When Price Stops Clearing: Valuation Disagreement and Housing Illiquidity*

Tai Lo Yeung†

August 10, 2026

Latest version

Abstract

When a housing market turns, transactions collapse while prices barely move: after Beijing's 2017 credit tightening, sales fell sixty-two percent, prices a tenth. Across 586,000 listings, homes in the same community and month that are hard to value from pre-listing comparables sell less often, wait longer, and disproportionately leave the platform unsold; the realized dispersion of asking prices, in turn, draws committed attention and widens bargaining. Listings priced near comparables sell readily where value is clear; where it is not, the price-execution schedule flattens. In the no-short-sale asset, disagreement should raise the price; borne in search, it is illiquidity.

Keywords: Housing illiquidity, idiosyncratic valuation uncertainty, divergence of opinion, search and matching, time on market, China housing market.
JEL Codes: G12, G14, R21, R31.

*Presented at the USI Economics Brown Bag Seminar and the Urban Economics Association Summer School at the LSE. I thank Lorenz Kueng, Francisco Amaral, Alberto Plazzi, Jessica Wachter, and participants at both venues for helpful comments and suggestions. All errors are my own.

†Yeung: Università della Svizzera italiana, Via Giuseppe Buffi 13, tai.lo.yeung@usi.ch.

1 Introduction

When a housing market turns, the transactions collapse before the prices move. Beijing's March 2017 credit tightening is a textbook case: within six months, resale transactions fell by sixty-two percent and stayed down for three years; the average price of the homes that still traded gave back a tenth from its peak and settled slightly above where it had started. The canonical explanations for this asymmetry are seller-side: owners hold out because their equity is thin (Stein, 1995; Genesove and Mayer, 1994) or because selling below the purchase price is painful (Genesove and Mayer, 2001; Andersen et al., 2022). This paper asks a different question of the same fact: not which owners hold out, but which homes stop trading. The answer, traced listing by listing: the homes that stop trading are, disproportionately, the ones whose value buyers cannot agree on: the price mechanism fails first where valuation is hardest, so the adjustment runs through quantities.

Disagreement among buyers should raise an asset's price. In housing it does the reverse: the home sells for less, and more often does not sell at all. The logic it overturns is plain. When an asset cannot be sold short, the most optimistic buyer sets the price, so the more buyers disagree, the higher it should go (Miller, 1977). No asset fits that premise better than a home, impossible to short and idiosyncratic in value, since a household owns one home, not a portfolio (Giacoletti, 2021; Peng and Thibodeau, 2017; Chetty, Sándor and Szeidl, 2017). But the optimist sets the price only if the market puts him in front of the asset. For a single home he must be found, one buyer at a time, and while the seller waits, disagreement is paid out in time, in a lower chance of sale, and in price, rather than banked into a higher one. Across 586,000 Beijing apartments, I compare units listed in the same community and month by how hard each is to value from contemporaneous local comparables. The dispersion behind that difficulty is visible directly: when two owners of the same apartment type list in the same month, their asking prices sit about four log points apart on average (Online Appendix OA.D.9). The hardest-to-value unit is less likely to sell, slower when it does, and sells at a discount, not a premium. On every attribute the data observe it is no worse than its neighbor; buyers simply cannot settle on its value. Disagreement is borne, not priced.

This paper measures what that costs, listing by listing, and Figure 1 is the central estimate. The mechanism runs through the one price a seller controls, the asking price set against comparable homes. Where comparables are clear, pricing near them sharply raises the chance of sale; where buyers cannot agree on what a home is worth, the same lever is dead, its slope falling monotonically from $+1.66$ in the clearest submarkets to a statistically-zero $+0.12$ in the noisiest (Section 5.2.2). Sellers do try it: they price low at listing, revise as often as anyone, and hold on longer before delisting. What fails is the conversion, and price ceases to clear the market.

The failure's terminal state is withdrawal. The hardest-to-value homes do not just sell slowly; many never sell at all. Comparing two units in the same community and month, the one buyers disagree about most is markedly less likely to sell, and roughly ninety-five percent of that shortfall appears as platform exit without an observed sale rather than as right-censoring: the home goes dark, withdrawn unsold, instead of waiting out a longer search (Section 5.1.2). The stakes are not marginal. The never-sold carry a combined asking value of roughly half a trillion yuan over 2016–2021; because they linger, more than a third of the listings a buyer browses on an average day will never transact on the platform; and when the market turned in 2017, the differential removal of hard-to-value homes is, as an accounting decomposition, equivalent to about twelve percent of the missing trades, an asset-side margin orthogonal to the seller-side down-payment channel to which Andersen et al. (2022) attribute most of the price–volume correlation.

These homes are not ignored: committed attention rises with the realized dispersion of pricing even as the sale fails.[1] They simply cannot be priced into a trade. The idiosyncratic second moment of housing value is therefore borne as illiquidity, in a discount the owner never earns back, instead of the return premium that an undiversifiable risk is supposed to command: where the premium reading of the same measure implies tens of basis points per year of compensation, the tail-robust estimate here is zero (Section 5.1.4).

The textbook prediction does hold at the market level, where a pool of optimists self-selects into trading and still sets prices, as speculative housing booms attest (Gao, Sockin and Xiong, 2021; Nathanson and Zwick, 2018). It inverts at the individual occupied home, the asset those same frictions wall off from speculators, because there the optimist has to be found rather than assumed. A search-and-matching framework makes the inversion precise. It also shows that optimist pricing is unreachable here: the price gradient stays negative across the whole empirically relevant range of meeting rates and well past it, so what governs the pricing of disagreement in housing is search frictions, not the short-sale constraint.

I study the Beijing resale apartment market from 2015 to 2021, using listing and transaction records for roughly 586,000 units on a major platform. Three features make it the right laboratory. Developers build residential communities (*xiaoqu*) that bundle closely comparable apartments under one set of schools, transit links, and management, so many near-identical units trade each month inside a stable boundary. The platform records the search a listing draws, its views and followers and days on market, next to its price. And every listing is followed to its terminal platform state (sale, withdrawal, or still-fresh censoring), so the probability of sale, the margin on which illiquidity ultimately bites, is observed rather than inferred. I measure a unit's valuation uncertainty as the dispersion of its hedonic pricing residual within community and month, how far its price sits from contemporaneous local

[1] Zillow confirmed the cost at industrial scale, winding down its algorithmic home-buying arm in 2021 after a year in which buying homes at model-set prices produced an $881 million loss (Zillow Group fourth-quarter and full-year 2021 results).

comparables, and I identify the friction by comparing units in the same community and month, which face the same demand, the same schools, and the same policy. The results open with the pattern in the raw data (Figure 7): across the distribution of valuation uncertainty, time on market rises and the probability of sale falls, and trimming the heavy upper tail of the proxy leaves the gradients intact. A cross-fitted, predetermined-feature index closes the loop: built only from pre-listing comparables (their number, recency, similarity, and dispersion, and the unit's configuration rarity), it reproduces the execution gradients with room to spare (sale -2.4 percentage points against -1.3, withdrawal $+2.4$ against $+1.2$; per within-cell standard deviation, about twice the realized gradient), so no result on whether and when a home sells leans on the seller's own ask (Section 2.4.3). The design's weight is carried by a falsification battery: Section 5.4 confronts thirteen alternative readings of the pattern (seller mispricing, unobserved quality, agent quality, loss aversion, transaction costs, and others), and Table 18 maps each to the discriminating evidence. A policy episode adds direction: when Xicheng's 2020 multi-school reform turned school assignment into a lottery, the probability of sale fell 8.5 percentage points against seven quarters of flat pre-trends, an estimate in the extreme tail of placebo-district assignments, though one treated district limits formal inference (Section 5.3).

A search framework with risk-averse, disagreeing buyers (Section 3) explains the pattern through two channels the data can tell apart. Valuation difficulty moves two moments of buyers' valuations at once. It lowers their mean, a risk discount that shades every bid, and it widens their dispersion, because buyers disagree more about what a hard-to-value home is worth. A single co-movement shows both are at work: the transaction price falls even as the disagreement margins rise, a wider negotiated spread and more committed attention than the longer spell would mechanically produce. Neither channel alone can generate that pattern, and it pins down disagreement on both sides of the market. The framework rests on one information primitive, holds up under a fully optimal seller who never anchors, and its quantitative version is estimated, fitting five liquidity and funnel gradients with three parameters.

The contribution is one mechanism, read off margins a price-level regression cannot see. The closest measurement papers infer valuation uncertainty from the dispersion of prices among homes that actually transact (Giacoletti, 2021; Sagi, 2021; Amaral, 2024); I read it off the clearing function itself, from whether and at what price a home trades at all. Three results follow. The friction is extensive, the mark of a home that withdraws rather than one that merely waits. Valuation uncertainty governs the slope of the seller's demand curve. And it splits into a risk discount and two-sided disagreement, recovered on both sides of the market where the closest prior study recovers neither. This is a within-market measurement rather than a natural experiment, so it earns its causal language sparingly, holding the community and month fixed and leaning on falsification, decomposition, and within-unit tests in place of a single shock. The setting is Beijing, but nothing in the framework is: its primitives fit any

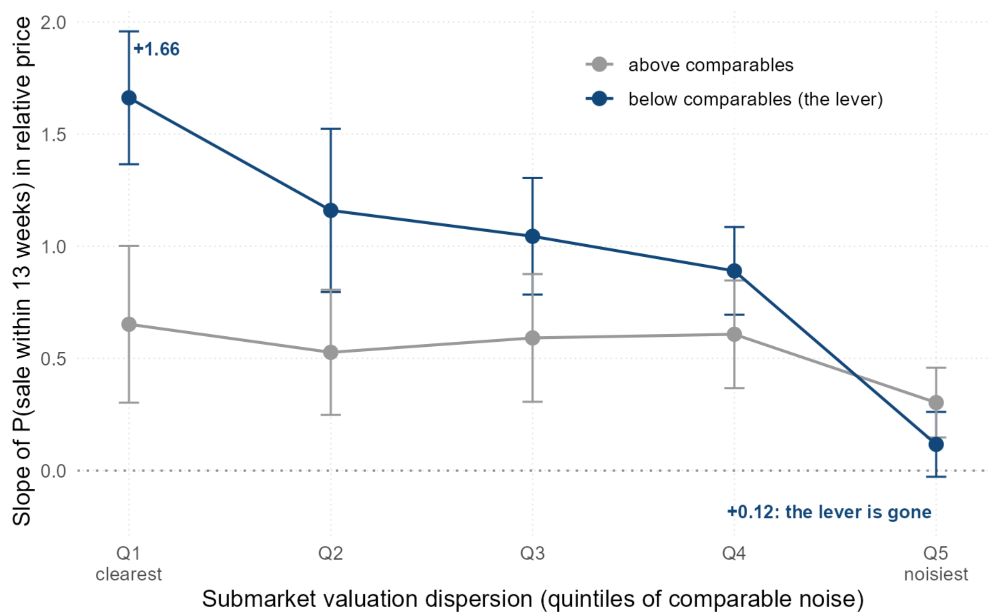


Figure 1: The Price Lever Dies Where Homes Are Hard to Value

Notes. The slope of the probability of sale within thirteen weeks in the listing's price relative to comparable homes (the within-community–month quality-adjusted residual of the log asking price), estimated separately within each quintile of submarket valuation dispersion with community×month fixed effects. Navy: the discount margin, the lever; grey: the overpricing margin, sign flipped so that moving toward comparables reads positive on both. Bars are 95 percent confidence intervals clustered by community. The lever falls monotonically from $+1.66$ in the clearest-comparable fifth to $+0.12$, statistically indistinguishable from zero, in the noisiest. Submarkets are quintiles of the leave-one-out community–month mean of the squared hedonic residual; sample: Beijing listings posted from July 2016 with thirteen-week follow-up, in community–months with at least eight listings ($N = 82{,}400$). Section 5.2.2 reports the estimates (Table 13) and the binned curves behind them.

decentralized listing market without a short side, and the mechanism reappears across four more provinces and some thirty-six cities, with only the probability-of-sale margin confined to Beijing (Section 7).

Related literature and contribution. These results speak to four literatures. A literature on idiosyncratic housing risk establishes that it is large and realizes at sale (Giacoletti, 2021; Sagi, 2021; Han, 2013); both facts reproduce here: Giacoletti's predictive evidence replicates on Beijing's ex-post dispersion, and a Sagi decomposition puts the at-sale variance component at 0.0197 over the full sample, rising to 0.0296 in the pre-tightening subsample closest to his market regime, against his 0.0326 (Online Appendices OA.C.5 and OA.D.7). What that literature cannot observe without listing records is the anatomy of the at-sale realization: the search, the price cuts, and the withdrawals through which it arrives. Those margins are this paper's object. Where the housing-search canon models risk-neutral buyers and behavioral sellers (Wheaton, 1990; Genesove and Mayer, 2001; Andersen et al., 2022), I supply the buyer-side valuation object (risk-averse, disagreeing buyers whose uncertainty moves two moments at once) and the co-movement that identifies both.

First, the mechanism is a search-market expression of divergence of opinion (surveyed in Hong and Stein, 2007). Miller (1977)'s logic (no shorting, so the optimist sets the price) grew into priced disagreement in speculative markets and credit (Harrison and Kreps, 1978; Scheinkman and Xiong, 2003; Xiong, 2013; Buraschi, Trojani and Vedolin, 2014), carries strong equity support (Diether, Malloy and Scherbina, 2002; Stambaugh, Yu and Yuan, 2015), and has a housing counterpart in the dispersion of professional forecasts (Li, Van Nieuwerburgh and Wang, 2026). The binding friction here is locating the counterparty, as in over-the-counter markets (Duffie, Gârleanu and Pedersen, 2005), and the distinction is market-wide versus unit-level disagreement: a market-wide pool of optimists self-selects into trading and still sets prices (Piazzesi and Schneider, 2009), while the buyer who values a particular home must be found listing by listing.[2] The market's clearing technology decides where that disagreement leaves its mark: in a centralized market it surfaces as portfolio flows, whose smallness against volatile prices implies a large price impact (van der Beck, Bretscher and Fu, 2026), whereas in a search market, where trade requires a match rather than a price, it surfaces instead as illiquidity. That belief dispersion is informationally distinct from uncertainty, and need not map to prices with the same sign, recurs across asset classes: Badidi (2026) finds forecast dispersion dampens the price impact of macroeconomic news (opposite to a pure-uncertainty reading), which is why I recover disagreement as an object separate from the risk discount rather than treat one as a proxy for the other. The paper thus adjudicates whether valuation disagreement is borne as a price (Buraschi, Trojani and Vedolin, 2014), as return compensation (Carlin, Longstaff and Matoba, 2014), or, as Zhang (2006) finds for equity returns, not unconditionally priced; here it is borne as illiquidity.

Second, an asset-pricing literature asks how idiosyncratic second moments are borne. In equities the volatility–return relation is contested (Ang et al., 2006), tail-driven where positive (Bali, Cakici and Whitelaw, 2011), and negative where hard-to-value stocks are overpriced under arbitrage asymmetry (Stambaugh, Yu and Yuan, 2015), with institutions managing the exposure on the quantity margin (Di Maggio et al., 2023).[3] In housing, idiosyncratic price risk is large and holding-period dependent (Giacoletti, 2021; Sagi, 2021; Campbell et al., 2001); Peng and Thibodeau (2017) measure it as hedonic-residual dispersion (the object I use) and find it uncompensated by appreciation, leaving open how it is borne. Closest is Amaral (2024), who reports the same gradients in German apartments as a check

[2]Segmented-search models in which buyer composition shapes price dispersion (Piazzesi, Schneider and Stroebel, 2020) offer a complementary route to the same dispersion.

[3]The title nods to the "hard to value, hard to arbitrage" cross-section of Baker and Wurgler (2006). In housing there is no arbitrageur, which is why the same valuation difficulty surfaces as illiquidity rather than as the mispricing it produces in equities. The contrast sharpens where the second moment is directly tradable: in equity options, anticipated volatility has an instrument, a market price, and a paid absorber in the market makers who profitably warehouse retail's demand for it (de Silva, So and Smith, 2026), so the cost of valuation difficulty is a transfer between investors. A home has none of the three, and the cost is deadweight: the dark shelf of Section 5.1.2, which no counterparty is paid to absorb.

on a rental-yield premium. I depart in four ways: illiquidity is the object rather than a mechanism check; committed attention rises with uncertainty, a demand-side margin his design does not separate; I separate the two channels and measure disagreement on both sides of the market; and the tail-immune rank overturns the premium reading: realized resale returns show no robust premium, with any positive tilt arising through selection on completed sales (Section 5.1.4). His yield compensation is largely the discount re-expressed (the same rent over a lower price; Himmelberg, Mayer and Sinai, 2005) and is weakest where rental and sale markets are segmented; in Beijing, rental yields sat below the one-year risk-free rate (Liu and Xiong, 2018). The framing of housing as a concentrated, undiversifiable household asset draws on Chetty, Sándor and Szeidl (2017); Cocco (2005); Adelino, Schoar and Severino (2018); ownership hedges the aggregate rent–price risk (Sinai and Souleles, 2005), the market-level risk whose pricing Han (2013) shows hedging demand shapes, so what remains for an owner-occupier is the idiosyncratic component studied here, in the joint price–quantity treatment Goetzmann, Spaenjers and Van Nieuwerburgh (2021) call for.

Third, a microstructure literature reads liquidity off the dispersion of pricing errors: in Treasuries (Hu, Pan and Wang, 2013), internationally (Malkhozov et al., 2017), in U.S. housing at the county level (Jiang, Kotova and Zhang, 2024*a*), and in Israeli listings at the segment level (Ben-Shahar and Golan, 2022). I move the measure to the individual home within a community and month, identifying which homes are illiquid, not which markets, and replace the tail-driven level with its within-cell rank. The within-cell signature is not an artifact of the squared-residual proxy: it survives replacing the residual with a *price-free* measure built from physical attributes alone (a within-community configuration surprisal, in the spirit of Amaral's characteristic-thickness instruments and Haurin's atypicality), which reproduces the full signature price-free (Section 5.4.2), and with a transaction-only comparable-scarcity count that travels to four further provinces. The construction adapts Jiang and Zhang (2025); the unobserved-quality concern of Halket, Nesheim and Oswald (2020) is differenced out within community.

Fourth, the paper draws on search and attention in housing. Decentralized-search models predict that harder-to-value listings search longer (Wheaton, 1990; Allen, Clark and Houde, 2019; Head, Lloyd-Ellis and Sun, 2014; Han and Strange, 2015; Goetzmann, Spaenjers and Van Nieuwerburgh, 2021); Jiang, Kotova and Zhang (2024*b*) show time on market is interpretable only jointly with prices and dispersion (the logic the signature test applies listing by listing), and Badarinza, Balasubramaniam and Ramadorai (2024) estimate the matching function across observed funnel stages. The empirical ancestor is Haurin (1988): atypical homes draw a wider variance of offers and market longer; I isolate the idiosyncratic component orthogonal to that observable atypicality. Intermediaries' revealed preference points the same way: iBuyers transact only the easiest-to-value, cookie-cutter homes (Buchak et al., 2025).[4] Attention, finally, is usually proxied (Gabaix, 2019; Barber and Odean, 2008);

[4]A related literature studies seller behavior and the listing–transaction wedge: Genesove and Mayer (1994,

recent work measures it directly for stocks (Da, Engelberg and Gao, 2011; Barber et al., 2022). I observe it directly as saved-listing followers, and its rise alongside a falling price separates the dispersion channel from price-pressure stories in which attention precedes a run-up (Barber et al., 2022). The pull of an anticipated second moment is not housing-specific: retail option purchases rise monotonically in expected announcement volatility, measured before the buying it attracts, and the buyers so drawn lose ten to fourteen percent on the highest-anticipation events (de Silva, So and Smith, 2026). A search market changes the outlet, not the attraction: the buyer drawn to a hard-to-value home has no contract to buy, so the engagement ends as committed following, not a position.[5]

Roadmap. Section 2 introduces the *xiaoqu* laboratory and the listing, transaction, and attention records; Section 3 builds the model that turns one information primitive into the two channels the data can separate; Section 4 states the design and the assumption it rests on. Section 5 then presents the evidence in the order a skeptic would ask for it: the liquidity gradients, the abandonment decomposition, the flattening clearing function, the mechanism tests, and the thirteen alternative readings Table 18 closes. Section 7 adds robustness and the cross-province replication; Section 8 concludes with the cost of the friction and the reform that could price it.

2 Institutional Setting and Data

This section provides institutional context for the feature of China's resale housing market that is central to the empirical design: the *xiaoqu* (residential community) system, which creates tight local submarkets with stable boundaries and dense within-market comparables. I then describe the data, the construction of the key variables, and how the institutional features connect to the model developed in Section 3. Two background facts about the broader market matter for reading the results (Glaeser et al., 2017; Liu and Xiong, 2018): purchases are equity-financed (down payments of thirty percent or more, low household leverage), so the buyer of an apartment bears its idiosyncratic value risk with own wealth; and the primary land and new-home markets are dominated by institutional sellers and actively price-managed (Chang, Wang and Xiong, 2025). The decentralized, household-to-household resale market studied here is where valuation uncertainty and search frictions operate and can be observed.

2001) on equity and loss aversion in time to sale; Andersen et al. (2022) on reference dependence; Glower, Haurin and Hendershott (1998); Genesove and Han (2012); Anenberg (2016); Carrillo (2013) on seller motivation, search, and information; Garnache, Schmalz and Sommervoll (n.d.); Gargano and Giacoletti (2021); Giacoletti and Parsons (2022) on sticky prices, auction underquoting, and reference-point spillovers.

[5] Local price experiences, social networks, and learning from noisy local prices also shape housing demand (Bailey et al., 2018; Gargano, Giacoletti and Jarnecic, 2023; Gao, Sockin and Xiong, 2021).

2.1 Residential Communities (*Xiaoqu*)

A key feature that makes this research feasible is the distinctive organization of residential housing in Chinese cities. Since the 1980s, urban housing has been delivered through *xiaoqu* (*residential communities*): gated, multi-building complexes developed at scale by large, often highly levered developers. Xiaoqu evolved from the socialist *danwei* (work-unit) housing system, which organized daily life around geographically concentrated bundles of housing and services (security, maintenance, small retail, nearby schools and clinics). With housing-market liberalization, xiaoqu construction continued under market-based differentiation: today, apartments within xiaoqu are privately owned, and compounds are managed by professional property-management companies. Xiaoqu scale varies widely, from single-building compounds to very large developments such as Tiantongyuan in Beijing, which houses hundreds of thousands of residents.[6]

Two implications of the xiaoqu system are central to the empirical design.

Xiaoqu as economically meaningful submarkets. Units inside a community typically share the same school-district assignment, property-management quality, access to transit and commercial amenities, and local governance environment. Because developers brand and market communities as integrated products, the community identity functions as a durable bundle of location and quality attributes that persists over time. This makes the community the natural unit for defining local comparables and fixed effects: within a given community–month cell, listings are benchmarked against contemporaneous transactions of units that share the same bundle of neighborhood characteristics.

