, timing of development, and vacant land prices: evidence for real options in Seattle. Journal of Urban Economics 59: 1–31.

DeCoster, G. P., & Strange, W. C. (2012). Developers, herding, and overbuilding. *The Journal of Real Estate Finance and Economics*, *44*(1-2), 7-35.

Despa, S. (2010). What is survival analysis. *Cornell University Statistical Consulting Unit*.

Edelstein, R., Liu, P., & Wu, F. (2012). The market for real estate presales: a theoretical approach. *The Journal of Real Estate Finance and Economics*, *45*(1), 30-48.

Grenadier, S. R. (1996). The strategic exercise of options: Development cascades and overbuilding in real estate markets. *The Journal of Finance*, *51*(5), 1653-1679.

Grenadier, S. R. (1999). Information revelation through option exercise. *The Review of Financial Studies*, *12*(1), 95-129.

Grenadier, S. R. (2002). Option exercise games: An application to the equilibrium investment strategies of firms. *The Review of Financial Studies*, *15*(3), 691-721.

Haurin, D. (1988). The duration of marketing time of residential housing. *Real Estate Economics*, *16*(4), 396-410.

Huang, Y. (2004). Housing markets, government behaviors, and housing choice: a case study of three cities in China. *Environment and Planning A*, *36*(1), 45-68.

Jorgensen, B. N., & Kirschenheiter, M. T. (2003). Discretionary risk disclosures. *The Accounting Review*, *78*(2), 449-469.

Kauffman, R. J., & Li, X. (2003, October). Payoff externalities, informational cascades and managerial incentives: a theoretical framework for IT adoption herding. In *Proceedings of the 2003 INFORMS Conference on IS and Technology, Atlanta, GA, October 2003*.

Lai, N., & Wang, K. (1999). Land-supply restrictions, developer strategies and housing policies: the case in Hong Kong. *International Real Estate Review*, *2*(1), 143-159.

Lai, R. N., Wang, K., & Zhou, Y. (2004). Sale before completion of development: pricing and strategy. *Real Estate Economics*, *32*(2), 329-357.

Lai, R. N., Zhang, Y., & Zhou, T. (2009) Observational influence on sales launch behavior of residential developersxxx . Working paper..

Lu, C. L., Liao, W. C., & Peng, C. W. (2016). Developer's Perspectives on Timing to Build: Evidence from Microdata of Land Acquisition and Development.

Povel, P., Sertsios, G., Kosová, R., & Kumar, P. (2016). Boom and gloom. *The Journal of Finance*, *71*(5), 2287-2332.

Quigg, L. (1993). Empirical testing of real option-pricing models. *The Journal of Finance*, *48*(2), 621-640.

Scharfstein, D. S., & Stein, J. C. (1990). Herd behavior and investment. *The American Economic Review*, 465-479.

Shiller, R. J. (1994). *Macro markets: creating institutions for managing society's largest economic risks*. OUP Oxford.

Shiller, R. J. (1995). Conversation, information, and herd behavior. *The American Economic Review*, *85*(2), 181-185.

Somerville, C. T. (2001). Permits, starts, and completions: structural relationships versus real options. *Real Estate Economics*, *29*(1), 161-190.

Tang, W., & Wang, Y. (2017). Incomplete information and real estate development strategy: Evidence from Hangzhou, China. *Habitat International,63*, 1-10.

Titman, S. (1985). Urban land prices under uncertainty. *The American Economic Review*, *75*(3), 505-514.

Trueman, B. (1994). Analyst forecasts and herding behavior. *The Review of Financial Studies*, *7*(1), 97-124.

Wang, Y., Tang, W., & Jia, S. (2016). Uncertainty, Competition and Timing of Land Development: Theory and Empirical Evidence from Hangzhou, China. *The Journal of Real Estate Finance and Economics*, *53*(2), 218-245.

Williams, J. T. (1993). Equilibrium and options on real assets. *The Review of Financial Studies*, *6*(4), 825-850.

Williams, J. T., 1991. Real estate development as an option. Journal of Real Estate Finance and Economics 4(2), 191-208.

