Matching People to Properties: A Matching Function for the Housing Market*

L. Rachel Ngai Felix Iglhaut Kevin D. Sheedy

Fei Wu Shuang Zhang

April 2026

Abstract

Housing markets are subject to search frictions. The matching function is a widely used tool in macroeconomics to summarize labour-market search frictions and has been estimated in an extensive literature. Limited data on home-buyers has precluded estimation of housing-market matching functions. This paper fills that gap using a novel dataset identifying buyers and their search behaviour in viewing properties. An event study based on the staggered removal of restrictions on housing purchases provides the first causal estimation of the elasticity of matches with respect to buyers for a constant-returns-to-scale Cobb-Douglas matching function, which is found to be around 0.5.

Keywords: search frictions, matching function, housing market

JEL codes: D83, R21, R31

*We thank Qian Zhang and Xiangyu Zhou for excellent research assistance, and Ting Chen for sharing information on Chinese housing-market regulations. We thank Russell Cooper, Maarten De Ridder, Essi Eerola, Hanming Fang, Eric French, David Genesove, Jim Hamilton, Leo Kaas, Barbara Petrongolo, Chris Pissarides, Valerie Ramey, Wenxuan Wang, Shu Yang, Eran Yarshiv, and Weilong Zhang for helpful discussions. We have benefited from the comments of participants at the Bristol Macro Workshop 2025, Cambridge University, EUI, Helsinki GSB, LSE, Rady School of Management UCSD, and the Workshop of the Australasian Macroeconomics Society 2025.

L. Rachel Ngai: London School of Economics, CfM, IZA, and CEPR

Felix Iglhaut: London School of Economics

Kevin D. Sheedy: London School of Economics and CfM

Fei Wu: Shanghai Advanced Institute of Finance

Shuang Zhang: Imperial College London and CEPR

1 Introduction

Search frictions are central to markets for labour, goods, money, marriage, credit, and housing where buyers and sellers do not match instantaneously. The concept of a matching function is a powerful tool to capture the complexity of the search process on both sides of the market in a tractable way. The matching function describes the number of transactions between buyers and sellers with positive total surplus that are realized in a period of time given the search technology, independent of how the surplus is divided. As formalized by Pissarides (1979), it acts as a production function for matches conditional on the numbers of buyers and sellers, analogous to how factor inputs are combined to produce output. This key innovation established search models as the canonical framework for studying the labour market in macroeconomics to address issues such as unemployment, business cycles, and policy.[1] The objective of this paper is to provide empirical evidence on the housing-market matching function.

The housing market shares the heterogeneity of buyers and sellers and imperfect information seen in the labour market (Clark, 1982), and since the seminal work of Wheaton (1990), a growing literature has adopted the search framework to study interactions between housing markets and macroeconomics. A matching function between home-buyers and properties for sale determines the number and speed of transactions.[2] This is important because transactions represent gains from trade that reduce housing mismatch, and the speed with which they occur affects the duration of mismatch. Typically, researchers assume a Cobb-Douglas, constant-returns-to-scale matching function, often imposing an elasticity with respect to buyers of 0.5 or calibrating indirectly. However, empirical evidence on the housing-market matching function remains scarce due to the lack of direct data on buyers. Genesove and Han (2012) infer the ratio of buyers to sellers (market tightness) indirectly from surveys of successful buyers and sellers, while recent work using land registry data or online platforms captures only a partial component of buyer activity.[3]

[1]An influential survey by Petrongolo and Pissarides (2001) of more than 40 papers that investigated empirically the functional form or parameters of the labour-market matching function concluded that it has a Cobb-Douglas functional form, constant returns to scale, and an elasticity with respect to vacancies (buyers) of between 0.3 and 0.5.

[2]Recent papers that use a matching function for the housing market include Novy-Marx (2009), Piazzesi and Schneider (2009), Díaz and Jerez (2013) Head et al. (2014), Burnside et al. (2016), Guren (2018), Gabrovski and Ortego-Marti (2019), Garriga and Hedlund (2020), Guren and McQuade (2020), Anenberg and Bayer (2020), Piazzesi et al. (2020), Han et al. (2025), and Kaas et al. (2026).

[3]Grindaker et al. (2025) use Norwegian Land Registry data to infer changes in tightness from the fraction of movers who buy first. Examples from online platforms include the use of clicks on property listings in van Dijk and Francke (2018) for the Netherlands, email alerts in Piazzesi et al. (2020) for the San Francisco Bay Area, clicks on the contact button in Kaas et al. (2026) for Germany, and user IP addresses in Badarinza et al. (2024) for the UK.

This paper makes two contributions in estimating the housing-market matching function. First, it exploits a unique dataset that combines information from an online platform together with offline viewing data in a large housing market. Data on in-person viewings of properties directly identifies buyers and their search effort without relying on proxies commonly used in the literature such as online clicks or surveys. Second, the paper provides the first causal estimation of the matching-function elasticity in the housing market. It uses an event study that leverages the staggered removal over time of legal restrictions on housing purchases across districts in two large provinces of China. This policy experiment provides a source of variation in buyer activity that is plausibly exogenous to local market conditions. The results support an elasticity of matches with respect to buyers of approximately 0.5 for a Cobb-Douglas matching function with constant returns to scale.

The data are from the largest real-estate transaction platform in China, which integrates both online listings and in-person viewings. The platform's online app is similar to Zillow in the US and Rightmove in the UK. A key difference is that all offline viewings are recorded on the platform's integrated system, and in-person viewings are linked to unique buyer and listing IDs. Therefore, the focus is on users who actually view properties and engage as active buyers in the housing market, as opposed to users who merely browse online.

Accounting for search effort is an important challenge when estimating matching functions. Here, we make progress because the data allow us to construct a measure of effective buyers based on the number of viewings made by active buyers in a given period of time. The output of the matching function is the number of transactions, and data on this are also available based on the date a contract is signed.[4] Hence, together with the stock of listings, we have data on all the variables appearing in the matching function.

The data come from 15 major cities in two provinces of China with a total population of over 100 million people. Listings are grouped geographically by districts, which are subregions of cities. The sample used in our analysis includes 116 districts, with a median population of 785,000 and a median area of 1,440km^2. Since we can show that buyers' search radius is largely confined to a single district, we aggregate listings, effective buyers, and transactions within districts. Our interest is to estimate an aggregate matching function, so we do not estimate separate matching functions for different types of housing, such as large or small apartments within a district. A more fundamental reason is that housing has multiple attributes that are substitutable for buyers, for example, a small apartment in a certain neighbourhood could be as attractive as a large apartment in another neighbourhood. This means it is not appropriate to estimate independent matching functions for different

[4]It is well known that there can be a substantial difference between the contract date and the closing date when ownership of a property is transferred (Anenberg and Laufer, 2017).

housing-market segments within a district.[5] The time period covered by the data is November 2021 to December 2024, and variables are aggregated within months.

First, following earlier literature, we estimate the functional form of the matching function using OLS panel regressions that include both district and time fixed effects. The fixed effects remove common macroeconomic trends and fluctuations and control for heterogeneity across districts. This is to address concerns about unobserved variation in matching efficiency that might be correlated with the numbers of buyers or sellers, but we are aware of the limitations of this approach and discuss them later. Starting with a trans-log specification, we show that second-order terms are statistically insignificant, and thus fail to reject a Cobb-Douglas (log linear) functional form. We also do not reject the hypothesis that the matching function has constant returns to scale (CRS). Therefore, we proceed by assuming a CRS Cobb-Douglas matching function.

Our main empirical contribution leverages changes in legal restrictions on buying properties in China that vary at the district level to conduct an event study. The Chinese government has long imposed tight controls on the housing market. Traditionally, there are restrictions on who can purchase property where, and how many properties they can own. While such restrictions have been in place for a long time, recent years have seen changes in policy that led to several of these restrictions being lifted across China. Importantly, while the initial trigger for this change was a central announcement to stimulate housing demand at the national level, the timing of the legislative lifting of restrictions varied across cities and districts. We provide evidence that the district-level policy timings are not correlated with local market conditions. Instead, the pattern of the staggered roll-out is consistent with China's broader practice of policy experimentation, in which reforms are first introduced in lower-stakes areas before wider adoption (Wang and Yang, 2025). Suburban districts, where the investment case for housing is relatively weak, removed restrictions first, with urban districts following later. While this sequencing raises a natural question about whether the two types of districts were on common trends prior to treatment, we find no evidence of differential pre-trends in buyers, inventory, or transactions.

We use a staggered differences-in-differences (DID) research design in which the treatment is the removal of housing purchase restrictions at the district level. Since all districts that had restrictions eventually removed them, but the timing of removal varied across districts, this is a staggered and absorbing treatment. The control group consists of districts that have not yet removed their restrictions at a given point in time, with districts that never had

[5] A matching function is an essential ingredient for any equilibrium search model, including both random-search and directed-search approaches. Although the latter allows for different sub-markets, the same matching function is typically assumed for all sub-markets.

restrictions being excluded. This approach does not require assuming that districts without restrictions are comparable to those with them. Instead, identification comes from variation in the timing of removal among districts that all eventually lift their restrictions.

The event studies show an absence of pre-trends for buyers, inventory, and transactions, supporting the key identification assumption of parallel trends in absence of the treatment. We find that from around one to one-and-half years after the lifting of restrictions, both buyers and transactions increase significantly. We find no significant effect on inventories. We then exploit the heterogeneity in treatment effects by cohort and time to obtain an estimated elasticity with respect to buyers of 0.499 for a CRS Cobb-Douglas matching function. This is around 30 percent larger than the point estimate obtained from an OLS panel regression.

The outline of the paper is as follows. After discussing related literature below, we briefly review the concept of a matching function, the importance of incorporating search effort, and the estimation strategy in section 2. Section 3 describes the data and defines the key variables, and section 4 presents the OLS panel regression results. Section 5 describes the event study and reports the causal estimates of the matching function elasticity. Section 6 concludes.

Related literature High-quality data on home-buyers are scarce. A key advantage of our paper is the direct measurement of buyers and their search effort that comes from the data on in-person viewings.

There are very few papers that aim to estimate a matching function for the housing market. The most well known is Genesove and Han (2012), which uses the U.S. National Association of Realtors Survey of Buyers and Sellers who have successfully completed a transaction in 11 of the years between 1987 to 2008. From data on times taken to sell and buy and assuming a CRS matching technology, they infer market tightness. Together with data on the number of homes visited, they estimate the elasticity of the contact (viewing) hazard with respect to market tightness. There are two advantages of our data. First, we can directly identify buyers and their search effort. Second, our data include buyers and sellers during their process of search, not only those who successfully completed a transaction, a retrospective sample.[6]

Since the rise of online real-estate platforms in the late 1990s, there is now more information available about buyers to facilitate estimation of a matching function. Vidal (2022) identifies buyers from users of an online valuation tool in France. He estimates a Cobb-

[6]Using micro data from Los Angeles, Clark and Smith (1982) compare households during the search process (a longitudinal sample) with those who have already bought (a retrospective sample). They find that on average, those in the longitudinal sample visit 50% more houses and search for three times longer, and many of them withdraw from search without buying.

Douglas matching function and finds that CRS is rejected. However, his study has a small coverage of buyers actively searching for a property because the identified stock of buyers is seven times smaller than the flow of transactions on average. Compared to Genesove and Han (2012) and Vidal (2022), we also have an event study that gives an exogenous source of variation in buyers.

A study with a more comprehensive coverage of buyers is Badarinza et al. (2024), which uses data from the largest online platform in the UK together with Land Registry data. In their paper, they use IP addresses as a proxy for unique users to estimate the number of buyers, and the numbers of online clicks to browse properties and clicks to contact agents as proxies for search effort. They estimate a CRS Cobb-Douglas matching function using a natural experiment based on the 2022 UK 'mini budget', which has a significant effect on market tightness. However, it has no significant effect on transactions. Our staggered roll-back of housing purchase restrictions gives statistically significant effects on both buyers and transactions, which we use to estimate the matching-function elasticity.[7]

2 Framework

There are many factors that determine the number of properties for sale and the number of buyers, such as prices, interest rates, and the state of the economy. In a frictionless market, transactions coincide with demand and supply in equilibrium. However, with frictions arising from imperfect information and heterogeneity on both sides of the housing market, even once the numbers of properties and buyers are determined, matching is not instantaneous, resulting in the coexistence of unmatched sellers and buyers. A matching function serves as a powerful tool to summarize the complex process of matching, allowing researchers to focus on the implications of search frictions rather than on their underlying causes.