Within-community heterogeneity motivates hedonic adjustment. Despite sharing community-level attributes, individual units still differ in size, floor, orientation, layout, renovation status, and building vintage. Hedonic adjustment is therefore needed to make listings and transactions comparable. The residual dispersion after conditioning on community–month fixed effects and an extensive set of unit attributes captures the apartment-specific component of pricing deviation: exactly the object that the theoretical framework identifies as idiosyncratic valuation uncertainty. Relative to settings where comparables are sparse (for example, the German apartment markets studied by Amaral 2024, where repeat-sales identification requires tracking the same unit across decades), the xiaoqu structure provides a much tighter comparison set at monthly frequency, mitigating a central identification challenge in illiquid asset pricing.

[6]Analogous institutions exist in other former socialist countries, for example the Soviet *microraion*, but China's version is distinctive in its combination of large developer-led construction, private ownership, and active resale markets on digital platforms.

2.2 Data and sample

The empirical analysis uses listing-level records from one of China's largest online real estate platforms, matched to subsequent transaction outcomes and to platform engagement metrics. The sample focuses on Beijing, China's policy center and most closely watched resale market, with outcomes observed through January 2021: 92 percent of listings are posted from January 2015 onward, posting activity ends in December 2020, and the platform records a thinner backlog of 2011–2014 postings whose exclusion leaves every estimate unchanged or slightly stronger (Section 7).[7] The unit of observation is a listing i posted in month t in community $c(i)$. The full listing sample is used for listing-side objects (prices, attention, and listing-time uncertainty); the matched listing–transaction subsample is used for deal-side outcomes (realized gains and the listing–deal wedge). Selection into transaction is endogenous, and Section 7 confirms that the main results are robust to restricting attention to matched transactions only.

Sample construction. I drop listings missing the posting date, posting price, size, or construction year, and units larger than 1,800 square meters as likely commercial or data-entry errors. For listings that match to a completed transaction, I follow standard practice and retain pairs whose deal price lies within 50–150% of the listing price (Amaral, 2024), excluding non-arm's-length transactions. The resulting Beijing sample contains 586,312 listings across 16 districts, of which 71 percent (414,979) match to a completed transaction within the sample window. The unsold remainder is not an omission. It is what makes the probability-of-sale margin estimable: deal-side objects (transaction prices, the listing–deal wedge, days on market) are defined on the matched subsample, while the sale indicator, listing prices, and attention are defined for every listing. Three districts account for 58% of the sample (Chaoyang with 181,000 observations or 31%, Haidian with 80,000 or 14%, and Fengtai with 78,000 or 13%). The median apartment is 76 square meters, built in 2002, and listed at approximately 3.98 million RMB. Table 1 reports the Beijing-wide summary statistics, with district-level panels in Online Appendix OA.C.2. Figure 2 maps average listing prices across Beijing's administrative divisions at the start (2015) and toward the end of the sample (2017).

Key measurement objects. The paper's primary outcomes are the illiquidity margins observed for each listing: an indicator for whether the unit sells, its time on market, and the platform attention (followers and views) it accumulates before clearing. The uncertainty discount is the cross-sectional slope of an outcome on valuation uncertainty, estimated within community and month.

[7] Robustness exercises in Section 7 extend the analysis to Guangdong, Henan, Zhejiang, and Shandong.

Table 1: Summary statistics, Beijing

	N	Mean	SD	P25	Median	P75
Prices, characteristics, and the uncertainty measure						
Listing Price (ten thousand RMB)	586 312	471	327	280	398	560
Deal Price (matched sales)	414 979	446	272	280	390	540
Size (m^2)	586 312	85.5	41.6	58.4	76.3	99.8
Construction year	586 312	2000	9.28	1994	2002	2007
Valuation Uncertainty	586 312	0.0138	0.1640	0.0006	0.0028	0.0093
Listing outcomes: the illiquidity margins						
Sold within window ($= 1$)	586 312	0.71	0.45	0	1	1
Days on market (sold)	414 978	70.4	110.0	6	30	87
Followers	586 312	35.4	88.2	3	17	44
Views	416 435	1 961	3 219	50	936	2 640
Number of districts	16					

Notes. The table reports Beijing-wide summary statistics for the sample period January 2015 to January 2021. Valuation uncertainty is the squared residual from the hedonic specification described in Section 2.4.2. All prices are nominal and reported in ten-thousand RMB; deal prices and days on market are defined on the 414,979 listings matched to a completed transaction, while the sale indicator, listing prices, and follower attention are defined for every listing (view counts are recorded by the platform for a subset of listings). The median listing price for Beijing is approximately 3.98 million RMB ($550,000 at the average sample-period exchange rate), and the median apartment is 76 m^2 built in 2002. Per-district panels for Chaoyang, Haidian, and Fengtai appear in Online Appendix OA.C.2.

Table 2: Analysis Samples at a Glance

Sample	N	Outcomes defined on it
All listings	586,312	Sale indicator, listing price, followers, uncertainty
with recorded view counts	416,435	Views, follower-to-viewing ratio (FVR)
Matched to a completed transaction	414,979	Deal price, days on market, list–deal wedge
with showings records	250,825	Agent-accompanied showings (*daikan*)
Attention-screened baseline (FVR ≤ 1)	267,911	Headline gradients (Table 6)
Three-fate accounting sample	468,698	Sale / withdrawal / censoring split (Table 7)

Notes. Each analysis object is estimated on the largest sample on which it is defined, and every results table states its sample in its notes. The attention-screened baseline drops listings with a follower-to-viewing ratio above one, which are typically driven by missing or miscoded view counts (Section 2.4.4), and its duration outcomes further restrict to cohorts with recorded (not backfilled) posting dates. The three-fate sample assigns each listing to exactly one terminal state, with the withdrawal state requiring the listing's last recorded day to fall at least sixty days before the data end (Section 5.1.2).

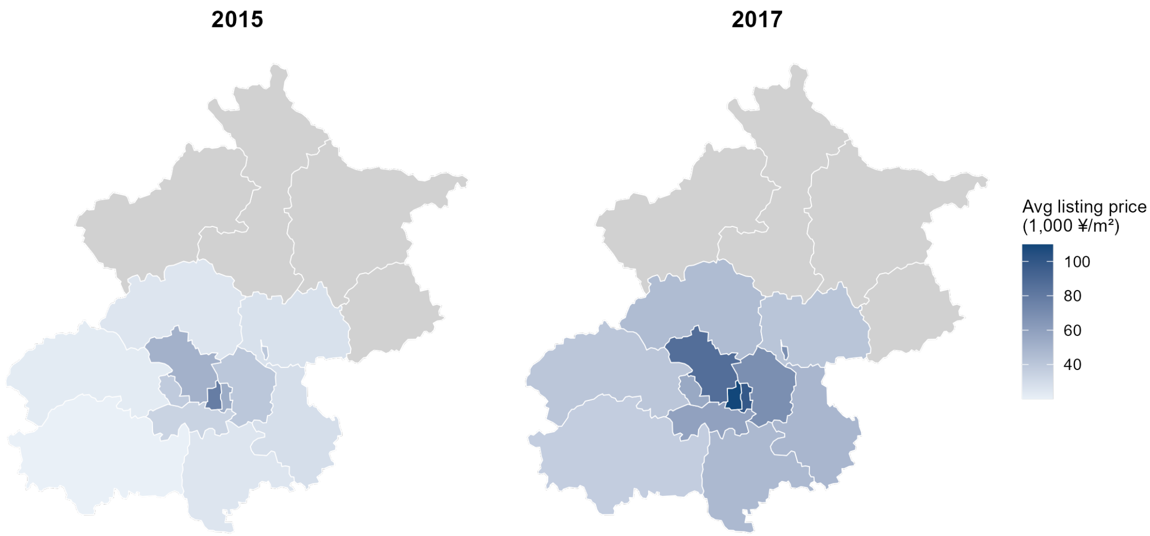


Figure 2: Average Listing Prices in Beijing Across Administrative Divisions

Notes. This figure maps the average listing price per m^2 across Beijing's administrative divisions in 2015 and 2017, on a common color scale. Data are from the online real estate platform (analysis sample, Table 1); divisions with fewer than twenty listings in a year are shown in grey.

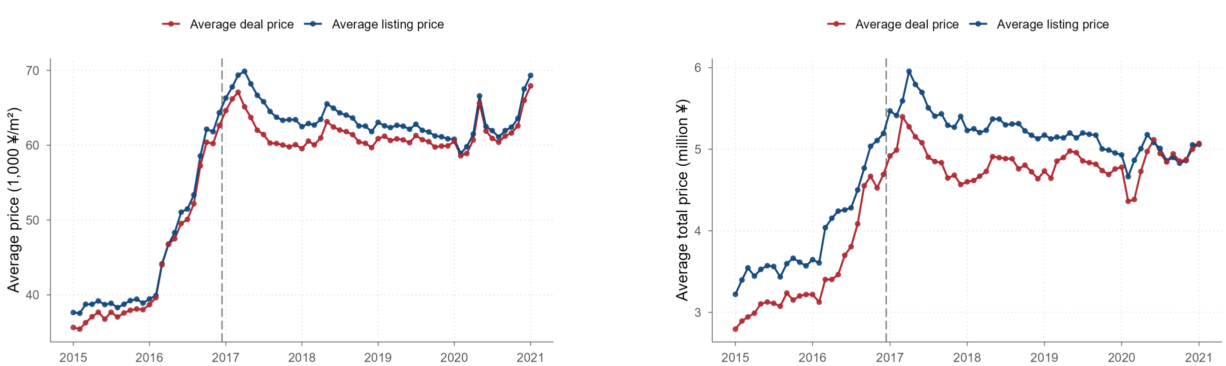
2.3 Market-level context

Figure 3 plots the monthly average listing price, matched deal price, and implied listing–deal wedge. Listing and deal prices co-move, but the wedge varies meaningfully over the 2015–2021 sample. In the model, this wedge is the difference between a listing price anchored on recent community comparables and the transaction price struck with the buyer actually met; Proposition OA.5 relates the wedge to a property's valuation uncertainty. Figure 4 adds the liquidity benchmark. Monthly average days-on-market and transaction volume co-move with the pricing series, consistent with a market in which liquidity and pricing are jointly determined rather than a stable market with a fixed listing composition (Wheaton, 1990; Genesove and Mayer, 1994).

The sign of that wedge is set by how the market clears, not by the balance of supply and demand. Beijing's resale market clears through bilateral search and bargaining: a seller posts an asking price, an anchor set above her reservation, and the platform and agent bring buyers to view, follow, and request showings one at a time, after which the transaction price is negotiated with the single buyer who is matched and lands below the ask.[8] A positive listing–deal wedge is therefore a structural feature of the clearing mechanism,

[8]Outside the judicial-auction segment, no purely online marketplace exists in Chinese housing: mainstream platforms circulate listing information while inspection and bargaining remain physical (Xu, Zheng and Wu, 2024). The platform record analyzed here is therefore the search log of a negotiated market, not an order book.

present in hot and cold markets alike, not a symptom of weak demand. Markets that clear by auction invert the sign: there the asking price is an invitation to bid, excess demand resolves into competitive over-bidding above the ask, and the winning bid lands above it, as documented for the Australian open-outcry market, where sale prices routinely exceed the quote (Gargano and Giacoletti, 2021). China runs the counterfactual clearing technology on the same asset class: judicial foreclosure auctions, online-only nationwide since 2017, pool registered bidders in open competition, answer a failed auction with a mandated twenty-percent cut to the starting price, and eventually clear more than sixty-five percent of listed dwellings at prices averaging within one percent of third-party appraisal (Xu, Zheng and Wu, 2024). Where an institution forces the price margin to absorb failure, the home clears; in the bilateral market studied here, the seller who controls that margin withdraws the home instead (Section 5.1.2). This institutional contrast is the one the paper turns on: when the buyer who most values a particular unit must be found, one meeting at a time, rather than delivered to the seller by an auction that lets the optimist set the price, divergence of opinion is borne in the negotiated discount and the search required to clear, not in a winning bid above the ask. Even at the peak of the 2015–16 boom the typical deal stayed below the ask (above-ask sales rose only from roughly fifteen to twenty percent of matched transactions and receded after the 2017 tightening; Section 5.2.5), because the resale market offers no open-bid venue in which competition could push the price past the ask; competition surfaces instead as a faster sale at a smaller discount.



(a) **Listing vs. matched deal prices**

(b) **Total apartment prices over time**

Figure 3: Listing prices, deal prices, and the listing premium over time

Notes. Panel (a) plots the monthly average nominal listing price (blue) and the average nominal deal price (red) per m^2, indexed by the listing date. Panel (b) plots the monthly average nominal listing price and deal price as absolute total apartment prices (in RMB) over time. The vertical dashed-dot line marks the December 2016 Central Economic Work Conference. Sample: Beijing apartment listings, January 2015 to January 2021. Panel (a) prices are nominal RMB per m^2; Panel (b) prices are nominal RMB total apartment prices.

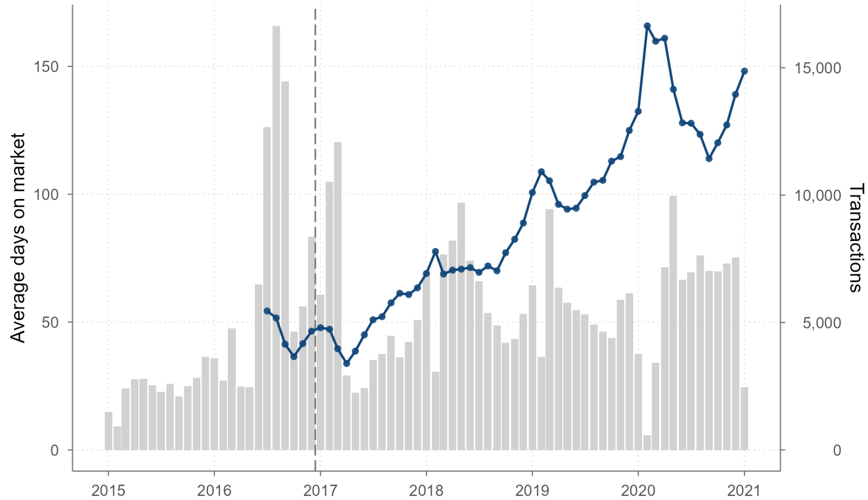


Figure 4: Average Days on Market and Transaction Volume Over Time

Notes. Bars show monthly transaction counts (right axis); the line shows the average days on market of transactions completed in each month (left axis). The days-on-market series begins in July 2016: earlier transaction records carry posting dates backfilled to the deal date (the platform's listing history begins in mid-2016), so durations are not measured before then. Indexing by deal month avoids the right-edge censoring of a posting-cohort index, in which the most recent cohorts contain only their fastest sales; each plotted month is the month's recorded transactions among cohorts posted from 2015, so the late-sample rise is the market slowing, not an endpoint artifact. The vertical line indicates the timing of the 2016 Central Economic Work Conference.

2.4 Variable construction

2.4.1 Quality-adjusted prices

To make prices comparable within local markets, I residualize observed prices on a hedonic specification with community fixed effects, year–month fixed effects, and unit attributes (size, construction year, floor group, construction type, orientation, layout, building type), separately by market segment $m \in \{\text{list}, \text{deal}\}$. Let $\widetilde{P}^m_{i,t}$ denote the resulting quality-adjusted price per square meter. The listing-implied pricing wedge $\text{ExpRaw}^{(1m)}_{i,t}$ and the realized capital gain built from these prices are secondary, price-side objects used only in the supporting analysis; their definitions, and the time series that validates the wedge against realized gains, are collected in Online Appendix OA.C.6.

2.4.2 Ex-ante valuation uncertainty

I quantify apartment-level ex-ante valuation uncertainty using a flexible hedonic specification that isolates idiosyncratic pricing dispersion at listing and at sale, following Amaral (2024) and Jiang and Zhang (2025) adapted to the *xiaoqu* structure described in Section 2.1. The construction implements, at the level of a single home, the standard notion of uncertainty as the dispersion of what observables cannot forecast (Jurado, Ludvigson and Ng, 2015), and it formalizes an old real-estate intuition: when comparables are noisy or scarce, the value of a property is a signal-extraction problem (Quan and Quigley, 1991). Because the residual is computed from the seller's own posted price, Online Appendix OA.D.8 re-estimates the liquidity margins on a choice-free counterpart, the scarcity of comparable transactions by other units before the listing, and finds the same gradients, so the illiquidity results are not an artifact of the seller's own pricing. For each market segment $m \in \{\text{list}, \text{deal}\}$, the specification

$$\ln P_{i,t}^{m} = \sum_{y} \Big[\beta_{1,y}^{m}\,\text{SIZE}_i + \beta_{2,y}^{m}\,\text{SIZE}_i^2 + \omega_{1,y}^{m}\,\text{BUILD\_YEAR}_i + \omega_{2,y}^{m}\,\text{BUILD\_YEAR}_i^2\Big]\mathbb{1}\{\text{YEAR} = y\}$$
$$+ \mu_{c(i)}^{m} + \tau_t^{m} + \mathbf{X}_i'\boldsymbol{\theta}^{m} + u_{i,t}^{m} \tag{1}$$

interacts second-order polynomials in size and construction year with calendar-year dummies, includes discrete attributes (floor category, construction type, facing, layout, building type) and CZ subdistrict dummies, and absorbs community fixed effects $\mu_{c(i)}^{m}$ and year–month fixed effects τ_t^{m}. Uncertainty is the squared residual:

$$\text{Unc}_{i,t}^{m} \equiv \left(\widehat{u}_{i,t}^{m}\right)^2, \qquad m \in \{\text{list}, \text{deal}\}. \tag{2}$$

In the model, $\text{Unc}_{i,t}^{\text{list}}$ proxies for σ_P^2, the idiosyncratic valuation uncertainty that enters the buyer's certainty equivalent (Equation (8)) and drives the uncertainty discount in Proposition OA.3. Jiang and Zhang (2025) show that the raw squared residual is the conservative estimator under classical measurement error in the latent uncertainty proxy. Online Appendix OA.C.3 states the full specification, discusses why omitted unit quality does not threaten the cross-sectional uncertainty coefficient under community, floor, and construction-type fixed effects, and discusses the incidental-parameters bias for thin community cells. Figure 5 plots the aggregate time series: valuation uncertainty declines over the sample, reaching levels in the later years roughly half of those in 2015, indicating that price discovery in the Beijing resale market tightened over the period studied.

Validation. Three diagnostics support interpreting $\text{Unc}_{i,t}^{\text{list}}$ as an idiosyncratic valuation-risk measure rather than a proxy for common shocks or a mechanical artifact of pricing noise. First, prices co-move strongly across communities while uncertainty does not. The median pairwise time-series correlation within a CZ is 0.936 for community-level prices but only 0.136 for community-level uncertainty (0.818 and 0.128 respectively at the coarser district level). This is the divergence one would expect if prices load on a common component while uncertainty is dominated by listing-specific dispersion. Second, in a community–month panel with at least five listings per cell, lagged community-level uncertainty forecasts next-month absolute price movements after controlling for recent volatility and community and year–month fixed effects; Online Appendix OA.C.4 reports the full specification. Third, the

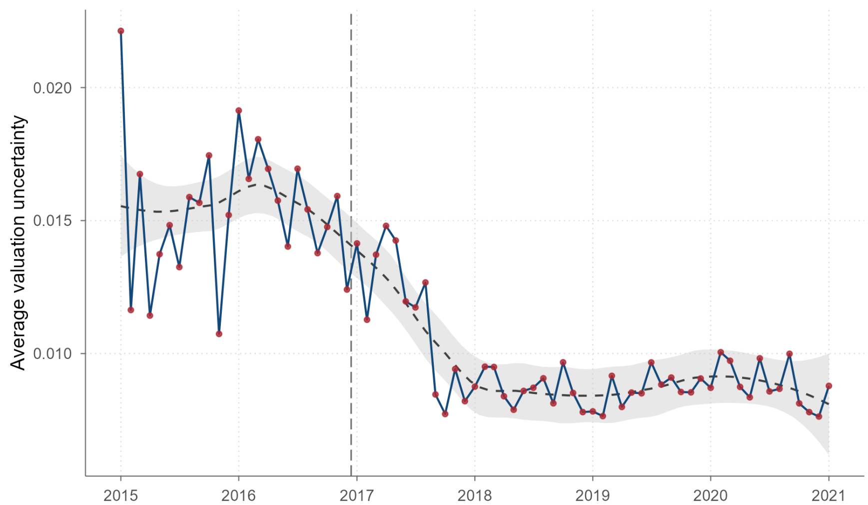


Figure 5: Valuation Uncertainty in Beijing Housing Market Over Time

Notes. Valuation uncertainty is represented by average squared residuals from the hedonic model in Equation (1). The vertical line indicates the timing of the 2016 Central Economic Work Conference.

disagreement is visible in people, not only in residuals. Among attribute-key twins listed concurrently, two owners of the same unit type post asks 5.3 log points apart on average (4.2 among same-month pairs, median 2.7), and the gap widens where the configuration has fewer comparables in its community, a rarity measure built from physical attributes alone ($+0.4$ log points per standard deviation; $t = 2.1$, and $t = 2.4$ controlling for the days between the twin postings). The gap also traces the aggregate decline in Figure 5: it peaks near eleven log points in the early-2016 run-up, when comparables aged fastest, and halves after the 2017 tightening (Figure OA.4). Online Appendix OA.D.9 reports the construction and the checks that separate the gap from market drift and from any mechanical link to the residual. A separate comparison to the ex-post hedonic dispersion measure of Giacoletti (2021) (Online Appendix OA.C.5) shows that the ex-post measure replicates Giacoletti's priced idiosyncratic-risk premium in Beijing while compressing by 8.1% over the sample, versus 34.0% for the forward-looking measure. The two capture complementary quantities. The forward-looking measure is the appropriate variable for capturing the uncertainty that participants perceive at the moment of listing. Online Appendix OA.C.1 reports listing characteristics by uncertainty quartile: high-uncertainty units are modestly larger and take longer to sell, but are similar in district mix and price per square meter, consistent with hard-to-value units rather than systematically lower quality.

2.4.3 A predetermined difficulty index

Because the squared residual is computed from the seller's own posted price, the measurement program also builds a difficulty measure that never sees that price. Using only information dated before the listing (the number of comparable transactions in the community over the trailing twelve months, their recency, their similarity in size, the dispersion of their quality-adjusted prices, and the rarity of the unit's physical configuration within its community, the surprisal of Section 5.4.2), I estimate the predicted absolute out-of-sample error of the hedonic (1) by five-fold cross-fitting with community-blocked folds, and rank the prediction within community and month. Nothing in the construction touches the seller's ask: the hedonic is estimated on training folds only, so a listing's own price never enters its own index, and every ingredient is fixed before the listing date. Two limits are worth stating: the folds are cross-sectional rather than temporal, and the out-of-sample residual uses the held-out community's own within-community mean (its fixed effect is not estimable out of fold), so the object is cross-fitted and predetermined rather than strictly real-time. A strictly forward-looking version, trained on listings posted through 2017 and predicting 2018 onward, reproduces the validation and the execution gradients at about ninety percent of their size (Online Appendix OA.G.5).

