



The role of agents in fraudulent activities: Evidence from the housing market in Beijing[☆]

Sumit Agarwal^a, Weida Kuang^b, Long Wang^c, Yang Yang^{d,*}

^a NUS Business School, National University of Singapore, 15 Kent Ridge Drive, 119245, Singapore

^b School of Business, Renmin University of China, 59 Zhongguancun St, Haidian District, Beijing, 100872, China

^c School of Economics, Fudan University, 600 Guoquan Road, Yangpu District, Shanghai, 200433, China

^d CUHK Business School, The Chinese University of Hong Kong, 12 Chak Cheung Street, N.T., Hong Kong Special Administrative Region

ARTICLE INFO

JEL classification:

R21
R28
H26
R30

Keywords:

Tax evasion
Real estate agent
Housing market
Learning-by-doing

ABSTRACT

This study examines the role played by real estate agents in fraudulent activities in China's housing market. We show that agents contribute to the formation of *Yin–Yang* contracts and the magnitude of resulting tax evasion through two possible mechanisms: the learning-by-doing effect and the peer effect. Agents' cumulative experience allows them to discover local registration authorities' monitoring capability and strategically register prices close to the internal guideline prices. Moreover, agents' involvement in tax evasion is significantly affected by the tax evasion behaviors of their peers. The difference-in-differences analyses across two policy shocks show that the involvement of experienced real estate agents exacerbates the magnitude of tax evasion during periods of frequent government policy adjustments.

1. Introduction

Agents play a critical role in facilitating transactions across diverse business activities, including investment management (Dvořák, 2005), insurance services (Eckardt and Rähke-Döppner, 2010), and real estate transactions (Levitt and Syverson, 2008; Agarwal et al., 2019b). Existing research has primarily concentrated on the principal–agent problem, which arises from moral hazard and conflicts of interest. For instance, agents may prioritize promoting high-priced products over more suitable alternatives to clients and encourage frequent transactions to boost their commissions (Mehran and Stulz, 2007; Bolton et al., 2007). Moreover, in the real estate market, agents' expertise and access to information may enable them to acquire properties at lower prices for personal investment and sell them at higher prices (Agarwal et al., 2019a; Levitt and Syverson, 2008). It is worth noting that agents, in

their efforts to enhance transaction efficiency and speed, may employ various tactics, some of which are legitimate while others may skirt the boundaries of legality. For instance, agents may engage in fraudulent activities such as fabricating resumes for clients seeking job search or school application services, or suggesting methods for clients to report lower income in the tax reporting industry. However, identifying such practices is challenging as they often occur covertly and are difficult to detect, and local authorities may not be motivated to intervene due to the perceived low net benefits associated with intervention. Therefore, empirical research on the questions of whether and how agents use their expertise and information advantage to facilitate fraudulent or illegal activities remains limited.¹

This study fills the literature gap by examining the role of agents in the formation of *Yin–Yang* contracts, which are closely linked to tax evasion in China's resale housing market. China serves as an ideal

[☆] We thank the Editor Stuart Rosenthal, two anonymous referees, James Conklin, Yongheng Deng, Hanming Fang, Tsur Somerville, Tien Foo Sing, Wenlan Qian, Yu Qin, and Jing Wu for helpful comments. We also benefited from discussants and participants at the AREUEA-ASSA meeting 2020, China Meeting of the Econometric Society (2019), 2019 Chinese Economic Association China Conference, AREUEA-International Conference 2019, AsRES Conference 2019, China Meeting of Econometric 2019 and Asian Meeting of Econometric 2019. All remaining errors are our own. Yang gratefully acknowledges the support from the National Natural Science Foundation of China (Grant No: 72203192). Wang gratefully acknowledges support from the National Natural Science Foundation of China (Grant No: 72303035). These authors contributed equally to this work.

* Corresponding author.

E-mail addresses: bizagarw@nus.edu.sg (S. Agarwal), kuangweida@rmbc.ruc.edu.cn (W. Kuang), wanglong@fudan.edu.cn (L. Wang), zoeyang@cuhk.edu.hk (Y. Yang).

¹ A recent study by Dimmock et al. (2018) shows that the probability of an advisor committing financial misconduct increases in accordance with the experience of misconduct among their new co-workers.

setting for our empirical analyses for several reasons. First, China's resale housing market involves buyers and sellers agreeing to sign *Yin–Yang* contracts, where the *Yin* contract reflects the actual transaction price of a property, while the *Yang* contract, with a much lower price, is submitted to local authorities to reduce tax payments. The widespread use of *Yin–Yang* contracts leads to short-term tax revenue losses, increased legal disputes between buyers and sellers, and distortion in the housing market. Second, intermediary services have experienced high demand in China's resale housing market over the past decade. A substantial amount of anecdotal evidence suggests that creating *Yin–Yang* contracts is a prevalent tactic utilized by agents to boost their business prospects (see further details in Section 2.1). Third, unlike other illegal activities where involved parties often conceal evidence, real estate agencies in China routinely record both the registered price in the *Yang* contract and the actual transaction price in the *Yin* contract, providing researchers with a valuable opportunity to investigate tax evasion (Agarwal et al., 2020; Dai and Xu, 2022; Fan et al., 2022). This study aims to address three key questions: Do agents engage in tax evasion in China's resale housing market? If so, what are the potential mechanisms through which agents facilitate tax evasion? Moreover, what unintended consequences arise from agents' involvement when the government frequently implements macro-prudential policies to stabilize the housing market?

We examine these questions using a novel and comprehensive transaction dataset provided by the largest real estate brokerage firm in China. The dataset offers several advantages: it provides precise information on actual transaction prices agreed upon by both parties and the registered prices submitted to registration authorities, allowing us to identify *Yin–Yang* contracts and quantify the extent of tax evasion. Additionally, it includes exceptionally detailed information on the real estate agents, providing a valuable opportunity to investigate their role in forming *Yin–Yang* contracts. Furthermore, the dataset covers a long period from January 1, 2014, to June 1, 2017, enabling us to study the role of real estate agents in tax evasion in an environment with frequent policy adjustments.

Specifically, consistent with existing literature and anecdotal evidence, we find that approximately 97% of transactions involve *Yin–Yang* contracts, with the registered price averaging 32% lower than the actual price. Importantly, our analysis reveals that the inclusion of agent fixed effect leads to a 37.4 percentage points increase in the explanatory power (R^2) for variation in tax evasion (i.e., the registered-to-actual price ratio), indicating that real estate agents might affect the severity of tax evasion. Further investigation uncovers two mechanisms: the learning-by-doing effect and the peer effect, both identified as ways through which agents gain information and acquire tax evasion tactics. First, we demonstrate agents' learning-by-doing effect, showing that a 100% increase in the number of past transactions leads to a 4.4 percentage points decrease in the average registered-to-actual price ratio, equivalent to a CNY 186,335 (around USD 25,882) in terms of the total transaction price. Second, we find that agents' cumulative experience enables them to discern the monitoring capabilities of local registration authorities and strategically register prices as close as possible to the government's internal guideline prices. Third, utilizing instrumental variable analysis, we identify the peer effect among agents: a 100% increase in the average registered-to-actual price ratio of peers causes an agent's registered-to-actual price ratio to increase by 31.5 percentage points. Notably, the peer effect does not take away the explanatory power of the learning-by-doing effect, suggesting that both mechanisms jointly contribute to explaining agents' tax evasion behavior.

Moreover, to study the unintended economic consequences of agents' involvement, we employ the difference-in-differences approach and exploit two plausibly exogenous housing policies introduced on

March 30, 2015, and September 30, 2016 (see the details in Section 2.2).² The policy analyses highlight three key implications: (1) the reduction of down-payment reduces the tax evasion, and vice versa.; (2) the effect on extensive margin dominates; and (3) with policy adjustments, agents' experience exacerbates the magnitude of tax evasion.

Our results survive a series of robustness checks, including tests for parallel pre-trend assumptions and falsification tests using alternative placebo policy dates. In addition, heterogeneity tests show that the learning-by-doing effects are pronounced for transactions involving mortgages, and that peer effects are stronger in larger branches. Furthermore, our analysis indicates that experienced male agents exhibit more aggressive tax evasion behavior compared to their experienced female peers.

The study contributes to three strands of literature, the first of which is agents' behavior. Agents are commonly engaged for their expertise, specialized knowledge, and established relationships with relevant parties (Levitt and Syverson, 2008; Benjamin et al., 2009; Eckardt and Rätke-Döppner, 2010; Barwick and Pathak, 2015; Barwick et al., 2017; Agarwal et al., 2019b). Most studies emphasize the occurrence of the “principal–agent problem” when an agent's incentive is not aligned with the client's, particularly when the agent's actions are not easily monitored. In contrast to literature focusing solely on the agency problem, this study investigates the role of agents in fraudulent activities, specifically tax evasion. Although tax evasion associated with the *Yin–Yang* contracts is a unique phenomenon in China's housing market, our findings can be generalized to other illegal activities involving intermediaries across various industries, such as underground banking, illegal recruitment and immigration, and insider trading.

This study also contributes to the tax evasion literature (Fisman and Wei, 2004; Marion and Muehlegger, 2008; Chetty, 2009; Merriman, 2010; Balafoutas et al., 2015; Artavanis et al., 2016; Agarwal et al., 2020). Tax evasion has been documented in various contexts, such as car imports (Fisman and Wei, 2004), income reporting (Erard, 1993; Chetty, 2009; Artavanis et al., 2016), diesel fuel purchasing (Marion and Muehlegger, 2008), and non-uniform tax-rates (Merriman, 2010). Keen et al. (2015) estimate that global tax evasion amounts to 5% of the global economy. In the context of China's housing market, recent studies demonstrate the prevalence of *Yin–Yang* contracts and substantial magnitudes of tax evasion (Agarwal et al., 2020; Dai and Xu, 2022; Lu et al., 2022; Fan et al., 2022). Our study differs from these existing studies in several key aspects. First, the most closely related work by Agarwal et al. (2020) assesses the impact of capital gains taxes on tax evasion, focusing on policy evaluation and the heterogeneous effects on different groups of buyers. Our study, however, is the first to focus on the role of agents in forming *Yin–Yang* contracts in the housing market, addressing a notable gap in existing literature due to the absence of demographic information and performance data regarding agents. Moreover, unlike Agarwal et al. (2020), which focus on the housing policies related to capital gains taxes, our study utilizes the housing policies that change minimum down payment to understand the unintended economic consequences resulting from agents' involvement in *Yin–Yang* contracts, rather than solely focusing on the direct impact of housing policies on tax evasion. Moreover, in terms of data advantages, our study uses a substantially larger sample, covering the period from January 1, 2014 to June 1, 2017, containing important information on real estate agents, which has not been explored in previous research.

² The government policies introduced on March 30, 2015 (Decree No. 39 [2015] and Decree No. 98 [2015]) lowered the minimum down-payment for a second home from 70% to 40%, and canceled the sales tax (5.6% of the total transaction price) for homes transacted after a two-year holding period. Another policy introduced on September 30, 2016 (Decree No. 46 [2016]) raised the minimum down-payment from 30% to 35% for the first home, and from 40% to 50% for the second home.

Our study also relates to the learning literature. Many studies show that learning from prior experience enables agents to acquire information and improves their performance (Nadler et al., 2003; Barkema and Schijven, 2008; Van Nieuwerburgh and Veldkamp, 2009). For instance, Feng and Seasholes (2005) show that investor sophistication and trading experience alleviate the disposition effect. Benjamin et al. (2009) illustrate a positive relationship between agents' earnings and their work experience. Barwick and Pathak (2015) suggest that agents' productivity depends significantly on their level of experience.

Overall, this paper makes a significant contribution by empirically examining the involvement of agents in tax evasion in the housing market. Moreover, it highlights that real estate agents not only learn from their own working experience but also from their peers. The findings have broad applicability across various industries such as manufacturing and the service sector, where *Yin–Yang* contracts are common and agents play active roles. Moreover, they carry substantial policy implications, as understanding the factors and mechanisms driving the formation of *Yin–Yang* contracts could aid local governments in navigating the challenge of addressing these contracts. By shedding light on this issue, the findings may offer insights into potential solutions and strategies, particularly regarding the imposition of penalties on intermediaries. To the best of our knowledge, this is one of the first attempts to conduct a systematic empirical analysis of agents' involvement in fraudulent and illegal activities.

The remainder of the paper is organized as follows. Section 2 provides the background on *Yin–Yang* contracts and policy interventions in the housing market in China. Section 3 presents the data and summary statistics. Section 4 presents the empirical methodology and results, followed by the heterogeneity and other tests in Section 5. Section 6 concludes the study.

2. Background

2.1. *Yin–Yang* contracts in the housing market

In China's real estate market, an agent typically works as a dual agent, representing both the buyer and the seller. In a property transaction, a home buyer/seller first creates a listing through a real estate brokerage, which is a firm that employs many real estate agents with expertise and specialized knowledge in the local housing market. Then, a real estate agent steps in and searches their client pool for potential sellers/buyers. In particular, differently from the practice of Western countries where signing an exclusive listing is the most common option for selling homes, non-exclusive (open) listings are prevalent in China's secondary housing market, where a seller/buyer usually asks multiple agents/brokers to find a potential buyer/seller. For non-exclusive listings, once a seller/buyer posts the request to sell/buy a property online/in-store, he/she can be approached by multiple agents from different branches or brokerage firms. Across branches and brokerage firms, simultaneous negotiations can be conducted by different agents in order to complete a transaction faster or at a better price. Only the agent who successfully sells/buys the home is paid commission. Therefore, clients represented by multiple agents are unaware of who will ultimately close the deal, which largely alleviates any concerns related to possible non-random matching between clients and agents. To facilitate the deal, the agents will use a series of marketing skills, such as showing properties to prospective buyers, mediating negotiations between buyers and sellers, and advising clients on mortgages and pricing decisions. Thus, a real estate agent plays a crucial role as a conduit between buyer and seller in the secondary housing market.