Xu, X. E., & Chen, T. (2012). The effect of monetary policy on real estate price growth in China. *Pacific-Basin Finance Journal*, *20*(1), 62-77.

Zhang, J., & Liu, P. (2012). Rational herding in microloan markets. *Management science*, *58*(5), 892-912.

Appendix

Figure 1: Number and average price of projects on an annual basis

This figure shows the number of projects and the average price of projects in the full sample (1997-2009) on an annual basis. The average price is adjusted for inflation (2000=100).

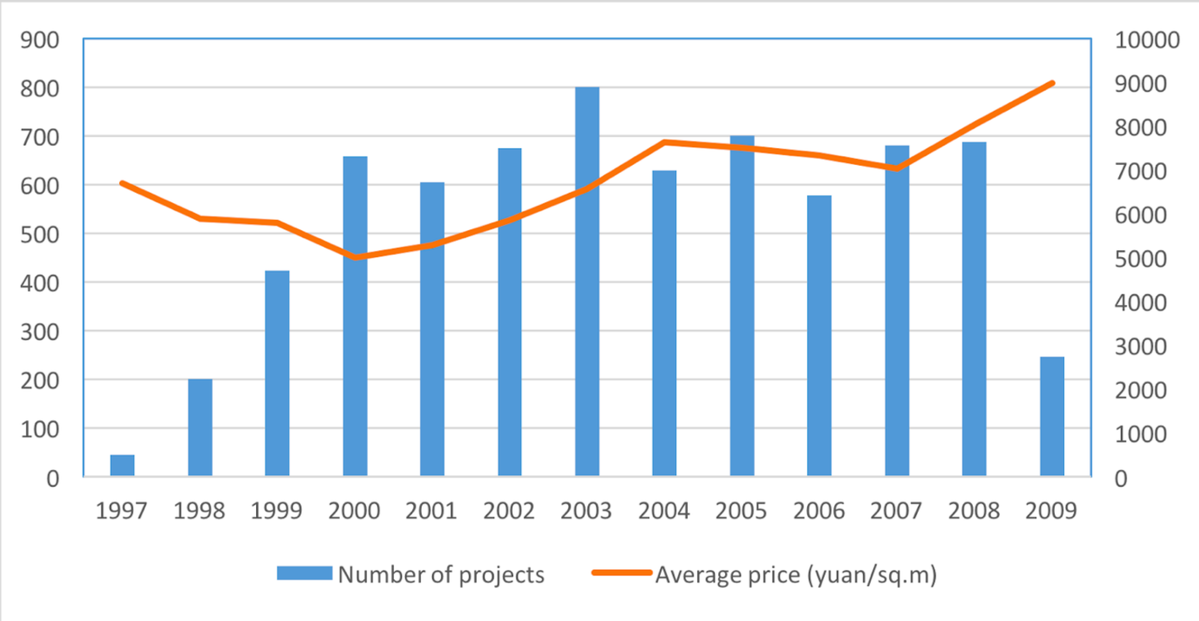

Figure 2: Number and average price of projects on a citywide basis

This figure shows the number of projects and the average price of projects in the full sample (1997-2009) on a citywide basis. The first four cities are tier 1 cities, then followed by new tier 1 cities and other cities. The average price is adjusted for inflation (2000=100).