The matching function describes the number of transactions with positive total surplus that are realized in a period of time given the search technology and the numbers of buyers and sellers searching.[8] Formally, the flow of transactions Q in a period of time is a function of the stock of properties for sale U and demand D coming from the stock of buyers searching for properties:

$$Q = m(D, U)\,. \tag{1}$$

[7]There are two other papers that also estimate the elasticity of a CRS matching function as part of research investigating the dynamics and spatial dispersion of house prices. Grindaker et al. (2025) infer changes in market tightness using the share of movers who buy first in Norway and estimate the elasticity of time-to-sell with respect to tightness. Kaas et al. (2026) infer the number of buyers using data on contacts (clicking on the 'contact button') from an online platform in Germany.

[8]This does not depend on knowing prices, which govern the division of the surplus. Prices might be of interest in their own right as an outcome of matching and bargaining, but that is not the focus of this paper.

If a time period is one month, the function specifies the number of successful transactions in that month given demand and supply during the month. What the matching function captures is not the number of transactions that will *ultimately* result from D and U, but the speed at which buyers and sellers can find a successful match given the search frictions. Even if each buyer and seller ultimately finds a match with certainty, the matching function is important in describing how changes in D and U affect the speed of transactions.

In specifying a matching function, Pissarides (2000) emphasizes that transaction speed depends on search effort, not just the numbers of buyers and sellers.[9] In the context of the housing market where buyers typically take the initiative, it is crucial to consider buyers' search effort. If B is the raw number of buyers and the average search effort per buyer is e then effective buyer input into the matching function is

$$D = eB\,, \tag{2}$$

which represents demand D from 'effective' buyers.[10] The variable B can be thought of as capturing the 'extensive margin' and e the 'intensive margin'.

In the housing literature, there is plenty of data on properties for sale, especially after the rise of online property listings.[11] But data on buyers' demand are more scarce, unlike in labour-market matching-function estimation where researchers have access to both data on unemployment and vacancies explicitly. Online platforms provide information on buyers who are searching, such as the number of clicks, the number of people who subscribe to email alerts, users' IP addresses, and contact requests (van Dijk and Francke, 2018; Piazzesi et al., 2020; Badarinza et al., 2024; Kaas et al., 2026). Our data provide a measure of buyers undertaking more serious search effort, namely in-person property viewings. This is a better measure given that significant time is required to make a viewing, whereas clicks, signing up to email alerts, etc., require almost no time or effort and can be done by those with no intention of buying, such as those browsing for curiosity, or even sellers themselves. We measure raw buyers B as the number of people who view at least one property in a period of time. Proxying search effort e by the number of viewings made per buyer, our overall measure of effective buyers D in (2) is the total number of viewings undertaken by all buyers.

Researchers using the matching function in equation (1) would like to know its functional form, degree of returns to scale, and elasticities of transactions with respect to buyers

[9]For example, in the labour market, Davis et al. (2013) highlights the importance of recruiting intensity (buyers' search effort) when specifying a matching function.

[10]See Gavazza et al. (2018) for a similar specification in the labour market, where they define effective vacancies (buyers) as the input into the matching function.

[11]The first website offering online access to U.S. Multiple Listing Service properties began in 1995 (Home-Seeker.com).

and sellers. One fundamental challenge to identification is that observed matching-function inputs (buyers and properties) are likely correlated with unobserved matching efficiency, defined in the same way as total factor productivity for a production function.[12] This resembles the challenge faced in the literature on estimating production functions (see Ackerberg et al., 2015, for a review) and labour-market matching functions (see Yashiv, 2000, for an early example). Addressing this concern is important because, as demonstrated by Borowczyk-Martins et al. (2013), there are theoretical arguments why matching efficiency and the number of buyers and sellers in a market should be correlated.

As first attempt, we followed the literature by exploring whether imposing a Markov structure on matching efficiency enables us to use GMM estimators (Arellano and Bond, 1991; Blundell and Bond, 2000) to obtain identified results on the parameters of interest. However, based on standard diagnostic tests, neither difference GMM nor system GMM yields reliable estimates of the matching function in our setting.[13]

We then turned to what has been the most common approach in the housing-search and earlier labour matching-function literatures of using panel estimation. These OLS regressions do not reject a constant-returns-to-scale Cobb-Douglas matching function. We acknowledge that OLS estimation with fixed effects may still suffer from potential omitted variable bias. We proceed following the housing search literature by assuming the matching function takes a CRS Cobb-Douglas form.

Our key innovation is to use an event study to obtain exogenous shifts in buyer demand to identify and estimate the elasticity of the matching function. This parameter represents the elasticity of transaction speed (the inverse of time taken to sell) with respect to changes in demand relative to the inventory of properties listed for sale. In a range of applications (Piazzesi and Schneider, 2009; Head et al., 2014; Gabrovski and Ortego-Marti, 2019; Guren and McQuade, 2020; Piazzesi et al., 2020; Han et al., 2025), one would like to know the response of housing-market liquidity as measured by transaction speed when demand increases, for example, due to changes in macroeconomic conditions such as mortgage rates, or the entry of new buyers, or as supply increases, for example, due to changes in new construction or zoning.

[12] Lange and Papageorgiou (2020) raises a different concern that when matching efficiency responds to the outcome of the matching process, no instrument can simultaneously have an impact on market tightness and be exogenous with respect to matching efficiency. However, most applications of a matching function would be concerned with the total effect on the output of transactions from given inputs, including any indirect effect on matching efficiency.

[13] The likely reason is a degree of persistence in properties listed for sale in our data beyond what the system GMM correction can handle. We are therefore not able to use instruments that are sufficiently strong to yield reliable empirical results on the functional form or degree of returns to scale with either difference GMM or system GMM.

3 Data and key variables

Data source We use data from the largest real-estate transaction platform in China. This platform integrates an online listing app with a large network of agents handling viewings and transactions, so all stages of the search process run through a single system. This unique feature is missing in other housing datasets used in the literature. In the U.S. or UK housing markets, apps such as Zillow or Rightmove record only online activity. In-person viewings are organized by agents and do not enter those platforms' databases, which makes it difficult to measure the number of buyers and their search effort. Our setting is different because the platform requires agents to record all viewings in real time, and these records are stored in the platform's database.

This feature enables us to fill an important gap in the literature. We observe in-person viewings each linked to a unique user ID, and hence can track both who is searching and how intensively they search. These advantages yield a better measure of buyers than has been available in the existing literature, which is of particular importance when for estimating the housing-market matching function.

The platform is mainly used for listings of existing second-hand homes and therefore our data exclude newly built properties.[14] Our data are from 15 major cities in two Chinese provinces, Jiangsu and Shandong. These cities are the most populous regions of the two provinces, covering over 100 million people. Each city includes several urban, suburban, and rural districts. We analyse data by district and month from November 2021 to December 2024. This time period does not overlap with Covid lockdowns in the cities of our sample.[15]

Definition of a housing market and key variables A matching function captures how quickly buyers and sellers can successfully match and transact. While buyers and sellers may remain searching for several months, the matching function itself can be estimated with higher frequency data. We use one month as our time period because even higher frequencies would result in a series for transactions with too many zeroes.[16] Nonetheless, it is important that properties transacted in a monthly period result from viewings made by their buyers within the same month. This is confirmed by Figure A.3, which shows that almost all transactions occur within 30 days of a property being viewed by its eventual buyer, and the

[14]New builds are typically marketed and sold directly by developers, not through a transaction platform. This means unoccupied newly built property developments in China, sometimes referred to as 'ghost towns', are not found in our data.

[15]For a summary of the Covid lockdown policies in Chinese cities, see `https://en.wikipedia.org/wiki/COVID-19_lockdown_in_China`.

[16]We verify robustness by re-estimating the matching function using quarterly aggregated data.

calendar month of the eventual buyer's viewing mostly coincides with the calendar month in which the transaction is recorded.

Our data also inform how we define a housing market in the spatial dimension. Across space, we define a market as a district because Figure A.4 shows that most buyers make viewings in only one district.[17] We therefore describe a housing market by a district-year-month, indexed by district i and time t.

For each (i, t), we have aggregate data on transactions Q_{it}, inventory U_{it}, and effective buyers D_{it}. Transactions Q_{it} denotes total sales agreed (contracts exchanged) during the month. Inventory U_{it} is any listing that shows up as available on any day in a given month.[18] Effective buyers D_{it} combines a measure of the number of active buyers B_{it} and average search effort e_{it} per buyer, as defined in equation (2). We measure active buyers B_{it} as the number of user IDs with at least one in-person viewing in a given month, and average search effort e_{it} as the number of viewings per active buyer, thus effective buyers D_{it} is the total number of viewings done during a given month.[19] Our measures of buyers and properties are not based on a snapshot on any particular day of the month, which mitigates temporal aggregation bias. For example, if inventory were recorded only at the beginning of a month, some transactions from that month could be for properties that were first listed during the month, but do not appear in the measure of inventory.

Analysis sample and summary statistics We impose some restrictions on the sample used in our main analysis. First, we exclude observations where the housing platform was little used or likely only just entered a local market. This is done by dropping observations where no inventory, viewings, or transactions are recorded, and also observations at the lower 5^{th} percentile of the ratio of inventory to the district population.

Second, we calculate for each observation to what degree the observed data violates an accounting identity tracking inventory, inflows, and outflows. The dynamics of the stock of properties on the market should satisfy

$$U_{i,t+1} - U_{it} = N_{i,t+1} - Q_{it} - W_{it} \,, \tag{3}$$

where changes in inventory between months t and $t+1$ come from the inflow $N_{i,t+1}$ of newly listed properties during month $t+1$, and outflows of existing inventory because of transactions Q_{it} in month t, or outflows for any reason other than a property being sold ('withdrawals')

[17] This finding uses micro data on all viewings in 2024 in the city of Nanjing.

[18] Agents visit and verify every property listing that goes live on the app. Properties that cannot be verified cannot be listed.

[19] Agents rarely decline viewing requests. If an agent is unavailable, they will arrange for a colleague in the same office to conduct the viewing on their behalf.

W_{it}.[20] Our data provide direct measures of all variables appearing in equation (3), including N_{it} and W_{it}. New listings N_{it} are defined as properties that become available for the first time in month t. Withdrawals W_{it} are properties delisted in month t for any reason other than a sale.

A violation of (3) results when there is a non-zero value after subtracting the right-hand side from the left-hand side. We normalize this error by $U_{i,t+1}$ and exclude observations at the top and bottom 5^{th} percentile of the distribution. Figure A.7 visualizes the error before and after we apply all of the previously described restrictions. As can be seen from the figure, imposing the described restrictions leads to the sample of analysis which is much more compliant with (3) than before our filtering.

Our main analysis sample includes 116 districts from November 2021 to December 2024. Figure A.1 shows the geographical coverage. Figure A.2 presents the district-level distributions of population and area. The median population of districts is about 785,000, and the median area is about 1,440 square kilometers. Table B.1 presents summary statistics for the variables used in the analysis.

4 Baseline OLS estimation

We consider a Cobb-Douglas functional form $Q_{it} = Z_{it} D_{it}^{\eta_d} U_{it}^{\eta_u}$ of the matching function (1). The parameters η_d and η_u are the constant elasticities of transactions Q_{it} with respect to buyers searching D_{it} and inventory U_{it}, and Z_{it} denotes matching efficiency. The following log-linear equation is estimated:

$$\log Q_{it} = \eta_d \log D_{it} + \eta_u \log U_{it} + \mu_i + \zeta_t + \epsilon_{it}\,. \tag{4}$$

where μ_i is a district fixed effect, ζ_t is a time fixed effect, and ϵ_{it} is an error term. The econometric challenge in estimating (4) is that a shift variable in the error term that captures the efficiency of matching Z_{it} is not directly observed. Here we assume that, once time and district fixed effects remove changes in $\log Z_{it} = \mu_i + \zeta_t + \epsilon_{it}$ that are common to all districts and differences across districts that are common over time, the remaining variation in Z_{it} is uncorrelated with D_{it} and U_{it}.