The index passes the validation a difficulty measure owes (Table 3). That it predicts the out-of-sample pricing error it was trained on ($t = 23.7$) is first-stage relevance, not independent validation; the external steps are the two dispersions it never saw. Where the index is high, two owners of the same apartment type listing in the same month post asks further apart ($t = 7.9$), and the negotiated list-to-deal wedges of the cell are more dispersed ($t = 9.4$): an object assembled from comparable scarcity alone forecasts the cross-seller and cross-bargain dispersion that disagreement produces. The forecast itself is deliberately weak (a mean out-of-sample R^2 of 0.023 on absolute error): the point is not that each home's latent variance is precisely measured, but that even a weakly predictive, predetermined index carries strong execution relevance. And the index is close to orthogonal to the realized squared residual (correlation 0.064), so the two measures are complements, not restatements: the results use the index for the execution margins and the realized residual for realized-dispersion objects (attention, revisions, and bargaining), a division Section 5.4.2 documents.

2.4.4 Platform attention

The platform records three engagement margins that form a commitment ladder. Total views (VIEWING$_{i,t}$) capture broad, low-cost browsing; followers (FOLLOWER$_{i,t}$) are users who save a listing to track it: higher intent, but still a low-cost click; and agent-accompanied showings (VISITING$_{i,t}$), the number of times an agent physically took a buyer to the unit (a

Table 3: The Predetermined Index Predicts Pricing Error and Dispersion It Never Saw

Validation step	Coef.	(SE)	N
(i) $\lvert$OOS hedonic error$\rvert$ $\sim$ index rank	$+0.0122^{***}$	(0.0005)	468,284
(ii) Twin $\lvert$log ask gap$\rvert$ $\sim$ pair index (z)	$+0.0122^{***}$	(0.0015)	318,949
(iii) Cell SD(wedge) $\sim$ cell index (z)	$+0.0013^{***}$	(0.0001)	47,486
Cross-fit OOS R^2: squared-error model 0.004; absolute-error model 0.023 (5-fold means).			

Notes. The index is the five-fold cross-fitted prediction of the absolute out-of-sample hedonic error from pre-listing features only: log one-plus-comparable-count in the community over the trailing twelve months, mean comparable recency, mean size gap to comparables, the standard deviation of comparables' quality-adjusted deal prices (five or more comparables), indicators for no and few comparables, and the configuration surprisal of Section 5.4.2. Folds are community-blocked and the hedonic is re-estimated on training folds, so no listing's own price enters its own index. Row (i) regresses the absolute out-of-sample hedonic error on the within-cell index rank with community×month fixed effects. Row (ii) regresses the absolute log ask gap between same-community, same-type, same-month listing pairs on the pair's standardized index. Row (iii) regresses the within-cell standard deviation of the list-to-deal wedge on the standardized cell-mean index. Standard errors clustered by community. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

daikan), are the costliest pre-offer engagement, one step short of a bid.[9] Follower counts are recorded for every listing and view counts for a subset; the showing count is recovered from the platform's raw transaction exports and is available for the matched-sold subsample ($250{,}825$ listings), with the merge and coverage detailed in Online Appendix OA.F.12.

I use levels as attention outcomes, together with two composition ratios that track how attention converts from one rung of the ladder to the next. The follower-to-viewing ratio $\text{FVR}_{i,t} \equiv \text{FOLLOWER}_{i,t}/\text{VIEWING}_{i,t}$ measures the conversion of casual browsing into committed tracking; the showings-per-follower ratio $\text{VISITING}_{i,t}/\text{FOLLOWER}_{i,t}$ measures the conversion of committed tracking into costly inspection. In the model the two ratios separate cheap monitoring from the costly visit gate: harder-to-value units draw more committed, high-intent attention before they clear (Proposition OA.6), yet convert a smaller share of it into a physical showing (Proposition OA.8), the funnel split documented in Section 5.2.1 and Online Appendix OA.F.12. Because both ratios have heavy-tailed denominators, I read them alongside the levels rather than on their own, and I trim observations with $\text{FVR}_{i,t} > 1$, which are typically driven by missing or miscoded view counts; Online Appendix OA.C.6 reports robustness to alternative trimming.

[9]The same ladder is visible in China's online judicial housing auctions, where free alert sign-ups measure initial interest and bidder registrations backed by deposits of five to twenty percent of the starting price measure commitment; listing information that barely moves sign-ups still moves registrations and completed sales (Xu, Zheng and Wu, 2024).

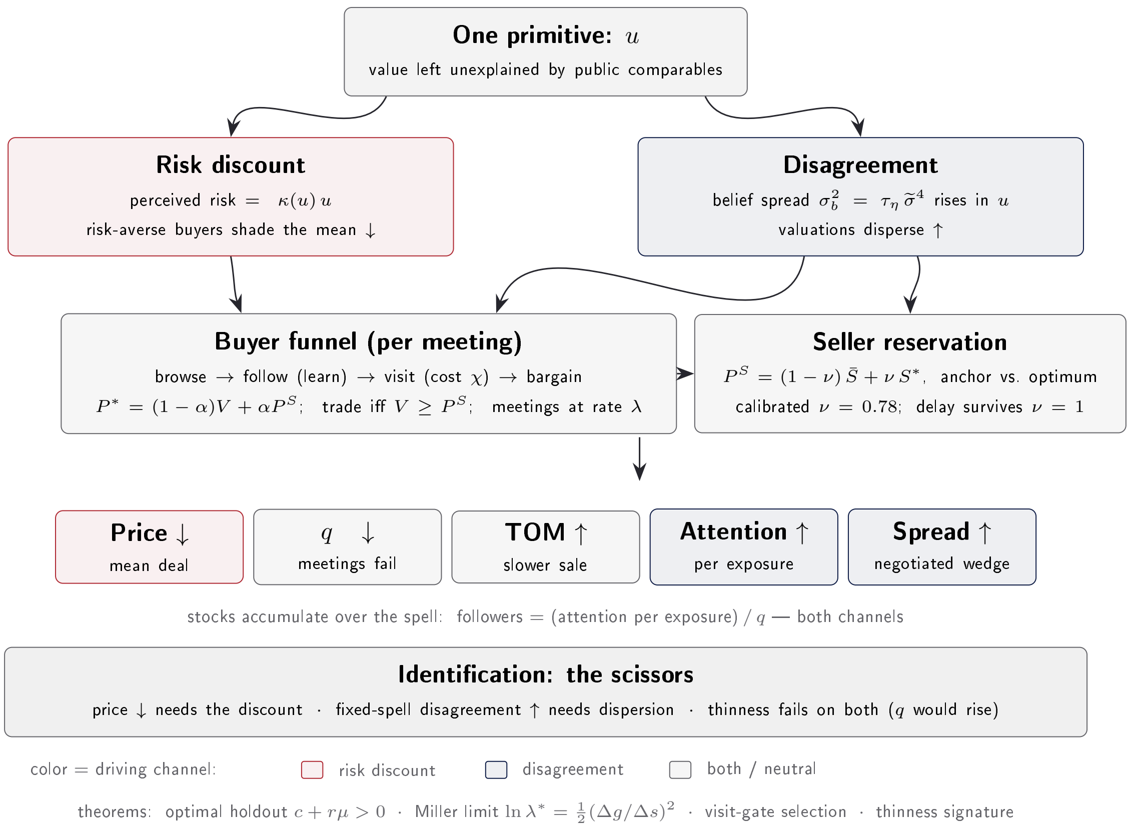
3 Theoretical Framework

The paper's main findings (Section 5) are that harder-to-value homes are less likely to sell, more often vanish from the market than wait on it, and trade at a discount their owners never recoup. This section states, in brief, the framework that explains them; Section 6 develops it in full after the evidence. The core is Nash bargaining between a risk-averse buyer with CARA utility and normally distributed payoffs (Nash, 1950) and a seller with an anchored reservation value; idiosyncratic valuation risk matters because a household holds a single house and cannot diversify across many (Chetty, Sándor and Szeidl, 2017; Cocco, 2005). Risk aversion lowers what a buyer will pay for a hard-to-value unit (the uncertainty discount, as in the static setting of Amaral (2024)), and when the seller's reservation value does not fall in step, the gap between the two widens, so a smaller share of the buyers a listing meets are willing to transact. The unit therefore takes longer to sell, sells with lower probability, and must draw more search to clear. A fuller search-and-matching foundation, with endogenous search intensity, seller listing strategy, and steady-state equilibrium, is developed in Online Appendix OA.A. This section states the two channels and the predictions they generate; Section 6, placed after the evidence, develops the pricing, acceptance, and attention logic in full, and its Section 6.2 grounds it: both channels follow from one information primitive, and the delay margin survives a fully optimal seller. The model is deliberately partial-equilibrium. It holds the local market fixed and is silent on construction, entry, and the level of the aggregate cycle: the margins that quantitative general-equilibrium treatments of housing search are built to explain (Head, Lloyd-Ellis and Sun, 2014; Ngai and Tenreyro, 2014). What it does carry to the cycle is composition: which trades survive a downturn (Section 5.2.7). That narrowing is what the design buys: those models would absorb the cross-listing variation studied here into a market fixed effect, whereas the within-community–month comparison isolates it. The two are complements: an aggregate question and a unit-level one.

Two channels run from valuation uncertainty to illiquidity, and the data can tell them apart. A *risk-discount* channel lowers the mean of buyers' willingness to pay for a hard-to-value unit, so fewer meetings clear and the unit sells for less. A *valuation-dispersion* channel widens the spread of buyers' valuations, so negotiated outcomes disperse, buyers evaluate more before committing, and a seller whose reservation rises with the option value that dispersion creates holds out longer. The two leave different fingerprints: the discount on the mean transaction price, dispersion on the spread of negotiated outcomes and on the committed evaluation a meeting draws. Proposition 1 shows that the pattern in the data, a falling price alongside rising disagreement, requires both.

Why a model at all—and why not an existing one? Because no off-the-shelf framework can produce the joint pattern the data show. Housing-search models take buyers to be risk-neutral, so valuation uncertainty cannot shade the mean of their bids; frictionless disagreement models predict the wrong sign, a price premium set by the optimist; and a pure information-asymmetry model moves neither the negotiated dispersion nor per-exposure attention at a fixed spell. The model exists to supply the two missing ingredients, risk-averse buyers and unit-level disagreement inside a matching market, and to be falsifiable by the funnel facts it predicts. Figure 6 maps the architecture.

Figure 6: The Model in One Picture: One Primitive, Two Channels, Signed Outcomes



Notes. One information primitive—the value a unit's public comparables leave unexplained, u—generates both channels (Proposition 3): a risk discount that shades the mean of buyer valuations, and divergence of opinion that widens their spread. The two feed a search market in which a buyer funnel (follow, visit, bargain) meets a partially anchored seller reservation; outcomes inherit their signs from the channels (red: discount; navy: disagreement; grey: both). Identification uses the joint pattern: a falling mean price requires the discount, because pure dispersion raises the transacted mean through selection on the accepting tail; rising fixed-spell disagreement requires dispersion; and segment thinness fails on both margins—it would raise the per-meeting acceptance rate (Propositions 1 and OA.1).

3.1 Testable predictions

The framework yields six predictions that map to the empirical sections; formal statements and proofs are collected in Online Appendix OA.B. Within a community and month:

(i) harder-to-value units are less liquid: lower probability of sale, longer time on market, and more search to clear (Equation (13); Proposition OA.6); and because exit competes with a sale that arrives ever more slowly, the shortfall realizes as withdrawal, with no faster pull-back conditional on not selling (Corollary 1)—the central prediction;

(ii) higher uncertainty lowers the mean transaction price, the uncertainty discount (Equation (11); Proposition OA.3);

(iii) higher uncertainty widens the dispersion of negotiated outcomes (Equation (12); Proposition OA.7);

(iv) committed buyer attention rises in uncertainty: the accumulated stock through both channels, the per-exposure component through dispersion alone (Equations (16)–(17); Proposition OA.6);

(v) the two-channel test: a falling mean price identifies the risk discount, because pure dispersion, through selection on the accepting tail, weakly raises the transacted mean; rising spell-adjusted disagreement (the negotiated spread, per-exposure attention) identifies divergence of opinion, because a pure discount moves neither once the spell is held fixed; the joint pattern in the data therefore requires both channels (Proposition 1);

(vi) the clearing function flattens: the slope of the sale probability in the seller's relative asking price shrinks toward zero where comparables are noisy, so a price cut buys almost no additional probability of sale (Proposition 2), the central margin the paper's title names.

The listing-implied pricing wedge and the listing–transaction wedge load on uncertainty as in Propositions OA.4 and OA.5.

Proposition 1 (Two channels: the scissors)**.** *Suppose the transacted mean price falls in uncertainty while the spell-adjusted disagreement margins, per-exposure committed attention (Equation (16)) and the dispersion of negotiated outcomes (Proposition OA.7), rise. Then both channels are active. With only the risk-discount channel ($\sigma_b^{2\prime} = 0$, so Var_V does not respond), neither spell-adjusted margin rises in σ_P^2; per-exposure attention is flat and the conditional dispersion weakly falls through selection: a pure discount lengthens the spell and accumulates stocks (Equation (17)), but raises neither margin at a fixed spell. With only the dispersion channel ($\bar{\gamma}\kappa = 0$, so μ_V is constant), the transacted mean is strictly increasing in σ_P^2 by Lemma OA.2(ii), selection on the accepting tail, so pure dispersion cannot lower the price. The jointly observed pattern therefore requires $\bar{\gamma}\kappa > 0$ and $\sigma_b^{2\prime} > 0$. Conversely, when both channels operate, the spell-adjusted disagreement margins rise, and the transacted mean falls provided the discount outweighs the selection effect,*

$$\frac{\rho^2\bar{\gamma}\kappa}{2}\big(1-\delta(z)\big) \;>\; \sigma_V'\big(M(z) - z\,\delta(z)\big)$$

(Equation (OA.8); the calibrated parameters satisfy it), so the scissors is observed. The identi-

fying direction (pattern observed implies both channels active) holds without the dominance condition.

Proposition 2 (The clearing function flattens: price stops clearing)**.** *A meeting clears when a buyer's valuation exceeds the seller's reservation at the posted relative price P^S, so the per-meeting acceptance rate of Equation* (13)*, read as a function of that price, traces the seller's clearing function $q(P^S) = 1 - \Phi\big((P^S - \mu_V)/\sigma_V\big)$, with μ_V the mean and $\sigma_V = \rho^2\sigma_b$ the standard deviation of buyer valuations. Its slope is $\partial q/\partial P^S = -\phi(z)/\sigma_V$, $z \equiv (P^S - \mu_V)/\sigma_V$, so the sensitivity of the sale to price near comparables ($z \approx 0$) is $\phi(0)/\sigma_V$ and is strictly decreasing in σ_V. Because disagreement makes the dispersion of valuations rise with uncertainty ($\sigma_b^2 = \tau_\eta\widetilde{\sigma}^4$, Equation* (14)*), the clearing function flattens as a home grows harder to value: where comparables are noisy, price barely moves the probability of sale, and a seller cannot discount the unit into a trade.*

The slope is one derivative of q; the comparative static $\partial[\phi(0)/\sigma_V]/\partial\sigma_V = -\phi(0)/\sigma_V^2 < 0$ is immediate. This is the structural content of the paper's title: where buyers disagree, the price mechanism stops clearing the market. It governs the slope of demand, not its level: the discount the friction imposes is real but, by Proposition 4(c), passes almost fully into price rather than quantity, so the curve flattens without shifting bodily.

Remark (curvature). Differentiating the clearing function a second time gives $\partial^2 q/\partial(P^S)^2 = z\,\phi(z)/\sigma_V^2$, so q is an S-curve in the price level: concave where the seller prices below comparables ($z < 0$), convex where she prices above ($z > 0$), with an inflection at the mean valuation μ_V. The slope is steepest near comparables and its magnitude is symmetric in $|z|$: a single Gaussian survival curve, not the globally concave relative-price curve of Guren (2018), whose asymmetry (overpricing punished sharply, underpricing rewarded little) is generated by a signal-noise distribution this model does not contain. What this paper takes from Guren is the empirical object, not the curve's shape; what it adds is how the slope near comparables, $\phi(0)/\sigma_V$, falls with valuation dispersion.

Table 4 collects the predictions and the estimates that test them, each borne out in sign and significance.

Table 4: **The Model's Predictions and Their Empirical Counterparts**

Prediction	Model	Estimate	Source
Sale probability falls	–	-0.0081^{***}	Table 6
Time on market rises	$+$	$+0.0534^{***}$	Table 6
Lost sale realizes as withdrawal, not delay	$+$	$+0.0124^{***}$	Table 7
Mean transaction price falls	–	-0.0074^{***}	Table 10
Negotiation spread widens	$+$	$+0.0013^{***}$	Table 10
Follower stock accumulated by a listing rises	$+$	$+0.0428^{***}$	Table 10
Costly inspections (showings) fall	–	-0.0204^{***}	Table 10
Clearing function flattens (price lever weakens)	–	-0.18^{***}	Table 13
Price falls and attention rises	both	confirmed (§5.2.1)	

Notes. Each row pairs a prediction of the framework (Section 3) with the estimate that tests it, measured by the within-community–month rank of valuation uncertainty on the mature 2016–2019 cohorts; standard errors and full specifications are in the cited tables. The "Model" column gives the predicted sign. The follower stock rises through both channels (Equation (17)); its spell-adjusted component, the dispersion fingerprint, is reported in Online Appendix OA.F.11. The showings sign is the model's race-signing margin rather than an unconditional implication: a pure discount would push it positive (worse conversion per visit, more showings before a sale), so the negative estimate both confirms the funnel split and identifies disagreement as dominant at the conversion stage (Proposition OA.8). The withdrawal row is Corollary 1: sale and exit are competing risks, so the sale shortfall accrues to withdrawal while the conditional exit timing does not accelerate. The final row is the two-channel test of Proposition 1: a falling mean price beside rising spell-adjusted disagreement, a pattern neither channel produces alone. The clearing-function row reports the discount-slope$\times$uncertainty interaction, -0.18 ($t = -3.9$); equivalently, the discount-side slope of the thirteen-week sale probability falls from $+1.45$ in transparent to $+0.17$ in opaque submarkets (Table 13), the flattening Proposition 2 predicts.

4 Identification Strategy

The design is reduced-form and cross-sectional. Every specification compares listings observed in the same community and the same month and asks how outcomes (whether and when a unit sells, and the search it attracts) vary with how hard the unit is to value. Identification comes entirely from within community–month variation: community and year–month fixed effects (and, in the saturated specification, community$\times$month fixed effects) absorb every force common to a local market at a point in time, so two listings enter the comparison only through the dispersion of their hedonic pricing residuals. The baseline design needs no policy break and no parallel-trends assumption.

4.1 Estimating equation

For listing i in community $c(i)$ and month t, I estimate

$$y_{i,t} = \beta\, r_{i,t}^{\text{unc}} + \mu_{c(i)} + \tau_t + \Xi_i + \varepsilon_{i,t}, \tag{3}$$

where $y_{i,t}$ is an illiquidity or attention outcome (an indicator for sale, log days on market, or log followers) and $r_{i,t}^{\text{unc}} \in [0, 1]$ is the within-community–month percentile rank of idiosyncratic valuation uncertainty $\hat{u}^2_{i,t}$, the squared hedonic residual defined in Section 2.4.2. The terms $\mu_{c(i)}$ and τ_t are community and year–month fixed effects and Ξ_i collects property controls entered as fixed effects (floor, construction type). Standard errors are clustered at the community level (5,153 communities in the baseline estimation sample). The coefficient β is the object of interest: the difference in the outcome between the hardest- and easiest-to-value unit in the same community and month.

Two measurement choices make the estimate robust. First, I rank uncertainty within each community–month cell rather than using its level. The squared residual has a heavy upper tail (Section 2.4.2); the rank is invariant to it, so no estimate is driven by a handful of extreme listings. Because uncertainty enters as this within-cell rank, the additive community and year–month fixed effects identify a genuinely within-community–month comparison (two listings enter only through their rank inside the same *xiaoqu*–month cell), and replacing the additive effects with interacted community$\times$month fixed effects leaves every headline estimate essentially unchanged: the sale-probability coefficient is -0.0060 vs. -0.0061, days on market $+0.0512$ vs. $+0.0511$, followers $+0.0393$ vs. $+0.0395$, and the wedge, revision, and withdrawal gradients likewise move only in the fourth decimal (Online Appendix OA.G.4, Table OA.38). Second, every result is re-estimated after dropping the top one and five percent of the uncertainty distribution and re-ranking within cell; the coefficients are essentially unchanged (Section 5). Under this trimming, a return-premium reading of the same measure does not survive (Section 5.1.4).

4.2 Identifying assumption

The design is a within-market measurement, not a quasi-experiment: no shock to a single home's valuation uncertainty is available, so identification comes from holding the local market fixed and letting only how hard a unit is to value vary. The aim is to read the liquidity margins a price regression cannot see (whether and when a home sells, and the search it takes) cleanly, not to mimic an experiment the setting does not provide. The identifying assumption is that, within a community and month and conditional on the property controls, a listing's rank in the valuation-uncertainty distribution is unrelated to unobserved determinants of its liquidity. Two units in the same *xiaoqu* listed in the same month face the same local demand, the same schools and transit, the same credit and macro conditions, and the same policy environment; the design asks whether the one whose value is harder to pin down from comparable sales is slower and less likely to sell.

Four features support the assumption. First, uncertainty is built from within-community hedonic residuals, so any community-level or market-level force (a neighborhood amenity, a citywide credit shock) is absorbed by the fixed effects and cannot drive the within-cell ranking. Second, the rank measure is distribution-free, so the result does not depend on the shape or scale of the residual variance. Third, the comparison holds the transacting market fixed, so it is not contaminated by the spatial and cyclical sorting that drives liquidity differences across markets. That sorting is documented, not hypothesized: within Beijing, constant-quality price levels differ several-fold between central and satellite submarkets, local trends and their turning points are unsynchronized, and local price shocks diffuse at a scale of one to two kilometers per month (Zhu et al., 2022), so pooling communities would compare submarkets at different phases of their own cycles. Fourth, the measure is not a relabeling of seller mispricing or omitted quality: the within-cell signature survives replacing the price residual with a price-free measure built from physical attributes alone (a within-community configuration surprisal) and with a transaction-only comparable-scarcity count, neither of which uses the seller's posted price; a cross-fitted index that aggregates those pre-listing features strengthens the execution gradients outright (Section 2.4.3, Table 5). What the assumption cannot exclude by construction, the falsification battery of Section 5.4 confronts directly (Table 18).