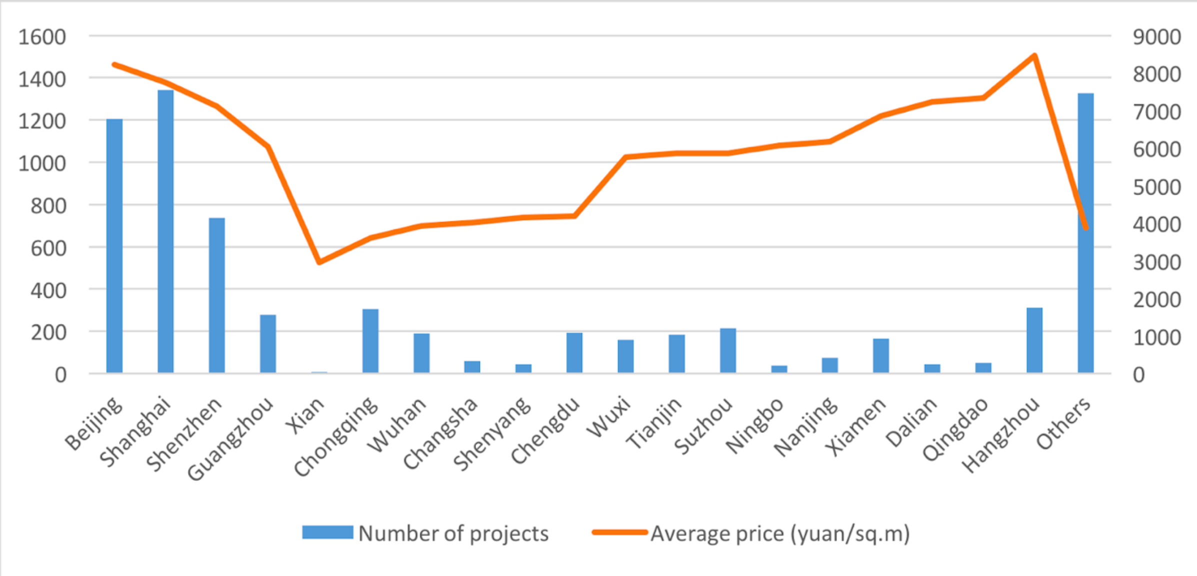

Figure 3: Survival probability curve of time to launch public sales

This figure shows the survival probability curve of time to launch public sales. Time to launch public sales is the period from commencement of construction to the action of launching public sales.