Meanwhile, considerable anecdotal evidence suggests a unique phenomenon in China's housing market, wherein real estate agents play a pivotal role in shaping the formation of *Yin–Yang* contracts.³ The

primary motivation behind this practice lies in agents' inclination to recommend such contracts to expedite transactions, thereby enabling both buyers and sellers to evade significant tax liabilities: the *Yin* contract stating the actual price is kept under the table, while the *Yang* contract under-reporting the actual price is submitted to the registration office. However, submitting an appropriate registered price to the local registration authorities is not straightforward, as these authorities typically establish an internal guideline price at the district level, which is periodically updated but not disclosed to the public.⁴ The registered price would be rejected by the local registration authorities if it is below the guideline price. Therefore, agents, especially those experienced ones well-informed about the market, can estimate the minimum guideline prices set by the local housing authorities and assist the buyers/sellers in submitting the most suitable registered price. To this end, it is important to understand how and to what extent real estate agents influence the formation of *Yin–Yang* contracts.

Although such *Yin–Yang* contracts are illegal, they remain widespread in China's housing market, with both buyers and sellers eager to cooperate (Agarwal et al., 2020; Dai and Xu, 2022; Fan et al., 2022). For sellers, signing such a *Yin–Yang* contract is a tempting choice because it helps them to pay lower capital gain taxes. For buyers, participation in *Yin–Yang* contracts is incentivized by market dynamics, perceived costs, and regulatory factors. During our study period, in cities such as Beijing, the housing market functions as a seller's market, characterized by strict purchase regulations and limited housing supply. Consequently, buyers are compelled to shoulder expenses such as capital gains taxes (Agarwal et al., 2020). Moreover, in the absence of effective regulatory oversight or enforcement concerning *Yin–Yang* contracts, buyers may believe that penalties are manageable and underestimate risks. Additionally, buyers may anticipate transferring future tax burdens or rely on property value appreciation to offset tax liabilities.

As shown in Fig. 1, the *Yin–Yang* contracts, with the registered prices 32% lower than the actual transaction prices, prevail in the market. Additionally, the distributions of registered-to-actual price ratio, actual price, and registered price are presented in Fig. A.1. Nevertheless, *Yin–Yang* contracts have been blamed for distorting property prices and leading to disputes between sellers and buyers since the official record of housing prices no longer provides a reliable and accurate measure of the fair value of a property.⁵

In many cities, local housing authorities have become aware of the problem and implemented specific policies and monitoring guidelines. On January 20, 2011, the Ministry of Housing and Urban–Rural Development, the National Development and Reform Commission, and the Ministry of Human Resources and Social Security jointly issued the “Real Estate Brokerage Administration Measures”, which took effect on April 1 of the same year. According to Article 25 (5) and Article 37 of these measures, *Yin–Yang* contracts are illegal, and real estate brokerage firms are prohibited from facilitating the *Yin–Yang* contracts for tax evasion purposes.⁶ For any violation of the articles, an agent shall receive a fine of 10,000 RMB (≈1500 USD), and a brokerage shall receive a fine of ranging from 10,000 to 30,000 RMB. Nonetheless, the regulations have limited effects due to the substantial benefits associated with *Yin–Yang* contracts, coupled with the negligible penalties they may incur.

Despite being aware of the issues, local governments choose to implement limited interventions due to three primary reasons when assessing the costs and benefits of various intervention levels. First,

⁴ Source: <http://house.people.com.cn/n1/2017/0418/c164220-29217345.html>.

⁵ Source: <https://www.scmp.com/article/705315/yin-yang-contracts-must-be-stamped-out>.

⁶ Source: http://www.mohurd.gov.cn/fqjs/jsbgz/201101/t20110125_202140.html.

³ Source: <https://www.reuters.com/investigates/special-report/china-risk-mortgages/>; <http://www.globaltimes.cn/content/560560.shtml>.

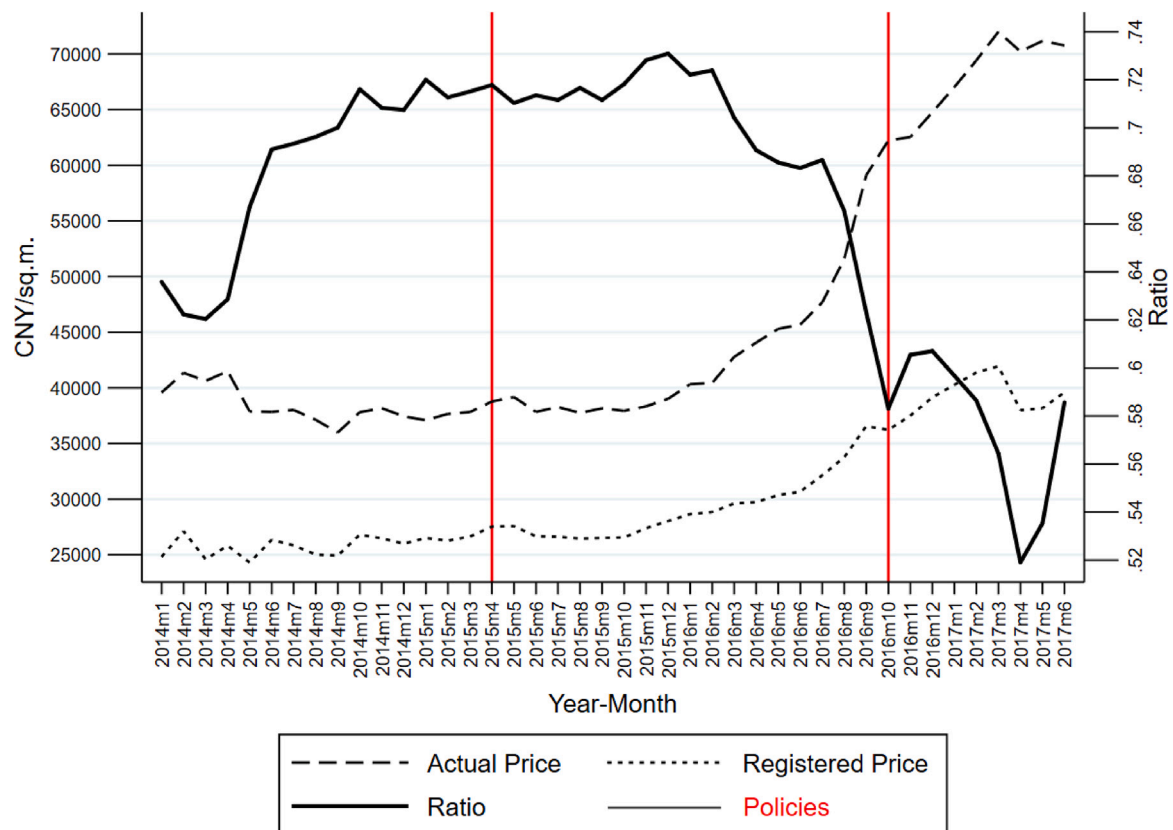


Fig. 1. The trends of prices and ratio. *Notes:* This figure presents the trends of actual unit price, registered unit price, and registered-actual price ratio in Beijing's resale residential housing market through the sample period. The two red vertical lines indicate two waves of policies introduced in Section 2.2.

determining the actual value of real estate presents a significant challenge due to various factors. Property prices fluctuate considerably over time, influenced by market dynamics and economic conditions. Even among properties sharing identical characteristics, owners often adjust prices in response to various factors such as shifting demand dynamics, unique property attributes, and prevailing market trends. This inherent volatility complicates the accurate assessment of real estate asset values. Moreover, buyers and sellers often employ another common strategy, which involves claiming that the transaction price comprises both the value of the property and an additional property improvement allowance, the latter of which is tax-exempt but prone to significant over-claiming. As a result, identifying *Yin-Yang* contracts becomes exceptionally difficult and costly for governments when collusion occurs between buyers and sellers to falsify prices.

Second, permitting *Yin-Yang* contracts in the housing market can result in potential short-term benefits. While reporting actual prices increases costs for buyers, which slows down housing market transactions, *Yin-Yang* contracts could reduce transaction costs, ultimately boosting market activity by incentivizing more transactions and thereby improving liquidity and investment in the real estate sector. This increased activity has the potential to foster economic expansion and employment opportunities, particularly within industries closely associated with real estate. Moreover, an active real estate market not only bolsters land sales and GDP but also significantly impacts the performance evaluations of government officials, as their performance is often evaluated based on economic growth metrics. As highlighted in Pan et al. (2015), Wang and Hui (2017) and Agarwal et al. (2020), local governments heavily rely on revenues from land sales and are thus reluctant to cool down housing markets, as the strict prohibition of *Yin-Yang* contracts would significantly reduce the number of housing transactions.

Third, the impact of lost tax revenue from *Yin-Yang* contracts may not be as severe as initially perceived in the long term. Consider a

scenario where a seller, A, sells a property to a buyer, B, with an actual transaction price of 10 million but reports it to the local government as 5 million through such a contract. As property prices gradually appreciate over time, the property's value may rise to 12 million after several years. Now, when B decides to sell the property to the next buyer, C, the taxes owed are calculated based on the difference between the current and the last transaction prices. However, since B previously paid taxes based on the reported transaction price of 5 million, taxes are now calculated based on the difference between the current transaction value and 5 million. Consequently, the previously unpaid tax amount, which was calculated based on an added value of 5 million (=10 million - 5 million), has now been effectively reclaimed. Nonetheless, it is important to note that the tax burden originally intended for B is now transferred to C in this process. Given that local governments can eventually recover previously lost tax revenues in future transactions, they have less incentive to invest huge resources in combating *Yin-Yang* contracts at present.

2.2. Two policy shocks in the housing market

In response to the rapid and nationwide housing price appreciation, the Chinese government has been implementing various control measures since 2010, such as purchase restrictions, financing constraints, and housing transaction taxes (Wu et al., 2012; Fang et al., 2016; Somerville et al., 2019; Wang and Yang, 2021). Fig. 2 summarizes the details of housing market policies announced by the central government from 2010 to 2017, providing opportunities for researchers to study China's housing market. Agarwal et al. (2020) examine the effect of capital gains tax changes occurred on March 30, 2013 on tax evasion. From the policy perspective, our study differs from prior research by analyzing how changes in down-payment requirements affect agents' tax evasion behavior, utilizing two policy shocks: one on

Date	Policy	Purchase Restriction	Financing Constraint	Tax Policy
17-Apr-10	[NATIONAL POLICY] No. 10 [2010] of the State Council			
08-May-10	No. 13 [2010] of the Municipal People's Government of Beijing: <i>Notice on Resolutely Curbing the Soaring of Housing Prices in Some Cities</i>	Each family can purchase one additional property	The down payment of the first loan shall not be less than 30% of the total price; for families who purchase a second unit with mortgage, the down payment of loan shall not be less than 50%, and the mortgage rate shall not be less than 1.1 times of the benchmark rate;	-
26-Feb-13	[NATIONAL POLICY] No. 17 [2013] of the State Council			
31-Mar-13	No. 17 [2013] of the Municipal People's Government of Beijing: <i>Notice on Further Improving Regulation of the Real Estate Market</i>	-	For families who purchase a second unit with mortgage, down payment of the loan increases from 50% to 70%	20% of capital gain tax for properties with holding period less than five years
30-Sep-14	No. 382 [2014] of the Beijing Municipal Commission of Housing and Rural and Urban Construction: <i>Notice on Adjusting the Preferential Policies on Deed Tax and Individual Income Tax during Real Estate Transactions</i>	-	-	Redefine ordinary housing that entitle for preferential tax treatment: properties with holding period longer than five years do not have to pay for the sales tax (5.6% of total price)
30-Mar-15	[NATIONAL POLICY] No.98 [2015] People's Bank of China and No. 39 [2015] of the Ministry of Finance			
30-Mar-15	No.98 [2015] People's Bank of China: <i>Notice of the People's Bank of China, the Ministry of Housing and Urban-Rural Development and the China Banking Regulatory Commission on Issues concerning Individual Housing Loan Policies</i> No. 39 [2015] of the Ministry of Finance: <i>Notice of the Ministry of Finance and the State Administration of Taxation on Adjusting the Business Tax Policies on Individual Housing Transfers</i>	-	For families who purchase a second unit with mortgage, down payment of the loan decreases from 70% to 40%	Properties with holding period longer than two years do not have to pay for the sales tax (5.5% of total price)
30-Sep-16	No. 46 [2016] of the Municipal People's Government of Beijing: <i>Notice on Promoting the Sustained and Healthy Development of the Real Estate Market</i>	-	Down payment for the first home loan increases from 30% to 35%, and down payment for the second home loan increases from 40% to 50%; down payment for luxury (non-ordinary) properties is 70%.	-

Fig. 2. Summary of the major policies from 2010 to 2017 in Beijing. Notes: This figure summarizes major policy changes from 2010 to 2017.

March 30, 2015, which aimed to decrease the minimum down payment, and another on September 30, 2016, which aimed to increase it.