In Table 1, we start by comparing two functional forms. Column (1) considers a translog functional form that allows for non-constant elasticities of transactions with respect to buyers and inventory. This can be interpreted as a second-order Taylor expansion of a general

[20] In the U.S., around half of properties are unsold at the point of delisting (Anenberg and Laufer, 2017).

matching function.[21] Column (2) uses the Cobb-Douglas functional form specified in (4). The results in column (1) show that the second-order terms are all statistically insignificant, supporting a log-linear functional form. We therefore assume a Cobb-Douglas matching function in the remainder of this paper.

An important characteristic of a matching function is the degree of returns to scale, which is measured by the sum of the coefficients η_d and η_u. In column (2), we also test whether $\eta_d + \eta_u = 1$. The point estimate of the sum of the coefficients is very close to 1 and we do not reject the hypothesis that the matching function has CRS. Imposing this assumption allows the matching function to be written as

$$\log\left(\frac{Q_{it}}{U_{it}}\right) = \eta_d \log\left(\frac{D_{it}}{U_{it}}\right) + \mu_i + \zeta_t + \epsilon_{it}\,, \tag{5}$$

and column (3) of Table 1 reports the elasticity η_d estimated using this form. The point estimate is 0.382.

As discussed in section 3, the data suggest our choice of monthly frequency is appropriate for estimating the matching function. To verify robustness, we re-estimate with a quarterly aggregation of the data. Appendix Table B.2 shows that quarterly rather than monthly aggregation makes almost no difference to the baseline results.

The importance of search effort In the absence of data on search effort, researchers are constrained to estimate a matching function with the raw number of buyers B as an input:

$$Q = \tilde{m}\,(B, U)\ . \tag{6}$$

Estimation of equation (6) suffers from the problem of search effort being an omitted variable given that it may be correlated with matching-function inputs. Our data provide a direct measure of search effort to address this concern, allowing us to use effective buyers D as an input to the estimated matching function (1).

Table 2 presents evidence in column (1) that search effort is positively correlated with the number of raw buyers. This shows how estimation of equation (6) could lead to an upward bias of the estimated elasticity with respect to raw buyers. Indeed, column (2) confirms that the coefficient of raw buyers in the estimated matching function is larger than the estimated

[21] The translog functional form is

$$\log Q_{it} = \log Z_{it} + \eta_d \log D_{it} + \eta_u \log U_{it} + \eta_{dd}(\log D_{it})^2 + \eta_{uu}(\log U_{it})^2 + \eta_{du}(\log D_{it}) \times (\log U_{it})\,.$$

Table 1: OLS estimation

	log(transactions)		log(transactions/inventory)
log(effective buyers)	0.080	0.375***	
	(0.372)	(0.056)	
log(inventory)	1.026	0.718***	
	(1.246)	(0.145)	
log(effective buyers)^2	0.023		
	(0.019)		
log(inventory)^2	-0.017		
	(0.099)		
log(effective buyers) × log(inventory)	0.001		
	(0.070)		
log(effective buyers/inventory)			0.382***
			(0.052)
Observations	3692	3692	3692
Districts	116	116	116
R^2	0.67	0.66	0.48
District FEs	✓	✓	✓
Time FEs	✓	✓	✓
H0: $\eta_d + \eta_u = 1$			
F-Statistic		0.713	
p-Value		0.400	

Notes: effective buyers = (average viewings per buyer) × (raw number of buyers) = total viewings. Standard errors clustered at district level in parentheses. Significance level indicated according to * p<0.1, ** p<0.05, *** p<0.01.

coefficient of effective buyers in Table 1.[22]

Measures of search effort are also used in the two existing papers on estimating the housing-market matching function (Vidal, 2022; Badarinza et al., 2024). Different from our approach, they estimate (6) and use search effort as a control variable. However, a fundamental concern is that search effort may be decided alongside other choices affecting the numbers of buyers and sellers, rather than being an independent variable that can be held fixed. This issue is often referred to as a 'bad control' problem (see, for example, Angrist and Pischke, 2009). Even if buyers were as good as randomly assigned, holding fixed search effort and inventory when increasing the number of buyers likely leads to biased estimates. An increase in the number of buyers, which under constant inventory is not met with an increase

[22] This difference is significant at the 99 percent confidence level based on a joint variance-covariance matrix for the two regression models (via seemingly unrelated regression).

Table 2: The importance of search effort

	log(search effort)	log(transactions)
log(raw-buyers)	0.126***	0.590***
	(0.040)	(0.082)
log(inventory)	0.079	0.562***
	(0.080)	(0.144)
Observations	3692	3692
Districts	116	116
R^2	0.29	0.69
District FEs	✓	✓
Time FEs	✓	✓
H0: $\eta_d + \eta_u = 1$		
F-Statistic		2.974
p-Value		0.087

Notes: Search effort is measured by viewings per buyer. Standard errors clustered at district level in parentheses. Significance level indicated according to *p<0.1, **p<0.05, ***p<0.01.

in effort per buyer, is likely correlated with something unobserved that is simultaneously correlated with transactions and buyers, introducing a problem of endogeneity. To avoid this problem, but nonetheless account for search effort, it is advantageous to estimate (1) using effective buyers D from (2) as an input, a combined measure of search effort and the raw number of buyers.

5 Event study

To identify the elasticity of the matching function, we proceed by assuming a CRS Cobb-Douglas matching function and exploit a change in housing policy that removed restrictions on housing purchases. Crucially, the roll-back of these restrictions was staggered across districts. Using this quasi-experimental variation in treatment timing, we conduct event studies using the staggered difference-in-differences (DID) method. Following de Chaisemartin and D'Haultfœuille (2024), we implement an estimation method that allows for heterogeneous treatment effects across districts.

5.1 Policy background and data

Before 2022, many large Chinese cities had two major restrictions on housing purchases. First, people's eligibility to purchase properties depended on their *hukou* status. The Chinese *hukou* system, a household registration system, links residents to their registered birthplace, granting local *hukou* holders access to social benefits. Under this restriction, non-local *hukou* holders (essentially internal migrants) were required to provide proof of local tax payments, social security contributions, or residency permits for at least two years to be eligible to purchase properties. Second, there was a limit on the number of properties people can purchase, applying to both local and non-local *hukou* holders. In our data, four cities had both restrictions before 2022.

In 2022, the Ministry of Housing and Urban-Rural Development issued central guidelines to stimulate housing demand. This was motivated by a range of factors including the slowdown in the property sector, widespread developer distress following the collapse of major developers such as Evergrande, and the resulting threat to broader economic stability.[23] In 2023, the Politburo meeting, where China's top leaders set the policy agenda, sent a message that cities should relax restrictions on housing purchases to increase demand. Following this directive from the central government, cities that had restrictions gradually removed the two restrictions described above between 2022 and 2024. This roll-back of restrictions varied not only across cities, but also across districts within a city.

Among the 15 cities in our data, four had housing purchase restrictions in place before 2022, covering a total of 34 districts. These restrictions were gradually removed across all 34 districts during the period 2022–2024. To construct the timing of removals at the district level, we conducted a comprehensive review of local housing policies based on two sources. The first source is an extensive collection of policy documents issued by the Chinese central and local governments since 1949, compiled by PKU Law (`PKULaw.com`), an online platform hosted by Peking University Law School. The second source is local government websites, including city governments, development and reform commissions, and housing bureaus.

Our event-study analysis focuses on these 34 districts. The start of the roll-back of restrictions in a district is defined as the month when either eligibility based on *hukou* status or the limit on the number of properties that can be purchased is relaxed, whichever comes first. We use the timing of the first removal in a binary estimation because the timing is relatively less likely to be anticipated. Our review of the details of the policy changes also found that the announcement date and the enforcement date were the same in all districts.

[23] Evergrande was one of China's largest property developers, which defaulted on its offshore bonds in December 2021 following the government's introduction of borrowing limits on highly leveraged developers, triggering a broader crisis of confidence across the Chinese property sector.

Figure A.5 shows over time the cumulative fraction of districts that had begun rolling back housing-purchase restrictions. The first was in April 2022 and the last in June 2024. There is substantial variation in the timing of the policy roll-back across districts.

While the central government's motivation to stimulate demand was common to all districts, this does not explain the local variation in timing. The staggered roll-back of housing-purchase restrictions reflects the interplay between central directives and local political economy. Since 2017, the central government has promoted the principle that "houses are for living in, not for betting", discouraging households from buying homes primarily as financial investments rather than as places to live. Lifting purchase restrictions was necessary to stimulate demand, but risked being seen as encouraging investment purchases that the central government had discouraged. Suburban districts, where prices are lower and the investment case weaker, could relax restrictions with little risk of increasing investment purchases and hence did so first. Each round of suburban relaxation legitimized the next step, generating implicit central endorsement for further action. Urban districts, where investment purchases are expected mostly to occur, could follow once the political ground had been cleared. Figure A.6 confirms this geographical pattern in which roll-back began in suburban districts first and then spread to urban districts. This staggered roll-back across districts is consistent with China's broader practice of policy experimentation, in which reforms are first introduced in lower-stakes areas before wider adoption (Wang and Yang, 2025).

Table B.3 shows results consistent with this interpretation. Urban districts began removing restrictions approximately 11 months later than suburban districts. Importantly, neither the pre-existing ratio of buyers to inventory nor the district-level exposure to developer distress (measured by the number of Evergrande projects by district) are correlated with the timing of removal.[24] Other district characteristics including population, hukou composition, and homeownership rates are similarly uncorrelated with the roll-back timing. Together, these results suggest that the variation in timing reflected political considerations rather than underlying market conditions. The suburban-first sequencing, however, raises a natural question about whether the two types of districts were following common trends prior to treatment, a concern we address directly in what follows.

5.2 Identification and estimation

Event study We begin by estimating the average treatment effect on treated districts (ATT) of the policy change (i.e. the first removal of housing restrictions) on buyers, inventory,

[24]Data on district-level Evergrande projects were collected from Chinese real-estate listing websites, cross-referenced with Evergrande regional platforms and local-government housing bureau records. Projects were assigned to districts based on their listed addresses as of September 2021, prior to the policy roll-back.

and transactions over event time using a staggered DID design. Due to the recent findings of a growing literature on the pitfalls of estimating the dynamic effects of a binary (absorbing) treatment indicator W_{it}

$$y_{it} = \alpha_i + \alpha_t + \sum_{k=a}^{b} \phi_k W_{it,k} + \vartheta_{it} \tag{7}$$

via a two-way-fixed-effects estimation, we follow de Chaisemartin and D'Haultfœuille (2024) to estimate ATTs over event-time.[25] The primary coefficients of interest in (7) is given by the set of ϕ_k, which provide an ATT estimate of the difference in log outcome between k event months after treatment and the last month before treatment. Identification comes from contrasting this change with a control group. Since all 34 districts that had housing restrictions eventually removed them, the control group consists of districts in which restrictions have not yet been lifted at a given point in time (the "not-yet-treated" group). Districts that never had housing restrictions are excluded from all event-study analyses.

We need identification assumptions that are standard in the difference-in-differences method: most importantly the parallel trends assumption, i.e. in the absence of treatment, ϕ_k should be zero. The validity of this identification assumption is untestable, but we can assess whether our data are consistent with the parallel trends assumption in the pre-treatment period.

Figure 1 visualizes the event-study estimates on the number of effective buyers, inventory, and transactions. This figure provides two key results. First, we find no statistically significant differences in the trends of effective buyers, inventory and transactions between the treatment and control groups before the event of treatment. Second, the estimated coefficients in the first and third subfigures show that effective buyers and transactions increase to a relevant degree only from around 1 to 1.5 years after the removals of housing restrictions.[26] While the estimated coefficients are aligned, we acknowledge that the event study on effective buyers is slightly noisier between 11 to 18 months after policy removal, likely due to the absence of an effect on search effort, as discussed below. In contrast, we find no significant effect on inventory even two years after the policy change.