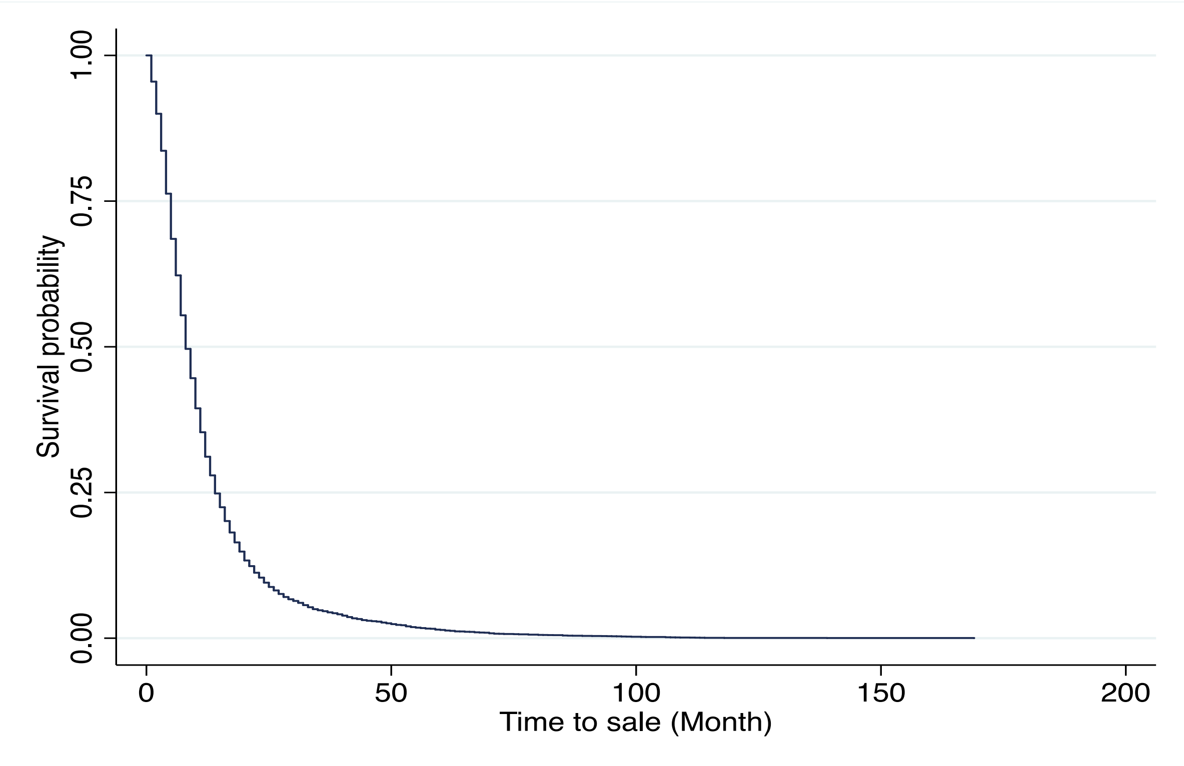

Table 1
Variable definitions

Variable	Definition
Outcome variable	
TIME_TO_SALE	The time that a project takes from construction commencement to public sales (in months)
Herd parameter	
SALES_COUNT	Number of same-city projects that launches public sales in previous 3 months within a 5-mile radius for project i
TOP500_COUNT	Number of same-city projects developed by top 500 developers that launches public sales in previous 3 months within a 5-mile radius for project i
PUBLIC_COUNT	Number of same-city projects developed by public developers that launches public sales in previous 3 months within a 5-mile radius for project i
Market characteristics	
TIER_1_CITY	Dummy, =1 if the project is located in either first-tier or new first-tier city
APPRECIATION	The annual average commercial houses price change compared to previous year on a national basis (%)
PRICE_LEVEL	The average commercial houses price on a year-city basis (in logarithm)
Hedonic variables	
DOMESTIC	Dummy, =1 if the project is sold in domestic market only
DECORATION	Dummy, =1 if the project is decorated
MIXED_BUILDING	Dummy, =1 if the project contains at least two types of buildings (eg. villa and low-rise building, medium-rise building and high-rise building)
PHASED	Dummy, =1 if the project is phased
VILLA	Dummy, =1 if the property type is villa
PLOT_RATIO	The plot ratio of the project, defined as the gross floor area (GFA) of the project divided by the area of the plot
GREEN_RATIO	Ratio of green space, defined as the ratio of public green space to the plot area of the property (%)
GFA	The total floor area contained within the building (in logarithm)

Table 2
Summary statistics

This table reports the summary statistics for all variables in the full sample. Variables *APPRECIATION* and *PRICE_LEVEL* are adjusted for inflation (2000=100). All continuous variables are winsorized at the 1% and 99% levels.

Variables	Obs	Mean	Median	25th percentile	75th percentile	Std. Dev
TIME_TO_SALE	6930	12	8	5	14	13
SALES_COUNT	6930	9.61	6	3	13	10.3
TOP500_COUNT	6930	0.592	0	0	1	1.01
PUBLIC_COUNT	6930	0.516	0	0	1	0.925
TIER_1_CITY	6930	0.807	1	1	1	0.395
APPRECIATION	6930	0.054	0.045	0.025	0.095	0.066
PRICE_LEVEL	6925	8.65	8.7	8.34	8.98	0.419
DOMESTIC	6930	0.824	1	1	1	0.381
DECORATION	6930	0.287	0	0	1	0.452
MIXED_BUILDING	6930	0.235	0	0	0	0.424
PHASED	6930	0.071	0	0	0	0.257
VILLA	6930	0.049	0	0	0	0.216
PLOT_RATIO	5415	3.12	2.48	1.6	3.8	2.33
GREEN_RATIO	5713	38	46.2	30	41.5	11
GFA	6263	11.2	11.1	10.4	11.9	1.08

Table 3
Correlation matrix

This table reports the Pearson correlation matrix of all the variables in the full sample. * Significant at 5%; ** Significant at 1%; *** Significant at 0.1%.