5 Results

This section presents the paper's empirical results, each reported holding the local market fixed: every specification includes community and year–month fixed effects, so identification compares units listed in the same complex in the same month that differ only in how hard they are to value. Four facts emerge: harder-to-value units are less liquid, slower and less certain to sell; the lost sale is a platform withdrawal rather than a longer wait; they trade at a discount, in listing and transaction prices; and that discount is not compensated by higher returns to the owner. Because the squared-residual proxy is heavy-tailed, every fact is measured by the within-cell rank of uncertainty and shown to survive trimming the extreme upper tail. Section 5.2 then traces these facts to the model's two channels.

5.1 Main findings

Four facts about how valuation difficulty surfaces at the point of sale follow in turn.

5.1.1 Harder-to-value homes are illiquid

Table 5 reports the paper's headline estimates on the largest sample each outcome admits, with both difficulty measures ranked within community and month and the ranks frozen on the full listing universe before any outcome-specific screening. Harder-to-value units are illiquid on every margin: moving from the least to the most difficult unit in a community and month lowers the probability of sale by 1.3 percentage points on the realized measure and 2.4 on the predetermined index; the censoring-proof outcome (sold within 365 days of posting, defined identically for every cohort) puts the same moves at 2.1 and 3.1 points; withdrawal rises by 1.2 and 2.4 points; and time on market among sold units lengthens by six and nine log points. Sellers feel the friction on the demand side as well: on the attention-valid sample, harder-to-value listings must accumulate significantly more followers (saved-listing attention) to clear, a margin that belongs to the realized measure alone (Section 5.4.2). Figure 7 shows the realized-rank gradients in the raw (residualized) data under the earlier attention-screened convention, and Table 6 retains that convention with its two threat checks: Column (2) drops the top one percent of the uncertainty distribution and re-ranks within cell, guarding against the heavy tail; Column (3) restricts to mature 2016–2019 posting cohorts that had time to transact before the data snapshot, guarding against right-censoring; every coefficient is essentially unchanged or slightly larger. The screen behind that convention, which requires recorded views and is needed only for attention outcomes, roughly halves the sale gradient (Table 5, robustness rows), so the earlier convention is a conservative floor rather than the paper's estimate of the friction. The within-cell rank is also invariant to the scale of first-stage estimation error, which caps the scope for generated-regressor bias, and Appendix B confirms the cap directly: a community pairs-cluster bootstrap that rebuilds the hedonic, the squared residual, and the within-cell rank on every one of three hundred draws leaves each headline standard error within eleven percent of its analytic value, across the liquidity outcomes, the price-side slopes, and the demand-curve lever (Table 22); a nested bootstrap that rebuilds the predetermined index inside every draw shows the same for the index gradients (Online Appendix OA.G.5).

Reading the magnitudes. The sale-probability gradient appears in this paper at several sizes, and the differences are samples and outcome definitions, not instability. On the full listing universe, with ranks frozen before any screening, the move from easiest to hardest prices at -1.31 percentage points (Table 5), the same sample and figure as the three-fate decomposition below (Table 7); the censoring-proof outcome (sold within 365 days of posting, defined identically for every cohort) puts it at -2.08; the attention-screened convention, which requires recorded views and is retained for the attention outcomes, roughly halves it (-0.55; Table 6), so the earlier convention is a conservative floor; the predetermined index roughly doubles each figure (-2.41 on the full universe); and the comparable-scarcity measure, which contains no choice of the seller, spans 2.3 points from fewest to most comparables (Online Appendix OA.D.8). The larger estimates arrive exactly where measurement is cleaner: unscreened samples, censoring-proof outcomes, and measures with no seller choice in them.

5.1.2 The lost sale is abandonment, not delay

The sale-probability gradient of Section 5.1.1 understates what it measures, because "less likely to sell" can mean two different things (a home that sells later, or a home that never

Table 5: Execution Margins on the Full Listing Universe: Realized and Predetermined Difficulty

Panel A. Realized valuation uncertainty (frozen within-cell rank of $\hat{u}^2$)

	Additive FE		Comm.×month FE		Horse race (int.)		
Outcome, sample	Coef.	(SE)	Coef.	(SE)	Coef.	(SE)	N
P(sale), all listings	-0.0131^{***}	(0.0016)	-0.0130^{***}	(0.0016)	-0.0116^{***}	(0.0016)	468,698
P(sale $\leq$ 365d), mature cohorts	-0.0208^{***}	(0.0018)	-0.0207^{***}	(0.0018)	-0.0190^{***}	(0.0018)	417,399
Withdrawal, all listings	$+0.0124^{***}$	(0.0016)	$+0.0123^{***}$	(0.0016)	$+0.0110^{***}$	(0.0016)	468,698
log DOM, sold, real durations	$+0.0581^{***}$	(0.0064)	$+0.0592^{***}$	(0.0068)	$+0.0539^{***}$	(0.0068)	263,783
log followers, FVR$\leq$1	$+0.0498^{***}$	(0.0063)	$+0.0391^{***}$	(0.0067)	$+0.0402^{***}$	(0.0067)	285,235
List–deal wedge, sold	$+0.0011^{***}$	(0.0001)	$+0.0012^{***}$	(0.0001)	$+0.0010^{***}$	(0.0001)	331,965
Revisions $\log(1+\#)$	$+0.0076^{***}$	(0.0021)	$+0.0075^{***}$	(0.0021)	$+0.0064^{***}$	(0.0021)	346,574
P(sale), FVR$\leq$1 [robustness]	-0.0055^{***}	(0.0020)	-0.0074^{***}	(0.0021)	-0.0058^{***}	(0.0021)	285,235
log DOM, FVR$\leq$1 [robustness]	$+0.0769^{***}$	(0.0068)	$+0.0592^{***}$	(0.0075)	$+0.0535^{***}$	(0.0075)	216,327

Panel B. Predetermined difficulty index (frozen within-cell rank)

	Additive FE		Comm.×month FE		Horse race (int.)		
Outcome, sample	Coef.	(SE)	Coef.	(SE)	Coef.	(SE)	N
P(sale), all listings	-0.0241^{***}	(0.0018)	-0.0250^{***}	(0.0018)	-0.0242^{***}	(0.0018)	468,284
P(sale $\leq$ 365d), mature cohorts	-0.0309^{***}	(0.0019)	-0.0322^{***}	(0.0020)	-0.0309^{***}	(0.0020)	417,017
Withdrawal, all listings	$+0.0236^{***}$	(0.0017)	$+0.0243^{***}$	(0.0018)	$+0.0235^{***}$	(0.0018)	468,284
log DOM, sold, real durations	$+0.0896^{***}$	(0.0069)	$+0.0989^{***}$	(0.0074)	$+0.0952^{***}$	(0.0074)	263,572
log followers, FVR$\leq$1	-0.0075	(0.0066)	-0.0103	(0.0071)	-0.0132^{*}	(0.0072)	284,988
List–deal wedge, sold	$+0.0022^{***}$	(0.0001)	$+0.0025^{***}$	(0.0002)	$+0.0024^{***}$	(0.0002)	331,676
Revisions $\log(1+\#)$	$+0.0173^{***}$	(0.0022)	$+0.0179^{***}$	(0.0022)	$+0.0175^{***}$	(0.0022)	346,301
P(sale), FVR$\leq$1 [robustness]	-0.0246^{***}	(0.0022)	-0.0268^{***}	(0.0023)	-0.0263^{***}	(0.0023)	284,988
log DOM, FVR$\leq$1 [robustness]	$+0.0931^{***}$	(0.0071)	$+0.1014^{***}$	(0.0079)	$+0.0977^{***}$	(0.0079)	216,165

Notes. Each row regresses the outcome on the within community×month rank of the panel's measure, computed once on the full eligible listing universe (no outcome-specific screening) and frozen before merging into each outcome's maximal sample. Additive fixed effects are community, year–month, floor, and construction type; the interacted columns replace the additive pair with community×month effects; the horse-race columns enter both frozen ranks jointly under interacted effects and report the panel measure's coefficient. Samples: all listings for the sale and withdrawal rows (withdrawal under the three-fate conventions of Table 7); cohorts with at least 365 days of follow-up for the censoring-proof outcome; sold units with recorded post–July-2016 posting dates for days on market; the attention-valid sample (FVR ≤ 1) for followers; sold units for the wedge. The final two rows reproduce the earlier attention-screened convention as robustness. Standard errors clustered by community. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

sells at all), and they carry different mechanisms. A friction that lengthens the wait is not a friction that ends it. The platform's listing histories separate the two. Each listing ends in exactly one of three states that partition unity: it sells (a deal date is recorded); it is withdrawn (no deal, and the listing's last recorded day (posting date plus the observed listing window) falls at least sixty days before the data end, 9 January 2021, so the post went dark while the panel was still running); or it is right-censored (no deal, but still fresh within sixty days of the data end). Because the three indicators sum to one, their coefficients on the within-cell uncertainty rank sum to zero: the drop in the probability of sale is split, as an accounting identity, between the homes that give up and the homes that are merely still waiting. Across all listings, about 71 percent sell, 29 percent are withdrawn, and under

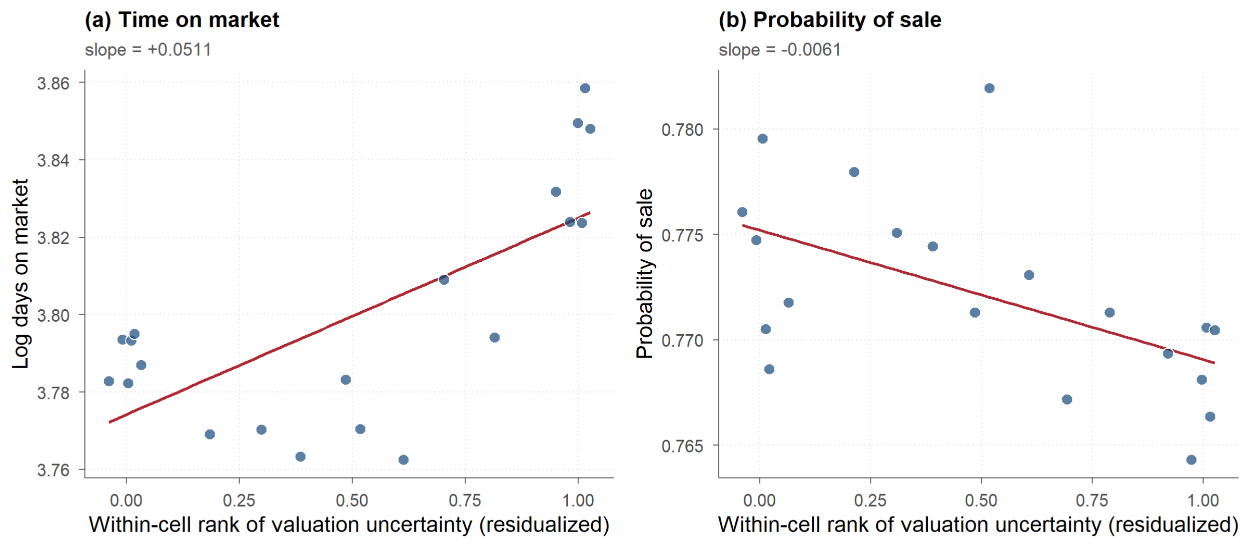


Figure 7: Harder-to-Value Homes Are Less Liquid: The Headline Gradients in the Raw Data

Notes. Binned scatter plots of the two headline illiquidity margins against the within-community–month rank of valuation uncertainty. Panel (a): log days on market (sold listings). Panel (b): probability of sale (all listings). Both axes are residualized on community, year–month, floor, and construction-type fixed effects, with sample means added back; each panel groups listings into twenty equal-count bins of the residualized rank, and the rank is computed within cell on each panel's estimation sample. The fitted lines have slopes $+0.0511$ and -0.0061, the attention-screened convention of Table 6, column (1); the full-universe headline estimates are in Table 5. Sample: Beijing listings, 2015–2021, follower-to-viewing ratio at most one. Durations for cohorts posted before July 2016 are backfilled (Figure 4); Table 6, column (3) shows the slopes on the clean window.

1 percent are censored, so the sixty-day rule sorts a withdrawal state that is already the second-largest fate, not a knife-edge sliver.

Table 7 reports the split. Moving from the easiest to the hardest unit to value in a community and month lowers the probability of sale by 1.31 percentage points (s.e. 0.16); of that, 1.24 points are additional withdrawals ($t = 7.8$) and only 0.07 points are slower sales still on the shelf ($t = 3.0$). Ninety-five percent of the observed sale shortfall appears in the withdrawal state rather than the still-waiting one: abandonment, not delay. The share is the same after the top one percent of $\hat{u}^2$ is trimmed and ranks rebuilt (95 percent) and under interacted community$\times$month fixed effects (95 percent); it rises, if anything, when the 2020–21 window is dropped (97 percent), so the split is not an artifact of the truncated COVID tail; and moving the withdrawal cutoff from sixty days to thirty or ninety holds it between ninety-three and one hundred percent. Two checks meet the censoring concern, that the withdrawn are merely late listings that ran out of observation window, head on. Restricting to the post–July 2016 window, where listing durations are recorded rather than backfilled, leaves the share at 93 percent ($N = 336{,}271$); and a competing-risks hazard that treats still-fresh listings as censored confirms it under the textbook correction. Uncertainty lowers both cause-specific hazards, but depresses the sale hazard about a third harder ($\mathrm{HR} = 0.91$

Table 6: Valuation Uncertainty and Illiquidity

	(1) Baseline		(2) Top-1% trimmed		(3) Mature 2016–2019	
Outcome	Coef.	(SE)	Coef.	(SE)	Coef.	(SE)
Probability of sale	-0.0061^{***}	(0.0020)	-0.0066^{***}	(0.0020)	-0.0081^{***}	(0.0023)
Days on market (log)	$+0.0511^{***}$	(0.0067)	$+0.0524^{***}$	(0.0067)	$+0.0534^{***}$	(0.0074)
Followers (log)	$+0.0395^{***}$	(0.0063)	$+0.0418^{***}$	(0.0063)	$+0.0428^{***}$	(0.0071)
Observations	267,911		265,015		217,967	

Notes. Each row is a separate regression of the row outcome on the within-community–month rank of idiosyncratic valuation uncertainty $r_{unc} \in [0, 1]$ (the squared hedonic residual $\hat{u}^2$, ranked within cell so the extreme upper tail cannot drive the estimates), with community, year–month, floor, and construction-type fixed effects and standard errors clustered by community. The outcomes are the sale probability, log days on market (sold units), and log followers. Column (1) is the baseline; Column (2) drops the top one percent of $\hat{u}^2$ and re-ranks within cell; Column (3) restricts to mature cohorts posted July 2016–December 2019—the window with real posting dates (durations before July 2016 are backfilled; Figure 4) and the least exposure to right-censoring—and re-ranks within cell; Columns (1)–(2) include the earlier cohorts, whose backfilled zero durations slightly attenuate the days-on-market row, so Column (3) shows the gradients without them. The reported observation count is for the sale and followers regressions; the days-on-market row conditions on sale and so uses fewer observations (193,815, 191,712, and 159,888 across the three columns). $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

against 0.93), so the cumulative incidence of withdrawal rises with uncertainty while the cumulative chance of sale falls. This is not mechanical truncation. Withdrawn listings sit on the platform longer than sold ones (a mean of 93 against 71 days), so they are homes that tried and gave up, not posts that expired before they had a chance to sell. The intensive margin is real but small by comparison: among homes that do sell, the hardest-to-value wait about five percent longer (Table 6). The friction is overwhelmingly extensive. This is the reduced-form counterpart of the withdrawal rate the structural extension estimates ($\delta_w = 0.88$ per year, Online Appendix OA.A.9): disagreement does not make the hard-to-value home wait its turn; the home ends up off the market because the sale never arrives, not because its owner pulls it sooner: conditional on not selling, harder-to-value listings if anything delist later ($t = 14.7$). Figure 8 shows this across posting cohorts: the withdrawn share is a stable band, flat to declining toward the data end rather than fanning out, while the still-fresh censored listings appear only as a thin sliver in the final months.

This extensive margin is not the loss-aversion holdout of Genesove and Mayer (2001). That channel works through the seller's gain–loss position relative to her purchase price, an axis orthogonal to the within-cell valuation difficulty that the community–month comparison differences out, and Proposition 4 shows the within-cell gradient survives a fully optimal, un-anchored seller, so the abandonment here does not rest on it. The two compound only where they coincide: a hard-to-value home held by an underwater seller in a cold market, where loss aversion deepens the withdrawal the friction already produces (Online Appendix OA.A.8).

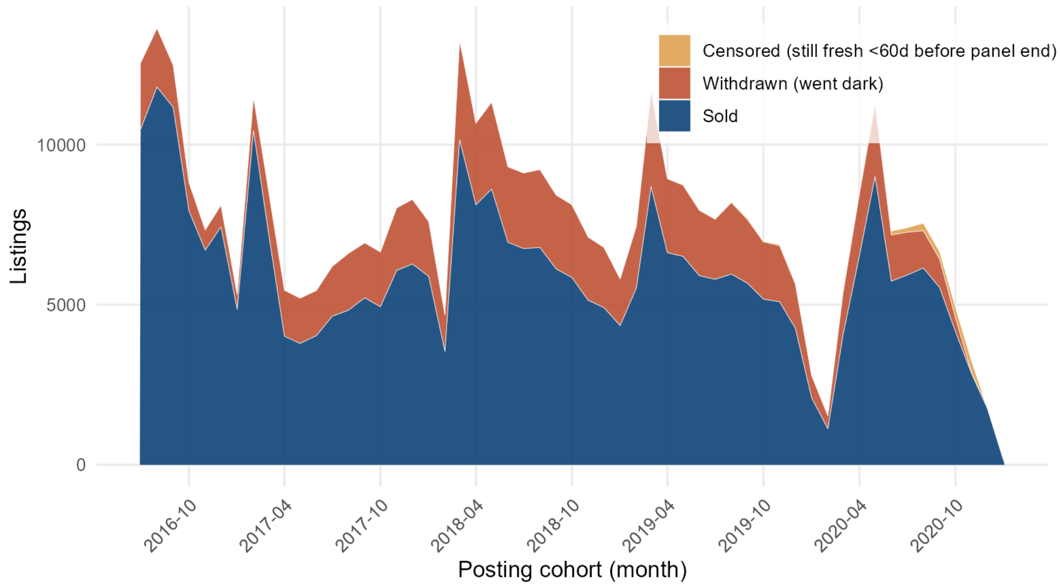


Figure 8: Abandonment Is Not Right-Censoring: Listing Fates by Posting Cohort

Notes. Each posting-month cohort from July 2016 (durations recorded, not backfilled) through the data end, by terminal fate: sold (a deal date is recorded), withdrawn (no deal, and the listing went dark at least sixty days before the data end of 9 January 2021), and censored (no deal, still fresh within sixty days of the end). The withdrawn band is a roughly stable share across cohorts and does not widen toward the panel end—if anything it narrows through 2020—while censored listings form a thin sliver confined to the final months, so the never-sell outcome is not an artifact of recent listings lacking time to realize.

Nor is a withdrawal a pause before a quiet relisting. The platform assigns no persistent property identifier, but the exact attribute key of Online Appendix OA.D.10 (community, size, floor band, facing, layout, build year, and building height) identifies the same unit or its identical twin, and so bounds relisting from above. Under that key, at most 14 percent of withdrawn listings ever reappear, and crediting every later same-key sale to the withdrawn home, at least 89 percent never transact on the platform at all. Counting a sale in any subsequent spell as a sale sharpens the extensive margin rather than softening it (from -0.012 for the single listing to $+0.013$, $t = 8.0$, for never selling in any spell). The minority that does return is not a repricing story either: returning sellers post prices no closer to comparables than the mechanical mean-reversion benchmark of homes that sold and later relisted ($t = -0.7$), and neither the change in the uncertainty rank nor the move toward comparables predicts whether the second spell sells. Conditional on the rank, a post-withdrawal relisting is no less likely to sell than a fresh listing ($+0.006$, $t = 1.5$). The barrier is the home's difficulty, not its listing history, the same verdict the price-free measure of Section 5.4.2 delivers from the other direction.

Two magnitudes give the never-sell margin its market-level weight. The 87,414 withdrawn listings of the clean window carry a combined asking value of about 517 billion yuan (a mean of 5.9 million per home). And because a home that will never sell stays on the shelf

Table 7: The Lost Sale Is Abandonment, Not Delay

Fate of the listing	(1) Baseline Coef.	(SE)	(2) Top-1% trimmed Coef.	(SE)	(3) Ex-COVID 2015–19 Coef.	(SE)
Probability of sale	-0.01307^{***}	(0.00161)	-0.01424^{***}	(0.00160)	-0.01482^{***}	(0.00173)
$\hookrightarrow$ withdrawal (no later observed sale)	$+0.01238^{***}$	(0.00160)	$+0.01354^{***}$	(0.00159)	$+0.01431^{***}$	(0.00172)
$\hookrightarrow$ slow / late sale (censored)	$+0.00069^{***}$	(0.00023)	$+0.00070^{***}$	(0.00023)	$+0.00051^{***}$	(0.00013)
Withdrawal share of lost sale	95%		95%		97%	
Observations	468,698		463,452		416,482	

Notes. Each listing ends in exactly one of three mutually exclusive states that sum to one: sold (a deal date is recorded); withdrawn (no deal, and the listing's last recorded day (posting date plus the observed listing window) falls at least sixty days before the data end, 9 January 2021); and slow / late sale (censored) (no deal, but still within sixty days of the data end). Each row regresses the state indicator on the within-community–month rank of valuation uncertainty $r_{unc} \in [0, 1]$, with community, year–month, floor, and construction-type fixed effects and standard errors clustered by community. No follower-to-viewing screen is applied: the screen conditions on the view count, whereas the extensive margin is defined for every listing. Because the three indicators sum to one, the three coefficients in each column sum to zero; the row "Withdrawal share of lost sale" reports the withdrawal coefficient as a fraction of the (negative) sale-probability coefficient. Column (1) is the baseline; Column (2) drops the top one percent of $\hat{u}^2$ and re-ranks within cell; Column (3) excludes 2020–21. ${}^{*}p < 0.10$; ${}^{**}p < 0.05$; ${}^{***}p < 0.01$.

while the ones that clear leave it, the never-sellers accumulate in the stock a buyer actually faces: among listings active on an average day of 2017–2019, about 37 percent will end withdrawn: the browsing shelf overweights precisely the homes the market cannot clear.