It is worth noting that the visualized chronology of events is consistent with the nature of housing search. At the time when housing restrictions are removed, potential buyers are likely locked into rental contracts, which typically last one to two years in Chinese cities. While these buyers may begin browsing available listings on the app soon after the policy

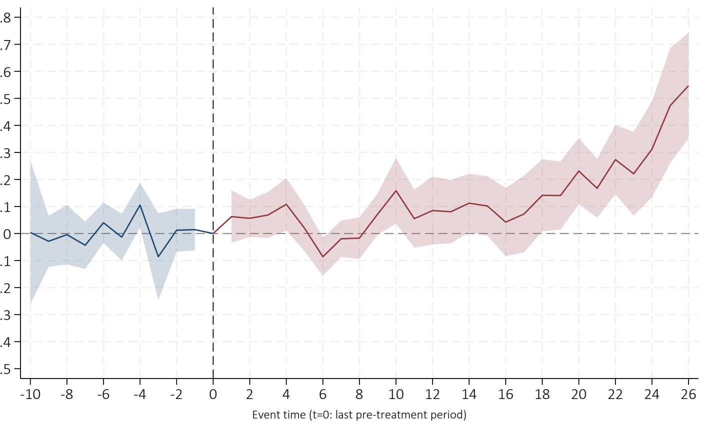
[25] We emphasize that with our binary treatment definition, the event-time coefficients after treatment from de Chaisemartin and D'Haultfœuille (2024) (DID_ℓ in their notation) are identical to ones by Callaway and Sant'Anna (2021) ($\theta_{es}(\ell - 1)$ in their notation). See the former paper for details.

[26] In order to put these magnitudes into context, we note that the magnitude of the changes in transactions is similar to the U.S. housing boom and bust documented in Ngai and Sheedy (2020).

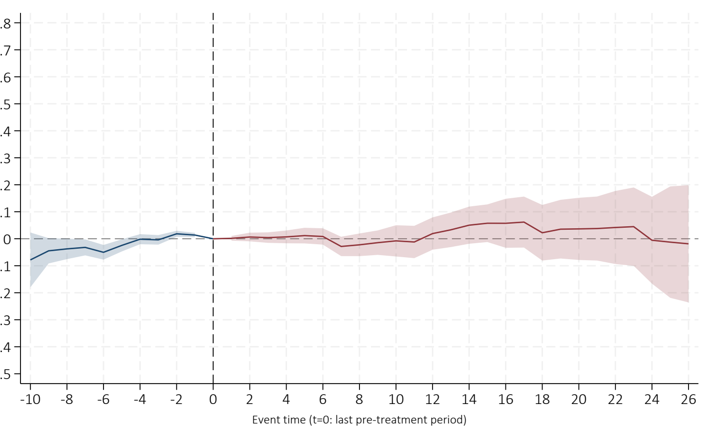
change, they are unlikely to arrange in-person viewings until their rental terms approach expiration. Indeed, Figure A.8 shows that the number of clicks on the app increases right after the restrictions are removed.

Figure 1: Event study

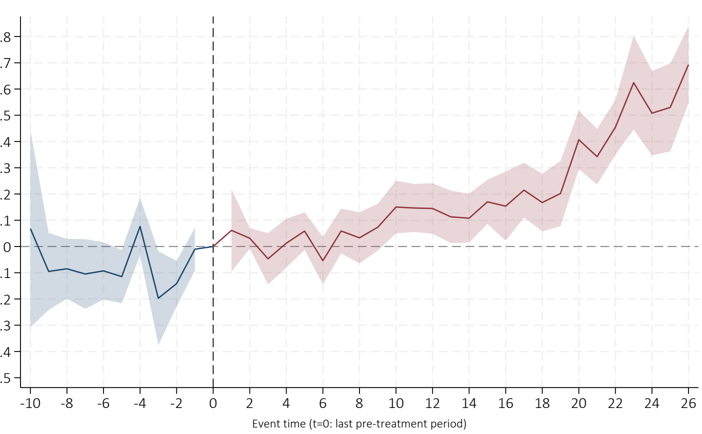
(a) log(effective buyers)



(b) log(inventory)



(c) log(transactions)



Notes: The bars indicate the 95 percent confidence intervals. Standard errors are clustered at the district level.

Assessing the validity of the identification assumptions The validity of the event-study estimation is subject to a set of identification assumptions for the staggered DID design. A key assumption is parallel trends in the counterfactual, untreated outcome: in the absence of the treatment, the trajectory of the outcome variable has to be parallel between the treatment and control groups. We discuss two potential threats to this assumption below, and show that the pre-treatment evidence in Figure 1 addresses each directly.

The first concern follows from our finding in section 5.1 that urban districts removed restrictions approximately 11 months later than suburban districts. If urban and suburban districts were on different trajectories before the policy removals, we would expect diverging pre-trends between early-treated suburban districts and the not-yet-treated urban districts that serve as their control group. Figure 1 shows no such divergence: the estimated coefficients are statistically indistinguishable from zero throughout the pre-treatment period for buyers, inventory, and transactions. The parallel trends evidence rules out the concern that the suburban-to-urban sequencing coincides with differential underlying market dynamics.

The second concern is that spillovers from treated to not-yet-treated districts could contaminate the control group. Once suburban districts removed restrictions, local residents in not-yet-treated urban districts might anticipate their own forthcoming relaxation, or find urban properties relatively cheaper, and begin responding before their own restrictions were lifted. However, a buyer entering the market in a treated suburban district was unlikely to simultaneously search in an untreated urban district, because as Figure A.4 shows, most buyers confine their search to a single district. Any spillovers operating through anticipation effects are likely non-uniform, for example, they would disproportionately affect districts close to already-treated areas. Non-uniform spillovers to the control group would generate detectable violations of parallel trends in Figure 1. The absence of pre-trends in Figure 1 is therefore inconsistent with non-uniform spillovers contaminating our results. The only possibility that cannot be ruled out by pre-trends alone is one of perfectly uniform spillovers that uniformly inflate outcomes in all control districts. Under uniform spillovers, all estimated ATTs are scaled down by the same factor, but as we discuss below, this leaves the estimated elasticity unaffected, since a uniform level shift in both the outcome and regressor ATTs does not change the estimated slope.

The remaining assumption is that other housing policies during our study period do not confound our findings in Figure 1. To assess this assumption, we conducted a comprehensive search of all other housing policies in our area of study during 2022-2024, from the same sources described in section 5.1 (i.e. local government websites and PKU law). First, national policies to reduce mortgage interest rates did not vary across regions and are therefore controlled for by the time fixed effects. Second, no other housing policies had a staggered roll-out within-city and between districts. Third, two policies that had a staggered roll-out by city are potentially concerning: one provided lower mortgage rates to buyers who used their provident fund for the down payment, and the other subsidized a small number of buyers with talent status in their housing purchases. In robustness tests, we control for the city-level roll-out of these two additional policies. Figure A.9 shows that all results of the event study are robust to including other policies.

Estimation of the elasticity As an evaluation of the policy change is not the primary focus of this paper, we will explain in the following how we use the event studies to provide a more reliable estimate for the elasticity of the CRS Cobb-Douglas matching function. In a quasi-experimental design, let Q be defined as the observed transactions and Q' as the counterfactual transactions had there been no removals of housing restrictions, which can be written as a ratio over inventory as follows:

$$\frac{Q}{U} = Z\left(\frac{D}{U}\right)^{\eta_d} \quad \text{and} \quad \frac{Q'}{U'} = Z'\left(\frac{D'}{U'}\right)^{\eta_d}. \tag{8}$$

The staggered roll-back of the policies enables us to identify and estimate ATTs by cohort-time, where a cohort is defined as the set of districts sharing the first period of treatment:

$$ATT_{cohort,time} = E[Y_{cohort,time} - Y'_{cohort,time} | \text{ Treatment }], \tag{9}$$

where Y could represent the natural log of transactions Q, or effective buyers D, or inventory U. Treatment status is equal to one from the month of the first removal of a housing-purchase restriction, and zero otherwise.

We focus on equation (10) below where the ratio of transactions to inventory is the outcome variable and the estimand is the elasticity of interest η_d. Omitting cohort-time notation for brevity, we can write $\mathbb{E}\left[\log(Q/U) - \log(Q'/U') \mid \text{Treatment}\right]$ as:

$$\underbrace{\mathbb{E}\left[\log\left(\frac{Q/U}{Q'/U'}\right) \mid \text{Treat.}\right]}_{\text{ATT of } Q/U} = \mathbb{E}\left[\log\left(\frac{Z}{Z'}\right) \mid \text{Treat.}\right] + \eta_d \underbrace{\mathbb{E}\left[\log\left(\frac{D/U}{D'/U'}\right) \mid \text{Treat.}\right]}_{\text{ATT of } D/U}. \tag{10}$$

A regression of cohort-time ATTs of $\log(Q/U)$ on cohort-time ATTs of $\log(D/U)$ identifies η_d under weaker assumptions than a simple regression of equation (4). Rather than requiring district-time-specific changes in matching efficiency to be uncorrelated with changes in buyers and inventory, the key identifying assumption here is that the policy's effect on D/U is uncorrelated with its effect on matching efficiency Z. Moreover, if spillovers from treated to not-yet-treated districts uniformly reduce all estimated ATTs, identification of η_d remains unaffected, since a uniform level shift in both the outcome and regressor leaves the estimated slope unchanged.[27] To control for potential sources of cohort-specific and time-varying confounders in matching efficiency, we include cohort and time fixed effects. We therefore estimate:

$$(\text{ATT of Q/U})_{cohort,t} = \mu_{cohort} + \zeta_t + \eta_d \, (\text{ATT of D/U})_{cohort,t} + \epsilon_{cohort,t} \,, \tag{11}$$

where we include for each cohort all estimated ATTs from the period of policy change onward.[28]

A remaining concern is that the reform might change buyer composition that shifts Z at the cohort-time level beyond what the cohort and time fixed effects can absorb. For example, one might be concerned that newly eligible buyers such as non-local *hukou* holders (essentially internal migrants) likely differ systematically from incumbent buyers in their ability to afford properties, generating a policy-induced shift in Z that is correlated with the shift in D/U. We assess this concern by testing whether the reform shifts Z through buyer composition, since any such shift could generate a correlation between the policy's effects on D/U and Z. In the absence of access to the social benefits tied to local *hukou* status, non-local *hukou* holders likely have lower average incomes and wealth than local buyers, making them more likely to afford only smaller and less expensive properties. If this is the case, we would expect the distribution of floor area of transacted properties to change and likely shift towards smaller units following the policy removals, since floor area is the most salient housing characteristic and is positively correlated with other property attributes such as the number of rooms.[29] Using transaction microdata to construct quantiles and average floor area by district and month, Figure A.10 shows no detectable change to the distribution of floor area of sold properties after the policy removal, inconsistent with compositional change along this dimension.

The results are provided in Table 3 and visualized in Figure 2.[30] We estimate $\eta_d = 0.499$.

[27] In Appendix C we discuss how under different assumptions than we impose on our main specification, an alternative way to yield an estimate of η_d is given by dividing aggregate weighted averages of ATTs of the left-hand side by aggregate weighted averages of ATTs of the right-hand side of (10).

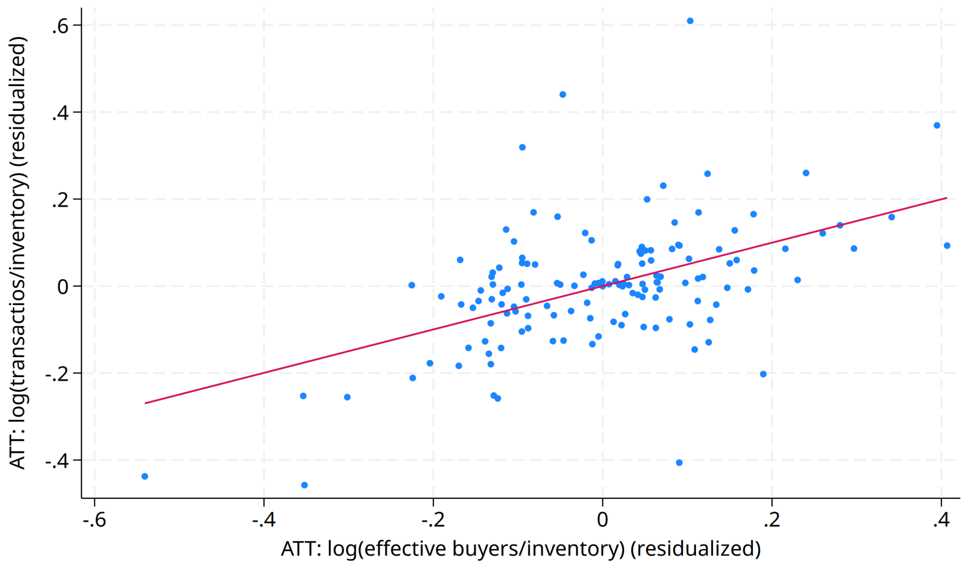
[28] We restrict to periods after treatment because under the assumption of parallel trends, adding the pre-treatment periods for both the dependent and independent ATTs would imply adding (noisy) zeros on both sides of the estimating equation, which serves no purpose. For a test of parallel trends, we refer to Figure 1.