On March 30, 2015, the People's Bank of China and the Ministry of Finance issued Decree No. 98 [2015] and Decree No. 39 [2015], which lowered the minimum down-payment for a second home from 70% to 40% and canceled the sales tax (5.5% of the total transaction price) for homes transacted after a two-year holding period.⁷ Obviously, the policy of March 30, 2015 is a bundled policy that affects the housing market in two dimensions: decreasing the minimum down-payment and reducing transaction tax. To precisely estimate the effects of a down-payment decrease on tax evasion and avoid the possible contamination from the effects of a tax reduction, we restrict the sample to properties with holding periods of less than two years, which are not affected by tax changes.

On September 30, 2016, the central government issued another Decree, No. 46 [2016], to modify home financing constraints. The policy raised the minimum down-payment from 30% to 35% for a first home purchase, from 40% to 50% for a second home, and mandated a 70% down-payment for homes not classified as ordinary property.⁸ In contrast to the March 30, 2015 policy, the September 30, 2016 policy allows us to study the effects of a down-payment increase on tax evasion.

As detailed in Section 4.4, the change in minimum down-payment would affect the buyers' tax evasion behavior both at the extensive and intensive margins, but in opposite directions. Moreover, since the down-payment changes in both policies only affect mortgage buyers, it provides an opportunity for us to utilize a difference-in-differences

method with mortgage buyers and cash buyers consisting of the treatment and control groups, respectively, to estimate the impacts of down-payment change on tax evasion.⁹

3. Data and summary statistics

We merge a micro-level housing transaction dataset from two data sources of the largest real estate brokerage firm, which employs approximately 120,000 employees in about 6000 branches in over 25 cities in China. The first data source provides the universe of the sales records in the secondary housing market of Beijing from January 1, 2014, to June 1, 2017 (i.e., 299,115 transactions).¹⁰ Of particular importance, the first data source contains the agents' information for each transaction. The second data source contains 179,580 transaction records with additional key information that is available only to the brokerage firm, such as actual transaction price and registered price. Fortunately, we can merge the data using the unique housing ID that exists in both data sources. Therefore, our merged dataset contains rich information not only on transaction prices, properties, buyers and sellers, but also on agents. Thus, it provides us with a unique opportunity to investigate the role of intermediaries in aiding tax evasion, a question that has never been addressed in existing studies.

Fig. 3 presents the geographical distributions of housing resales, real estate agency branches, and real estate registration offices across 13 administrative districts in Beijing. Although our main analysis is based on data from Beijing, it is reasonable to assume that agents in this city would demonstrate behavior similar to those in other major cities in China. This assumption is supported by anecdotal evidence suggesting that agents' active involvement in *Yin-Yang* contracts is a prevalent phenomenon across the country.

⁹ The term "mortgage buyers" refers to those who finance the purchase of a home via a mortgage loan, whereas the term "cash buyers" describes those who pay in full for a home.

¹⁰ Agarwal et al. (2020) document the pervasive tax evasion in China by first utilizing the micro-level data from Beijing and then generalizing the analysis to all 35 major Chinese cities in 2013. In this study, we further extend the sample period from 2014 to 2017 to examine the role of agents in tax evasion and their behavioral responses to two policy shocks.

⁷ The policy was issued by the People's Bank of China and the Ministry of Finance, titled "Notice on Issues Concerning Individual Housing Loan Policies", and "Notice on Adjusting the Business Tax Policies on Individual Housing". See <https://wiki.mbalib.com/wiki/3.30%E6%96%B0%E6%94%BF> for more details of the policies.

⁸ Source: <http://www.waizi.org.cn/law/13390.html>.

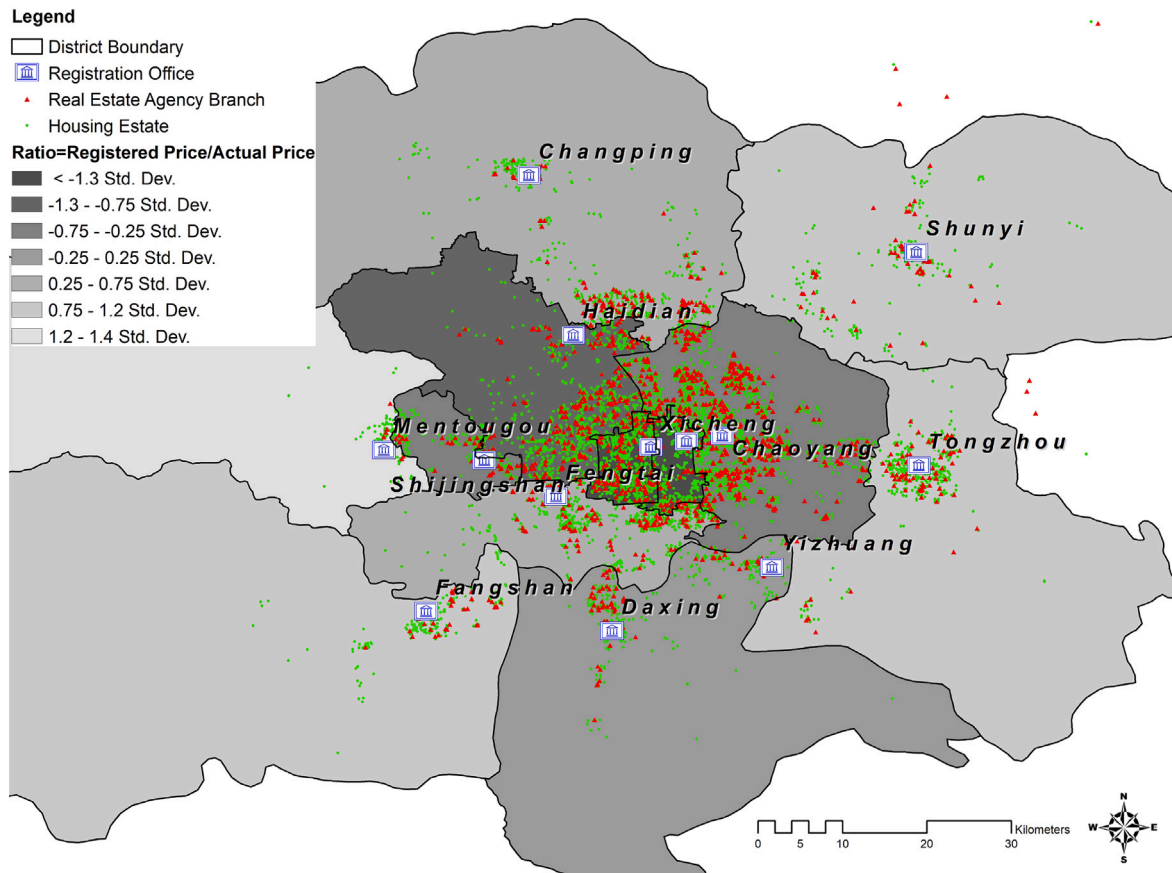


Fig. 3. The sample distribution. Notes: This figure presents the sample distributions of housing complexes, real estate agency branches, and corresponding registration offices across 13 administrative districts in Beijing.

Our data has numerous strengths for analysis. First, it contains exceptionally detailed information on real estate agents, including their demographics (such as gender and age), portraits, and the total completed transactions up to September 2018. In particular, since the data provides a full sample of transaction records for every agent, we can calculate an agent's experience upon a specific transaction. For instance, if an agent i completed 50 transactions before September 30, 2018, and 10 of those transactions occurred between January 1, 2015, and September 30, 2018, then agent i 's experience for a specific transaction on January 1, 2015, is 40. This approach allows us to investigate the impact of an agent's experience on tax evasion.

Second, the dataset provides comprehensive information about both the listed and transacted properties, including address, housing characteristics (e.g., story, unit size, number of bedrooms, number of living rooms, and number of bathrooms), number of pre-sale visits, listing and sale prices, financing type (mortgage or cash), holding periods since the last transaction date (categorized as below two years, two-to-five years, and over five years), and key dates related to home sales (such as listing date, days on the market, and contract date). Notably, over half of the dataset (179,580 transactions) contains additional proprietary information accessible only to the brokerage firm, such as the actual price. Moreover, with the precise information available, we can compare the registered price with the actual transaction price, to identify the existence of *Yin-Yang* contract and quantify the magnitude of tax evasion. Specifically, we identify a contract to be a *Yin-Yang* contract if the registered price is lower than the actual transaction price. The magnitude of tax evasion is precisely measured as the registered-to-actual price ratio, where a lower ratio indicates a higher level of tax evasion.

Third, the dataset also provides detailed information on both buyers and sellers, including gender, age, and place of birth.¹¹ The information regarding the place of birth allows us to distinguish between local and non-local buyers and sellers.

However, we should consider three caveats regarding the dataset. First, the dataset excludes new sales in the primary market, where properties are typically sold directly by real estate developers without the involvement of third-party agents. Second, it also excludes properties in the secondary market that are sold without assistance from real estate agents. Nevertheless, we believe the dataset remains representative, as it encompasses all transactions facilitated by China's largest brokerage firm during the sample period, accounting for approximately one-third of total transactions in Beijing's secondary market. Third, only 179,580 out of 299,115 transactions can be used for our analyses due to missing housing ID. As shown in Table A.1, we observe no significant differences in housing characteristics and actual prices between the regression sample and the unmatched sample. This mitigates concerns regarding potential differences between the 119,535 unmatched observations and our regression sample.

Table 1 presents summary statistics for key variables used in the paper. Panel A reports transaction information. Surprisingly, 97% of contracts have registered prices lower than the actual transaction

¹¹ The dataset contains the first six digits Citizen Identity Number of clients. The first six digits are an address code pinpointing the place of birth. Specifically, the first two digits denote the province, the following two digits denote the city, and the final two digits represent the district or county. A client is classified as local if the first two digits of their ID number are 11, which corresponds to Beijing.

Table 1
Summary statistics (between January 1, 2014 and June 1, 2017).

	Observations	Mean	S.D	Min	Max
Panel A. Transaction information					
Actual total price (thou.)	179,580	3833.25	2558.05	305	250,000
Registered total price (thou.)	179,580	2478.08	1609.65	0	250,000
Listed total price (thou.)	179,580	4063.97	68,356.82	260	302,500
Actual unit price	179,580	46,033.65	16,916.25	0	615,000
Registered unit price	179,580	30,841.94	20,684.09	3938	150,000
Listed unit price	179,580	49,177.80	76,679.69	3937	794,299
Registered-Actual Price Ratio	179,580	0.68	0.20	0.10	1.65
<i>YinYang</i>	179,580	0.97	0.16	0.00	1.00
Mortgage	179,580	0.85	0.36	0.00	1.00
LTV	94,299	0.63	0.13	0.00	1.00
Daily volume	1223	160	181.05	1	1454
Panel B. Housing characteristics					
Unitsize (sq.m)	179,580	85.04	39.42	5	1810
Days on Market	179,580	20.60	46.69	0	1464
Distance (km)	179,580	7.43	5.54	0.02	58.15
Level of storey	179,580	7.44	6.07	-2	40
# of visits	179,580	29.28	32.71	1	567
# of bedrooms	179,580	2.05	0.78	0	9
# of livingrooms	179,580	1.18	0.53	0	6
# of bathrooms	179,580	1.20	0.46	0	9
Decoration type					
<i>Fine decoration</i>	92,406				
<i>Simple decoration</i>	62,316				
<i>None decoration</i>	4457				
<i>Other</i>	20,401				
Panel C. Agent characteristics					
# of Total Transactions	179,580	21.28	18.28	1	144
Experience	179,580	3.83	5.82	1	128
Agent age	179,580	26.57	4.39	18	49
Male agent	179,580	0.72	0.45	0	1
Ratio of Peers	179,580	0.71	0.06	0	1.16
Tenure of Peers (# of month)	179,580	28.84	9.28	0	66.59
Panel D. Buyer characteristics					
Buyer age	179,580	35.90	9.91	10	109
Male buyer	179,580	0.52	0.50	0	1
Local buyer	179,580	0.36	0.48	0	1
Panel E. Seller characteristics					
Seller age	179,580	47.08	14.30	10	106
Male seller	179,580	0.55	0.50	0	1
Local seller	179,580	0.56	0.50	0	1
Panel F. Other information					
# of administrative districts	13				
# of complexes	6195				
# of agents	15,802				
# of below two-year	33,400				
# of two-to-five-year	23,277				
# of over five-year	122,903				

Notes: This table presents the summary statistics for the variables used in this paper. *YinYang*, equals 1 if the Registered-Actual Price Ratio is less than 1; *Mortgage*, equals 1 if the buyer uses a mortgage loan; *LTV*, Loan-to-value ratio; *Daily Volume*, the number of transactions on the daily basis; *Days on Market*, the number of days that a property stays on the market before sold; *Distance (km)*, the physical distance between the complex and the corresponding registration office; *# of Total Transactions*, the number of completed transactions for an agent as of September 2018; *Experience*, the number of completed transactions for an agent upon the focal transaction; *Ratio of Peers*, the average Registered-Actual Price Ratio of peers in a branch; *Tenure of Peers (# of month)*, the average length of service of peers in a branch.

prices, with the registered-to-actual price ratio averaging at 0.68, highlighting the prevalence of *Yin-Yang* contracts and tax evasion in the secondary housing market.¹²

Panel B presents the summary statistics of housing attributes. The average days-on-market is around 20 days, and the average size is 85 square meters (sq.m). The sample covers 6195 housing complexes and

15,802 agents.¹³ Around two-thirds of the houses are held for over five years since their last transactions. Panels C, D, and E respectively report the demographic characteristics of agents, buyers, and sellers. Agents have an average of 21.28 total transactions as of 2018 and 3.83 transactions on average upon each deal. The average ages for agents, buyers, and sellers is 26, 36, and 47, respectively. Among agents, 72% are male, while the gender distributions are quite even among buyers

¹² Buyers in those 3% of transactions often set the registered prices at or above the actual transaction prices, primarily with the intention of obtaining a higher mortgage loan from the bank.