	Time to Sale	Sales Count	Top500 Count	Public Count	Tier 1 City	Appreciation	Price Level	Domestic	Decoration	Mixed Building	Phased	Villa	Plot Ratio	Green Ratio	GFA
Time to Sale	1.000														
Sales Count	0.110***	1.000													
Top500 Count	-0.025	0.497***	1.000												
Public Count	-0.025	0.570***	0.753***	1.000											
Tier 1 City	0.004	0.160***	0.163***	0.174***	1.000										
Appreciation	-0.010	-0.089***	0.001	-0.018	0.001	1.000									
Price Level	0.166***	0.096***	0.216***	0.215***	0.492***	0.164***	1.000								
Domestic	-0.035*	0.154***	0.057***	0.101***	0.157***	-0.009	0.245***	1.000							
Decoration	0.041**	0.093***	0.102***	0.089***	0.096***	0.017	0.127***	0.044**	1.000						
Mixed Building	0.001	-0.065***	-0.014	-0.011	-0.068***	0.067***	0.014	-0.012	-0.121***	1.000					
Phased	0.054***	0.054***	0.056***	0.049***	0.050***	0.043**	0.117***	0.052***	-0.011	0.080***	1.000				
Villa	0.068***	-0.071***	-0.050***	-0.029*	0.047**	0.019	0.100***	0.028	-0.037*	-0.083***	0.004	1.000			
Plot Ratio	-0.019	0.143***	0.068***	0.054***	0.095***	-0.055***	-0.075***	-0.020	0.098***	-0.237***	0.071***	0.231***	1.000		
Green Ratio	0.041**	0.001	0.036*	0.034*	0.026	0.001	0.084***	0.022	-0.064***	0.120***	0.095***	0.281***	-0.296***	1.000	
GFA	0.173***	-0.154***	0.001	-0.022	-0.091***	0.036*	0.082***	-0.048***	-0.033*	0.239***	0.058***	0.045**	-0.134***	0.167***	1.000

Table 4
Hazard regression: Hypothesis 1

This table shows the results from the Cox proportional hazard model in the full sample. Reg. (1) tests the herd parameter. Reg. (2) tests herd parameter and market characteristics. Reg. (3) tests all variables, which is also the baseline regression. The estimated hazard model is $h(t|X_i) = h_0(t)exp\{X_i\beta\}$. Coefficients are reported in real form (β) and one unit change in X_i leads to a $(e^{\beta} - 1)$ percent change in hazard rate $h(t)$. * Significant at 5%; ** Significant at 1%; *** Significant at 0.1%.

	Reg. (1)	Reg. (2)	Reg. (3) Baseline
Sales Count	0.013***	0.014***	0.012***
	(11.51)	(12.32)	(7.54)
Tier 1 City		0.177***	0.250***
		(4.73)	(5.17)
Appreciation		0.251	0.583**
		(1.52)	(3.07)
Price Level		-0.465***	-0.541***
		(-13.41)	(-12.16)
Domestic			0.160***
			(3.88)
Decoration			-0.059
			(-1.76)
Mixed Building			0.079*
			(2.25)
Phased			-0.150**
			(-2.73)
Villa			-0.249***
			(-3.50)
Plot Ratio			-0.028***
			(-3.78)
Green Ratio			0.001
			(0.38)
GFA			-0.132***
			(-9.14)
Observations	6930	6925	4752

Table 5
Hazard regression: Hypothesis 2

This table shows the results from the Cox proportional hazard model in the full sample. Reg. (1) is the baseline regression. In Reg. (2), Reg. (3) and Reg. (4), the herd parameter *SALES_COUNT* are calculated using 3 months and 10 miles, 6 months and 5 miles, 6 months and 10 miles separately. The estimated hazard model is $h(t|X_i) = h_0(t)exp\{X_i\beta\}$. Coefficients are reported in real form (β) and one unit change in X_i leads to a ($e^{\beta} - 1$) percent change in hazard rate $h(t)$. * Significant at 5%; ** Significant at 1%; *** Significant at 0.1%.