[29] Transaction prices are not an appropriate outcome for this test since they are equilibrium outcomes determined by both demand and supply as well as bargaining. Floor area, as a physical characteristic of the transacted property, provides a more useful test of whether the composition of buyers has shifted toward those with systematically different affordability constraints.

[30] Standard errors are bootstrapped by resampling treated districts and re-computing all ATTs. See appendix D for more details.

The number of observations in this regression refers to the number of time periods we observe each cohort after treatment. The number of cohorts used for estimation is equal to seven, while the maximum number of periods is 26. All of these cohort-time specific ATTs are visualized in Figure 2 (after residualizing using as regressors only fixed effects). The linear slope in this figure is identical to the estimated coefficient from the table based on the Frisch–Waugh–Lovell theorem.

Figure 2: Scatterplot of ATT estimates by group and time



Notes: The data points refer to residualized estimated ATTs on cohort-time level. After estimating the effect of the policies on either log-transformed transactions over inventory (vertical axis) or log-transformed effective buyers over inventory (horizontal axis), these dynamic effects on cohort-time level were regressed on cohort and time fixed effects. The figure visualized the residuals of these regressions together with a linear fit, which based on the Frisch–Waugh–Lovell theorem, has a slope identical to the estimated coefficient shown in Table 3.

Table 3: DID estimation: estimating elasticity

	$\text{ATT}_{cohort,t}$ log(transactions/inventory)
$\text{ATT}_{cohort,t}$ log(effective buyers/inventory)	0.499***
	(0.111)
Observations	140
Treat. Groups	7
Time Periods	26
R^2	0.804
Cohort FEs	✓
Time FEs	✓

Notes: effective buyers = (average viewings per buyer) × (number of buyers) = viewings. Dependent and independent variable are both estimates of ATTs over cohort and time. A cohort consists of all districts sharing the same first time period of treatment. Bootstrap standard errors based on resampling treated districts and re-computing all ATTs in parentheses (see appendix D for details). Significance level indicated according to * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

6 Conclusion

The modelling of search frictions in macroeconomics has built on the key innovation of the matching function, which has become pervasive in analyses of the labour market. This theoretical research has been supported by an extensive empirical literature that has estimated labour-market matching functions. Research on the interactions between frictional housing markets and macroeconomics has often assumed a CRS Cobb-Douglas matching function, but has lacked empirical evidence on the key elasticity parameter.

Our paper fills this crucial gap in the literature using a novel dataset with information on buyers and their search behaviour as well as an event study that provides exogenous shifts in buyers. We find evidence that the housing-market matching function has an elasticity with respect of demand of around 0.5 for a CRS Cobb-Douglas function:

$$\boxed{Q = ZD^{0.5}U^{0.5}}.$$

Knowledge of this matching function is important for papers that incorporate search frictions in the housing market. Future research could expand on the aggregate matching function by exploring the search process in greater detail at the micro level.

Appendix

A Additional figures

Figure A.1: Geographical coverage of this study

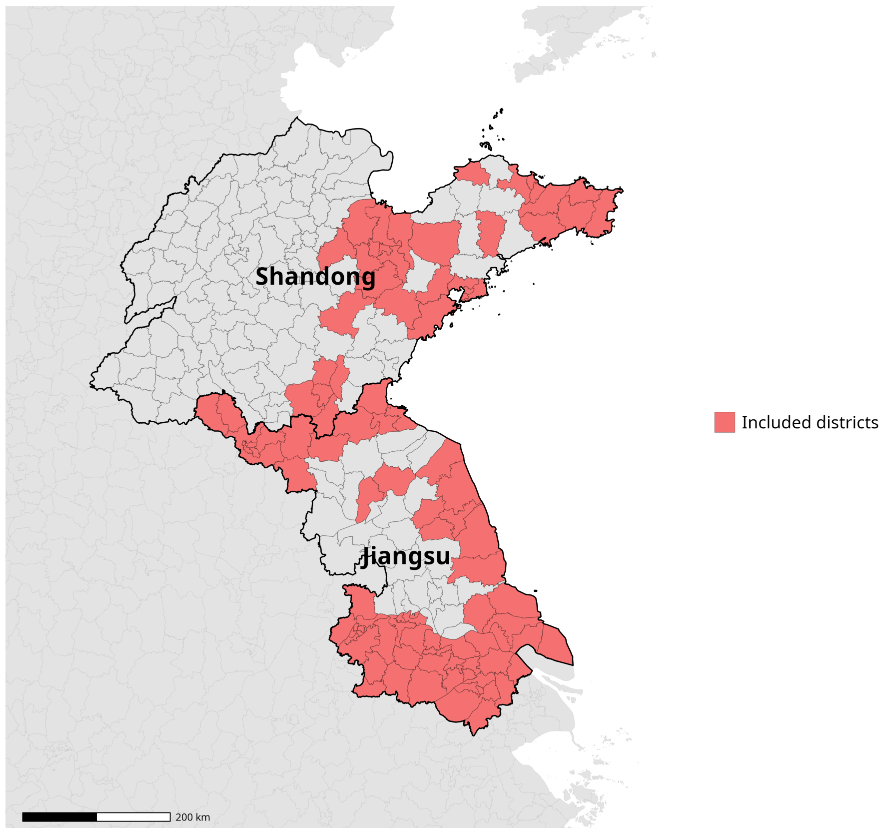
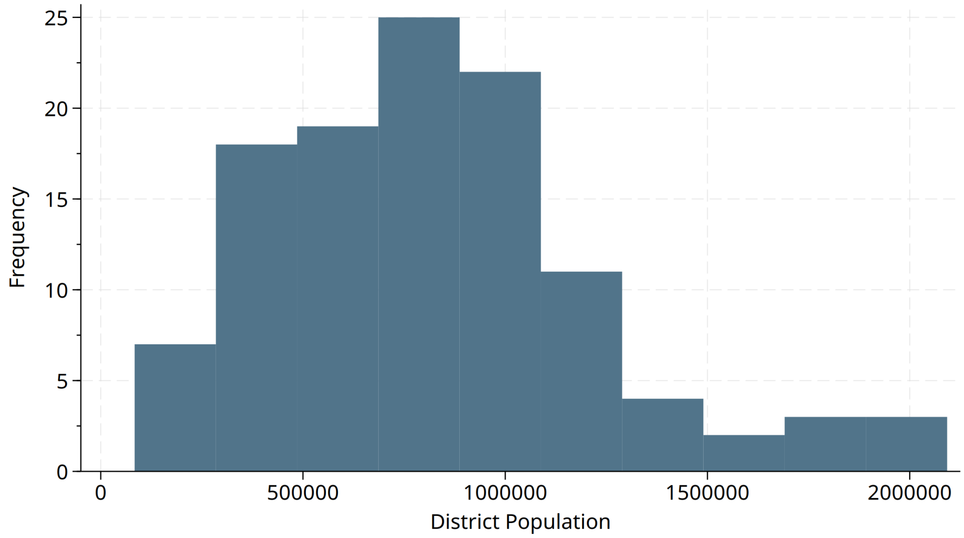


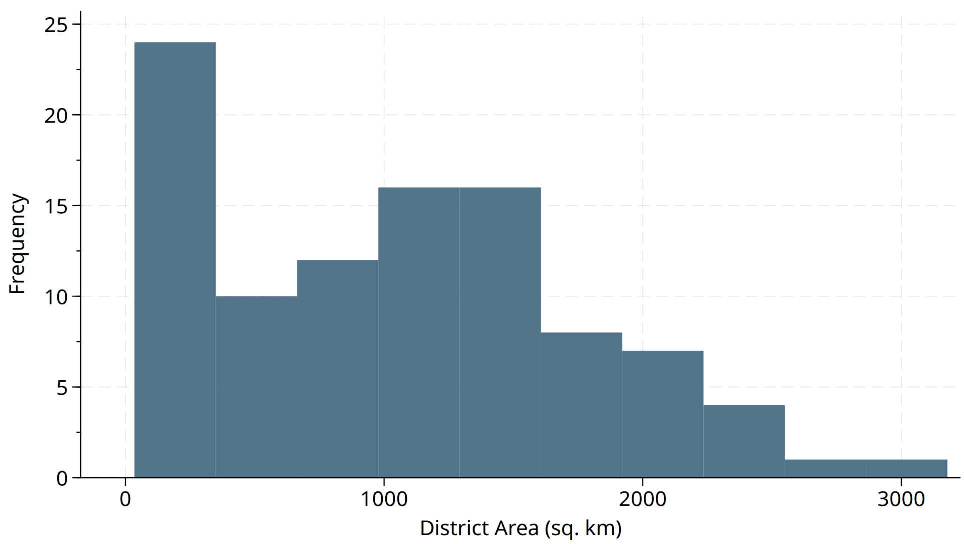
Figure A.2: District population and area

(a) District population



Median Population: 785092

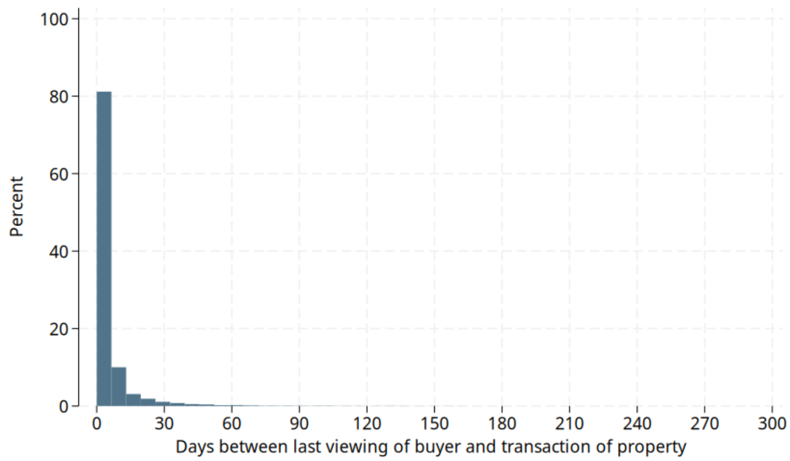
(b) District area



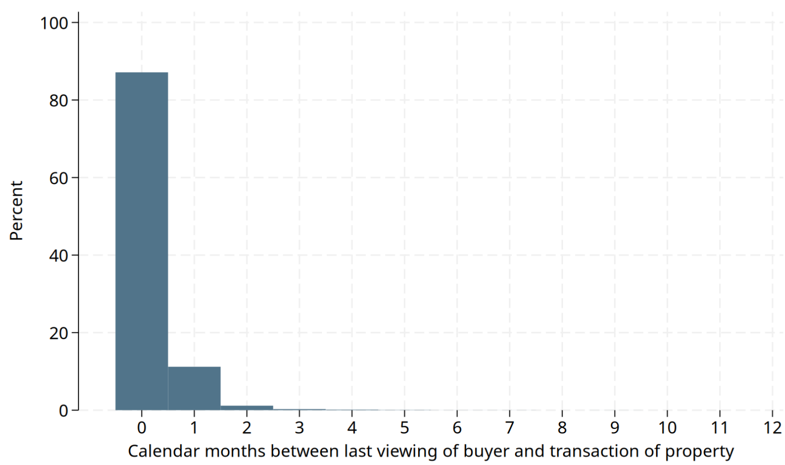
Median Area: 1441 km^2

Figure A.3: Time between a property transaction and the viewing made by its buyer