How far can withdrawal be pushed toward a later trade? Table 8 reports progressively more generous creditings of subsequent platform activity for the 87,414 withdrawn listings of the clean window. Under the exact attribute key, 13.7 percent reappear and 10.9 percent later sell, the source of the ≥ 89 percent never-transacts-again bound. The platform records no unit identifier, so the exact key is the tightest available match, and every coarser key mechanically credits identical-twin activity in large communities: a tolerant key "finds" subsequent activity for two-thirds of withdrawn listings because some similar unit eventually trades, not because the withdrawn home returned. Even that most promiscuous crediting leaves a third of withdrawn homes with no subsequent same-segment sale at all. The economic claim the paper needs, that the platform never observes a later transaction for the great majority of withdrawn homes, sits between those bounds and closest to the exact-key row; what happens off the platform the data cannot say.

5.1.3 Uncertainty and price levels

Harder-to-value units also trade at a discount. Regressing log deal price per m^2 on valuation uncertainty with community, year–month, floor, and construction-type fixed effects, the saturated coefficient is -0.37: roughly 6 percent lower price per square meter for a one-

Table 8: Relisting and Later-Sale Bounds for Withdrawn Listings

Later platform event for a withdrawn listing	Share	Key
(1) Same listing identifier relisted	not computable — export carries no unit ID	
(2a) Exact attribute key relisted (paper key)	13.7%	community|size|floor|facing|layout|yr|height
(2b) Coarse key relisted	42.2%	community × layout × floor band × size $\pm 1\text{m}^2$
(3) Tolerant key relisted	65.5%	community × layout × adjacent band × $\pm 3\text{m}^2$
(4a) Later same-key sale (paper key)	10.9%	
(4b) Later tolerant-key sale	59.9%	
(5) Later similar-unit sale in community	66.1%	layout × $\pm 3\text{m}^2$, any floor

⇒ P(never relists) $\geq 34.5\%$; P(never transacts on the platform again) $\geq 33.9\%$.
Withdrawn listings of the clean window (N=87,414); key matches include identical twins, so every share is an upper bound on relisting.

Notes. Each row credits a withdrawn listing with a later platform event under a progressively coarser match. The exact attribute key is the paper's relisting convention; coarser keys overstate relisting because they match identical twins within the community, so each share is an upper bound on the true relisting rate under that key. The final rows give the implied lower bounds on the share of withdrawn listings with no subsequent observed platform event of any similar kind.

standard-deviation rise in uncertainty, comparable to the discount documented for German apartments in Amaral (2024) (Online Appendix OA.E.2, Table OA.15). The listing-price counterpart is nearly identical (-0.3652 versus -0.3667), so the discount is set at listing and survives into the negotiated price.

Two trim-robust statements anchor the discount against the heavy upper tail of $\hat{u}^2$. The list-to-deal wedge, which nets out the hedonic residual from which the uncertainty measure is built, falls with uncertainty (-0.0011, $t = -6.0$) whether or not the extreme tail is trimmed. And matching the specification to the discount's shape (the portfolio sort, Figure OA.6, shows it concentrated among the hardest-to-value listings in a cell) makes the transaction discount trim-robust: the top within-cell uncertainty decile transacts 0.73 percent lower ($t = -6.0$), and after removing the extreme top one percent of $\hat{u}^2$ the remainder still transacts 0.22 percent lower ($t = -2.4$). The linear rank gradient on the deal price thins under trimming (-0.0059, $t = -4.3$, insignificant once the top one percent is removed), but that reflects the wrong functional form for a tail-concentrated discount, not an artifact.

5.1.4 Is the idiosyncratic risk compensated? Repeat sales

One distinction keeps the paper's asset-pricing claim precise: the discount is the price at which a hard-to-value home changes hands, not a return its owner earns for bearing the idiosyncratic risk. That such homes transact at lower prices is part of how the friction surfaces at the point of sale, and the paper documents it; what the paper denies is that the buyer is compensated in returns for holding the risk. Across the repeat sales the attribute keys identify, the answer is no on every test (Table 9): no reliable return premium, no recoupment of the entry discount at resale, and a cell-price position that travels with the unit from one owner to the next.

The repeat-sale pairs test this directly: across $12{,}445$ pairs in which both spells sold (median holding 1.6 years), the resale carries a positive uncertainty-rank persistence (0.12,

$t = 10.1$; the unit tends to remain harder to value than its cell), and annualized realized returns show no robust premium: the raw gradient on the purchase uncertainty rank is $+0.35$ percentage points per year ($t = 1.3$; Table 9, column 1). Partialling out the signed entry deviation, which itself carries the textbook transitory-price reversion, leaves only a marginally positive, conventionally insignificant tilt that moves across reasonable pair constructions.

Without the screen the sample widens to $20{,}364$ both-sold pairs and the raw gradient turns significantly positive ($+1.3$ percentage points per year, $t = 3.9$). But the added pairs are precisely the high-attention units the screen removes, and they are upward-selected on resale: their mean annualized return is 19.6 percent against -0.6 percent for the screened pairs. The gradient lives entirely in the tail: trimming the top and bottom five percent of returns collapses it to $+0.2$ percentage points ($t = 1.2$) without the screen and to zero ($t = 0.1$) with it. The realized-return premium does not survive tail-robust measurement; in both samples the tail-robust gradient is statistically indistinguishable from zero. The tilt, where it appears, is selection rather than ex-ante compensation. A return regression sees only the homes that found a buyer, and a hard-to-value home needs an unusually high buyer to close, so the resales that appear are selected on a lucky high draw, more so the more dispersed the home's valuations. Formally, realized returns condition on a completed resale, and Lemma OA.2(ii) says accepted prices select upward more strongly for more dispersed units, the force Sagi (2021) documents in commercial real estate. The waiting, the failed showings, and the sales that never happen, the liquidity outcomes this paper measures, are censored out of any return regression, and they are what the ex-ante buyer pays.

Two further cuts of the same attribute-key pairs (under a marginally tighter key construction that yields $12{,}445$ both-sold pairs) sharpen the no-recoupment reading. First, a unit's position in its cell's price distribution is sticky across the two sales (slope 0.44, $t = 16.4$), and the change in that position between purchase and resale is unrelated to the purchase-time uncertainty rank (-0.0004, $t = -0.2$): whatever discount the first transaction embeds, the resale does not converge it away. The burden is borne by each owner in turn, not banked by the buyer. Second, the resale's own spell prices in exactly as the conceding-seller mechanics of Online Appendix OA.F.12 imply: each log point of resale time on market is associated with a 0.92 percentage-point lower annualized return ($t = -6.8$), while the length of the purchase spell carries no offsetting premium ($t = -0.8$): the concession paid at the exit gate is visible in returns, and the concession captured at the entry gate is not harvested, because the position it buys travels with the unit. (The fast-resale selection interaction implied by Lemma OA.2(ii) carries the predicted positive sign but is imprecise, $t = 1.0$.) The listing-implied premium reading, for its part, does not survive the same tail-robust measurement (Amaral, 2024 for the contrasting reading).

Table 9: The risk is borne at the point of sale, not compensated in the holder's return

	(1) Annual. return	(2) Annual. return	(3) Δ position	(4) Exit position
Purchase uncertainty rank r	$+0.0035$	$+0.0040$	-0.0004	
	(0.0028)	(0.0028)	(0.0022)	
$\log$ days on market, purchase		-0.0008		
		(0.0010)		
$\log$ days on market, resale		-0.0092^{***}		
		(0.0014)		
Entry position (purchase)				$+0.4438^{***}$
				(0.0270)
Community + month FE (additive)	Yes	Yes	Yes	Yes
Observations	11,697	11,697	11,697	11,697

Notes. Quasi-repeat sales: listing pairs sharing an attribute key (community $\times$ rounded size $\times$ floor band $\times$ facing $\times$ floor plan $\times$ build year $\times$ building height), exactly two listings on the key, both spells sold, holding above three months ($12{,}445$ pairs; $11{,}697$ after community and month singletons are absorbed). Columns (1)–(2): annualized log resale return $(\log p_2 - \log p_1)$/years. Column (3): the change between purchase and resale in the unit's position in its community $\times$ deal-month price distribution (demeaned log deal price), so ≈ 0 means the entry discount is not converged away. Column (4): exit position on entry position. All prices condition on a completed sale; the follower-to-viewing screen and additive community and month fixed effects are applied throughout; standard errors clustered by community. Read across: no reliable return premium on the purchase-time uncertainty rank (Column 1; the listing-implied premium likewise does not survive tail-robust measurement, Section 5.4.3); the resale concession is visible in returns while the purchase wait is not (Column 2, the two gates); the entry discount is not recouped at resale (Column 3); and a unit's price position carries from one owner to the next (Column 4). $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

5.2 Tests of the mechanism

The main findings establish what happens; the model says why, and it makes sharper predictions a price-level regression cannot deliver. I take them up in turn. Valuation uncertainty moves two distinct moments of buyer valuations at once, visible as a falling price alongside rising attention (the two-channel scissors); the seller's clearing function flattens where comparables are noisy, so the asking price barely moves the chance of sale; and the friction intensifies when the market is thin.

5.2.1 Two channels: risk discount versus valuation dispersion

The price discount in Section 5.1.3 is a statement about the mean of the price distribution, but the illiquidity of hard-to-value homes can arise through two economically distinct channels that act on different moments. Under a risk-discount channel, valuation uncertainty lowers the mean of buyers' willingness to pay (cautious buyers shade their bids), so harder-to-value units sell for less and clear more slowly. Under a valuation-dispersion channel, uncertainty widens the spread of buyer valuations: the seller must sample more buyers to find a high draw, which lengthens search and raises the attention a listing draws, while the realized price reflects selection on the upper tail rather than a lower mean. The two channels imply different signatures (in the price level, in the attention buyers pay and the way sellers revise their asking prices, and in the dispersion of negotiated outcomes), so the data can say which is active even without a structural model.

Table 10 reports the fingerprints, each measured by the within-cell rank of uncertainty as in Table 6; they point to both channels operating. The mean log deal price falls with uncertainty (column 1): the risk discount. The dispersion channel cannot be read from the raw variance of prices: the uncertainty measure is itself the squared price residual, so the two are mechanically linked (their correlation exceeds 0.98). I therefore measure disagreement with objects that difference out the price level, and three independent ones all rise with uncertainty. The dispersion of the negotiated outcome, the spread of the list-to-deal wedge, widens (column 2); buyers accumulate more committed attention before the unit clears (column 3); and, on the other side of the market, sellers accumulate more asking-price revisions over the listing's longer life (column 4). The costly stage of the funnel moves the opposite way: physical showings fall (-0.0204, $t = -2.9$; column 5)—cheap evaluation up, costly inspection down, the split the visit gate predicts. The split survives holding the spell fixed: on the matched showing sample, followers still rise ($+1.5$ percent, $t = 2.2$) while showings fall 4.4 percent ($t = -7.2$) at the same log days on market, and showings per follower fall 6.1 percent (Table OA.36, Online Appendix OA.F.12).

Each of columns 2–4 is essentially uncorrelated with the uncertainty measure itself (correlations below 0.03), so none is a price-residual tautology. One mechanical channel remains: followers and revision counts are stocks that grow over the listing spell, and any force that slows clearing, the risk discount included, accumulates them (Equation (17)). I therefore re-estimate each margin at a fixed spell (Online Appendix OA.F.11, Table OA.33). The split is sharp: the negotiated-wedge dispersion is untouched ($+0.0013$, $t = 10.2$); a third of the follower gradient survives ($+0.017$, $t = 2.5$, and $+0.006$ per day of exposure; $+0.023$, $t = 3.6$ without the FVR screen) while casual page views do not rise (falling significantly on the analysis screen, flat without it): per unit of exposure, hard-to-value homes draw more committers but no more browsers, a composition shift that spell length cannot manufacture; revision counts do not survive ($t = 0.5$), so I read column (4) as accumulation rather than independent disagreement evidence.

The two channels are identified by the combination: the mean price falls while the fixed-spell disagreement margins rise, a co-movement no single channel produces. A pure risk discount lowers the price but gives a buyer nothing more to evaluate per visit, so it moves neither fixed-spell margin (in the model, per-exposure attention loads on the valuation spread alone); pure dispersion raises disagreement and attention but, through selection on the upper tail, weakly raises the realized price. Only the two channels together reproduce a falling price alongside rising fixed-spell disagreement (Proposition 1).

I read this as evidence that the illiquidity operates through both a risk-discount and a valuation-dispersion margin, not one alone, and that the uncertainty measure captures genuine divergence of opinion: it shows up independently on the buyer side (per-exposure attention and the follower–showing split), in the negotiation (the wedge dispersion), and, as an accumulation over the longer spell, in seller revisions, not merely in the price residual from which it is built. A transparent, though non-causal, timing decomposition lends further, weaker support: these disagreement gradients steepen sharply after the December 2016 "houses for living, not for speculation" turn, from a thin, imprecisely estimated pre-period (13 percent of listings; Online Appendix OA.G.7). The cross-sectional facts themselves hold over the full sample and within the mature 2016–2019 cohorts; nothing in the paper rests on the break.

This also sharpens the interpretation of the discount in Section 5.1.3: it is neither solely a risk premium (which would move none of attention, revisions, or the negotiated spread) nor solely mismeasured quality (which would not raise buyer attention). Estimating the two channels' relative magnitudes would require a structural search model mapping dispersion to the arrival and attention channels and the discount to the offer mean; the cross-sectional design does not sharply identify that decomposition, so I leave it to future work.

Three alternatives, tested. Three rational alternatives could in principle mimic the scissors, and all are testable. First, sellers can buy attention with a low ask (Anundsen et al., 2022): underpricing alone produces rising attention and a falling price with no role for valuation difficulty. But uncertainty is not underpricing. The top within-cell uncertainty decile is posted below comparables almost exactly half the time (50.3 percent; the measure is symmetric in the sign of the residual), and controlling for the signed relative ask leaves the attention gradient essentially unchanged ($+0.047$ to $+0.044$) even though the ask itself attracts attention strongly: the underpricing channel is present in the data, and it does not carry the uncertainty effect (Online Appendix OA.F.11).

Second, under common-value uncertainty rational buyers shade their bids, more deeply the more competitors they face (Milgrom and Weber, 1982), so winner's-curse shading with endogenous entry also pairs a lower price with more interest. Bid shading, however, is a buyer-side force: it cannot widen the dispersion of the negotiated wedge (the two-sided margin of column 2, which survives every spell adjustment) nor drive the seller-side asking-price revisions of column 4, both margins that a one-sided force on the winning bid leaves untouched. The discount is indeed concentrated among high-attention listings (Online Appendix OA.F.11); the dispersion channel itself predicts that concentration (the listings that need the most search are those with the most disagreement), so that moment alone cannot separate shading from dispersion, and I rest the separation on the seller-side margins, which shading does not produce.

Third, sellers who post noisier asks would mechanically realize more variable list-to-deal wedges, mimicking the column-2 dispersion with no buyer disagreement. But the deal price tracks the signed list residual nearly one-for-one within community and month (pass-through 0.997, and 0.994 in the top uncertainty quartile), so the ask's idiosyncratic variance enters the wedge with weight $(1-0.997)^2 < 10^{-5}$: the negotiated-wedge dispersion is precisely the part of price formation the ask does not pin down. The same noisy-ask-then-concede story also fails the funnel: at a fixed spell it has no mechanism to draw more cheap monitors while drawing fewer costly showings (followers $+1.5$ percent, showings -4.4 percent; Table OA.36), the split that requires information demand itself to respond to disagreement. A complementary panel exercise points the same way, though I read it as suggestive: the sale hazard rises after comparable transactions close nearby, and the rise appears weaker for hard-to-value units, consistent with idiosyncratic valuation difficulty that public comparable information does not resolve, and that search must (Online Appendix OA.F.13). A recovered showing count sharpens the split between cheap and costly engagement: harder-to-value homes draw more followers but fewer physical showings, and markedly fewer per follower, so disagreement raises evaluation while the shaded mean deters the costly visit, just as it deters the purchase (Online Appendix OA.F.12). This two-channel pattern is not confined to Beijing: committed attention, seller revisions, and the negotiated wedge all rise with valuation uncertainty in four further provinces spanning some thirty-six cities, while the sale-probability margin, which requires the unsold listings only Beijing records, stays the city's to test (Section 7, Table OA.12).

Figure 9 collects the signature in one place. The margins that move are the quantity and disagreement margins: time on market, showings, followers, and revisions; the transaction price barely moves, about -0.1 percent. The second moment is borne in how the home trades, not in what it fetches.

5.2.2 Where homes are hard to value, the price mechanism stops clearing

The margins so far (a lower probability of sale, withdrawal rather than delay, a discount that is never recouped) share one structural cause, visible in the object a list-price-setting seller actually controls: the demand curve mapping how a listing is priced relative to comparable homes into its probability of sale. Following Guren (2018), define a listing's *relative markup* as the within-community–month, quality-adjusted residual of its log posting price: how far above or below contemporaneous local comparables the seller asks. (This is the listing's price relative to its competitors, distinct from the list-to-deal wedge of Section 5.2.1, which compares a single unit's own asking and transaction prices.) The outcome is sale within thirteen weeks, a withdrawal counted as a non-sale. Guren (2018) shows for U.S. listings that this curve is concave: pricing above comparables cuts the sale probability sharply while pricing below helps little. The Beijing curve is likewise hump-shaped, peaking near

Table 10: Two Channels: A Risk Discount in the Price Level, Disagreement on Both Sides

	(1) Price level log deal price	(2) Negotiation |wedge| spread	(3) Buyer attention log followers	(4) Seller revisions price changes	(5) Costly visits log showings
Uncertainty rank r	-0.0074^{***} (0.0015)	$+0.0013^{***}$ (0.0001)	$+0.0428^{***}$ (0.0071)	$+0.0358^{***}$ (0.0070)	-0.0204^{***} (0.0070)
Observations	169,967	169,967	217,967	217,967	168,648

Notes. Each column is a separate regression on the within-community–month rank of valuation uncertainty $r^{\text{unc}} \in [0, 1]$ (mature 2016–2019 posting cohorts, follower-to-viewing ratio at most one, with community, year–month, floor, and construction-type fixed effects; standard errors clustered by community). Column (1) is the log deal price per m^2, whose mean carries the risk-discount fingerprint. Two caveats apply, because the uncertainty measure is itself the squared price residual. First, the raw variance of prices is mechanically tied to it (correlation above 0.98), so I never use it. Second, the price level in Column (1) partly inherits the sign asymmetry of the residual: $\hat{u}^2$ is left-skewed, so high-uncertainty listings are disproportionately priced below their hedonic benchmark (the rank correlates -0.34 with the signed residual), which depresses both the list and the deal price. The residual-free part of the discount is the list-to-deal wedge, which nets out the hedonic residual entirely and still falls with uncertainty (-0.0011, $t = -6.0$); I read it as the clean evidence for the risk-discount channel. Columns (2)–(4) likewise use disagreement signals that difference out the price level, each essentially uncorrelated with the measure (correlations below 0.03): the absolute deviation of the list-to-deal wedge from its fixed-effects benchmark, a centered measure of negotiated-outcome dispersion robust to the wedge's mean shift (column 2); log saved-listing followers accumulated by the listing, buyer attention (column 3); and the count of asking-price revisions, the seller side (column 4). Followers and revision counts are stocks that grow with the listing spell; Online Appendix OA.F.11 (Table OA.33) reports each column at a fixed spell. Column (5) is the recovered count of physical showings (*daikan*) on the matched sold subsample—the costly-visit stage of the funnel, a count with no mechanical tie to the price residual: harder-to-value homes draw fewer showings even as they accumulate more followers, the split the visit-gate selection of Proposition OA.8 predicts (Online Appendix OA.F.12). The mean price falls, the disagreement signals rise, and the costly visit falls—the pattern that requires both channels. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

comparables (Figure 10); the slopes on each side, and how they vary with submarket opacity, are what follow.

What is new is what governs the curve's slope. I split community–months at the median of a leave-one-out measure of submarket valuation dispersion (the average squared hedonic residual of the other listings in the cell, so a home's own price never enters its own submarket's uncertainty), and trace the curve in transparent (clear-comparable) and opaque (noisy-comparable) submarkets separately (Figure 10, Table 13). Where comparables are clear, the curve is steep: moving the asking price toward comparables raises the sale probability sharply, a discount-side slope of $+1.45$ ($t = 15.2$). Where comparables are noisy, the curve is nearly flat, a discount-side slope of $+0.17$ ($t = 3.3$): in raw terms a price cut moves the probability of sale about eight times less. Opaque submarkets do carry wider price dispersion, so most of that raw gap is the wider variance; but in comparable within-cell standard-deviation units the lever is still about 1.4 times steeper in transparent submarkets ($+0.082$ versus $+0.060$; interaction -0.017, $t = -4.0$), a more modest but genuine flattening, not a pure variance artifact. The flattening is concentrated on the discount margin (the interaction of

Figure 9: Borne as Illiquidity, Not Priced

Notes. Effect of a move from the least- to the most-uncertain unit within a community and month on each outcome, as a percent change (point estimate and 95% confidence interval). The estimates are the headline within-cell-rank coefficients of Table 6 (probability of sale, time on market, followers) and Table 10 (asking-price revisions, physical showings, list-to-deal price wedge); the probability of sale and the revision count are scaled by their sample means for a common axis. The quantity and disagreement margins move several percent; the transaction price is essentially flat.

the discount slope with submarket uncertainty is -0.18, $t = -3.9$); the overpricing slope attenuates only modestly.[10]

In economically comparable units the contrast reads as follows. Restricting to common support (a relative markup within five percent of comparables) and pricing a five-percentage-point move toward comparable value, Table 11 reports the implied change in the thirteen-week sale probability by submarket-dispersion quintile: about six to seven percentage points in the two clearest-comparable quintiles, against two points, statistically indistinguishable from zero, in the noisiest, with the middle quintiles noisily ordered between. The common-support translation gives up the visually dramatic raw contrast for one a buyer or seller could act on: the same price concession buys roughly three times the execution where comparables are clear.

The split isolates the slope, not the level. Opaque submarkets if anything clear marginally faster (a thirteen-week sale rate of 63 versus 62 percent, an identical eventual sale rate, a composition difference the community$\times$month fixed effects absorb), which is why Figure 10

[10]What the lever looks like when buyers are pooled rather than searched for is visible in China's judicial housing auctions: participation falls by roughly 2.4 log points per unit of the starting-price-to-appraisal ratio (Xu, Zheng and Wu, 2024), the assembled crowd responding sharply to relative price. The flattening documented here is what remains of that lever when each buyer must be met one at a time and buyers disagree.

Table 11: The Price Lever in Economic Units: A Five-Point Move Toward Comparables

Dispersion quintile	Discount-side slope	(SE)	ΔP, 5pp toward comps	95% CI	N
Q1 (clearest)	$+1.207^{***}$	(0.367)	$+6.03$pp	$[+2.44, +9.63]$	10,755
Q2	$+1.399^{***}$	(0.444)	$+6.99$pp	$[+2.64, +11.35]$	9,731
Q3	$+0.250$	(0.396)	$+1.25$pp	$[-2.63, +5.13]$	9,023
Q4	$+0.923^{**}$	(0.377)	$+4.62$pp	$[+0.93, +8.31]$	7,925
Q5 (noisiest)	$+0.401$	(0.498)	$+2.01$pp	$[-2.87, +6.88]$	6,115

Common support $|\text{markup}| \leq 0.05$; Demand_Curve.R conventions (cells ≥ 8, 13-week horizon).