	Reg. (1) baseline 3 months 5-mile radius	**Reg. (2) 3 months 10-mile radius**	**Reg. (3) 6 months 5-mile radius**	**Reg. (4) 6 months 10-mile radius**
Sale Count	0.012***	0.007***	0.007***	0.004***
	(7.54)	(8.66)	(7.50)	(8.82)
Tier 1 City	0.250***	0.225***	0.248***	0.223***
	(5.17)	(4.61)	(5.12)	(4.56)
Appreciation	0.583**	0.623**	0.583**	0.624**
	(3.07)	(3.26)	(3.07)	(3.26)
Price Level	-0.541***	-0.564***	-0.549***	-0.577***
	(-12.16)	(-12.57)	(-12.33)	(-12.80)
Domestic	0.160***	0.146***	0.161***	0.146***
	(3.88)	(3.54)	(3.93)	(3.53)
Decoration	-0.059	-0.071*	-0.060	-0.073*
	(-1.76)	(-2.11)	(-1.79)	(-2.14)
Mixed Building	0.079*	0.081*	0.077*	0.079*
	(2.25)	(2.30)	(2.19)	(2.24)
Phased	-0.150**	-0.155**	-0.153**	-0.165**
	(-2.73)	(-2.83)	(-2.79)	(-3.00)
Villa	-0.249***	-0.244***	-0.245***	-0.241***
	(-3.50)	(-3.44)	(-3.45)	(-3.40)
Plot Ratio	-0.028***	-0.025***	-0.028***	-0.025***
	(-3.78)	(-3.45)	(-3.81)	(-3.47)
Green Ratio	0.001	0.000	0.001	0.000
	(0.38)	(0.18)	(0.38)	(0.16)
GFA	-0.132***	-0.130***	-0.131***	-0.128***
	(-9.14)	(-8.98)	(-9.00)	(-8.86)
Observations	4752	4752	4752	4752

Table 6
Hazard regression: Hypothesis 3

This table shows the results from the Cox proportional hazard model in the full sample. Reg. (1) is the baseline regression. In Reg. (2) and Reg. (3), the herd parameter is *TOP500_COUNT* and *PUBLIC_COUNT* separately. The estimated hazard model is $h(t|X_i) = h_0(t)exp\{X_i\beta\}$. Coefficients are reported in real form (β) and one unit change in X_i leads to a $(e^{\beta} - 1)$ percent change in hazard rate $h(t)$. * Significant at 5%; ** Significant at 1%; *** Significant at 0.1%.

	Reg. (1) baseline	**Reg. (2) Top 500 developers**	**Reg. (3) Public developers**
Sales Count	0.012***		
	(7.54)		
Top500 Count		0.059***	
		(3.96)	
Public Count			0.064***
			(3.77)
Tier 1 City	0.250***	0.264***	0.258***
	(5.17)	(5.47)	(5.35)
Appreciation	0.583**	0.485**	0.487**
	(3.07)	(2.60)	(2.61)
Price Level	-0.541***	-0.557***	-0.553***
	(-12.16)	(-12.50)	(-12.45)
Domestic	0.160***	0.195***	0.190***
	(3.88)	(4.79)	(4.66)
Decoration	-0.059	-0.057	-0.056
	(-1.76)	(-1.69)	(-1.67)
Mixed Building	0.079*	0.068	0.069
	(2.25)	(1.94)	(1.94)
Phased	-0.150**	-0.123*	-0.121*
	(-2.73)	(-2.26)	(-2.21)
Villa	-0.249***	-0.282***	-0.289***
	(-3.50)	(-3.99)	(-4.08)
Plot Ratio	-0.028***	-0.024***	-0.024***
	(-3.78)	(-3.36)	(-3.33)
Green Ratio	0.001	0.001	0.001
	(0.38)	(0.76)	(0.74)
GFA	-0.132***	-0.146***	-0.144***
	(-9.14)	(-10.17)	(-10.06)
Observations	4752	4752	4752

Table 7
Hazard regression: Robustness test 1

This table shows the results from the Cox proportional hazard model in the subsamples. Reg. (1) uses data from tier 1 cities. Reg. (2) uses data from Beijing. *RING* = i is defined as the project is located between the $i^{th}$and $(i+1)^{th}$ ring roads. The estimated hazard model is $h(t|X_i) = h_0(t)exp\{X_i\beta\}$. Coefficients are reported in real form (β) and one unit change in X_i leads to a $(e^{\beta}-1)$ percent change in hazard rate $h(t)$. * Significant at 5%; ** Significant at 1%; *** Significant at 0.1%.