¹³ A complex is a cluster of housing units surrounded by a private fence, and is also referred to as the project or “*loupán*” in Chinese. A housing complex may range from detached houses to high-density tower blocks with or without commercial facilities.

and sellers. Concerning the information on place of birth, 36% of buyers and 56% of sellers are local. Panel F presents other information.

4. Methodology and results

4.1. The influence of real estate agents on tax evasion

We begin the analysis by examining the determinants of tax evasion using the classic hedonic pricing model (Rosen, 1974). The specifications are as follows:

$$Y_{i,j,t} = \alpha + \beta * X_i + \theta_j + \gamma_{day} + \delta_{ym} + \epsilon_{i,t}, \quad (1a)$$

$$Y_{i,j,t} = \alpha + \beta * X_i + \theta_j + \gamma_{day} + \delta_{ym} + \lambda_{agent} + \epsilon_{i,t}, \quad (1b)$$

where i, j, t index the transaction, the housing complex, and the transaction date, respectively. The dependent variable $Y_{i,j,t}$ includes the logarithmic actual unit price, logarithmic registered unit price, and the registered-to-actual price ratio. X is a vector consisting of characteristics of housings, buyers, and sellers, which are summarized in Table 1.¹⁴ Housing complex fixed effect (denoted by θ_j) is included to eliminate the unobserved heterogeneity, such as location, quality, brand, etc., across housing complexes. Fixed effects for days of the week (γ_{day}) and year-month (δ_{ym}) are included to absorb the effects from time trend and seasonality. The standard errors are clustered at the complex level.

Notably, we study the impact of agents on tax evasions by examining the changes in R^2 with (Eq. (1b)) and without (Eq. (1a)) the agent fixed effect (λ_{agent}). Real estate agents can be considered to significantly explain variations in the outcome variable if the inclusion of the agent fixed effect substantially increases the R^2 .

Panels A and B in Table 2 present the results of estimating Eqs. (1a) and (1b), respectively. Although the inclusion of agent fixed effect barely changes R^2 for actual price (Columns 1 and 4), it significantly increases R^2 for registered price from 0.678 (Column 2) to 0.878 (Column 5), which further leads to a 37.4 percentage points increase ($0.759 - 0.385 = 0.374$) in the explanatory power for the variation in registered-to-actual price ratio (Columns 3 and 6). Similar changes are observed in adjusted R^2 . To ensure robustness, we exclude transactions involving 2543 agents who have only one transaction in our sample period and find consistent results in Table A.2. The substantial increases in R^2 for registered price and registered-to-actual price ratio after including agent fixed effect suggest that real estate agents play a significant role in producing *Yin–Yang* contracts.

In Table 2, our interest lies in understanding the effects of mortgage usage on the three outcome variables, as it sheds light on the differing behaviors of mortgage buyers and cash buyers in tax evasion. Specifically, the estimated coefficients on *mortgage* are significantly and positively correlated to registered price and the ratio, implying that buyers using mortgage loans (referred to as mortgage buyers) register higher prices and engage in less tax evasion compared to buyers not using mortgages (referred to as cash buyers). For instance, a coefficient of 0.094 in Column (3) indicates that properties with mortgages present a registered-to-actual price ratio of 9.4 percentage points higher than those without mortgages. Given the average under-reporting level of 32% (i.e., baseline registered-to-actual price ratio of 68%) in the sample, this constitutes a 29.4% increase ($9.4\%/32\% = 29.4\%$) in the intensive margin of under-reporting. On average, the under-reporting of the total price is CNY 398,420 larger for properties with mortgages than those without.¹⁵ However, coefficients of *mortgage* on actual transaction prices are insignificant in Columns (1) and (4), suggesting that mortgage usage does not affect the transaction price.

¹⁴ Housing characteristics include size, number of visits by potential buyers/sellers, floor level, number of bedrooms, number of living rooms, number of bathrooms, decoration condition, and holding period; characteristics of buyers and sellers include gender, age, and local identity.

¹⁵ $29.4\% * (3,833,250 - 2,478,080) = 398,420$.

The substantial power of agent fixed effect in explaining the variation in tax evasion leads us to further explore two important questions: First, what motivates real estate agents to engage in transactions involving *Yin–Yang* contracts? Second, what are the specific channels through which real estate agents affect the formation of *Yin–Yang* contracts and the magnitude of tax evasion?

To address the first question, we offer both anecdotal and empirical evidence. Anecdotal evidence from interviews with brokerage upper management sheds light on agents' salary structures and monthly performance reviews, indicating that real estate agents prioritize closing deals and earning commissions as their primary goals. Agents are required to actively attract clients and employ various strategies to expedite transactions. Intense competition among agents across branches and brokerages, driven by the nature of non-exclusive listings, has resulted in the involvement of nearly all agents in *Yin–Yang* contracts.¹⁶ In addition, as illustrated in Section 2.1, given the negligible punishments administered by local governments and the significant potential returns, almost all agents would take the risk of engaging in *Yin–Yang* contracts.

In terms of empirical evidence, we examine whether tax evasion correlates with properties' days-on-market (DoM) and agents' performances. Table A.3 presents the results. In Column (1), we regress the property's DoM on the registered-to-actual price ratio and find a significantly positive coefficient, which indicates that the registered-to-actual price ratio (a higher ratio refers to fewer tax evasions) is positively associated with a property's days-on-market. This implies that the higher the tax evasion in the *Yin–Yang* contracts, the shorter the time it takes for the properties to sell. These results align with the findings of Levitt and Syverson (2008), which show that agents convince clients to sell their houses quickly to expedite transactions and earn their commissions. To illustrate, with commission fees ranging from 1% to 3% of the total transaction price, for an average property in Beijing (worth 3.83 million yuan), an agent could earn an extra commission of 3833 to 11,499 yuan by negotiating a better price (10% change in price, for instance) for the client. However, prolonged negotiation could reduce the probability of closing the deal as clients may be engaging in simultaneous negotiations with other agents. Therefore, agents have a strong incentive to reduce properties' DoM by employing the strategy of *Yin–Yang* contracts.

To further examine agents' incentive for forming *Yin–Yang* contracts, we conduct a cross-section estimation and regress agents' total transactions on agents' average registered-to-actual price ratio of the transactions. Column (2) of Table A.3 reveals a negative correlation between the registered-to-actual price ratio and agents' transaction volume, indicating that higher levels of tax evasion are associated with increased completion of transactions by agents. Nonetheless, we want to emphasize that results in Table A.3 serve only as suggestive evidence that motivates us to investigate the impact of agents on *Yin–Yang* contracts, rather than as evidence for causal inference.

To answer the second question concerning the potential channels through which agents influence the formation of *Yin–Yang* contracts, which is the main focus of our study, we examine the learning-by-doing effect and the peer effect in Sections 4.2 and 4.3.

4.2. Learning-by-doing effect

We first study whether an agent's past experience (the learning-by-doing effect) affects tax evasion. Specifically, we use the number of transactions an agent has completed before each transaction as a proxy

¹⁶ For instance, according to the Economic Information Daily, a reputable Chinese newspaper, reported that "many real estate intermediaries in China are skillful in fabricating 'Yin–Yang' contracts and use this skill as a bait to attract business". Source: http://www.jjckb.cn/2013-07/22/content_456872.htm.

Table 2
Hedonic regressions with and without agent fixed effect.

Panel	Panel A. Without agent FE			Panel B. With agent FE		
Dep. variable	ln(A-Price)	ln(R-Price)	Ratio	ln(A-Price)	ln(R-Price)	Ratio
Model	(1)	(2)	(3)	(4)	(5)	(6)
Mortgage	0.000 (0.001)	0.158*** (0.003)	0.094*** (0.002)	0.002 (0.002)	0.152*** (0.006)	0.092*** (0.003)
Observations	179,580	179,580	179,580	179,580	179,580	179,580
R-squared	0.942	0.678	0.385	0.977	0.878	0.759
Controls	Characteristics of housing, buyer and seller					
Agent FE	No	No	No	Yes	Yes	Yes
Complex FE	Yes	Yes	Yes	Yes	Yes	Yes
Day of the week FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-month FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table examines changes of R-square with/without including the agent fixed effect. Characteristics for housing, buyer and seller, in association with fixed effects for agent, housing complex, days of the week, and year-month are included in all columns. Heteroscedasticity-consistent standard errors are clustered at the complex level as shown in the brackets. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level, respectively.

Table 3
Hedonic regressions with agent characteristics.

Panel	Panel A. Continuous learning effect			Panel B. Binary learning effect		
Dep. variable	ln(A-Price)	ln(R-Price)	Ratio	ln(A-Price)	ln(R-Price)	Ratio
Model	(1)	(2)	(3)	(4)	(5)	(6)
ln(Experience)	-0.054*** (0.008)	-0.125*** (0.019)	-0.044*** (0.011)			
First				0.006*** (0.001)	0.010*** (0.002)	0.003*** (0.001)
Observations	179,580	179,580	179,580	179,580	179,580	179,580
R-squared	0.979	0.882	0.763	0.979	0.881	0.763
Controls	Characteristics of housing, buyer and seller					
Fixed effects	Agent, Complex, DoW, Year-month					

Notes: This table examines the effects of agent characteristics on actual unit price, registered unit price, and registered-actual price ratio. *ln(Experience)* is continuous measure of agent learning effect used in Panel A; and *First* is the binary measure of agent learning effect, which is equal to 1 if the transaction is the first-time transaction of an agent, and 0 otherwise. The headers in the second row denote the dependent variable used in respective column. Characteristics for housing, buyer and seller, in association with fixed effects for agent, housing complex, days of the week, and year-month are included in all columns. Heteroscedasticity-consistent standard errors are clustered at the complex level as shown in the brackets. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level, respectively.

for their experience, following the existing literature (Nadler et al., 2003; Loewenstein and Thompson, 2006; Barkema and Schijven, 2008; Reuer and Ragozzino, 2008; Knill et al., 2015; Cuypers et al., 2017).¹⁷ As introduced in Section 2.1, the local tax authorities set a minimum price as an internal guideline for second-hand housing transactions in each neighborhood. Registration cannot proceed if the registered price falls below this guideline.¹⁸ Therefore, reporting a price just above the internal guideline price assists clients in minimizing their tax payments. We test the hypothesis that experienced agents could learn from their experience to appraise each district's guideline price and assist clients in registering relatively lower prices. In particular, we include measures of agent experience in Eq. (1b) and report the results in Table 3.

Panel A of Table 3 includes the number of historical transactions (indicated as *ln(Experience)*) of an agent upon transaction *i*. The coefficients on *ln(Experience)* are significantly negative in Columns (1) to (3). The coefficient of *ln(Experience)* in registered price (Column 2) is 2.3 times larger than that in actual price (Column 1), resulting in the negative coefficient in registered-to-actual price ratio (Column 3). Since a lower registered-to-actual price ratio corresponds to higher tax evasion, the results in Panel A of Table 3 suggest that experienced agents facilitate contracts with a higher level of tax evasion compared to their inexperienced peers. Specifically, it highlights the learning-by-doing effect among agents, demonstrating that a 100% increase in the

number of transactions leads to an average decrease in registered-to-actual price ratio of 4.4 percentage points, equivalent to approximately 186,335 CNY.¹⁹

For a robustness check, we include a binary measure of learning (denoted as *First*) in Panel B. *First* takes the value of 1 for the first transaction of an agent, and 0 otherwise.²⁰ The coefficients of *First* are significantly positive in Columns (4) to (6), with magnitude in Columns (5) being 0.004 units greater than that in Column (4), implying that the rookie agents report higher prices to local registration offices and produce fewer tax evasions.

The results showing a decrease in actual transaction price with agents' experience (as shown in Column 1) align with the findings of Levitt and Syverson (2008), which show that agents advise sellers to accept lower prices to expedite transactions. We also conduct additional tests to show that experienced agents sell houses faster at the cost of low transaction prices (see Table A.4). Since the transaction price does not affect the number of previous transactions of an agent, the reverse causality is less of a concern in Table 3.

Demonstrating that agents' work experience could assist them in reporting a lower price to the registration offices naturally prompts the question of what they have learned from their experience. To explore the possible learning mechanism, we first show that local registration

¹⁷ Barkema and Schijven (2008) provides a comprehensive review of firms' learning mechanisms.

¹⁸ "Yin-yang' contracts on the rise", Global Times, 2010. See details on <http://www.globaltimes.cn/content/560560.shtml>.

¹⁹ $\frac{4.4\%}{32\%} * (3\,833\,250 - 2\,478\,080) = 186\,335.875$.

²⁰ We identify the first transaction of an agent using another dataset from the brokerage firm, which contains agents' transaction records from the first day of work.

Table 4
Learning-by-doing effects — Monitoring capability.