(a) Lag in days

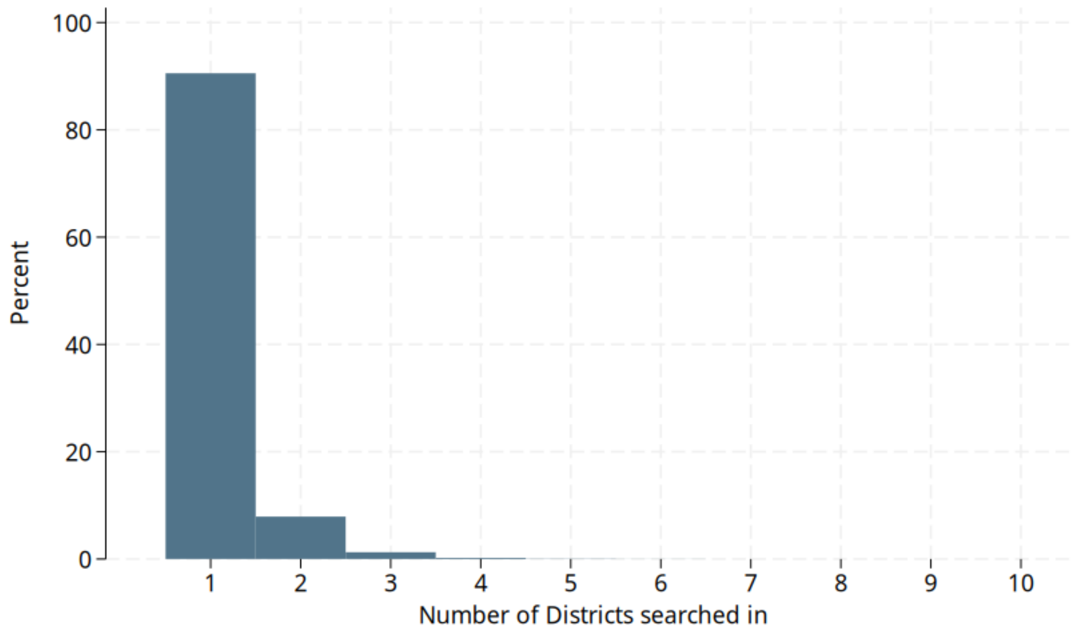


(b) Lag in calendar months



Notes: The figures show the lag in days and calendar months between a property transaction and the viewing by its buyer. This uses microdata from Nanjing City.

Figure A.4: Number of district where buyers made viewings



Notes: The figure shows the number of districts we observe individuals view properties. This uses microdata from Nanjing City.

Figure A.5: Roll-back of the housing-purchase restrictions

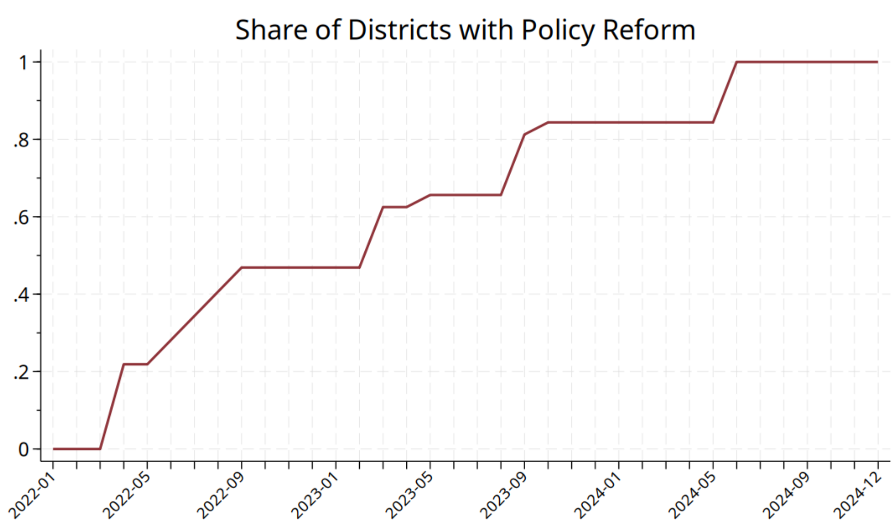


Figure A.6: Rollout of the removals by district

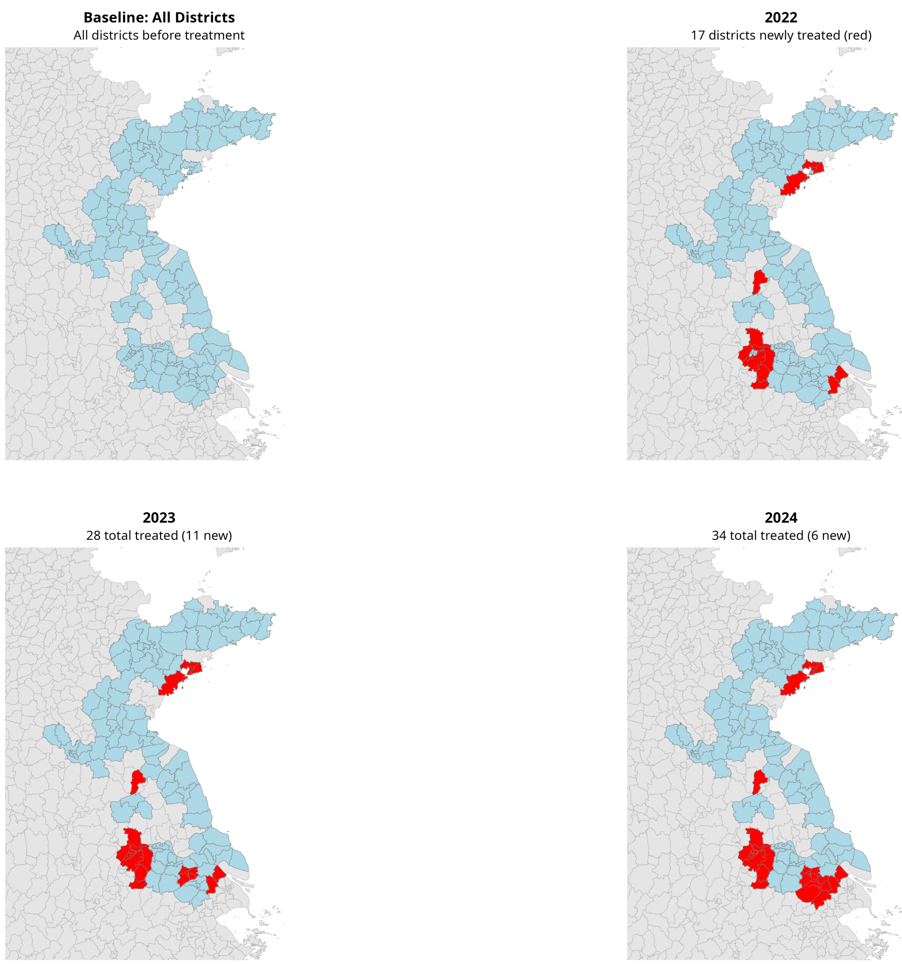
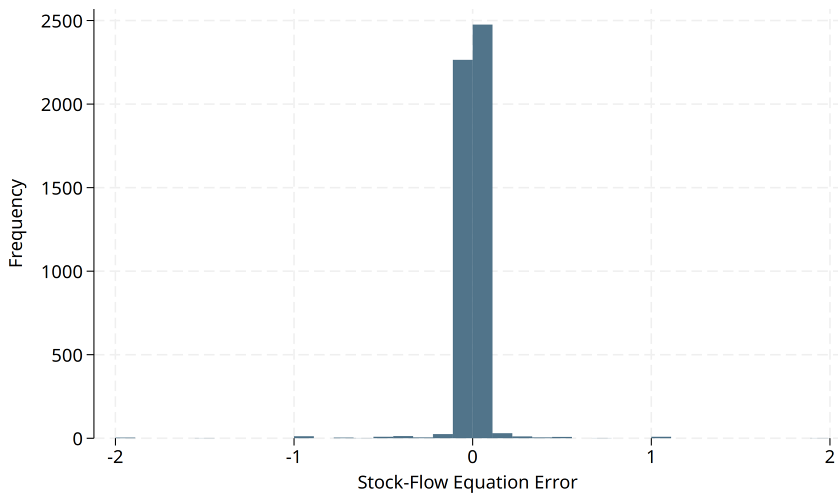
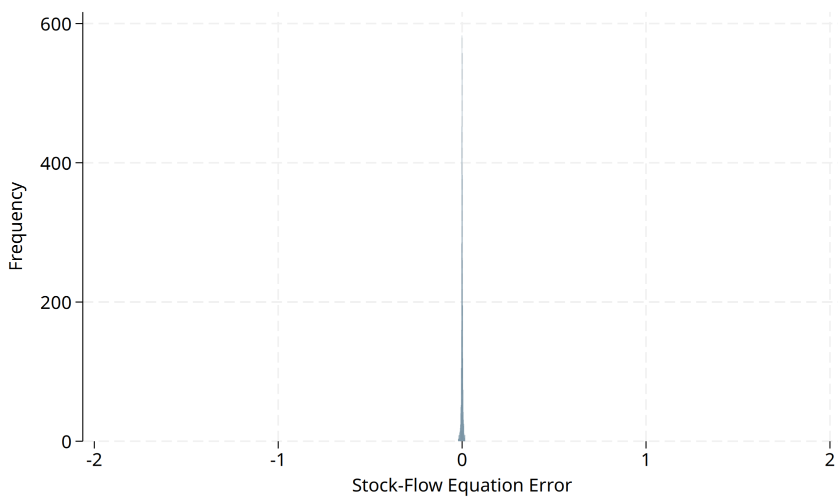


Figure A.7: Accounting identity errors

(a) Before filtering



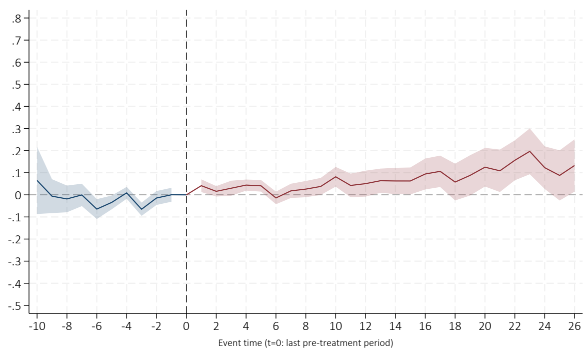
(b) After filtering



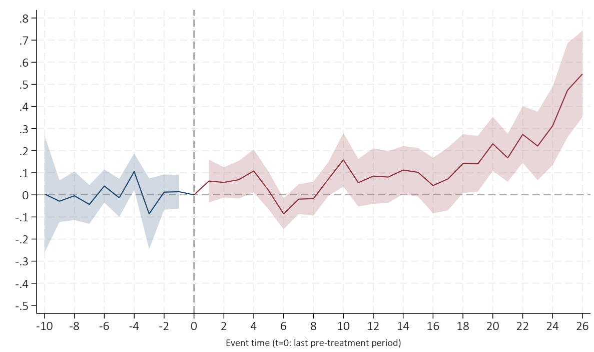
Notes: The figures visualize errors where there is a failure to satisfy the accounting identity (3) before and after we apply the restrictions described in section 3 to our data. The size of an error is given by subtracting the right-hand side from the left-hand side of (3) and normalizing by $U_{i,t+1}$.