Notes. Discount-side slopes of the thirteen-week sale probability in the listing's relative markup, estimated within submarket-dispersion quintiles on the common-support sample ($|\text{markup}| \leq 0.05$) under the demand-curve conventions (community$\times$month fixed effects, cells of at least eight listings, leave-one-out dispersion), with cluster-robust confidence intervals; ΔP prices a five-percentage-point move toward comparables. The middle quintiles are noisily ordered; the load-bearing comparison is the clearest against the noisiest. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

centers each curve on its own mean. The paper's within-cell penalty on the hard-to-value home survives inside both submarket types (-0.050 in transparent, -0.024 in opaque): what a noisy submarket changes is the seller's price lever, not the home's fate.

This is the mechanism beneath the paper's title, and it is a comparative static of the model rather than a reduced-form regularity. A meeting clears when a buyer's valuation exceeds the seller's anchored reservation P^S, so the per-meeting acceptance rate (and hence the probability of sale, a monotone transform of it) traces a demand curve in the relative price, $q(P^S) = 1 - \Phi\big((P^S - \mu_V)/\sigma_V\big)$, where μ_V is the mean and $\sigma_V = \rho^2\sigma_b$ the standard deviation of buyers' valuations (σ_b is belief dispersion; Equation (13)). Its slope at the reservation is $\partial q/\partial P^S = -\phi(z)/\sigma_V$, $z \equiv (P^S - \mu_V)/\sigma_V$: the effect of price on the chance of sale is the density of buyer valuations at the reservation, scaled by $1/\sigma_V$. Where buyers agree (σ_V small) that density is concentrated and price is a sharp lever near comparables; where they disagree (σ_V large) it is spread thin and price barely moves the sale. Because the disagreement channel makes σ_b, and hence σ_V, rise with valuation uncertainty ($\sigma_b^2 = \tau_\eta\widetilde{\sigma}^4$, Equation (14)), the model predicts exactly the ordering the data show (Proposition 2): a steep clearing function in transparent submarkets, a flat one in opaque ones. A seller of a hard-to-value home thus has little ability to price it into a sale: the discount does not recoup the friction (Section 5.1.4), and withdrawal rather than sale is the more common terminal state (Section 5.1.2). The one information primitive that fixes the risk discount ($\Delta g = 0.98$ percent of value) and the withdrawal rate ($\delta_w = 0.88$ per year) also implies the sign of the slope's response to dispersion (Online Appendix OA.A.9): the flattening is a prediction of the model, not a fact standing beside it.

Relation to Guren (2018). This paper adopts Guren's empirical demand-curve object and measurement convention (the probability of sale against the listing's relative markup, withdrawal counted as a non-sale) and adds a margin he holds fixed: the dependence of the curve's slope on cross-sectional valuation dispersion. Guren characterizes the shape of a fixed demand curve at the market level to generate strategic complementarity and house-price momentum in the time series; this paper shows how that curve's slope responds to dispersion σ_V to generate illiquidity in the cross-section. The two underlying models differ (risk-neutral monopoly list-pricing there, CARA–Nash bargaining here), so this is a shared and extended object, not a nested model. Both route price sensitivity through a second moment, but the analogy stops there: Guren's is the fixed variance of a signal-noise distribution that pins the curvature, whereas here it is the endogenous dispersion of buyer valuations σ_V, microfounded from the information primitive u (Equation (14)), that scales the slope. I do not import the strategic-complementarity and momentum payoff: that mechanism is a different, general-equilibrium architecture, and the momentum prediction does not survive in the 2015–21 Beijing sample.

Two notes bound the claim. The curve is a binned cross-section and, like any demand curve estimated this way, inherits the unobserved-quality bias Guren (2018) corrects with an instrument; that bias distorts the level and shape of a single curve far more than the difference in slope between transparent and opaque submarkets, which would require unobserved quality to load on price differently across the two. A pass-through check bounds that possibility directly: a listing's markup carries into its transaction-price residual at a rate of 0.98 in transparent submarkets and 0.99 in opaque ones, so the ask conveys the same information about value on both sides of the split, and what changes across submarkets is whether the ask moves the sale (Online Appendix OA.D.9). And the submarket measure is deliberately leave-one-out, because a home's own valuation uncertainty is by construction its squared relative markup, so only the dispersion of other homes can identify the flattening without tautology. Guren's appreciation-since-purchase instrument transplants only partially to the quasi-repeat pairs. In linear form it has no first stage on the demand-curve sample; recast as the seller's nominal-loss position (the kink reference dependence predicts when purchases are equity-financed), it clears relevance on a widened sample (first-stage $F \approx 22$) but moves the ask only in opaque submarkets, and there the instrumented price lever on the probability of sale is statistically indistinguishable from zero, consistent with the flattening. Because the instrument has no power where comparables are clear, it cannot identify the transparent-side lever or the flattening interaction, so I read the demand curve as a model-implied comparative static recovered from the cross-section, not a causally identified object.

The flattening also bears on the amplification mechanism of Guren (2018). There, the steepness of the demand curve is what disciplines list-price strategic complementarity: sellers shade toward comparables because mispricing against them is punished, and that complementarity is what turns small frictions into momentum. Where comparables are noisy the punishment weakens, the discount-side lever falling from $+1.66$ to $+0.12$, so the complementarity, and with it the momentum it transmits, should be weakest exactly where valuation disagreement is highest: a cross-sectional prediction about momentum

Table 12: The Flattening Is Not Thickness, Coldness, or Wider Variance

Panel A. The flattening survives market thickness and coldness
(discount-side slope $\times$ submarket dispersion; $P(\text{sale} \leq 13\text{wk})$, community$\times$month FE)

	Baseline	+ thickness	+ coldness	disp. $\perp$ both
Interaction	-0.182^{***}	-0.174^{***}	-0.199^{***}	-0.153^{***}
	(0.047)	(0.045)	(0.046)	(0.046)

Panel B. Discount-side slope in within-cell standard-deviation units

	Transparent		Opaque	
	Coef.	(SE)	Coef.	(SE)
Slope (SD units)	$+0.082^{***}$	(0.005)	$+0.060^{***}$	(0.006)

Notes. Raw-unit slope ratio 8.3×; in comparable standard-deviation units 1.37×, so the wider price dispersion of opaque submarkets accounts for most of the raw gap, but a genuine flattening survives (SD-unit interaction -0.017, $t = -4.1$). Standard errors clustered by community.

heterogeneity that the within-market design cannot test but a market-level one can.

Three checks address the concern that an opaque submarket is simply a thinner or colder market, rather than one where the home itself is hard to value. First, the flattening is monotone, not a median-split artifact: across quintiles of submarket dispersion the discount-side slope falls steadily, from $+1.66$ in the clearest-comparable fifth to $+0.12$ in the noisiest, where the price lever all but vanishes—the dying lever of Figure 1, the estimate this section constructs. Second, it is not market thickness or coldness: the slope$\times$dispersion interaction is essentially unchanged by controlling for the number of listings in the cell (a thickness proxy) and for the citywide monthly sale rate (a coldness proxy), and the component of dispersion orthogonal to both still drives it (-0.15 to -0.20, all significant; Table 12, Panel A). Third, opaque submarkets do carry wider price dispersion, so most of the raw eight-fold gap is that wider variance; but measured in comparable within-cell standard-deviation units the transparent lever is still about 1.4 times the opaque one ($+0.082$ versus $+0.060$; Table 12, Panel B), a genuine flattening rather than a variance artifact. The price lever weakens where the home, not the market, is hard to value.

5.2.3 Information arrives: a close comparable closes

The clearing-function evidence is cross-sectional; the data also contain a dated information event that moves the same object inside a listing's own spell: the closing of a close comparable (a same-community transaction within fifteen square meters and a similar layout), whose deal date the platform records. In a monthly panel of listing spells (1.52 million listing-months), event time runs from three months before to three months after a listing's first close-comparable arrival, with spell-age controls and community and calendar-month fixed effects; distant or less-similar arrivals serve as the comparison event (Table 14). Pre-arrival

Figure 10: The Binned Clearing Curves Behind the Lever of Figure 1

Notes. Binned probability of sale within thirteen weeks against the listing's price relative to comparable homes (the within-community–month quality-adjusted residual of the log posting price per square meter); withdrawals count as non-sales; each curve is centered on its own submarket's mean sale rate, so the comparison is of slopes, not levels, and submarkets split at the median of the leave-one-out community–month dispersion measure. These are the data behind the quintile slopes of Figure 1. The curves' level and shape away from comparables inherit the unobserved-quality composition discussed in the text (deeply discounted listings are disproportionately worse homes), which is why the paper claims only the difference in slopes across submarkets, backed by the equal pass-through of the ask into transaction prices (Online Appendix OA.D.9). Panel (b) shows the binned data behind the estimates: the probability of sale against the listing's price relative to comparables (the within-community–month quality-adjusted residual of the log posting price per square meter), withdrawals counted as non-sales, each curve centered on its own submarket's mean so the comparison is of slopes, not levels. The panel's level and shape away from comparables inherit the unobserved-quality bias discussed in the text (deeply discounted listings are disproportionately worse homes), which is why the paper claims only the slope difference; within a community and month the harder-to-value home is less likely to sell (Table 6). Submarkets split at the median of a leave-one-out community–month measure of valuation dispersion. Sample: Beijing listings posted from July 2016 with thirteen-week follow-up, in community–months with at least eight listings ($N = 82{,}400$).

coefficients are economically negligible (between -0.005 and -0.010); in the arrival month the monthly sale hazard jumps 4.2 percentage points ($t = 27.6$) on a base of 21 percent, decaying through the third month. In levels, a close arrival raises the post-arrival sale hazard by 3.1 points, roughly double the 1.6 of a distant arrival, and the withdrawal hazard also rises ($+0.3$ points, $t = 6.5$): resolution pushes both exits, the signature of information arriving rather than demand, which would raise sales alone. Two honest limits bound the reading. The arrival-month jump partly reflects correlated demand (the comparable and the focal home face the same buyers that month), which the decay profile and the close-versus-distant contrast mitigate but do not eliminate; and while the hazard's sensitivity to the listing's relative price shifts after a close arrival (-0.031, $t = -3.4$), the shift after distant arrivals is similar (-0.029), so the level response, not the slope response, is the informative margin.

Table 13: Price Is a Weaker Lever on Sale Where Comparables Are Noisy

	Transparent submarkets (clear comparables)		Opaque submarkets (noisy comparables)	
Slope of $P(\text{sale} \leq 13\text{wk})$	Coef.	(SE)	Coef.	(SE)
below comparables (discount margin)	$+1.454^{***}$	(0.096)	$+0.174^{***}$	(0.053)
above comparables (overpricing margin)	-0.635^{***}	(0.086)	-0.332^{***}	(0.066)
Observations	40,412		41,182	

Notes. Each column reports the slope of the probability of sale within thirteen weeks on the listing's relative markup, separately for the discount margin (priced below comparables) and the overpricing margin (priced above), with community $\times$ month, floor, and construction-type fixed effects and standard errors clustered by community. Transparent (opaque) submarkets are community–months below (above) the median of the leave-one-out average squared hedonic residual of the other listings in the cell. Withdrawals count as non-sales. The pooled interaction of the discount slope with submarket uncertainty is -0.18 ($t = -3.9$), and -0.017 ($t = -4.0$) with the markup expressed in within-cell standard-deviation units; the overpricing-slope interaction is -0.01 ($t = -0.1$). The level and shape of the curve inherit the unobserved-quality bias of Guren, 2018 and are not causally identified; the model-implied object is the cross-submarket difference in slope. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Within these limits the event says in time what the clearing function says in the cross-section: when public value information improves, execution responds.

5.2.4 Attention

The illiquidity of hard-to-value homes shows up in how much demand-side attention a listing must accumulate before it clears. Table 6 shows that moving from the least to the most uncertain unit in a community and month raises log followers (the count of users who save a listing, a marker of genuine purchase intent) by about four percent. Harder-to-value units must gather more committed, saved-listing attention before a buyer commits: exactly the search the model predicts when a smaller share of viewers clear the seller's reservation value (Proposition OA.6). Online Appendix OA.E.9 reports popularity and follower-to-viewing-ratio variants.

5.2.5 State-dependence: the illiquidity is stronger in thin markets

The illiquidity of hard-to-value homes is not constant; it is stronger when trading is thin. The acceptance margin predicts this: as buyer arrivals thin, each meeting matters more and the risk discount separating a hard-to-value unit's buyers from its anchored seller bites harder, so the within-market penalty should widen as the market cools. This is the unit-level counterpart of the market-level cyclicality of liquidity (sales slow and time on market lengthens in cold markets, Krainer, 2001; Novy-Marx, 2009; thick markets improve match quality, Ngai and Tenreyro, 2014; buyers turn choosier in downturns, Guren and McQuade, 2020). I split months by the Beijing-wide sale rate, a market-state measure exogenous

Table 14: The Arrival of a Close Comparable Raises Execution Within the Spell

A. Sale hazard, close-comparable arrival (ref. $\tau = -1$)	Coef.	(SE)
$\tau \leq -4$	-0.0095^{***}	(0.0013)
$\tau = -3$	-0.0073^{***}	(0.0020)
$\tau = -2$	-0.0045^{**}	(0.0018)
$\tau = 0$	$+0.0421^{***}$	(0.0015)
$\tau = +1$	$+0.0268^{***}$	(0.0015)
$\tau = +2$	$+0.0186^{***}$	(0.0016)
$\tau = +3$	$+0.0129^{***}$	(0.0018)
$\tau \geq +4$	-0.0048^{***}	(0.0014)
Distant arrival, post (level)	$+0.0157^{***}$	(0.0011)
B. Post-arrival indicators		
Sale hazard: close $\times$ post	$+0.0309^{***}$	(0.0010)
distant $\times$ post	$+0.0164^{***}$	(0.0010)
Withdrawal hazard: close $\times$ post	$+0.0033^{***}$	(0.0005)
distant $\times$ post	$+0.0016^{***}$	(0.0006)
C. Price-sensitivity steepening (fixed relative ask)		
relp	$+0.0879^{***}$	(0.0078)
close post $\times$ relp	-0.0311^{***}	(0.0093)
distant post $\times$ relp (separate reg.)	-0.0292^{***}	(0.0109)
1,524,013 listing-months; FE community, calendar month, spell month; cluster community.		

Notes. Monthly listing-spell panel; the event is the first same-community comparable transaction within $\pm 15\,\mathrm{m}^2$ and a similar layout closing within twenty-four months of the focal posting, defined unconditionally on the focal spell's fate so post-arrival status is not selected by the listing's own termination. Panel A traces the sale hazard in event time around the close arrival (reference month -1), with the distant-arrival post-period level shown for comparison; Panel B compares post-arrival level shifts for close and distant arrivals on both terminal hazards; Panel C interacts the post indicators with the listing's relative ask. Fixed effects: community, calendar month, and spell month; standard errors clustered by community. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

to any individual listing, and interact the within-cell uncertainty rank with an indicator for cold (below-median sale-rate) months. Table 16 reports the result for time on market, estimated under the paper's measurement conventions (the FVR screen, and cohorts posted from July 2016 so every duration is real): the uncertainty–time-on-market slope is about 60 percent steeper in cold months ($+0.032$, $t = 2.5$), the steepening survives trimming the top one percent of the uncertainty distribution ($+0.028$, $t = 2.1$) and is marginal when the 2020–21 COVID window is excluded ($+0.027$, $t = 1.9$), and a continuous market-state measure carries the same sign (the interaction with the standardized monthly sale rate is -0.018, $t = 2.1$). Against a mean time on market near seventy days, the hardest-to-value homes wait roughly three to four extra days in a normal month and about six in a thin one.

I read this as suggestive, not dispositive. Its precision is modest, it appears only for turnover-measured market states and on the duration margin rather than the probability of a quick sale, and it concentrates in the thinnest months. The cross-sectional design cannot establish causal amplification over the cycle, so I present it as a state-dependent regularity, not a propagation channel.

5.2.6 The boom as a test of the disagreement channel

The model's disagreement channel makes the boom a sharp test: the homes buyers disagree about most have the most room for optimism to lift price, so if the market-level Miller premium ever surfaces unit by unit it should do so for the hardest-to-value homes at the peak. The 2015–16 Beijing boom lets me look, and the price cross-section bears a partial mark, mirroring the equity pattern in which overpricing concentrates in hard-to-value, high-sentiment assets (Baker and Wurgler, 2006; Stambaugh, Yu and Yuan, 2015), now at the unit level in housing. Among matched deals closed before December 2016, top-decile-uncertainty units sold at or above the asking price 4.6 percentage points more often than their easier cellmates (19.8 versus 15.2 percent), a differential that vanishes after the March-2017 tightening (6.0 versus 6.1 percent). Once the extreme top one percent of $\hat{u}^2$ is trimmed and ranks are rebuilt, the boom deal-price gradient turns mildly positive ($+1.0$ percent, $t = 2.6$) where the post-tightening gradient stays negative. The flip is partial: the extreme decile still transacts at a discount even at the peak (-2.8 percent, $t = -3.8$), and on the only boom cohorts with real posting dates (July 2016–February 2017) the sale-probability penalty persists (-1.5 percentage points, $t = -3.8$) while the time-on-market gradient flattens to zero, the compression in market thickness the calibration of Online Appendix OA.A.7 predicts. Two caveats temper the reading: the boom sample is small and selected on platform coverage ($6{,}241$ deals in the estimation sample), and pre-2016 posting dates are backfilled, so ask staleness in a rising market cannot be controlled, though staleness common to all units differences out of the top-decile comparison. I read the pattern as the market-level optimism of the boom literature (Gao, Sockin and Xiong, 2021) leaking into the unit-level cross-section through the homes with the most room for disagreement (the sentiment channel, not the meeting-rate limit of Proposition OA.2, which no feasible thickening delivers), and it carries the same suggestive status as the cold-market result. The search inversion the paper documents is the regime in which more than ninety-five percent of the matched sample trades.

Whether that boom imprint is transitory matters for the mechanism. A diagnostic-expectations account (buyers over-extrapolating noisy signals, as in Chodorow-Reich, Guren and McQuade, 2024) predicts that the cross-sectional premium on hard-to-value homes inflates in the boom and then unwinds as beliefs correct; permanent disagreement (Miller, 1977) predicts it persists. The data side with persistence. Building a quality-adjusted price

Table 15: The Re-Rating of Hard-to-Value Communities Is Permanent, Not Transitory

Community price change	Coef. on pre-boom valuation uncertainty (z)	(SE)
Boom (2015–2017 peak)	$+0.0069$	(0.0049)
Post-peak (peak–2020)	$+0.0092^{**}$	(0.0038)
Full (2015–2020)	$+0.0160^{***}$	(0.0047)
Communities	341	

Notes. Each row regresses the change in a community's quality-adjusted log price per square meter (the deal price residualized on size and build-year polynomials, floor, and construction type, averaged by community and quarter) on the community's pre-boom valuation uncertainty (mean squared hedonic residual through 2016Q2, z-scored), across communities with at least six deals per quarter in each window. Boom: trough (2015) to peak (2016Q4–2017Q2). Post-peak: peak to 2020. Full: trough to 2020. A diagnostic or extrapolative account predicts a negative post-peak coefficient (the re-rating unwinds); the data show none. Heteroskedasticity-robust standard errors in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

index for each community and regressing its price change on the community's pre-boom valuation uncertainty (Table 15), the communities whose homes are hardest to value do not give back their re-rating: the post-peak change is, if anything, positive ($+0.009$ per standard deviation of pre-boom uncertainty, $t = 2.4$, not the reversal extrapolation predicts), and the full 2015–2020 re-rating ($+0.016$, $t = 3.4$) exceeds the boom-window move. The boom amplitude itself rises with uncertainty only weakly once the initial price level is absorbed ($t = 0.9$), so the robust fact is the absence of reversal. A permanent cross-sectional re-rating is the signature of persistent disagreement, not of a belief cycle. The test is between communities and on a price-derived measure, so it complements rather than re-identifies the within-cell results. But it is inconsistent with the sharpest behavioral alternative to the paper's reading, though the between-community design has limited power against a slow or partial belief correction.

5.2.7 Who stops trading when the market turns

A housing downturn arrives as a quantity event: transactions collapse while prices barely move (Stein, 1995; Genesove and Mayer, 2001), the asymmetry the hot–cold cycle literature treats as the market's defining regularity (Krainer, 2001; Novy-Marx, 2009; Ngai and Tenreyro, 2014). The mechanism here gives that fact a cross-sectional reading. If the price lever is weakest where valuation is hardest (Section 5.2.2), and the exit from an unpriceable listing is withdrawal rather than a deeper cut (Section 5.1.2), then when demand falls, the trades that disappear should be disproportionately the hard-to-value ones. Beijing's two in-sample contractions behave this way (Table 17). In the six months after the March 2017 credit tightening, volume fell by 62 percent. Transacted prices moved an order of magnitude less: up 7 percent on the same six-month windows, which abut the March peak; from that

Table 16: Cold-Market Amplification of the Uncertainty–Illiquidity Link

	(1) Baseline	(2) Top-1% trim	(3) Excl. 2020–21
Rank uncertainty r_{unc}	$+0.0514^{***}$ (0.0087)	$+0.0476^{***}$ (0.0087)	$+0.0569^{***}$ (0.0100)
$r_{unc}\times$ Cold market	$+0.0324^{**}$ (0.0132)	$+0.0275^{**}$ (0.0133)	$+0.0268^{*}$ (0.0144)
Community FE	Y	Y	Y
Year–Month FE	Y	Y	Y
Observations	204,557	190,173	169,967

Notes. Dependent variable: log days on market (sold units). The sample applies the paper's conventions throughout: FVR ≤ 1, and cohorts posted from July 2016 onward (earlier posting dates are backfilled; Figure 4); the cold/hot classification is computed on the same months. The regressor is the within-community–month rank of idiosyncratic valuation uncertainty $r_{unc} \in [0, 1]$, interacted with an indicator for cold months (Beijing-wide monthly sale rate below the sample median). All columns include community and year–month fixed effects, with standard errors clustered by community; the cold main effect is absorbed by the month fixed effects. Column (1) is the baseline; (2) drops the top one percent of $\hat{u}^2$ and re-ranks within cell; (3) excludes 2020–21. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

peak they gave back about a tenth by the second half of the year, settling 2 percent above the pre-tightening average.[11] The decline in trade was 56 percent in the easiest within-cell difficulty tercile against 66–68 percent in the upper two, and the mean difficulty rank of the homes that still traded fell by 0.048. A back-of-the-envelope share follows: had the upper terciles fallen at the easiest tercile's rate, about 12 percent of the missing trades would not have gone missing—the compositional excess this margin accounts for, the remainder being the common demand shock. The COVID freeze of early 2020 repeats the pattern monotonically across terciles (53, 57, 58 percent), with the same compositional tilt.