Dep. variable Model	ln(A-Price) (1)	ln(R-Price) (2)	Ratio (3)
ln(Distance)	0.141*** (0.054)	0.071 (0.052)	−0.043*** (0.006)
Observations	179,580	179,580	179,580
R-squared	0.913	0.801	0.652
Controls	Characteristics of housing, buyer and seller		
Fixed effects	Agent, DoW, Year-month		

Notes: This table examines the effect of monitoring capability of local tax authority, proxied by the distance between complexes and local tax authority, on actual unit price, registered unit price, and registered-actual price ratio. The headers in the first row of both panel denote the dependent variable used in the corresponding column. Characteristics for housing, buyer and seller, in association with fixed effects for agent, days of the week, and year-month are included in all columns. Heteroscedasticity-consistent standard errors are clustered at the complex level as shown in the brackets. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level, respectively.

offices possess different monitoring capabilities across different housing complexes. Coval and Moskowitz (2001) and Agarwal and Hauswald (2010) show that physical proximity facilitates monitoring and access to information; therefore, the physical distance serves as a standard measure of monitoring capability. In this section, we use the physical distance (denoted as *Distance*) between registration offices and housing complexes to proxy the monitoring capability of the local registration offices. Fig. 3 plots the geographic distributions of 13 property registration offices at the district level. To examine the impact of the monitoring capability of local registration offices on the transaction price and tax evasion, we include $\ln(\text{Distance})$ in Eq. (1b) and report the results in Table 4. Notably, we do not control for the complex fixed effect in Table 4 as doing so would omit the variable of our interest — $\ln(\text{Distance})$. To eliminate the impacts of complex, we include instead the characteristics of complex (such as building age, distance to the CBD, distance to the nearest subway station, and developer fixed effect) in the regression.

The results show that although the estimated coefficient on $\ln(\text{Distance})$ is significantly positive, indicating that properties located farther away from registration offices tend to be more expensive, the registered prices do not increase proportionally and significantly with the proximity to local registration offices, which leads to a significantly negative estimated coefficient of −0.043 on $\ln(\text{Distance})$ in Column (3). The results deliver an important message: the farther a housing complex is from a registration office, the lower a registration office's monitoring capability, resulting in higher levels of tax evasion. This prompts us to investigate whether agents leverage the monitoring capability of local registration authorities to help clients evade as much tax as possible. To gain deeper insights into how agents game the system and how the pricing patterns evolve, we conduct three heterogeneous tests based on the results in Table 4. We examine whether the learning-by-doing effects vary across developers' state ownership, the age of housing complexes, and the average price level of complexes. Table A.5 presents the results, followed by detailed explanations in Appendix C.

We carry out tests on the learning mechanism by replicating the regressions in Table 4 with sub-samples. The results are reported in Table 5. In each column of Table 5, only the k_{th} -time ($k = 1, 2, \dots, 5+$) transaction of an agent is included. For instance, Columns 1 and 2 contain only the first and second transactions of each agent, respectively. Results in Panel A demonstrate a significant positive correlation between distance and actual transaction prices, and this effect is significant across all columns. In Panel B, the coefficients on $\ln(\text{Distance})$ are statistically insignificant on registered price except for transactions completed by fresh agents. Coefficients on $\ln(\text{Distance})$ are insignificantly correlated with registered price from the second transaction onwards, supporting the hypothesis that agents report the lowest “acceptable” prices to the registration offices regardless of the

Table 5
Learning-by-doing effects — Learning the guideline prices.

Panel A. ln(Actual Price)					
Experience Model	1st (1)	2nd (2)	3rd (3)	4th (4)	5th and + (5)
ln(Distance)	0.123*** (0.025)	0.142** (0.063)	0.213*** (0.081)	0.169*** (0.015)	0.174*** (0.005)
Observations	15,802	13,259	10,650	7841	132,028
R-squared	0.958	0.964	0.968	0.971	0.941
Controls	Characteristics of housing, buyer and seller				
Fixed effects	DoW, Year-month				
Panel B. ln(Registered Price)					
Experience Model	1st (1)	2nd (2)	3rd (3)	4th (4)	5th and + (5)
ln(Distance)	0.076** (0.031)	0.078 (0.068)	0.085 (0.105)	0.015 (0.019)	0.018 (0.019)
Observations	15,802	13,259	10,650	7841	132,028
R-squared	0.743	0.805	0.823	0.850	0.709
Controls	Characteristics of housing, buyer and seller				
Fixed effects	DoW, Year-month				
Panel C. Registered-Actual Price Ratio					
Experience Model	1st (1)	2nd (2)	3rd (3)	4th (4)	5th and + (5)
ln(Distance)	−0.029 (0.018)	−0.040 (0.030)	−0.078*** (0.015)	−0.101*** (0.009)	−0.069*** (0.004)
Observations	15,802	13,259	10,650	7841	132,028
R-squared	0.515	0.616	0.659	0.703	0.464
Controls	Characteristics of housing, buyer and seller				
Fixed effects	DoW, Year-month				

Notes: This table examines the effect of monitoring capability of local tax authority, proxied by the distance between complexes and local tax authority, on actual price, registered price, and registered-actual price ratio conditioning on the number of agents' transaction experience. The header in the first row of each panel denotes the dependent variable. Column (1), Column (2), ..., and Column (5) include the first-time transactions, the second-time transactions, ..., and the fifth-time and onward transactions of all agents, respectively. Characteristics for housing, buyer and seller, fixed effects for days of the week, and year-month are included in all columns. Heteroscedasticity-consistent standard errors are clustered at the complex level as shown in the brackets. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level, respectively.

actual property prices. In Panel C, the coefficients on distance show no statistical significance in the first two transactions but become significantly negative from the third-time transaction onwards. The results point to a mechanism whereby, as agents gain experience, they gradually learn the monitoring capability of the local registration offices, and strategically create a lower registered-to-actual price ratio for housing complexes located farther away from local registration offices.

4.3. Agents' peer effect

Peer effects refer to externalities in which peers' behaviors or characteristics affect an individual's behavior (Manski, 2000; Arcidiacono and Nicholson, 2005; Bayer et al., 2009; Dahl et al., 2014). Given that real estate agents working in the same branch can exchange information on internal guideline prices, discuss new listings, and update clients' needs, it is worth examining the impacts of peer effects on tax evasion.

To begin, we repeat the regressions from Table 2 by replacing the agent fixed effect with the branch fixed effect. As shown in Table A.6, the statistical significance and economic magnitudes of estimated coefficients on the control variables remain consistent with those in Table 2. More importantly, including the branch fixed effect significantly increases R^2 of registered-to-actual price ratio from 0.385 (Column 3) to 0.598 (Column 6), indicating the influential role of peers within the same branch in explaining the variation in tax evasion.

Table 6

Peer effects.

Panel	Panel A. OLS	Panel B. IV	
Stage		First-stage	Second-stage
Dep. variable	Ratio	Ratio of Peers	Ratio
Model	(1)	(2)	(3)
Ratio of Peers	0.056*** (0.009)		0.315** (0.158)
Tenure of Peers		−0.063*** (0.004)	
ln(Experience)	−0.045*** (0.001)	0.005 (0.015)	−0.063*** (0.014)
Observations	158,479	158,479	158,479
R-squared	0.389	0.922	0.216
Wald test (F-Stat)		22.590	
Hausman test (p-value)			0.001
Controls	Characteristics of housing, buyer and seller		
Fixed effects	Agent, Complex, DoW, Year-month		

Notes: This table examines the peer effect on registered-actual price ratio. Panels A and B use the OLS estimation and Instrumental Variable estimation, respectively. The headers in the third row denote the dependent variable used in respective column. **Ratio of Peers**, is the measure of peer effect. **Tenure of Peers**, is the instrument variable. The definitions of **Ratio of Peers** and **Tenure of Peers** are presented in Table 1. Characteristics for housing, buyer, seller and agent, in association with fixed effects for agent, housing complex, days of the week, and year-month are included in all columns. Heteroscedasticity-consistent standard errors are clustered at the complex level as shown in the brackets. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level, respectively.

Table 7

The expected sign of the impact of down-payment change on the registered-actual price ratio.

	Decrease of down-payment in March 30, 2015 Policy	Increase of down-payment in September 30, 2016 Policy
Extensive margin	+	−
Intensive margin	−	+

Notes: This table summarizes the expected impacts of down-payment change on the registered-actual price ratio at the extensive margin and the intensive margin.

Table 8

DID analysis on policy impact.

Dep. variable	Ratio	
Policy Model	March 30, 2015 Policy (1)	September 30, 2016 Policy (2)
<i>Treat</i> * <i>After</i> ₂₀₁₅	0.031*** (0.004)	
<i>Treat</i> * <i>After</i> ₂₀₁₆		−0.045** (0.002)
Observations	76,569	74,560
R-squared	0.345	0.431
Controls	Characteristics of housing, buyer and seller	
Fixed effects	Agent, Complex, DoW, Year-month	

Notes: This table examines the effects of March 31, 2015 Policy and September 30, 2016 Policy on registered-actual price ratio. The sample period for March 30, 2015 Policy is between March 30, 2014 and March 31, 2016. The sample period for September 30, 2016 Policy is between April 1, 2016 and May 30, 2017. In Columns (1), the transactions of houses held between two and five years are excluded and *Treat* and is a dummy equal to 1 if the houses are bought via mortgage, and 0 otherwise. *After*₂₀₁₅ is a dummy equal to 1 if the transaction took place after March 30, 2015, and 0 otherwise. *After*₂₀₁₆ is a dummy equal to 1 if the transaction took place after September 30, 2016, and 0 otherwise. Characteristics for housing, buyer, seller, in association with fixed effects for agent, housing complex, days of the week, and year-month are included in all columns. Heteroscedasticity-consistent standard errors are clustered at the complex level as shown in the brackets. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level, respectively.

The comprehensive information on agents' transaction records enables us to further explore the peer effect, which has been difficult to identify and measure in previous studies. In this study, we construct a new variable, *Ratio_of_Peers*, which is the average registered-to-actual

price ratio of agent *k*'s peers in the same branch at agent *k*'s transaction *i*. Specifically, we examine whether the tax evasion of an agent *k*'s transaction *i* is affected by the tax evasion behaviors of agent *k*'s peers in the same branch. In Column (1) of Table 6, we estimate the relationship between an agent's tax evasion and *Ratio_of_Peers*.

The estimated coefficient on *Ratio_of_Peers* in Column (1) suggests a significant and positive peer effect, but the results suffer from the reverse causality problem because agent *k* and *k*'s peers' behaviors are interdependent. There are three common approaches to address the endogeneity concerns in the existing peer-effect studies (Manski, 1993; Dahl et al., 2014): (1). controlling for as many group characteristics as possible; (2). exploiting exogenous assignment to peer groups; (3). using instrumental variables that affect peers' achievement but do not directly affect an agent's achievement.

To establish a causal interpretation of peer effect on tax evasion, we employ an instrumental variable strategy. This requires an instrument that is correlated with the peers' tax evasion behavior but uncorrelated with agent *k*'s tax evasion behavior. We first identify agent *k*'s senior peers who joined the branch earlier than agent *k*. Then, we construct a measure, denoted as *Tenure_of_Peers*, representing the average number of months that agent *k*'s senior peers have served in the same branch. This approach ensures the exogeneity of the instrument because the entries of agent *k*'s senior peers in the same branch are pre-determined, and therefore should not be inversely affected by agent *k*'s behavior. Meanwhile, *Tenure_of_Peers* could significantly affect the tax evasion behaviors of those senior peers. We estimate the following two-stage least square (2SLS) equations:

$$\text{Ratio_of_Peers} = \alpha + \varphi * \text{Tenure_of_Peers} + \beta * X_i + \theta_j + \gamma_{day} + \delta_{ym} + \lambda_{agent} + \epsilon_{i,t} \quad (2)$$

$$Y_{i,j,t} = \alpha + \rho * \widehat{\text{Ratio_of_Peers}} + \beta * X_i + \theta_j + \gamma_{day} + \delta_{ym} + \lambda_{agent} + \epsilon_{i,t} \quad (3)$$

The results of estimating the 2SLS are reported in Panel B of Table 6. In Column (2), we observe the first-stage estimate from Eq. (2). The first-stage estimate shows that the service tenure of agent *k*'s senior peers, *Tenure_of_Peers*, is negatively associated with the tax evasion of agent *k*'s senior peers, *Ratio_of_Peers*. This result confirms the prediction that the average service tenure of peers is a statistically significant determinant of the extent of their tax evasion behavior. We also find that the F-statistic on the first stage is 22.5, which lends us confidence in rejecting the null hypothesis of a weak instrument (Angrist and Krueger, 2001).

Column (3) shows the second-stage estimate from Eq. (3) and reports a significantly positive effect of agent *k*'s peers on agent *k*'s tax evasion. Specifically, a 100% increase in the predicted registered-to-actual price ratio of agent *k*'s peers causes agent *k*'s registered-to-actual price ratio to increase by 31.5 percentage points. The second-stage estimation shows that peers' causal effect is around 6 times the OLS estimate of Column (1), suggesting that OLS estimates are downward biased. One potential explanation for why OLS underestimates the peer effect is that the average tax evasion of agent *k*'s peers increases with agent *k*'s tax evasion, and vice versa.

In summary, the IV estimation shows that agents can learn from their peers in evading taxes through Yin–Yang contracts. More importantly, we include the proxy of the learning-by-doing effect in Columns (1) and (3), and find that the coefficient on *ln(Experience)* remains negative and statistically significant. This suggests that both learning-by-doing and peer effects jointly affect tax evasion in the housing market.

Table 9
DID analysis on learning behavior.