Figure A.8: Event study

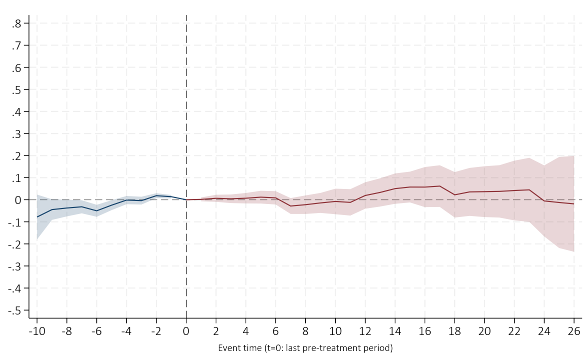
(a) log(clicks)



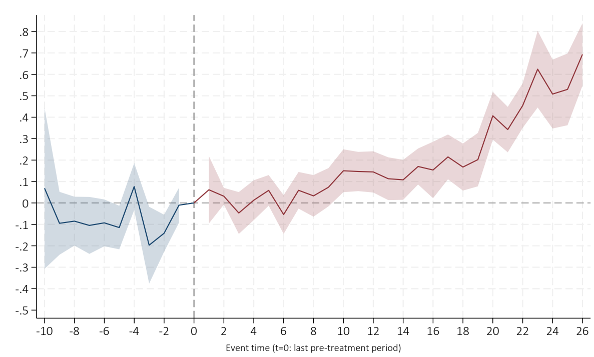
(b) log(effective buyers)



(c) log(inventory)



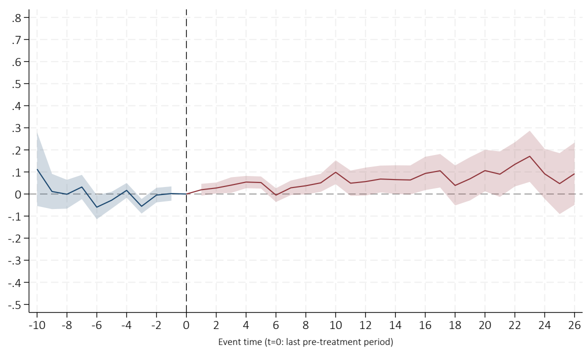
(d) log(transactions)



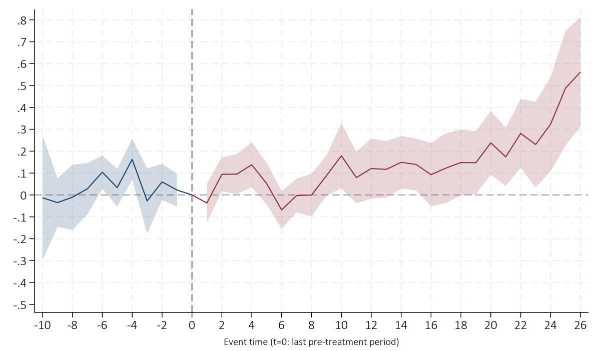
Notes: The bars indicate the 95 percent confidence intervals. Standard errors are clustered at the district level.

Figure A.9: Event study with policy control variables

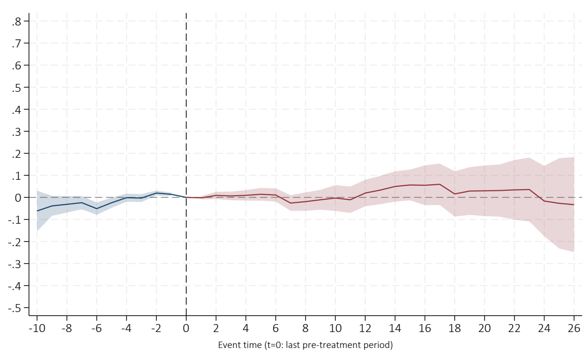
(a) log(clicks)



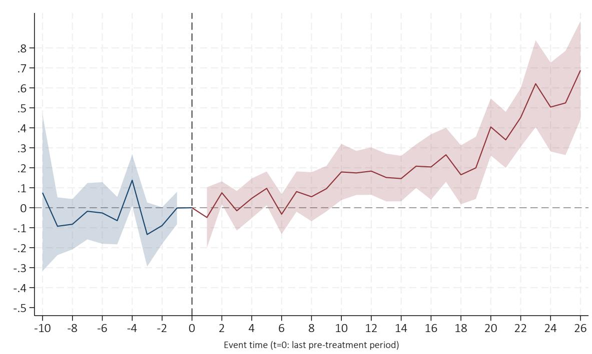
(b) log(effective buyers)



(c) log(inventory)



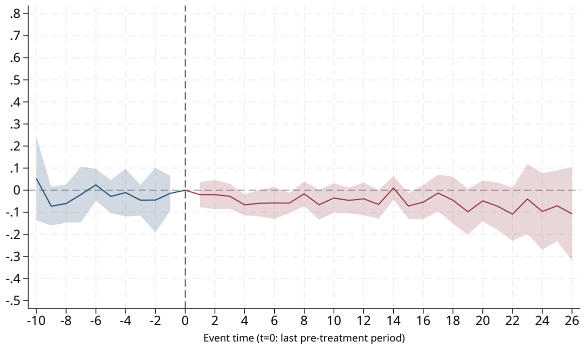
(d) log(transactions)



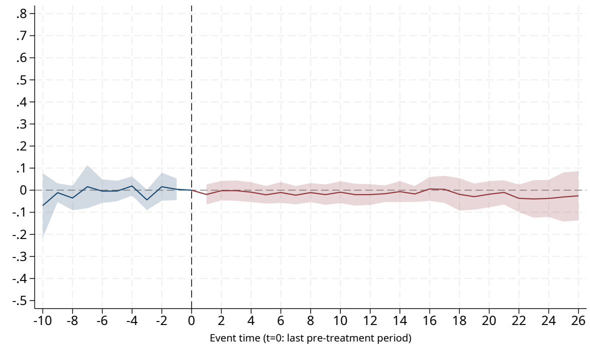
Notes: Event figures use binary policy control variables as explained in section 5.2. The bars indicate the 95 percent confidence intervals. Standard errors are clustered at the district level.

Figure A.10: Event study results on floor area

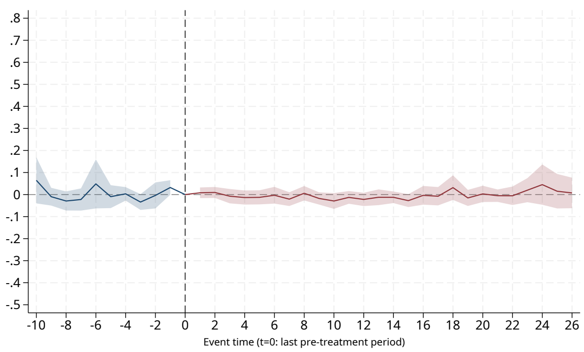
(a) 10th percentile



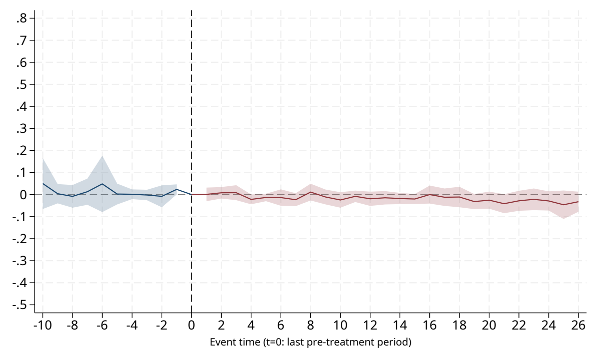
(b) 25th percentile



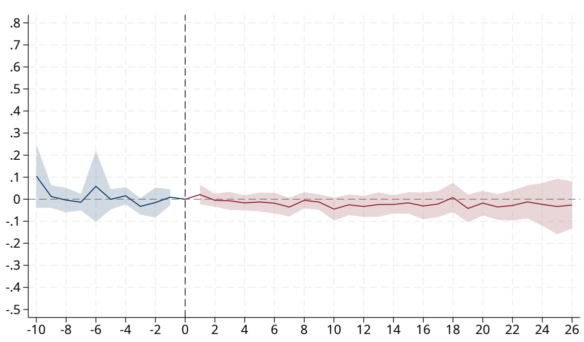
(c) Average



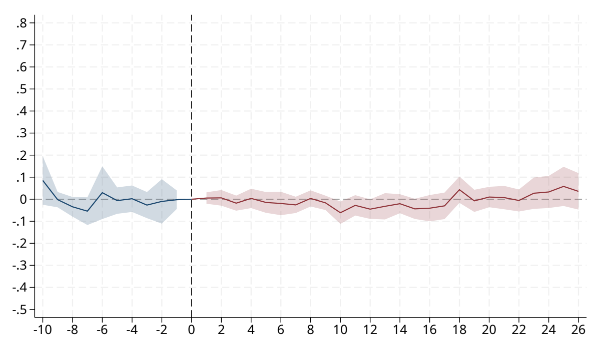
(d) Median



(e) 75th percentile



(f) 90th percentile



Notes: Event figures show the effect on quantiles and average floor area of sold properties, where quantiles and average are computed by district-month. The bars indicate the 95 percent confidence intervals. Standard errors are clustered at the district level.

B Additional tables

Table B.1: Housing data summary statistics

Variable	Mean	Std. Dev.
Inventory	11051	9482
Raw-Buyers	2824	3689
Viewings	8668	13363
Transactions	128	169
Withdrawals	1008	922
New Listings	1201	1162

Note: unit of observation is on district-year-month level. Time periods includes observations from 2021/11 to 2024/12. The table shows summary statistics after excluding district-year-month observations whenever one of the following conditions is true: a main variable of interest (buyers, transactions, or inventory) equals zero, the stock-flow equation error is in the top or bottom 5th percentile, or the inventory over population ratio is in the bottom 5th percentile. This leads to 116 unique districts.

Table B.2: Quarterly OLS estimation

	log(transactions)		log(transactions/inventory)
log(effective buyers)	0.391	0.351***	
	(0.430)	(0.065)	
log(inventory)	1.005	0.882***	
	(1.204)	(0.166)	
log(effective buyers)^2	0.028		
	(0.020)		
log(inventory)^2	0.018		
	(0.099)		
log(effective buyers) × log(inventory)	-0.056		
	(0.071)		
log(effective buyers/inventory)			0.389***
			(0.062)
Observations	1310	1310	1310
Districts	114	114	114
R^2	0.73	0.73	0.47
District FEs	✓	✓	✓
Time FEs	✓	✓	✓
H0: $\eta_b + \eta_u = 1$			
F-Statistic		3.373	
p-Value		0.069	

Notes: effective buyers = (average viewings per buyer) × (raw number of buyers) = total viewings. Standard errors clustered at district level in parentheses. Significance level indicated according to * p<0.1, ** p<0.05, *** p<0.01.

Table B.3: Determinants of policy timing

	Month relative to the first policy	
	(1)	(2)
urban district	11.025***	9.534***
	(2.555)	(1.819)
log(effective buyers/inventory)	1.155	-0.429
before April 2022	(1.354)	(1.144)
number of Evergrande projects	-0.421	0.030
per capita	(0.683)	(0.620)
population (in 1,000)	-0.416	-2.656
	(3.264)	(2.505)
fraction with local hukou	-62.454	-5.774
	(38.042)	(34.458)
fraction owning home	334.214	342.826
	(544.815)	(406.649)
number of rooms per household	-6.076	2.017
	(4.667)	(4.909)
per capita housing area	-0.309	-0.343
	(0.373)	(0.286)
years of education	-1.969	2.232
	(1.866)	(1.671)
Observations	34	34
R^2	0.69	0.88
City FE	N	Y

Notes: Significance level indicated according to *p<0.1, **p<0.05, ***p<0.01.

C Wald-style estimator

In this section we discuss an alternative route to identify the elasticity of interest by showing a Wald-style approach of taking the ratio of two weighted averages of cohort-time-specific ATTs is posing stronger assumptions on our identification.