These are descriptive decompositions, and I do not read them causally: a credit tightening moves demand through many channels, and the sale-probability gradient itself does not steepen significantly with citywide volume (interaction $t = 0.9$; the state dependence of Section 5.2.5 lives on the time-on-market margin). The content is compositional, and it is a different axis from the canonical accounts of the asymmetry, which are seller-state ones: owners hold out because their equity is thin (Stein, 1995; Genesove and Mayer, 1994) or because selling below the purchase price is painful (Genesove and Mayer, 2001), and Andersen et al. (2022) attribute most of the price–volume correlation (58.5 percent) to the down-payment channel. Those frictions select which owners withhold trade; the margin here is an attribute of the asset, selecting which homes stop trading in whoever's hands. The two are separable in these data by the sign of the closing price (the loss-position holdout sells slower and higher, the hard-to-value home slower and lower), and the seller-state

[11] Narrative evidence puts divided beliefs at the center of that regime: a text-mined index of directional housing expectations on Chinese social media sat near 0.50 nationally from May through July 2017, an even split between optimists and pessimists, after running near 0.65 in March (Zhu et al., 2023). What a mean-sentiment series reads as neutrality, the cross-section resolves as disagreement: the market did not stop expecting, it stopped agreeing.

Table 17: When the Market Turns, the Hard-to-Value Trades Disappear First

Episode	Volume	Decline in transactions, by difficulty tercile Easiest	Middle	Hardest	Δ mean rank, transacted	Transacted price
March 2017 tightening (6mo pre vs. post)	-62%	56%	68%	66%	-0.048	$+7.2\%$
COVID freeze (Feb–Apr 2020 vs. prior 5mo)	-55%	53%	57%	58%	-0.023	-2.9%

Notes. Each row compares transaction counts in the window before and after an in-sample contraction: the March 2017 credit tightening (September 2016–February 2017 versus April–September 2017; March skipped as the announcement month) and the COVID freeze (September 2019–January 2020 versus February–April 2020). Terciles of the within-community–month valuation-uncertainty rank are assigned at posting; columns report each tercile's own percentage decline in transactions. "Δ mean rank, transacted" is the change in the average difficulty rank of homes that still traded; "transacted price" is the change in the average deal price per m^2 among them, on the same windows; because the 2017 post window abuts the March price peak, the level had given back about ten percent from that peak by the second half of 2017, settling two percent above the pre-tightening average. Descriptive decompositions: a tightening moves demand through many channels, and no causal claim is attached.

channel is documented on this sample's own extensive margin in Online Appendix OA.A.8. The market that survives a downturn is thus selected toward the homes comparables can price. The selection barely moves the measured price level (quality-adjusted, transacted prices in late 2017 sat 1.6 percent above their pre-tightening level, 3.3 percent raw, and the within-cell price gradient prices the 0.048 composition shift at under a tenth of a percent), so the price index is not much biased by the exits; it is silent about them. An index reads only from the homes that still have prices, the sample-selection concern long recognized in the index literature (Clapp and Giaccotto, 1992; Gatzlaff and Haurin, 1997), and here the selection runs on valuation difficulty: the index is right about the survivors and says nothing about the growing stock of homes the market can no longer price. The tilt shows in flows as well: when a community's own market cools, more of its hard-to-value listings end withdrawn (the rank–lagged-coldness interaction on withdrawal is $+0.006$, $t = 3.1$), the flow counterpart of the shelf hardening, realized through the failing sale, consistent with Corollary 1, not through faster pull-back. The stock counterpart is the shelf of Section 5.1.2: a downturn does not just thin the market, it tilts what remains toward the tradable.

5.3 A shock to opacity: the multi-school zoning reform

On April 30, 2020, Beijing's Xicheng district announced that homes purchased after July 31 would no longer carry a fixed school assignment: enrollment would be allocated by lottery across multi-school zones. For the city's school-premium capital, the announcement converted the most valuable single attribute of a home from a certainty into a draw. In the language of the framework, this is a shock to opacity. It widens what public information leaves unresolved about a unit's value, and it does so through a component common to every unit in a community, precisely the component the within-community design nets out. The within-cell rank measure should therefore barely move while the liquidity margins, which respond to the buyer's total uncertainty, deteriorate; and the measured within-community dispersion indeed barely moves ($+0.006$ log points, $t = 0.1$) while the margins collapse.

I compare Xicheng listings with listings in districts never treated inside the sample window, excluding Haidian and Dongcheng, which adopted their own multi-school rules in January 2019 and July 2020. The window runs from May 2018 through December 2020, with community and year–month fixed effects, Xicheng-specific indicators for the February and March 2020 COVID freeze, and standard errors clustered by community. Relative to the announcement, the probability of sale within thirteen weeks falls 8.5 percentage points ($t = -7.1$) and committed follower attention rises 19 log points, about 21 percent ($t = 6.0$); a placebo treatment dated twelve months earlier, estimated on the pre-period alone, finds nothing (-0.004, $t = -0.3$). Figure 11 traces the path in quarters: seven flat pre-period quarters on both margins, then the sale margin drops ten points and the attention margin jumps twenty log points in the announcement quarter and stays there. The dip in quarter -1 on the sale margin is the announcement month itself (the rule was reported at the end of April), which I read as anticipation.

Two neighboring reforms discipline the design. Haidian and Dongcheng, estimated the same way against the same controls, produce "effects" that are statistically indistinguishable from their own pre-period placebos (Haidian $+0.067$ against a placebo of $+0.069$; Dongcheng -0.037 against -0.034), so the method flags differential trends where they exist and finds none in Xicheng. Three further audits temper the reading (Online Appendix OA.G.6). Permutation inference is the honest standard with one treated district: reassigning the treatment to each of fourteen never-treated districts and to shifted announcement dates, the real sale estimate sits third of seventeen valid placebo assignments (pseudo-$p \approx 0.22$; second of sixteen excluding a degenerate early date), so the estimate is in the extreme tail of the placebo distribution but the design cannot deliver conventional significance with this many districts, and the community-clustered t of seven overstates precision against that benchmark. The attention margin is weaker than its stock suggests: followers per day on market do not rise (-0.05 log points, $t = -1.7$), so the follower-stock jump substantially reflects listings lingering longer under the reform, not faster accumulation. And a school-premium intensity split within Xicheng is not monotone: the sale effect is largest in the middle tercile of pre-reform community ask premia, not the top (top $-$ bottom $= +0.035$, $t = 1.3$). The limits are therefore plain: one district, three post-announcement quarters for attention and two for the sale margin, an announcement rather than a randomized rule, and an event that moved expected school quality and buyer composition alongside opacity. I read the episode as directional corroboration on the quantity margin, the first in-sample slice of the design the conclusion describes, and I do not lean on its attention response.

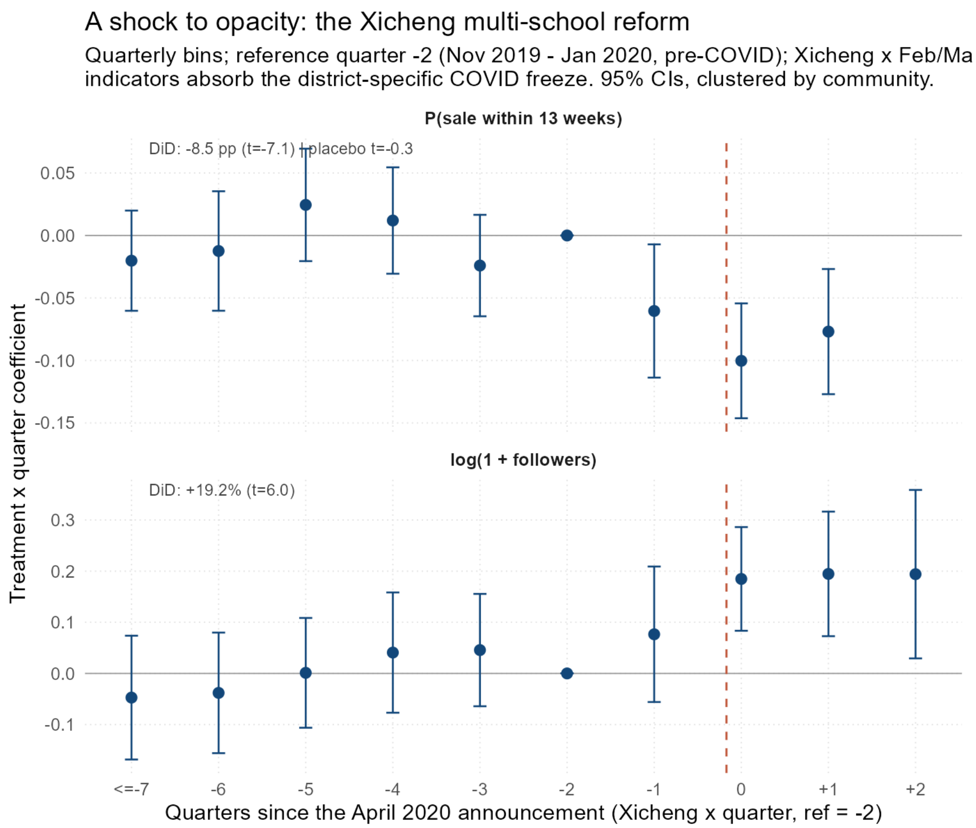


Figure 11: A Shock to Opacity: The Xicheng Multi-School Reform

Notes. Event-study coefficients on Xicheng × quarter relative to the April 30, 2020 multi-school announcement (dashed line), for the probability of sale within thirteen weeks (top) and log one plus followers (bottom). Sample: postings May 2018–December 2020 in Xicheng and never-treated districts; Haidian and Dongcheng are excluded (own multi-school rules). Quarterly bins with reference quarter −2 (November 2019–January 2020, pre-COVID); Xicheng × February and March 2020 indicators absorb the district-specific COVID freeze. The sale outcome requires a 91-day follow-up window, so its series ends in the second post quarter. Bars are 95 percent confidence intervals, clustered by community.

5.4 Ruling out alternatives

This section confronts the alternative readings of the uncertainty–illiquidity signature. Table 18 collects them: for each alternative, the margin on which it makes a different prediction, and the result that closes it. The subsections and online appendices report the estimates.

The findings could in principle reflect forces other than how hard a home is to value. This section rules out the leading alternatives in turn: threats to the within-market comparison, the possibility that the result is an artifact of price, and a time-on-market reading of the sale-probability gradient.

Table 18: Alternative Readings and the Evidence That Closes Them

Alternative reading	Discriminating evidence	Where
Panel A: Is the measure what it claims to be?		
Reverse causality: illiquidity widens dispersion	The rank is fixed at listing, from the unit's own posted price against contemporaneous comparables, not from any market outcome; supply-pressure diagnostics show no volume feedback.	§5.4.1
Seller mispricing: the residual is a choice	A price-free measure built from physical configuration alone, and a transaction-only comparable-scarcity count, reproduce the signature; both survive a horse race.	§5.4.2
Unobserved quality	Within identical-attribute units (the same home or its twin) the gradients persist; the predictable, atypicality-driven part of dispersion carries the opposite sign.	§5.4.1, OA OA.D.10
Agent quality, not the home	The price-free measure contains no agent; agents vary across a unit's spells; poorly marketed listings should draw less attention—these draw more.	§5.4.1
Computable ambiguity	A model-disagreement ensemble forecasts pricing errors well yet produces no illiquidity: what is priced is market-revealed difficulty.	OA OA.D.8
Panel B: Is the friction something else?		
Seller loss aversion, thin equity	Opposite exit signs: the loss-position holdout sells slower and higher; the hard-to-value home sells slower and lower, or never.	§5.2.7, OA OA.A.8
Transaction costs	The tax class facing the higher levy shows the mirror image—slower, dearer, less watched; the difficulty gradients are unchanged under tax-class controls.	§5.4.1
Adverse selection (lemons)	Uncertainty is two-sided—sellers revise as if they do not know the value either—and committed attention rises where lemons logic says buyers walk away.	§5.4.1
Winner's-curse shading	Rests on seller-side dispersion (revisions, the negotiated wedge) that bid shading cannot produce; in open-outcry auctions atypical homes earn premia, not discounts.	OA OA.F.11
Underpricing for traffic	Difficulty is symmetric in the signed ask (half post below comparables); the attention gradient survives an ask control that itself draws traffic.	OA OA.F.11
Duration stigma	The price discount survives conditioning on time on market.	OA OA.F.11
Thin submarkets, competition	Active same-segment competitors move no margin; the demand-curve flattening survives thickness and coldness controls.	§5.2.2
Restricted-segment policy artifact	Dropping every non-ordinary-residential listing (the segment hit by the March 2017 purchase ban) leaves the gradient slightly stronger.	§5.4.1

Notes. Each row states an alternative account of the signature, the margin on which it makes a distinct prediction, and the result that rejects it; the referenced sections and online appendices report the estimates. "OA" abbreviates the Online Appendix.

5.4.1 Threats to identification

Five concerns warrant treatment; each is addressed by the design or by extensions in Section 7.

1. **Reverse causality.** Dispersion might be an equilibrium outcome of illiquidity rather than a cause: thin trading can itself widen the residual variance. The within-community–month design mitigates this, because r^{unc} is measured from the hedonic residual of a listing's own posted price relative to contemporaneous local comparables: a property attribute fixed at the time of listing, not a market-clearing quantity realised after the fact. The supply-pressure and market-thickness diagnostics in Online Appendix OA.F show the relationship is not an artifact of local trading volume.

2. **Unobserved quality.** Community fixed effects absorb building-level amenities and the property controls absorb sorting on observables, but interior finish or view remain unobserved. Restricting the hedonic to progressively denser communities, where the within-community comparable set is larger and listing-level variation is more disciplined by competition, leaves the illiquidity relationship intact (Section 7). Decisively, decomposing $\hat{u}^2$ into a component predictable from observable characteristics (size, size-atypicality, floor, and construction type) and an orthogonal idiosyncratic component shows the illiquidity effect operates entirely through the idiosyncratic part: the predictable component carries the opposite sign (more atypical units, if anything, sell faster and are more likely to sell), so the result cannot be a relabelling of atypicality or unobserved quality (Online Appendix OA.D.5). A within-unit design corroborates this: fixed effects on exact attribute combinations (the same physical unit or its identical twin in the same building, so every time-invariant attribute, observed or unobserved, is differenced out) preserve the gradients. Within key, time on market rises 3.2 percent from the easiest- to the hardest-to-value spell ($t = 1.7$, about sixty percent of the baseline slope; $76{,}694$ listings on $31{,}094$ keys), and among clean sequential relistings (the first spell sold before the second began, a median of nineteen months apart), the probability of sale falls 1.3 percentage points ($t = -2.1$; Online Appendix OA.D.10). Within-key rank variation is limited, so I read these as calibrated support rather than headline estimates, but their value is what they exclude: a story in which the gradients reflect fixed unobserved quality would not survive a design that differences the unit's fixed attributes out entirely. A listing-side variant of the same concern is the agent rather than the home: Gilbukh and Goldsmith-Pinkham (2024) show listings held by inexperienced agents sell with markedly lower probability, a gap that widens in the bust, so bust selectivity could in principle reflect who sells the home rather than what the home is. Three results already in hand close this reading. The price-free

physical-rarity measure contains no listing choice (no price, no agent) and reproduces the gradients (Section 5.4.2); the within-key design just described differences out the home while agents vary across spells; and the attention margin runs the wrong way for an agent story: a poorly marketed listing should draw less attention, while hard-to-value listings draw more of it per exposure. Agent heterogeneity is a real margin of housing-market performance; it is not this margin.

3. **Selection into transaction.** The sale indicator is observed for every listing, but time on market is recorded only for homes that sell, so the hardest-to-value homes that never sell drop out of the duration regression and the surviving sample could understate how long hard-to-value homes take. Lee bounds on the extensive margin and an inverse-probability-weighting correction (Online Appendix OA.G) leave the sign and significance of the illiquidity results unchanged.

4. **Asymmetric information.** Adverse selection alone can generate the same illiquidity: if a hard-to-value home is one whose seller holds private information (interior condition, latent defects), then lemons logic delivers a lower price, a lower probability of sale, and a longer wait with risk-neutral buyers (Akerlof, 1970; Guerrieri and Shimer, 2014). Three features separate the results from a lemons reading. First, the uncertainty measure is built from the dispersion of contemporaneous comparables, not from the opacity of the unit itself, and the previous item shows it operates through the component orthogonal to observables, whereas the scope for seller private information loads on observable atypicality. Second, the disagreement the paper measures is two-sided: the dispersion of the negotiated wedge, where buyer and seller meet, widens with uncertainty even at a fixed spell (Section 5.2.1), consistent with symmetric, two-sided uncertainty (the seller does not know the value either), not the informed-seller asymmetry that drives Akerlof (1970). Third, under adverse selection, the information-asymmetry channel Kurlat and Stroebel (2015) document for housing, committed buyer interest should fall away from suspect listings; instead per-exposure committed attention rises with uncertainty while casual browsing does not (Section 5.2.1; Online Appendix OA.F.11). A lemons component may well coexist, but it cannot account for the joint pattern.

5. **Transaction costs.** A transfer levy is the textbook reason a home trades slowly, and levies demonstrably reduce volume (Dachis, Duranton and Turner, 2012; Deng et al., 2024); in disagreement models they are also the wrong friction in kind: trading costs do not overturn optimist pricing (Scheinkman and Xiong, 2003). Beijing prices the test inside the sample: resale taxes fall discretely with the seller's holding period, so every listing carries a tax class, and the class facing the higher levy (held two to five years) bundles a real cost wedge with a recently purchased seller. Within a

community and month that class is indeed less liquid: sale probability -0.7 percentage points ($t = 3.5$), time on market $+8$ percent ($t = 11.8$). But it sells higher ($+0.4$ percent, $t = 5.2$) and draws less committed attention (-4 percent, $t = 7.1$): longer, dearer, less watched, the signature of a priced cost borne by a holdout seller (Genesove and Mayer, 2001; Andersen et al., 2022), and the mirror image of the scissors. The difficulty gradients, estimated in the same rows, keep their own signs (attention $+4.7$ percent, $t = 7.8$), and controlling for the tax class leaves them unchanged. Costs are borne in prices; disagreement is borne in quantities. A related institutional check: the commercial-titled segment, whose purchase eligibility was separately restricted on 26 March 2017 (its transaction volume falls by two-thirds in the data), does not carry the result either: dropping every non-ordinary-residential listing and re-ranking within cell leaves the sale gradient slightly stronger (-0.013, $t = 7.4$).

5.4.2 The illiquidity is not an artifact of price

Because the uncertainty measure is a squared price residual, two doubts remain: that it captures seller mispricing rather than valuation difficulty, and that the residual reflects the seller's own posting price rather than an attribute intrinsic to the unit. A measure that never touches price answers both. From a unit's physical attributes alone (its parsed room and bath counts, size, orientation, floor band, building type, and elevator-to-household congestion) I compute how rare its physical configuration is within its community: the surprisal $-\log($share of community units of this physical type$)$. A unit whose layout is unusual where it sits is hard to price for the reason the residual measure captures, but the construction uses no price, and the rarity is intrinsic to the unit, not a seller decision. The same physical thinness underlies the characteristic-frequency instruments Amaral (2024) uses to price a return premium; I use it not as an instrument but as a direct measure, and ask whether it carries the illiquidity the price residual does.

Table 19 ranks this physical-rarity measure within community and month and runs it through the same regressions. It reproduces the entire signature: the rarest-configuration units are 2.3 percentage points less likely to sell ($t = -7.6$), wait about 21 log points longer, accumulate more followers and more revisions, and close at a wider list-to-deal wedge, each highly significant. In a horse race that enters both, the physical-rarity coefficients barely move, and the price-residual rank stays significant alongside it on time on market, followers, revisions, and the wedge ($|t|$ between 3.9 and 6.2), losing significance only on the noisy sale-probability margin (-0.0015, $t = -0.5$). The two measures correlate 0.35 within cell: two proxies for the same difficulty, one of them price-free. A measure built without any price reproduces the signature, which is difficult to reconcile with a pure mispricing or seller-choice account of the squared residual. The corroboration strengthens, not weakens, outside the COVID window.

Table 19: A Price-Free Measure of Valuation Difficulty Reproduces the Signature

	(1) P(sale)	(2) log TOM	(3) log foll.	(4) revisions	(5) wedge
Physical-rarity rank, entered alone					
Physical-rarity rank r^{phys}	-0.0233^{***}	$+0.2192^{***}$	$+0.0963^{***}$	$+0.1461^{***}$	-0.0036^{***}
	(0.0029)	(0.0100)	(0.0092)	(0.0092)	(0.0002)
Horse race against the price-residual measure					
Physical-rarity rank r^{phys}	-0.0226^{***}	$+0.2054^{***}$	$+0.0881^{***}$	$+0.1352^{***}$	-0.0032^{***}
	(0.0030)	(0.0102)	(0.0095)	(0.0097)	(0.0002)
Price-residual rank r (baseline)	-0.0015	$+0.0608^{***}$	$+0.0366^{***}$	$+0.0453^{***}$	-0.0011^{***}
	(0.0029)	(0.0098)	(0.0094)	(0.0099)	(0.0002)
Community $\times$ month FE	Yes	Yes	Yes	Yes	Yes
Observations	327,982	241,043	327,982	327,982	247,994

Notes. The physical-rarity rank is the within-community surprisal $-\log(\text{share})$ of a unit's physical configuration—parsed room and bath counts from the layout field, size decile, facing, floor band, building type, and elevator-to-household congestion from the elevator-ratio field—computed in communities of at least twenty listings and ranked within community $\times$ month to $[0,1]$. It uses no price information. The price-residual rank r is the baseline within-cell rank of the squared hedonic residual. Outcomes are as in Table 10: probability of sale; log days on market (sold units posted July 2016 onward); log followers; asking-price revisions; and the log list-to-deal wedge (sold units). All columns carry community and year–month fixed effects; standard errors clustered by community. The two ranks correlate 0.35 within cell. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

The predetermined index on the execution margins. The predetermined-feature index of Section 2.4.3 completes the program. On the execution margins it is stronger than the realized residual (Table 5): the full within-cell move lowers the probability of sale by 2.4 percentage points (against 1.3 on the realized rank), lengthens time on market by nine log points (against six), and raises withdrawal by 2.4 points (against 1.2), and in a horse race it dominates the realized rank on all three margins. Because a zero-to-one move in two different ranks is not a common scale, Table 20 restates the comparison in comparable units: per within-cell standard deviation of the underlying scores, the index gradient is about twice the realized one on sale and withdrawal (-1.36 versus -0.68 percentage points; $+1.32$ versus $+0.64$) and about 1.6 times on time on market. The margin the index does not carry is committed attention, which loads on the realized rank ($+0.040$, $t = 6.0$) and not on the index (-0.013, $t = -1.8$). The division of labor is informative, and it is the two-channel structure in one more guise: the predictable, scarcity-driven component of valuation difficulty is an execution friction, while the attention signature rides on realized dispersion, the component in which the disagreement realization lives. It also disciplines interpretation throughout the paper: every result on whether and when a home sells survives, indeed strengthens, on a measure containing no seller choice and no realized price, while the attention results are read as properties of realized dispersion. No single ingredient carries the index: rebuilt from transaction-comparable features alone, or from configuration rarity alone, it delivers about ninety percent of the execution gradients, and the two blocks are complementary in the variance they span (Online Appendix OA.G.5).