Dep. variable	Ratio	
Policy Model	March 30, 2015 Policy (1)	September 30, 2016 Policy (2)
$\ln(\text{Experience})$	-0.102*** (0.005)	-0.094*** (0.006)
$\text{Treat} * \text{After}_{2015}$	0.045*** (0.011)	
$\text{After}_{2015} * \ln(\text{Experience})$	-0.008 (0.011)	
$\text{Treat} * \text{After}_{2015} * \ln(\text{Experience})$	-0.012** (0.005)	
$\text{Treat} * \text{After}_{2016}$		-0.054** (0.022)
$\text{After}_{2016} * \ln(\text{Experience})$		0.006 (0.007)
$\text{Treat} * \text{After}_{2016} * \ln(\text{Experience})$		0.013** (0.006)
Constant	0.584*** (0.008)	0.469*** (0.018)
Observations	76,569	74,560
R-squared	0.355	0.431
Controls	Characteristics of housing, buyer and seller	
Fixed effects	Agent, Complex, DoW, Year-month	

Notes: This table examines agents' learning effect in the policy analysis. The definitions of the variables are the same as in Table 8. Characteristics for housing, buyer, seller, in association with fixed effects for agent, housing complex, days of the week, and year-month are included in all columns. Heteroscedasticity-consistent standard errors are clustered at the complex level as shown in the brackets. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level, respectively.

4.4. Examining agents' response to policy adjustments

In the previous sections, we show that agents' involvement significantly affects the formation of *Yin–Yang* contracts and the magnitude of tax evasion in the secondary housing market. Consequently, an important follow-up question arises: how does agents' participation in the housing market affect social welfare, market efficiency, or real estate development? Since it is impossible to explore the answers to all related questions in one paper due to data limitation and space constraints, we examine the unintended economic consequences stemming from agents' involvement in *Yin–Yang* contracts. To explore this, we utilize two housing policies that aimed at stabilizing the housing market but affecting the minimum down-payment in opposing ways. Specifically, we attempt to answer two questions: (1) how do agents respond to different policies that change buyers' incentive to evade taxes? (2) does agents' involvement lead to increased tax evasion?

4.4.1. A difference-in-differences approach

To establish a causal interpretation, we utilize a difference-in-differences estimation to examine agents' responses to two policy shocks: one implemented on March 30, 2015, and another on September 30, 2016, as discussed in Section 2.2.

The policy introduced in March 2015 is a bundled policy that includes both a 30% reduction in minimum down-payment for mortgage buyers and a 5.5% reduction in the tax rate for properties held for more than two years. To precisely estimate the effects of a down-payment reduction on tax evasion and avoid the possible contamination from the effects of a tax reduction, we exclude the transactions for properties held for over two years. Typically, cash buyers, who do not require mortgage loans, tend to register the price as low as possible to evade taxes. However, mortgage buyers face a trade-off between registering higher prices to secure larger loans from the bank, and registering lower prices to evade taxes. Thus, reducing minimum down-payment is expected to affect the buyers' tax evasion behavior on both the extensive and intensive margins, but in opposite directions. On the extensive margin, a lower required down-payment could convert some cash buyers into mortgage buyers, leading to an increase in registered price and a decrease in tax evasion. To illustrate this possibility, let

us consider two extreme cases: a buyer may choose to get a mortgage loan if the minimum down-payment is 1% only, and another buyer may choose a full payment if the minimum down-payment is 99%. On the intensive margin, among mortgage buyers, a lower down-payment requirement in the post-policy period allows them to register lower prices while still securing the same loan amount as in the pre-policy period, thus resulting in increased tax evasion. For instance, when the minimum down-payment decreased from 70% to 40% in the March 30, 2015 policy, to obtain a mortgage loan of CNY 0.9 million, a buyer needed to register a total price of CNY 3 million before the policy, but only needed to register a total price of CNY 1.5 million after the policy.²¹

In contrast, the September 30, 2016 policy increases the required minimum down-payment from 30%–40% to 35%–70%, offering an opportunity to examine the impact of an increase in minimum down-payment on tax evasion in the resale housing market. On the extensive margin, a higher minimum down-payment requirement could convert some mortgage buyers to cash buyers, leading to a decrease in registered price and an increase in tax evasion. On the intensive margin, among the mortgage buyers, a higher down-payment requirement in the post-policy period requires them to register higher prices to obtain the same loan amount as in the pre-policy period, leading to a decrease in tax evasion. To help readers better understand the potential impacts of down-payment changes on tax evasions, we outline the expected effects of down-payment decrease and down-payment increase on the registered-to-actual price ratio at the extensive and intensive margins in Table 7.

Since down-payment changes in the March 30, 2015 and September 30, 2016 policies targeted mortgage buyers, we can utilize a difference-in-differences method to estimate the impacts of down-payment change

²¹ Before the policy change, with a minimum down-payment of 70%, a property priced at CNY 3 million allowed for a mortgage loan of CNY 0.9 million (=3 million * 30%); and after the policy change, with a minimum down-payment of 40%, a house priced at CNY 1.5 million allowed for the same mortgage loan of CNY 0.9 million (=1.5 million * 60%).

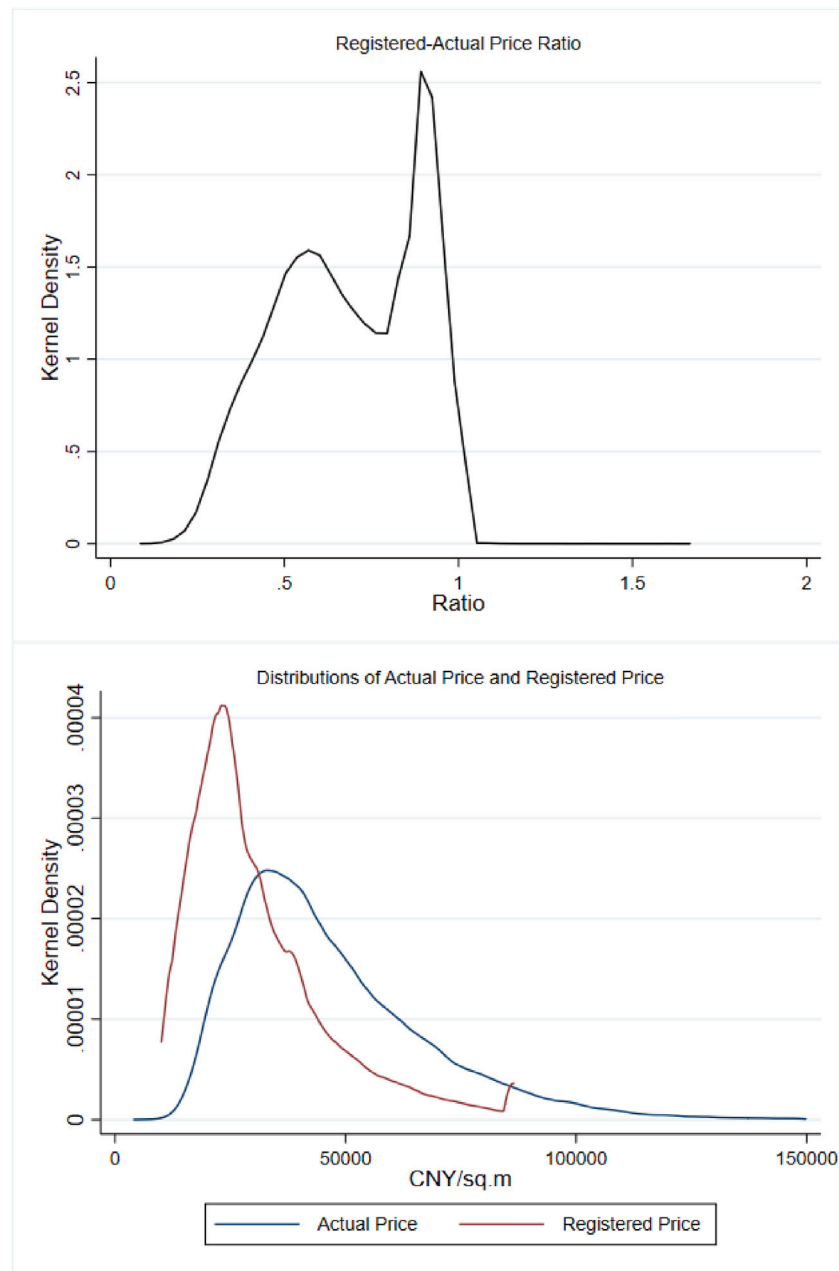


Fig. A.1. The trends of prices and ratio. Notes: This figure presents the distributions of registered-actual price ratio, actual price and registered price.

Table A.1

Comparison of used sample and dropped sample.

	Used sample		Dropped sample		Diff. of mean
	Mean	S.D	Mean	S.D	
Actual unit price	46,033.65	20,684.09	46,463.46	23,091.68	-429.81
Unit size (sq.m)	85.04	39.42	80.89	37.24	4.15
Days on Market	20.60	46.69	23.8	91.22	-3.20
Level of storey	7.44	6.07	7.11	6.91	0.33
# of bedroom	2.05	0.78	1.96	0.77	0.09
# of livingroom	1.18	0.53	1.12	0.51	0.06
# of bathroom	1.20	0.46	1.17	0.43	0.03

Notes: This table compares the characteristics of housings between used sample and dropped sample.

on tax evasion, with the mortgage buyers and cash buyers consisting of the treatment group and control group, respectively. The DID specification is as follows:

$$Y_{i,j,t} = \alpha + \beta_1 * Treat * After + \beta_2 * X_t + \theta_j + \lambda_{agent} + \gamma_{day} + \delta_{ym} + \epsilon_{i,t}$$

(4)
where $Y_{i,j,t}$ is the registered-to-actual price ratio. $Treat$ takes the value of 1 if a mortgage loan is used in the housing purchase, and 0 otherwise. $After$ takes two forms: $After_{2015}$ takes the value of 1 if the transaction

Table A.2

Hedonic regressions with and without agent fixed effect-excluding agents with only 1 transaction.

Panel	Panel A. Without agent FE			Panel B. With agent FE		
	ln(A-Price)	ln(R-Price)	Ratio	ln(A-Price)	ln(R-Price)	Ratio
Dep. variable	(1)	(2)	(3)	(4)	(5)	(6)
Model						
Mortgage	0.001 (0.001)	0.159*** (0.003)	0.095*** (0.002)	0.001 (0.003)	0.153*** (0.007)	0.092*** (0.004)
Observations	177,037	177,037	177,037	177,037	177,037	177,037
R-squared	0.943	0.676	0.388	0.979	0.886	0.776
Controls	Characteristics of housing, buyer and seller					
Agent FE	No	No	No	Yes	Yes	Yes
Complex FE	Yes	Yes	Yes	Yes	Yes	Yes
Day of the week FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-month FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table examines changes of R-square with/without including the agent fixed effect. Characteristics for housing, buyer, seller, in association with fixed effects for agent, housing complex, days of the week, and year-month are included in all columns. Heteroscedasticity-consistent standard errors are clustered at the complex level as shown in the brackets. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level, respectively.

Table A.3

Incentive of agents.

Dep. variable	ln(DoM)	ln(Total Transactions)
Model	(1)	(2)
Ratio	0.111* (0.059)	-1.266*** (0.049)
Observations	179,580	15,802
R-squared	0.808	0.101
Controls	Characteristics of housing, buyer and seller	Agent characteristics
Fixed effects	Agent, Complexes, DoW, Year-month	

Notes: This table examines the incentive of the agent in doing *Yin-Yang* contracts by regressing the properties' days on market against the registered-actual-price ratio in Column (1) and by regressing the agents' total transactions against their average registered-actual-price ratio in Column (2). Characteristics for housing, buyer, and seller, in association with fixed effects for agent, housing complex, days of the week, and year-month are included in Column (1); and characteristics for agents are included in Column (2). Heteroscedasticity-consistent standard errors are clustered at the complex level in Column (1). We use ***, **, and * to denote significance at the 1%, 5%, and 10% level, respectively.

Table A.4

The impact of experience on days on market.

Dep. variable	ln(DoM)	
Model	(1)	(2)
ln(Experience)	-0.126*** (0.045)	
First		0.014*** (0.004)
Observations	179,580	179,580
R-squared	0.807	0.808
Controls	Characteristics of housing, buyer and seller	
Fixed effects	Agent, Complex, DoW, Year-month	

Notes: This table reports the agent learning effect on properties' days on market (ln(DoM)). Characteristics for housing, buyer and seller, in association with fixed effects for agent, housing complex, days of the week, and year-month are included in all columns. Heteroscedasticity-consistent standard errors are clustered at the complex level as shown in the brackets. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level, respectively.

took place after March 30, 2015, and *After*₂₀₁₆ takes the value of 1 if the transaction took place after September 30, 2016. The other variables are defined the same as in Eq. (1). To minimize contamination across policies and ensure a long test window, the sample period for the analysis of March 30, 2015 policy is set between March 30, 2014 and March 31, 2016, and the sample period for the analysis of September 30, 2016 policy is set between April 1, 2016 and May 30, 2017.

Table 8 presents the DID estimation results. The estimated coefficient in Column (1) suggests that the reduction of down-payment in 2015 increases the registered-to-actual price ratio by 3.1 percentage points. In Column (2), the estimated coefficient indicates that the increase in down-payment in 2016 leads to a 4.5 percentage points reduction in the registered-to-actual price ratio. Recall that the extensive and intensive margins are affected in opposite directions by changes in down-payment policies, as shown in Table 7, the positive response of the registered-to-actual price ratio to the 2015 policy and the negative response to the 2016 policy in Table 8 demonstrate that the effect on the extensive margin dominates.