	Reg. (1) Tier 1 cities	Reg. (2) Beijing
Sales Count	0.013***	0.017**
	(6.91)	(3.14)
Appreciation	0.212	0.091
	(0.69)	(0.19)
Price Level	-0.663***	-0.626***
	(-7.85)	(-3.77)
Domestic	0.318**	0.293
	(2.76)	(1.95)
Decoration	-0.101*	-0.018
	(-2.31)	(-0.26)
Mixed Building	0.174***	-0.016
	(3.35)	(-0.12)
Phased	-0.089	-0.172
	(-1.35)	(-1.34)
Villa	-0.084	-0.270
	(-0.89)	(-1.64)
Plot Ratio	-0.011	-0.018
	(-1.13)	(-1.15)
Green Ratio	-0.001	-0.006
	(-0.76)	(-1.46)
GFA	-0.105***	-0.093**
	(-5.25)	(-3.08)
ring=1		-0.753***
		(-4.23)
ring=2		-0.516***
		(-3.32)
ring=3		-0.592***
		(-3.92)
ring=4		-0.570***
		(-3.92)
ring=5		-0.577***
		(-4.10)
Observations	2513	955

Table 8
Hazard regression: Robustness test 2

This table shows the results from the Cox proportional hazard model in the subsamples. Reg. (1) uses data from projects developed by public developers. Reg. (2) uses data from projects developed by non-public developers. The estimated hazard model is $h(t|X_i) = h_0(t)exp\{X_i\beta\}$. Coefficients are reported in real form (β) and one unit change in X_i leads to a $(e^{\beta} - 1)$ percent change in hazard rate $h(t)$. * Significant at 5%; ** Significant at 1%; *** Significant at 0.1%.

	Reg. (1) Public developers	**Reg. (2) Non-public developers**
Public Count	0.110	0.059***
	(1.73)	(3.38)
Tier 1 City	-0.434	0.287***
	(-1.78)	(5.79)
Appreciation	0.393	0.494*
	(0.60)	(2.52)
Price Level	-0.554**	-0.564***
	(-2.95)	(-12.28)
Domestic	0.115	0.202***
	(0.64)	(4.80)
Decoration	-0.009	-0.054
	(-0.07)	(-1.54)
Mixed Building	0.231	0.059
	(1.61)	(1.61)
Phased	0.160	-0.134*
	(0.72)	(-2.36)
Villa	-0.201	-0.290***
	(-0.68)	(-3.97)
Plot Ratio	0.052	-0.028***
	(1.40)	(-3.73)
Green Ratio	-0.001	0.001
	(-0.21)	(0.63)
GFA	-0.274***	-0.134***
	(-4.89)	(-8.82)
Observations	299	4453

Table 9
Hazard regression: Robustness test 3

This table shows the results from the Cox proportional hazard model in the subsamples. Reg. (1) uses data from phased projects. Reg. (2) uses data from single-phase projects. The estimated hazard model is $h(t|X_i) = h_0(t)exp\{X_i\beta\}$. Coefficients are reported in real form (β) and one unit change in X_i leads to a $(e^{\beta} - 1)$ percent change in hazard rate $h(t)$. * Significant at 5%; ** Significant at 1%; *** Significant at 0.1%.

	Reg. (1) Phased	Reg. (2) Single-phase
Sales Count	0.004	0.013***
	(0.80)	(7.74)
Tier 1 City	0.803***	0.213***
	(3.72)	(4.26)
Appreciation	1.833*	0.482*
	(2.32)	(2.46)
Price Level	-0.770***	-0.529***
	(-4.21)	(-11.51)
Domestic	-0.015	0.165***
	(-0.08)	(3.91)
Decoration	0.072	-0.063
	(0.54)	(-1.79)
Mixed Building	0.187	0.070
	(1.54)	(1.88)
Villa	-0.143	-0.242**
	(-0.55)	(-3.28)
Plot Ratio	-0.000	-0.029***
	(-0.00)	(-3.86)
Green Ratio	0.000	0.000
	(0.01)	(0.31)
GFA	-0.139*	-0.132***
	(-2.32)	(-8.78)
Observations	375	4377