Consider as starting point again

$$E\Big[\log\big(\frac{Q/U}{Q'/U'}\big)|Treatment\Big] = E\Big[\log\big(\frac{Z}{Z'}\big)|Treatment\Big] + \eta_d E\Big[\log\big(\frac{D/U}{D'/U'}\big)|Treatment\Big] \tag{12}$$

which we can re-write with some simplification to

$$\begin{aligned} ATT^{Q/U}_{cohort,time} &= ATT^{Z}_{cohort,time} + \eta_d ATT^{D/U}_{cohort,time} \\ \equiv \delta^{Q/U}_{g,t} &= \delta^{Z}_{g,t} + \eta_d \delta^{D/U}_{g,t} \end{aligned} \tag{13}$$

Our current approach is to get estimates on cohort-time level for $\delta^{Q/U}_{g,t}, \delta^{D/U}_{g,t}$ and then regress the previous equation with fixed effects. One should realize that for all (g,t) combinations for which it holds that $\delta^{Z}_{g,t} = 0$, we can also get

$$\eta_d = \frac{\delta^{Q/U}_{g,t}}{\delta^{D/U}_{g,t}} \tag{14}$$

Whenever in a given cohort-time combination $\delta^{Z}_{g,t} \neq 0$, this will not be true. The question is whether a weighted combination of $\delta^{Q/U}_{g,t}$ divided by a weighted combination for $\delta^{D/U}_{g,t}$ can also give us η_d. Let

$$\begin{aligned} \delta^{Q/U} &\equiv \sum_g \sum_t \omega(g,t)\delta^{Q/U}_{g,t} \\ \delta^{D/U} &\equiv \sum_g \sum_t \omega(g,t)\delta^{D/U}_{g,t} \end{aligned} \tag{15}$$

Then it holds that

$$\begin{aligned} \frac{\delta^{Q/U}}{\delta^{D/U}} &= \frac{\sum_g \sum_t \omega(g,t)\delta^{Q/U}_{g,t}}{\sum_g \sum_t \omega(g,t)\delta^{D/U}_{g,t}} \\ &= \frac{\sum_g \sum_t \omega(g,t)\big(\delta^{Z}_{g,t} + \eta_d \delta^{D/U}_{g,t}\big)}{\sum_g \sum_t \omega(g,t)\delta^{D/U}_{g,t}} \\ &= \frac{\sum_g \sum_t \omega(g,t)\delta^{Z}_{g,t}}{\sum_g \sum_t \omega(g,t)\delta^{D/U}_{g,t}} + \eta_d \frac{\sum_g \sum_t \omega(g,t)\delta^{D/U}_{g,t}}{\sum_g \sum_t \omega(g,t)\delta^{D/U}_{g,t}} \\ &= \frac{\sum_g \sum_t \omega(g,t)\delta^{Z}_{g,t}}{\sum_g \sum_t \omega(g,t)\delta^{D/U}_{g,t}} + \eta_d \end{aligned} \tag{16}$$

Therefore, we get that

$$\frac{\delta^{Q/U}}{\delta^{D/U}} = \eta_d \iff \sum_g \sum_t \omega(g,t)\delta^{Z}_{g,t} = 0 \tag{17}$$

This can be satisfied in two scenarios:

1. $\delta^Z_{g,t} = 0 \, \forall g, t$, that is, the average effect of the policy on Z is zero in all cohorts. In principle, one could still have effects in single districts, but only to the extent that these effects cancel themselves out to zero within cohorts.

2. In some cohort-time combinations, the effect of the policy is negative while in others it is positive and such that these effects perfectly cancel each other.

In contrast, in our main specification we are neither (1) assuming that the effect must be zero in all cohort-time periods, nor (2) that it must be negative for some cohorts and time periods and at the same time perfectly cancel with positive effects from other cohorts. However, we are assuming that the effects on Z and D/U are uncorrelated.

D Bootstrapped standard errors

The standard errors in Table 3 require special attention because both the outcome and the right-hand-side variable in equation (11) are themselves estimated. While bootstrapping standard errors is likely the appropriate response, in the following we contrast three different implementations thereof. Results are presented in Table D.1. While these approaches are different in nature, we emphasize that in practice the impact on the estimated standard errors of our main result is very minor, far from impacting conclusions on statistical significance.

The first approach (column 1 of Table D.1) treats the estimated ATTs as fixed and bootstraps only the standard errors for the regression of equation (11). In each of B replications, for G cohorts and T total time periods for which we have dynamic effect estimations, the $G \times T$ dataset of cohort-time ATT estimates is resampled with replacement, and equation (11) is re-estimated by OLS with cohort and time fixed effects. The standard error is the standard deviation of the B resulting coefficients. This procedure ignores that both the dependent and independent variables are estimates from the same underlying difference-in-differences structure, and therefore may be subject to correlated estimation error. The reported standard error captures sampling variability in the data of ATTs but not uncertainty transmitted from the estimation of these ATTs based on our underlying panel data on district-time level.

The second approach (column 2 of Table D.1) resamples the underlying district-month panel directly. In each replication, N district-time observations are drawn with replacement from the analysis panel of eventually-treated districts (that is, after sample restrictions are applied and never-treated districts have been removed). Since the same district-time pair may appear multiple times, duplicate pairs are assigned distinct unit identifiers so that the DID estimator can process them as separate panel units. Both ATT series are then re-estimated from the resampled panel, and the second-stage regression is run on the new ATTs. The standard error is again the standard deviation of the B stored coefficients. This method jointly re-estimates the full pipeline and therefore accounts for first-stage estimation error. Its considerable drawback is that resampling individual observations breaks the within-district time-series structure: a district whose observations are drawn at scattered time periods produces panel units with incomplete trajectories, undermining the pre- and post-treatment comparisons on which the DID estimator relies. If for a district the last period before treatment is not sampled, no ATT can be computed.[31]

[31] This problem occurs many times, because for a district-time period that is drawn multiple times, we must form separate (artificial) districts. For all of those for which we do not draw additionally the period before treatment, ATTs cannot be computed.

The third approach (column 3 of Table D.1) is a cluster bootstrap that resamples at the district level. In each replication, n district identifiers are drawn with replacement from the set of eventually-treated districts (after other sample restrictions have been applied), where n equals the number of such districts. For every drawn district – including duplicates – the district's complete time series of observations is retained and assigned a unique panel identifier. This produces a resampled panel of districts for which ATT trajectories of varying length can be computed. Both ATT series are re-estimated from this panel, and the second-stage regression yields a coefficient that is stored. The standard error is the standard deviation across replications. By keeping each district's full time series intact, this method preserves the within-district serial correlation and the panel structure required by the DID estimator. For this reason, we report this standard error for our main result.

Table D.1: DID estimation: estimating elasticity

	$\text{ATT}_{cohort,t}$ log(transactions/inventory)		
$\text{ATT}_{cohort,t}$ log(effective buyers/inventory)	0.499***	0.499***	0.499***
	(0.104)	(0.086)	(0.111)
Observations	140	140	140
Treat. Groups	7	7	7
Time Periods	26	26	26
R^2	0.804	0.804	0.804
Cohort FEs	✓	✓	✓
Time FEs	✓	✓	✓
Resampling	ATTs only	Whole panel	Districts
Used for main result (Table XY)			✓

Notes: effective buyers = (average viewings per buyer) × (number of buyers) = viewings. Dependent and independent variable are both estimates of ATTs over cohort and time. A cohort consists of all districts sharing the same first time period of treatment. Bootstrap standard errors in parentheses. See text for description of resampling schemes. Significance level indicated according to * p<0.1, ** p<0.05, *** p<0.01.

References

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer, "Identification Properties of Recent Production Function Estimators," *Econometrica*, 2015, *83* (6), 2411–2451.

Anenberg, Elliot and Patrick Bayer, "Endogenous Sources of Volatility in Housing Markets: The Joint Buyer-Seller Problem," *International Economic Review*, August 2020, *61* (3), 1195–1228.

_ **and Steven Laufer**, "A More Timely House Price Index," *Review of Economics and Statistics*, October 2017, *99*, 722–734.

Angrist, Joshua D. and Jörn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, December 2009.

Arellano, Manuel and Stephen Bond, "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *The Review of Economic Studies*, April 1991, *58* (2), 277.

Badarinza, Cristian, Vimal Balasubramaniam, and Tarun Ramadorai, "In Search of the Matching Function in the Housing Market," Working paper, ECB 2024.

Blundell, Richard and Stephen Bond, "GMM Estimation with persistent panel data: an application to production functions," *Econometric Reviews*, January 2000, *19* (3), 321–340.

Borowczyk-Martins, Daniel, Grégory Jolivet, and Fabien Postel-Vinay, "Accounting for endogeneity in matching function estimation," *Review of Economic Dynamics*, July 2013, *16* (3), 440–451.

Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo, "Understanding Booms and Busts in Housing Markets," *Journal of Political Economy*, August 2016, *124*, 1088–1147.

Callaway, Brantly and Pedro H.C. Sant'Anna, "Difference-in-Differences with multiple time periods," *Journal of Econometrics*, December 2021, *225* (2), 200–230.

Clark, W. A. V. and T. R. Smith, "Housing market search behavior and expected utility theory: 2. The process of search," *Environment and Planning A*, 1982, *14*, 717–737.

Clark, W.A.W., *Modelling Housing Market Search*, New York: Routledge, 1982.

Davis, Steven J., Jason Faberman, and John C. Haltiwanger, "The Establishment-Level Behavior of Vacancies and Hiring," *Quarterly Journal of Economics*, 2013, pp. 581–622.

de Chaisemartin, Clément and Xavier D'Haultfœuille, "Difference-in-Differences Estimators of Intertemporal Treatment Effects," *The Review of Economics and Statistics*, 02 2024, pp. 1–45.

Díaz, Antonia and Belén Jerez, "House Prices, Sales, and Time on the Market: A Search-Theoretic Framework," *International Economic Review*, August 2013, *54*, 837–872.

Gabrovski, Miroslav and Victor Ortego-Marti, "The cyclical behavior of the Beveridge Curve in the housing market," *Journal of Economic Theory*, May 2019, *181*, 361–381.

Garriga, Carlos and Aaron Hedlund, "Mortgage Debt, Consumption, and Illiquid Housing Markets in the Great Recession," *American Economic Review*, 2020, *110*, 1603–1634.

Gavazza, Alessandro, Simon Mongey, and Giovanni L. Violante, "Aggregate Recruiting Intensity," *American Economic Review*, 2018, *108* (8), 2088–2127.

Genesove, David and Lu Han, "Search and matching in the housing market," *Journal of*

Urban Economics, July 2012, *72*, 31–45.

Grindaker, Morten, Artashes Karapetyan, Espen R Moen, and Planmen Nenov, "Transaction Sequencing and House Price Pressures," Working paper, BI Norwegian Business School 2025.

Guren, Adam, "House Price Momentum and Strategic Complementarity," *Journal of Political Economy*, June 2018, *126* (3), 1172–1218.

Guren, Adam M and Timothy J McQuade, "How Do Foreclosures Exacerbate Housing Downturns?," *Review of Economic Studies*, 2020, *81*, 1331–1364.

Han, Lu, L. Rachel Ngai, and Kevin D. Sheedy, "To Own or to Rent? The Effects of Transaction Taxes on Housing Markets," *Review of Economic Studies*, 2025. forthcoming.

Head, Allen, Huw Lloyd-Ellis, and Hongfei Sun, "Search, Liquidity, and the Dynamics of House Prices and Construction," *American Economic Review*, April 2014, *104*, 1172–1210.

Kaas, Leo, Georgi Kocharkov, and Syrichas, "Understanding Spatial House Price Dynamics in a Housing Boom," *Economic Theory*, 2026. forthcoming.

Lange, Fabian and Theodore Papageorgiou, "Beyond Cobb-Douglas: Flexibly Estimating Matching Functions with Unobserved Matching Efficiency," Technical Report 26972, NBER April 2020.

Ngai, L Rachel and Kevin D Sheedy, "The Decision to Move House and Aggregate Housing-Market Dynamics," *Journal of the European Economic Association*, October 2020, *18* (5), 2487–2531.

Novy-Marx, Robert, "Hot and Cold Markets," *Real Estate Economics*, Spring 2009, *37*, 1–22.

Petrongolo, Barbara and Christopher A Pissarides, "Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature*, 2001, *39*, 390–431.

Piazzesi, Monika and Martin Schneider, "Momentum Traders in the Housing Market: Survey Evidence and a Search Model," *American Economic Review*, 2009, *99*, 406–11.

_ , _ , and Johannes Stroebel, "Segmented Housing Search," *American Economic Review*, March 2020, *110* (3), 720–59.

Pissarides, Christopher A, "Job Matchings with State Employment Agencies and Random Search," *Economic Journal*, 1979, *89*, 818–833.

_ , *Equilibrium Unemployment Theory*, 2nd ed., MIT Press, 2000.

van Dijk, Dorinth W. and Marc K. Francke, "Internet Search Behavior, Liquidity and Prices in the Housing Market," *Real Estate Economics*, 2018, *46* (2), 368–403.

Vidal, Pierre, "Estimating the housing market matching function through internet traffic analysis," Working paper, Meilleurs Agents Paris 2022.

Wang, Shaoda and David Y. Yang, "Policy Experimentation in China: The Political Economy of Policy Learning," *Journal of Political Economy*, 2025, *133* (7), 2180–2228.

Wheaton, William C, "Vacancy, Search, and Prices in a Housing Market Matching Model," *Journal of Political Economy*, December 1990, *98* (6), 1270–1292.

Yashiv, Eran, "The Determinants of Equilibrium Unemployment," *American Economic Review*, December 2000, *90* (5), 1297–1322.