The results in Table 8 may raise concerns regarding the potential composition change between treatment and control groups in the DID analysis: a decrease (increase) in down-payment would cause a shift of cash (mortgage) buyers into mortgage (cash) buyers. Although the composition change might lead to an overestimation of the change in transaction volume, as illustrated in Appendix B, it does not harm our analysis because our objective is to estimate the policy impact on tax evasion. Specifically, cash buyers, who do not require mortgage loans, tend to register the price as low as possible, so the average magnitudes of tax evasion within the control group should stay constant. Therefore, although the change in transaction volume may be unambiguous in the presence of the composition change, the result of registered-to-actual price ratio remains meaningful because the mean value of registered-to-actual price ratio has been changed by the down-payment policies, regardless of whether the change comes from the extensive margin or the intensive margin. Moreover, the results in Table 8 help us verify that the extensive margin is the main source of the impact of minimum down-payment change on tax evasion, even though the respective effects of extensive margin and intensive margin cannot be quantified at this stage.

4.4.2. Agent's experience and two policies

In this section, we delve deeper into whether the involvement of experienced real estate agents exacerbates or alleviates the tax evasion problem following the implementation of the two policies. We add a triple interaction term, *Treat * After * ln(Experience)*, to Eq. (4) and examine its impact. The results are reported in Table 9. The coefficient on *Treat * After*₂₀₁₅ * *ln(Experience)* in Column (1) appears to be significantly negative, indicating that the transactions involving agents with more experience tend to register lower prices or evade more taxes when the required down-payments are reduced. Conversely, as depicted in Column (2), the effect for the September 30, 2016 policy is symmetric: when the minimum down-payment increases, transactions involving agents with more experience tend to have higher registered prices or evade fewer taxes.

Thus far, results in Table 9 offer insights into the questions posed at the beginning of Section 4.4. First, experienced agents exhibit a

Table A.5
Heterogeneous tests of learning-by-doing effects.

Panel A:	(1)	(2)	(3)	(4)	(5)	(6)
	POE developer			SOE developer		
Dep. variable	ln(A-Price)	ln(R-Price)	Ratio	ln(A-Price)	ln(R-Price)	Ratio
ln(Distance)	0.051 (0.032)	0.014 (0.043)	−0.026** (0.011)	0.161*** (0.032)	0.091* (0.048)	−0.049** (0.024)
Observations	111,137	111,137	111,137	68,443	68,443	68,443
R-squared	0.956	0.754	0.526	0.951	0.721	0.465
Panel B:	Complex age ≥ 10 yrs			Complex age < 10 yrs		
Dep. variable	ln(A-Price)	ln(R-Price)	Ratio	ln(A-Price)	ln(R-Price)	Ratio
ln(Distance)	0.012 (0.037)	−0.032 (0.044)	−0.030** (0.012)	0.225*** (0.070)	0.138* (0.084)	−0.056** (0.027)
Observations	75,076	75,076	75,076	104,504	104,504	104,504
R-squared	0.961	0.717	0.441	0.943	0.718	0.520
Panel C:	Below average price			Above average price		
Dep. variable	ln(A-Price)	ln(R-Price)	Ratio	ln(A-Price)	ln(R-Price)	Ratio
ln(Distance)	0.056 (0.037)	0.012 (0.024)	−0.032** (0.015)	0.149*** (0.010)	0.082** (0.033)	−0.046** (0.022)
Observations	91,573	91,573	91,573	88,007	88,007	88,007
R-squared	0.874	0.652	0.474	0.909	0.564	0.466
Controls	Characteristics of housing, buyer and seller					
Fixed effects	Agent, DoW, Year-month					

Notes: This table examines the heterogeneous effects of the monitoring capability of local tax authorities, proxied by the distance between complexes and their nearest local tax authority, on actual unit price, registered unit price, and registered-to-actual price ratio. In Panel A, we categorize real estate developers of complexes into two groups: privately owned enterprises (POE) and state-owned enterprises (SOE); in Panel B, we group complexes based on building age: aged above 10 years and aged below 10 years; and in Panel C, we group complexes into two groups based on the average price levels: below and above the average transaction price. Characteristics for housing, buyer and seller, in association with fixed effects for agent, days of the week, and year-month are included in all columns. Heteroscedasticity-consistent standard errors are clustered at the complex level as shown in the brackets. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level, respectively.

Table A.6
Hedonic regressions with and without branch fixed effect.

Panel	Panel A. Without branch FE			Panel B. With branch FE		
Dep. variable	ln(A-Price)	ln(R-Price)	Ratio	ln(A-Price)	ln(R-Price)	Ratio
Model	(1)	(2)	(3)	(4)	(5)	(6)
Mortgage	0.000 (0.001)	0.158*** (0.003)	0.094*** (0.002)	0.001 (0.002)	0.155*** (0.004)	0.093*** (0.002)
Observations	179,580	179,580	179,580	179,580	179,580	179,580
R-squared	0.942	0.678	0.385	0.962	0.791	0.598
Controls	Characteristics of housing, buyer and seller					
Branch FE	No	No	No	Yes	Yes	Yes
Complex FE	Yes	Yes	Yes	Yes	Yes	Yes
Day of the week FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-month FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table examines changes of R-square with/without including the branch fixed effect. Characteristics for housing, buyer, and seller, in association with fixed effects for complex, days of the week, and year-month are included in all columns. Heteroscedasticity-consistent standard errors are clustered at the complex level as shown in the brackets. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level, respectively.

symmetric response to the decrease and increase in minimum down-payment. Second, the involvement of experienced agents in a housing transaction could may lead to unintended consequences of housing policies: their expertise contributes to more tax evasion when minimum down-payments are reduced, and *vice versa*.

4.4.3. Parallel pre-trends and falsification tests

We also examine the parallel trend assumption by adding an interaction term of $Treat * Before$ to Eq. (4). $Before$ is a dummy equal to 1 for the period four weeks before the policy date. The results are presented in Table A.7. For both policy analyses, the interactions of $Treat * Before$ are statistically insignificant and economically indifferent from zero, suggesting that the parallel trend assumption is satisfied in our DID setting.

We conduct a falsification test to investigate potential spurious results over time. We treat the pre-policy period as a new sample period

and create a fictitious policy date during the new sample period. Doing so allows us to examine whether the DID effects merely reflect more general time patterns, such as those associated with real estate cycles or macro-prudential restrictions on the housing market. We replace the actual policy implementation dates with the fictitious date, set four months earlier, and then re-estimate the models. The results are reported in Table A.8. The interaction terms are statistically insignificant, indicating that the observed changes in tax evasion are indeed triggered by the actual policy shocks.

5. Heterogeneity tests

We finally conduct the heterogeneity tests of the learning-by-doing effect and peer effect across different types of transactions. Table A.9 presents the heterogeneity tests for the learning-by-doing effect. More specifically, the learning-by-doing effect is significantly negative at

Table A.7
Tests for the parallel trend assumption.

Dep. variable	Ratio	
Policy Model	March 30, 2015 Policy (1)	September 30, 2016 Policy (2)
<i>Treat * Before</i> ₂₀₁₅	0.009 (0.008)	
<i>Treat * After</i> ₂₀₁₅	0.033*** (0.004)	
<i>Treat * Before</i> ₂₀₁₆		−0.014 (0.013)
<i>Treat * After</i> ₂₀₁₆		−0.045*** (0.002)
Observations	76,569	74,560
R-squared	0.345	0.431
Controls	Characteristics of housing, buyer and seller	
Fixed effects	Agent, Complex, DoW, Year-month	

Notes: This table examines the parallel trend assumption. *Before*₂₀₁₅ is dummy equal to 1 if the transaction took place during the period of four weeks before March 30, 2015. *Before*₂₀₁₆ is a dummy equal to 1 if the transaction took place during the period of four weeks before September 30, 2016. Other variables are the same as in previous tables. Characteristics for housing, buyer and seller, in association with fixed effects for agent, housing complex, days of the week, and year-month are included in all columns. Heteroscedasticity-consistent standard errors are clustered at the complex level as shown in the brackets. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level, respectively.

Table A.8
Placebo tests.

Dep. variable	Ratio	
Policy Model	March 30, 2015 Policy (1)	September 30, 2016 Policy (2)
<i>Treat * After</i> ₂₀₁₅	0.006 (0.011)	
<i>Treat * After</i> ₂₀₁₆		0.002 (0.003)
Observations	17,286	32,642
R-squared	0.741	0.372
Controls	Characteristics of housing, buyer and seller	
Fixed effects	Agent, Complex, DoW, Year-month	

Notes: This table reports the results of a placebo test. The pre-policy period is used as sample period. The artificial date is set as three months before the actual policy dates. Characteristics for housing, buyer and seller, in association with fixed effects for agent, housing complex, days of the week, and year-month are included in all columns. Heteroscedasticity-consistent standard errors are clustered at the date level as shown in the brackets. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level, respectively.

Table A.9
Heterogeneity tests of the learning-by-doing effect.

Dep. variable	Ratio			
Group Model	Cash Trans. (1)	Mortgage Trans. (2)	Male agent (3)	Female agent (4)
ln(Experience)	−0.027 (0.028)	−0.042*** (0.012)	−0.051*** (0.013)	−0.029 (0.021)
Constant	0.729*** (0.068)	1.241*** (0.026)	1.028*** (0.029)	1.066*** (0.043)
Observations	23,587	138,671	116,059	46,199
R-squared	0.539	0.342	0.393	0.437
Controls	Characteristics of housing, buyer and seller			
Fixed effects	Agent, Complex, DoW, Year-month			

Notes: This table examines the heterogeneity of agent learning effects across different agent groups. The dependent variable is registered-actual price ratio (Ratio). The headers in the second row denote the sample group used in respective column. Characteristics for housing, buyer and seller, in association with fixed effects for agent, housing complex, days of the week, and year-month are included in all columns. Heteroscedasticity-consistent standard errors are clustered at the complex level as shown in the brackets. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level, respectively.

Table A.10
Heterogeneity tests of peer effect (2SLS).

Dep. variable	Ratio			
Group Model	Male agent (1)	Female agent (2)	Small branch (3)	Big branch (4)
Tenure of Peers	0.698*** (0.152)	0.242** (0.120)	0.181 (0.360)	0.550* (0.297)
Observations	113,075	44,402	118,410	38,960
R-squared	0.177	0.228	0.213	0.247
Controls	Characteristics of housing, buyer and seller			
Fixed effects	Agent, Complex, DoW, Year-month			

Notes: This table examines the heterogeneity of peer effect across different agent groups. The dependent variable is registered-actual price ratio (Ratio). The headers in the second row denote the sample group used in respective column. Characteristics for housing, buyer and seller, in association with fixed effects for agent, housing complex, days of the week, and year-month are included in all columns. Heteroscedasticity-consistent standard errors are clustered at the complex level as shown in the brackets. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level, respectively.

−0.042 for transactions with mortgage loans, while bearing a small and insignificant coefficient (−0.027) for transactions with full payments. Additionally, we find that experienced male agents are more aggressively involved in creating tax evasion, while the learning-by-doing effect for female agents is insignificant.

Table A.10 reports the heterogeneity tests for the peer effect. The peer effects are statistically significant for both male and female agents, although the effects are more prominent in magnitude for the former group. In the last two columns, we find that the peer effects are more prominent in larger branches. Table A.11 investigates the impact of policy changes in various sub-groups based on the gender and residency status of buyers, and the results are statistically significant in all sub-groups.

6. Conclusion

While agents' information advantage and the principal-agent problem have been extensively discussed in the literature, the role of agents in fraudulent or illegal practices remains poorly understood. This study fills the gap by answering three questions: Do agents engage in fraudulent and illegal conduct? If so, what are the underlying mechanisms? Additionally, what are the economic consequences of agents' involvement in these fraudulent activities?

We employ China's resale housing market as a unique experimental setting, characterized by the widespread but illegal practice of the *Yin-Yang* contracts. Despite the efforts by the Chinese government to curb the prevalence of *Yin-Yang* contracts, progress has been limited. Buyers and sellers collaborate on price manipulation strategies to minimize transaction taxes. Meanwhile, local governments exhibit limited motivation to combat these illegal practices, as they may perceive limited intervention as a more favorable option after weighing the costs and benefits of various levels of intervention. Real estate agents, who have information advantages in the housing market, can advise buyers and sellers on an appropriate price to report to the local registration offices. Therefore, it is important to investigate the role of agents in the practice of *Yin-Yang* contracts and the resulting tax evasion.

We find that agents learn from their past transactions and peers in suggesting registered prices, leading to substantial tax evasion. We also examine the unintended consequences of agents' involvement by analyzing two policy shocks in March 2015 and September 2016. Our findings indicate that agents' experience significantly influences the reporting of lower prices to local registration authorities when down-payments are reduced, and *vice versa*. The study shows that with policy adjustments, agents' involvement magnifies the tax evasion problem. This research presents solid empirical evidence on agents' fraudulent and illegal behaviors in the real estate market, identifies their role in aiding tax evasion, and issues critical warnings to the tax and real estate regulatory authorities.

Table A.11
Heterogeneity tests of policy impact.

Panel	Panel A. March 30, 2015				Panel B. September 30, 2016			
Dep. variable	Ratio				Ratio			
Buyer group	Male	Female	Local	Non-local	Male	Female	Local	Non-local
Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Treat * After</i> ₂₀₁₅	0.031*** (0.006)	0.034*** (0.006)	0.018*** (0.006)	0.034*** (0.007)				
<i>Treat * After</i> ₂₀₁₆					−0.044*** (0.003)	−0.047*** (0.004)	−0.060*** (0.004)	−0.039*** (0.003)
Observations	39,999	36,570	27,852	48,717	38,131	36,429	25,817	48,743
R-squared	0.374	0.389	0.415	0.332	0.466	0.467	0.462	0.418
Controls	Characteristics of housing, buyer and seller							
Fixed effects	Agent, Complex, DoW, Year-month							

Notes: This table examines the heterogeneity of policy effect across different buyer groups. Panels A and B correspond to the March 30, 2015 policy, and the September 30, 2016 policy, respectively. The dependent variable is registered-actual price ratio (Ratio). The headers in the third row denote the sample group used in respective column. Characteristics for housing, buyer and seller, in association with fixed effects for agent, housing complex, days of the week, and year-month are included in all columns. Heteroscedasticity-consistent standard errors are clustered at the complex level as shown in the brackets. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level, respectively.

Table A.12
DID analysis on transaction volume.

Dep. variable	ln(Vol)	
Policy Model	March 30, 2015 Policy (1)	September 30, 2016 Policy (2)
<i>Treat * After</i> ₂₀₁₅	0.448*** (0.043)	
<i>Treat * After</i> ₂₀₁₆		−0.132* (0.077)
Observations	1429	837
R-squared	0.865	0.810
Daily FE	Yes	Yes

Notes: This table examines the effects of March 31, 2015 Policy and September 30, 2016 Policy on daily transaction volume. The sample period for March 30, 2015 Policy is between March 30, 2014 and March 31, 2016. The sample period for September 30, 2016 Policy is between April 1, 2016 and May 30, 2017. In Columns (1), the transactions of houses held between two and five years are excluded and *Treat* and is a dummy equal to 1 if the houses are bought via mortgage, and 0 otherwise. *After*₂₀₁₅ is a dummy equal to 1 if the transaction took place after March 30, 2015, and 0 otherwise. *After*₂₀₁₆ is a dummy equal to 1 if the transaction took place after September 30, 2016, and 0 otherwise. Fixed effects for days of the week and year-month are included in all columns. We use ***, **, and * to denote significance at the 1%, 5%, and 10% level, respectively.

CRedit authorship contribution statement

Sumit Agarwal: Conceptualization, Methodology, Project administration, Supervision, Validation. **Weida Kuang:** Conceptualization, Data curation, Validation. **Long Wang:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Yang Yang:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary figures and tables

See Fig. A.1 and Tables A.1–A.12.

Appendix B. The impact of two policies on transaction volume

In the difference-in-difference estimations, we examine the policy impact on daily transaction volume using the following specification:

$$Vol_{m,t} = \alpha + \beta_1 * Treat * After + \beta_2 * Treat + \delta_t + \epsilon_{m,t} \quad (5)$$

where $Vol_{m,t}$ in Eq. (5) represents the number of transactions for group m on date t . m takes the value of 1 for the treatment group, and 0 otherwise. The daily fixed effect, δ_t , is included in Eq. (5). *Treat* takes the value of 1 if a mortgage loan is used in the housing purchase; and 0 otherwise. *After*₂₀₁₅ takes the value of 1 if the transaction took place after March 30, 2015; and *After*₂₀₁₆ takes the value of 1 if the transaction took place after September 30, 2016. The other variables are defined as the same as in Eq. (1). To avoid contamination from other policies and assure a long test window, the sample period for the analysis of March 30, 2015 policy is from March 30, 2014 to March 31, 2016; and the sample period for the analysis of September 30, 2016 policy is from April 1, 2016 to May 30, 2017.

Table A.12 presents the estimation results. In Columns (2) and (3), the reduction of minimum down-payment increases the transaction volume of properties with mortgage loans by $(\exp(0.448) - 1) = 56.5\%$, and that the increase of minimum down-payment leads to a $(1 - \exp(-0.132)) = 12.4\%$ reduction of transactions with mortgage loans, compared to transactions in full cash payments.

We want to emphasize that we are cautious to interpret the DID estimation results in Table A.12 because the statistically significant change in transaction volume could be a result of composition change between treatment and control group rather than a genuine increase or decrease in the treatment group after the policy. For example, the cash buyers may switch to use mortgage in the presence of down-payment reduction in the 2015 policy, resulting in more mortgage buyers coming from cash buyers. If this is the case, then the estimated coefficients in Table A.12 are upward biased. However, we believe this potential composition change does not affect our interpretation of the tax evasion results in Table 8.

Appendix C. Heterogeneous tests of learning-by-doing effects

The results are presented in Table A.5. In Panel A of Table A.5, we categorize real estate developers of complexes into two groups: privately owned enterprises (POE) and state-owned enterprises (SOE); in Panel B of Table A.5, we group complexes based on building age: aged above 10 years and aged below 10 years; and in Panel C of Table A.5, we group complexes into two groups based on the average price levels: below and above the average transaction price.

We find that the monitoring capability (the coefficient on $\ln$ (distance)) is relatively weaker for SOE complexes, complexes aged less

than 10 years, and complexes that priced above market average, than for their respective counterparts. The results of Table A.5 suggest that the government imposes less supervision on the expensive complexes, as both the complexes with SOE developers and the new complexes are relatively more expensive. We would like to acknowledge that we do not know the specific reason why the monitoring capability is weaker for more expensive complexes. One possible explanation is that the market prices of expensive complexes increased at a faster rate than the internal guideline prices, which were not regularly updated by the registration offices. Nevertheless, since this part is beyond the scope of this paper, we leave this for future research when more detailed data becomes available.

References

- Agarwal, S., Choi, H.-S., He, J., Sing, T.F., 2019a. Matching in housing markets: The role of ethnic social networks. *Rev. Financ. Stud.* 32 (10), 3958–4004.
- Agarwal, S., Hauswald, R., 2010. Distance and private information in lending. *Rev. Financ. Stud.* 23 (7), 2757–2788.
- Agarwal, S., He, J., Sing, T.F., Song, C., 2019b. Do real estate agents have information advantages in housing markets? *J. Financ. Econ.* 134 (3), 715–735.
- Agarwal, S., Li, K., Qin, Y., Wu, J., Yan, J., 2020. Tax evasion, capital gains taxes, and the housing market. *J. Public Econ.* 188, 104222.
- Angrist, J.D., Krueger, A.B., 2001. Instrumental variables and the search for identification: From supply and demand to natural experiments. *J. Econ. Perspect.* 15 (4), 69–85.
- Arcidiacono, P., Nicholson, S., 2005. Peer effects in medical school. *J. Public Econ.* 89 (2–3), 327–350.
- Artavanis, N., Morse, A., Tsoutsoura, M., 2016. Measuring income tax evasion using bank credit: Evidence from Greece. *Q. J. Econ.* 131 (2), 739–798.
- Balafoutas, L., Beck, A., Kerschbamer, R., Sutter, M., 2015. The hidden costs of tax evasion: Collaborative tax evasion in markets for expert services. *J. Public Econ.* 129, 14–25.
- Barkema, H.G., Schijven, M., 2008. How do firms learn to make acquisitions? A review of past research and an agenda for the future. *J. Manag.* 34 (3), 594–634.
- Barwick, P.J., Pathak, P.A., 2015. The costs of free entry: an empirical study of real estate agents in Greater Boston. *Rand J. Econ.* 46 (1), 103–145.
- Barwick, P.J., Pathak, P.A., Wong, M., 2017. Conflicts of interest and steering in residential brokerage. *Am. Econ. J.: Appl. Econ.* 9 (3), 191–222.
- Bayer, P., Hjalmarsson, R., Pozen, D., 2009. Building criminal capital behind bars: Peer effects in juvenile corrections. *Q. J. Econ.* 124 (1), 105–147.
- Benjamin, J.D., Chinloy, P., Winkler, D.T., 2009. Labor supply, flexible hours and real estate agents. *Real Estate Econ.* 37 (4), 747–767.
- Bolton, P., Freixas, X., Shapiro, J., 2007. Conflicts of interest, information provision, and competition in the financial services industry. *J. Financ. Econ.* 85 (2), 297–330.
- Chetty, R., 2009. Is the taxable income elasticity sufficient to calculate deadweight loss? The implications of evasion and avoidance. *Am. Econ. J.: Econ. Policy* 1 (2), 31–52.
- Coval, J.D., Moskowitz, T.J., 2001. The geography of investment: Informed trading and asset prices. *J. Polit. Econ.* 109 (4), 811–841.
- Cuyper, I.R., Cuyper, Y., Martin, X., 2017. When the target may know better: Effects of experience and information asymmetries on value from mergers and acquisitions. *Strateg. Manag. J.* 38 (3), 609–625.
- Dahl, G.B., Løken, K.V., Mogstad, M., 2014. Peer effects in program participation. *Amer. Econ. Rev.* 104 (7), 2049–2074.
- Dai, Y., Xu, Y., 2022. Cheating under regulation: Evidence from “Yin-and-Yang” contracts on Beijing’s housing market. *Sustainability* 14 (20), 13346.
- Dimmock, S.G., Gerken, W.C., Graham, N.P., 2018. Is fraud contagious? Coworker influence on misconduct by financial advisors. *J. Finance* 73 (3), 1417–1450.
- Dvořák, T., 2005. Do domestic investors have an information advantage? Evidence from Indonesia. *J. Finance* 60 (2), 817–839.
- Eckardt, M., Rätthke-Döppner, S., 2010. The quality of insurance intermediary services—empirical evidence for Germany. *J. Risk Insurance* 77 (3), 667–701.
- Erard, B., 1993. Taxation with representation: An analysis of the role of tax practitioners in tax compliance. *J. Public Econ.* 52 (2), 163–197.
- Fan, Z., Wang, X., Zhang, H., 2022. Understanding misreporting: New evidence from responses to a housing transaction tax.
- Fang, H., Gu, Q., Xiong, W., Zhou, L.-A., 2016. Demystifying the Chinese housing boom. *NBER Macroecon. Annu.* 30 (1), 105–166.
- Feng, L., Seasholes, M.S., 2005. Do investor sophistication and trading experience eliminate behavioral biases in financial markets? *Rev. Finance* 9 (3), 305–351.
- Fisman, R., Wei, S.-J., 2004. Tax rates and tax evasion: evidence from “missing imports” in China. *J. Polit. Econ.* 112 (2), 471–496.
- Keen, M., Toro, J., Baer, K., Perry, V., Norregaard, J., Ueda, J., Brondolo, J., Cleary, D., Hutton, E., Luca, O., et al., 2015. Current Challenges in Revenue Mobilization: Improving Tax Compliance. *Staff Rep.*, pp. 1–79.
- Knill, A., Minnick, K.L., Nejadmalayeri, A., 2015. Experience, information asymmetry, and rational forecast bias. In: *Handbook of Financial Econometrics and Statistics*. Springer, pp. 63–100.
- Levitt, S.D., Syverson, C., 2008. Market distortions when agents are better informed: The value of information in real estate transactions. *Rev. Econ. Stat.* 90 (4), 599–611.
- Loewenstein, J., Thompson, L.L., 2006. Learning to negotiate: Novice and experienced negotiators. In: *Negotiation Theory and Research*. Psychology Press, New York, pp. 77–97.
- Lu, Z., Zhang, S., Zhao, D., 2022. Tax me if you can: Tax evasion on Chinese housing market.
- Manski, C.F., 1993. Identification of endogenous social effects: The reflection problem. *Rev. Econom. Stud.* 60 (3), 531–542.
- Manski, C.F., 2000. Economic analysis of social interactions. *J. Econ. Perspect.* 14 (3), 115–136.
- Marion, J., Muehlegger, E., 2008. Measuring illegal activity and the effects of regulatory innovation: Tax evasion and the dyeing of untaxed diesel. *J. Polit. Econ.* 116 (4), 633–666.
- Mehran, H., Stulz, R.M., 2007. The economics of conflicts of interest in financial institutions. *J. Financ. Econ.* 85 (2), 267–296.
- Merriman, D., 2010. The micro-geography of tax avoidance: evidence from littered cigarette packs in Chicago. *Am. Econ. J.: Econ. Policy* 2 (2), 61–84.
- Nadler, J., Thompson, L., Boven, L.V., 2003. Learning negotiation skills: Four models of knowledge creation and transfer. *Manage. Sci.* 49 (4), 529–540.
- Pan, J.-N., Huang, J.-T., Chiang, T.-F., 2015. Empirical study of the local government deficit, land finance and real estate markets in China. *China Econ. Rev.* 32, 57–67.
- Reuer, J.J., Ragozzino, R., 2008. Adverse selection and M&A design: The roles of alliances and IPOs. *J. Econ. Behav. Organ.* 66 (2), 195–212.
- Rosen, S., 1974. Hedonic prices and implicit markets: product differentiation in pure competition. *J. Polit. Econ.* 82 (1), 34–55.
- Somerville, T., Wang, L., Yang, Y., 2019. Using purchase restrictions to cool housing markets: A within-market analysis. *J. Urban Econ.* 103189.
- Van Nieuwerburgh, S., Veldkamp, L., 2009. Information immobility and the home bias puzzle. *J. Finance* 64 (3), 1187–1215.
- Wang, Y., Hui, E.C.-m., 2017. Are local governments maximizing land revenue? Evidence from China. *China Econ. Rev.* 43, 196–215.
- Wang, L., Yang, Y., 2021. Political connections in the land market: Evidence from China’s state-owned enterprises. *Real Estate Econ.* 49 (1), 7–35.
- Wu, J., Gyourko, J., Deng, Y., 2012. Evaluating conditions in major Chinese housing markets. *Reg. Sci. Urban Econ.* 42 (3), 531–543.