Table 20: Common-Scale Contrasts: Rank 10–90 Moves and Within-Cell Standard Deviations

Outcome	Measure, contrast	Effect	(SE)	N
sold	Realized rank (frozen), rank 10-90	-0.0131^{***}	(0.0016)	468,284
sold	Ex-ante rank (frozen), rank 10-90	-0.0250^{***}	(0.0018)	468,284
sold	price_unc, within-cell SD, 1 within-cell SD	-0.0068^{***}	(0.0008)	468,191
sold	index, within-cell SD, 1 within-cell SD	-0.0136^{***}	(0.0009)	463,682
withdrawn	Realized rank (frozen), rank 10-90	$+0.0124^{***}$	(0.0016)	468,284
withdrawn	Ex-ante rank (frozen), rank 10-90	$+0.0243^{***}$	(0.0018)	468,284
withdrawn	price_unc, within-cell SD, 1 within-cell SD	$+0.0064^{***}$	(0.0008)	468,191
withdrawn	index, within-cell SD, 1 within-cell SD	$+0.0132^{***}$	(0.0009)	463,682
log DOM	Realized rank (frozen), rank 10-90	$+0.0596^{***}$	(0.0068)	249,722
log DOM	Ex-ante rank (frozen), rank 10-90	$+0.0989^{***}$	(0.0074)	249,722
log DOM	price_unc, within-cell SD, 1 within-cell SD	$+0.0336^{***}$	(0.0033)	249,684
log DOM	index, within-cell SD, 1 within-cell SD	$+0.0551^{***}$	(0.0035)	248,136

Notes. Because a full within-cell rank move is not comparable across two differently distributed measures, each pair of rows restates the realized-uncertainty and predetermined-index gradients on a common scale: the effect of a 10th-to-90th percentile move in the frozen rank, and the effect of one within-cell standard deviation of the underlying (unranked) score. All regressions are on the shared full sample with interacted community×month, floor, and construction-type fixed effects; standard errors clustered by community. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

A second cut separates valuation disagreement from generic disamenity. A home that is simply undesirable (a bottom or basement floor, a crowded elevator bank) should be slow to sell, but it should not draw committed attention; valuation disagreement should slow the sale and draw attention, because buyers monitor an asset whose value they cannot pin down. Table OA.35 bears this out. A generic-disamenity indicator and the configuration-rarity rank both lengthen time on market and lower the sale probability, but their attention loadings have opposite signs: rarity raises followers ($+0.061$, $t = 6.2$), disamenity lowers them (-0.026, $t = -3.1$). The contrast sharpens against a genuine institutional friction: resale-restricted and formerly-public units (identified from the ownership-type field, not from price) are 12 to 16 percentage points less likely to sell, yet they too draw fewer followers (-0.05, $t = -3.0$ and -6.4), the wrong sign for an attention-drawing object. Committed attention is the fingerprint that separates this paper's mechanism from the broad class of homes that are merely costly to transact: only valuation disagreement makes a home both harder to sell and more closely watched. These cuts are associational and the disamenity proxies are coarse, so I read the sign split, not its magnitudes, as the load-bearing fact.

5.4.3 Time on market and sale probability

The time-on-market result of Section 5.1.1 also holds with uncertainty measured in levels: conditional on community, year–month, floor, and construction-type fixed effects, a one-unit increase in valuation uncertainty raises time on market by about 3.9 days (s.e. 1.4). The sale-probability margin, by contrast, is sign-sensitive to the heavy upper tail of $\hat{u}^2$. Measured in levels the coefficient is positive (higher-uncertainty listings are more likely to transact eventually, $+0.038$, s.e. 0.006), whereas under the tail-immune within-community–month rank used throughout the paper it is negative, and strongly so for clearing within a fixed window. This reversal is itself the sharpest justification for the rank measure: the level estimate is driven by a handful of extreme-residual listings, while the rank, which compresses that tail, delivers the economically coherent pattern in which harder-to-value units both sell more slowly and are less likely to clear within a given window. Online Appendix OA.E.7 reports the levels table and contrasts it with the rank specification.

A related concern is duration stigma: buyers rationally read a long spell as a bad-quality signal (Taylor, 1999), so harder-to-value units could close lower simply because they sit longer. The uncertainty measure is fixed at listing, before any spell accumulates, and the discount survives conditioning on realized days-on-market (Online Appendix OA.F.11), so the illiquidity is not a relabelling of stale-listing dynamics. A seller-side behavioral alternative fails on the price sign: loss-averse, reference-dependent sellers fish with high asks, producing longer spells that close higher conditional on sale (Genesove and Mayer, 2001; Andersen et al., 2022); hard-to-value homes wait longer and close lower. Loss aversion is nonetheless present in these data (it surfaces sharply in the quasi-repeat-sale pairs, where sellers facing a nominal loss pass only about a fifth of the market's decline into the ask; Online Appendix OA.A.8), but it operates on the seller's gain/loss position, an axis orthogonal to how hard the home is to value, and Proposition 4 shows the illiquidity gradient needs no anchor at all.

The platform also records the number of price revisions per listing. The sign of the uncertainty–revision relation is itself measure-dependent (negative in levels, positive under the within-cell rank used throughout the paper), another instance of the heavy upper tail driving the levels estimate (Online Appendix OA.E.1). What matters for identification is that the headline uncertainty result is unchanged when revisions are added as a control (Online Appendix OA.E.1): the illiquidity of hard-to-value homes does not reflect a hidden seller-revision channel.

5.5 The cost of illiquidity

How costly is the friction? Rather than estimate a structural welfare function, I use the reduced-form slopes as sufficient statistics for the cost a seller bears, in the spirit of Chetty

(2009). This number is the empirical counterpart of the model's central claim: it is what the idiosyncratic second moment costs when it is borne as an illiquidity quantity, a slower and lower-priced sale, rather than priced as a return premium. The owner of a hard-to-value home pays for illiquidity in two ways: a flow carrying cost c (the opportunity cost of equity locked in an unsold home, plus maintenance and fees) over a longer expected time to sale, and a lower sale price. With T the expected time to sale and δ_p the log price discount, the private cost relative to a benchmark unit is

$$W = \underbrace{c\,\Delta T}_{\text{carrying cost}} + \underbrace{|\delta_p|}_{\text{discount}}, \tag{4}$$

where ΔT, the uncertainty-induced increase in expected time to sale, is recovered from the sale-probability and time-on-market slopes, and δ_p is the discount of Section 5.2.1. Because the discount is concentrated in the tail rather than linear in the rank (Section 5.1.3), the cost scales with how hard the home is to value. For the median-to-ninetieth-percentile move, time to sale lengthens about two percent and the price falls a few tenths of a percent, so at a carrying cost of $c = 5$ percent per year the private cost is near 0.35 percent of value—about $17{,}000$ yuan for a median Beijing apartment. For a top-decile hard-to-value home the trim-robust transaction discount alone is 0.73 percent (0.22 percent even after removing the extreme one percent of residuals, both significant), so its private cost runs from roughly 0.3 to 0.8 percent of value per sale, larger again in the thin markets where the slopes steepen, and borne anew at each resale. Two benchmarks give the number scale. Per transaction, it is roughly one-eighth of the standard agent commission. In the aggregate, applying the gradient across the distribution of within-cell uncertainty to the platform's $59{,}000$ annual sales values the burden at about one billion yuan per year on this platform alone, relative to a benchmark in which every home were as easy to value as its cell's easiest: a back-of-envelope counterfactual, much of it transfer rather than deadweight, but borne sale after sale by whoever owns the hard-to-value stock. The discount is in part a transfer to buyers, so the deadweight component is the carrying cost together with the foregone gains from trade on units that never clear; both rise in the thin markets where the slopes steepen (Section 5.2.5). Small for the typical move but reaching most of a percent for the hardest-to-value homes and the thinnest markets, the cost is borne anew by every owner who must eventually sell: the quantity counterpart to the return premium that the same measure does not command. The form of the answer has an asset-pricing pedigree: when trade is costly, agents absorb the friction on the quantity margin, trading less and waiting longer, leaving only second-order price effects (Constantinides, 1986), and the option value lost to delayed marketability is increasing in the asset's volatility (Longstaff, 1995), the same comparative static the cross-section displays here. Where deep markets price illiquidity into average returns (Amihud and Mendelson, 1986; Amihud, 2002), an owner-occupied

Table 21: A Valuation-Transparency Counterfactual (Partial Equilibrium)

	Status quo	Halve difficulty	Full transparency
Time to sale (days)	106	104	102
Withdraw unsold (%)	25.5	25.0	24.5
Per-sale discount (% of value)	-0.76	-0.38	0.00

Notes. Model-implied partial-equilibrium comparative statics on the estimated extension (Online Appendix OA.A.9), for the hardest-to-value units. A disclosure reform shrinks the estimated valuation-difficulty increment $(\Delta g, \Delta s)$: "halve" resolves half of it, "full transparency" all of it, returning the unit to its cell's easiest. The buyer meeting rate λ is held fixed—there is no market-clearing or entry response, so these are partial-equilibrium magnitudes, not general-equilibrium welfare incidence. The per-sale discount is a transfer between seller and buyer; the efficiency gain is the recovered deadweight (foregone gains from trade on units that would withdraw unsold) plus the carrying cost saved from the faster sale, per listing.

housing market with no arbitrage capital leaves it where this section measures it: in the speed, certainty, and price of the individual sale.

The estimated model lets me ask what that friction's natural policy lever (comparable-information transparency, a public automated valuation or standardized disclosure that makes the hardest units easier to value) would buy. Holding the meeting rate fixed (a partial-equilibrium comparative static; the lean model has no market-clearing or entry block, so I claim no general-equilibrium incidence), I shrink the estimated difficulty increment $(\Delta g, \Delta s)$ of the hardest-to-value units and re-solve their acceptance, withdrawal rate, and price on the extension of Online Appendix OA.A.9. Table 21 reports the result: resolving half the excess difficulty halves the per-sale discount (from 0.76 to 0.38 percent of value), shortens the sale by about two days, and trims withdrawals; full resolution removes the discount. The decomposition behind those numbers is the point. The discount is overwhelmingly a transfer from seller to buyer, so even full transparency recovers little surplus: the genuine efficiency loss (gains from trade lost on homes that withdraw unsold, plus the carrying cost of the longer wait) is a small fraction of the discount it accompanies. That gap is not a weak counterfactual but the paper's claim made quantitative, and it is exactly what the canonical analysis of trading frictions predicts: a first-order friction absorbed on the quantity margin, trading less and waiting longer, leaves only second-order welfare effects (Constantinides, 1986). The second moment is borne (redistributed at the point of sale, from the seller of a hard-to-value home to its buyer), not burned as a loss to the economy. Who bears it is therefore the first-order question (Section 5.1.4); the case for a transparency intervention is distributional, not Pigouvian. The comparative static inherits the estimate's leave-one-moment-out stability (Online Appendix OA.A.9): the discount response is mechanical in Δg, and the quantity responses do not lean on the time-on-market–showings pair on which the over-identified fit is rejected.

5.6 Summary

The results establish five facts, each estimated within community and month and each robust to trimming the heavy upper tail of the uncertainty distribution.

1. Harder-to-value units are less liquid: a move from the least to the most difficult unit in a community and month lowers the probability of sale by about 1.3 percentage points on the realized measure (2.4 on the predetermined index) and lengthens time on market by six to nine percent (Proposition OA.6).
2. They must draw more committed buyer attention to clear: more saved-listing followers.
3. They trade at a discount, in transaction prices and in the listing-implied pricing wedge (Propositions OA.3 and OA.4).
4. The discount and the attention response are fingerprints of two distinct channels: uncertainty lowers the mean of buyer valuations (a risk discount) and widens their dispersion (raising search and attention). The realized price falls while attention rises (a pattern no single channel produces), so both margins are active (Section 5.2.1).
5. The illiquidity is stronger when trading is thin: the uncertainty–time-on-market slope is steeper in low-turnover months, a turnover-specific state-dependence I read as suggestive (Section 5.2.5).

A three-parameter calibration of the framework matches the attention, price, and time-on-market gradients and is then checked against moments it never saw: the funnel's opposite-signed split at the costly stage, the seller's within-spell concession path, the absence of recoupment at resale, the gain-domain anchoring weight in quasi-repeat relistings, of the same order as the calibrated $1 - \nu$ (Online Appendix OA.A.8), and the two measured congestion moments (Online Appendix OA.A.7). An estimated extension pushes further: adding a withdrawal margin and a costly inspection gate, three parameters fit five liquidity-and-funnel gradients to within two standard errors each, over-identified by two, and pass two untargeted checks: the share of listings that ever sell, and the accumulated-showings gradient (Online Appendix OA.A.9).

Section 7 documents the robustness of these findings to alternative samples, uncertainty proxies, selection into transaction, and replication across provinces.

6 The Model

This section develops in full the framework whose logic and predictions Section 3 stated and whose predictions the results have now tested: the pricing, acceptance, and attention machinery behind the lever of Figure 1, and the mapping from the estimated coefficients back to its parameters.

6.1 Setup

Time is indexed by t at the monthly frequency. Each listing corresponds to a house h in community $c(h)$ in month t. Trade takes place in three periods. In period 0, a seller with reservation value $P^S_{h,t}$ meets a buyer and they bargain over the transaction price. In period 1, the buyer receives a stochastic housing-services flow $r_{h,1}$. In period 2, the buyer resells the house at a stochastic price $P_{h,2}$. The buyer's total payoff from buying at price P_0 is

$$W_h = -P_0 + \rho\, r_{h,1} + \rho^2 P_{h,2}, \tag{5}$$

where $\rho \in (0, 1)$ is the discount factor. Period-1 and period-2 payoffs are normally distributed and independent,

$$r_{h,1} \sim \mathcal{N}(\mu_r,\, \sigma_r^2), \qquad P_{h,2} \sim \mathcal{N}(\mu_P,\, \sigma_P^2), \tag{6}$$

where σ_P^2 is the idiosyncratic valuation uncertainty measured empirically in Section 2.4.2. Normality is for tractability and delivers exact certainty equivalents under CARA utility; the qualitative results extend to any location–scale family, and independence between rent and resale is standard (Amaral, 2024) and can be relaxed at the cost of additional notation.

Two types of buyers populate the market, indexed by $\tau \in \{E, S\}$. End-users (E) are risk-averse households who value housing services and are sensitive to idiosyncratic price risk, with absolute risk aversion $\gamma_E > 0$. Speculative or short-horizon participants (S) have lower risk aversion γ_S, with $0 \leq \gamma_S < \gamma_E$. Both types have CARA utility over terminal wealth, $u_\tau(W) = -\exp(-\gamma_\tau W)$, so the certainty equivalent of $X \sim \mathcal{N}(\mu, \sigma^2)$ is exact,

$$\mathrm{CE}_\tau(X) = \mu - \frac{\gamma_\tau}{2}\sigma^2. \tag{7}$$

This is the key tractability advantage relative to the CRRA–lognormal specification in Amaral (2024), which yields closed forms only under log utility. Comparative statics in γ are transparent for any value of the risk-aversion parameter, which is essential for comparing valuations across the two buyer types. The type structure earns its place in the empirics: the discount slope in Equation (11) scales with the average risk aversion of the pool a listing meets, while the disagreement that drives dispersion is a property of the unit, so a buyer pool tilted toward low-γ short-horizon participants flattens the pricing and liquidity gradients but not the attention gradient, the composition test of the tax-favored small segment (Online Appendix OA.F).[12] The buyer's private valuation of house h discounts each period's payoff at its certainty equivalent,

$$PV^B_\tau = \rho\Big[\mu_r - \tfrac{\gamma_\tau}{2}\sigma^2_r\Big] + \rho^2\Big[\mu_P - \tfrac{\gamma_\tau}{2}\sigma^2_P\Big]. \tag{8}$$

Valuing each period's payoff at its own certainty equivalent and then discounting is a convention: applying CARA to the terminal payoff W_h instead would discount the two risk penalties by ρ^2 and ρ^4 rather than ρ and ρ^2, rescaling the structural slope but changing no sign or comparative static. Conditional on meeting, the buyer and seller engage in Nash bargaining with buyer bargaining weight $\alpha \in (0,1)$.[13] The standard solution yields a transaction price equal to a weighted average of the buyer's private valuation and the seller's reservation value,

$$P^*(\tau) = (1-\alpha)\, PV^B_\tau + \alpha\, P^S_{h,t}. \tag{9}$$

Buyers are risk-averse and, for a hard-to-value home, they disagree about its worth. Two forces therefore shape the valuations a listing meets. First, every buyer discounts the resale risk σ^2_P at the average rate of risk aversion $\bar{\gamma}$; risk aversion is homogeneous within the pool a given listing meets, with the two types differing across segments, which is what the composition test of Online Appendix OA.F uses. Second, each buyer i holds a private belief b_i about the resale value, with common mean μ_P but a dispersion $\sigma^2_b(\sigma^2_P)$ across buyers that is larger the harder the unit is to value ($\sigma^{2\prime}_b > 0$): divergence of opinion in the sense of Miller (1977), and the direct counterpart of the within-community disagreement this paper measures. Substituting into Equation (8), a meeting draws a private valuation whose first two moments across buyers are

$$\mu_V(\sigma^2_P) = \bar{a} - \frac{\rho^2}{2}\,\bar{\gamma}\,\sigma^2_P, \qquad \mathrm{Var}_V(\sigma^2_P) = \rho^4\,\sigma^2_b(\sigma^2_P), \tag{10}$$

where $\bar{a}$ collects the uncertainty-independent terms. A single primitive, the idiosyncratic valuation uncertainty σ^2_P, thus moves two moments of the offer distribution through two distinct forces: risk aversion lowers the mean valuation (the risk discount, through $\bar{\gamma}$), while disagreement widens its dispersion (divergence of opinion, through σ^2_b). These are the model's two channels, a risk discount and disagreement, and the rest of the section traces

[12] CARA implies that the risk penalty is independent of the apartment's value, so the model's RMB-denominated magnitudes should be interpreted with caution. The qualitative predictions, including the uncertainty discount, do not depend on CARA.

[13] Complete-information Nash bargaining is a deliberate simplification: it locates all of the disagreement across meetings (which buyer the seller draws) rather than within them. Two-sided private information would add a separate within-meeting bargaining-breakdown margin whose sign under a mean-preserving spread is ambiguous in general (Myerson and Satterthwaite, 1983; Chatterjee and Samuelson, 1983); the model does not rely on it. The dispersion-delays-trade result runs entirely through the seller's across-meeting acceptance margin (Proposition 4) and needs no within-meeting breakdown.

each to a distinct observable.

The risk-averse, two-moment object is what sets this framework apart from the housing-search literature, which allows heterogeneous valuations but takes buyers to be risk-neutral (Sagi, 2021; Wheaton, 1990; Allen, Clark and Houde, 2019), so uncertainty there cannot shade the mean of valuations as it does here. The disagreement channel inverts Miller's logic in a revealing way: where short sales are impossible, divergence of opinion would, in a Walrasian market, let the most optimistic buyer set a high price, but because the binding friction here is finding that buyer through costly search, the premium never materializes: the uncertainty that widens disagreement also carries a risk discount that lowers the average price, while the disagreement itself surfaces as delay and dispersion rather than as a higher price. The flip is driven by the extensive margin of matching, not by per-trade transaction costs: Scheinkman and Xiong (2003) show that higher trading costs do not overturn the disagreement-raises-price result, whereas here it is the finite probability of meeting the optimist, the counterparty-search friction that prices over-the-counter assets (Duffie, Gârleanu and Pedersen, 2005), that reroutes divergence of opinion from a price premium into illiquidity. Nor do search frictions alone suffice: when optimists are a market-wide pool, they self-select into the trading pool and still set high prices (Piazzesi and Schneider, 2009). The flip requires disagreement that is unit-level, so that the optimist for a particular home must be found listing by listing.

A calibration disciplines the claim quantitatively and an asymptotic result makes it sharp: under an optimal reservation the seller's selectivity grows only like $\sqrt{2\ln\lambda}$ in the meeting rate λ (optimists are exponentially scarce in a normal tail), so the price gradient turns positive only beyond $\ln\lambda^* \approx \frac{1}{2}(\Delta g/\Delta s)^2$, the squared ratio of the risk-discount increment to the dispersion increment. At the calibrated increments that is some thirty orders of magnitude above Beijing's meeting rate; no rate within five orders of magnitude comes close (Proposition OA.2; Online Appendix OA.A.7).[14]

Taking expectations over which buyer the seller meets, the expected transaction price is affine and decreasing in idiosyncratic valuation uncertainty,

$$\mathbb{E}\big[P^*\big] = a_{h,t} - (1-\alpha)\,\frac{\rho^2}{2}\,\bar{\gamma}\,\sigma_P^2, \tag{11}$$

where $a_{h,t}$ collects the terms that do not vary with σ_P^2 (the community–month valuation components and the seller-reservation term $\alpha P^S_{h,t}$, absorbed in estimation by the fixed effects and controls), and $\bar{\gamma}$ is the average risk aversion of the buyers a listing meets. Equation (11) is the uncertainty discount: harder-to-value units transact at lower prices, the risk-discount

[14]The threshold uses the calibrated increments; the estimated extension recovers a smaller dispersion increment Δs (Online Appendix OA.A.9), which only raises λ^*, so the thirty-orders-of-magnitude gap is a conservative bound. A mapping convention separates the two: the extension takes log followers to move one-for-one with the log valuation spread s, while the calibration maps attention to Var_V, an elasticity of two; under the calibration's convention the extension's Δs would double, which again only tightens the bound.

channel acting on the first moment of price.

The dispersion channel acts on the second moment. Because realized sales occur only when a buyer's valuation clears the seller's reservation, the variance of transaction prices inherits the widening of the valuation distribution in Equation (10):[15]

$$\mathrm{Var}\big[P^* \mid \text{sale}\big] = (1-\alpha)^2\ \mathrm{Var}\big[V \mid V \geq P^S\big], \qquad \frac{\partial}{\partial \sigma_P^2}\mathrm{Var}\big[P^* \mid \text{sale}\big] > 0 : \tag{12}$$

harder-to-value units transact at a wider spread of prices, even within a community and month. The inequality is exact for the dispersion channel; when the discount channel also operates, it tightens the accepting tail, and the spread still widens as long as the disagreement response is not too weak: a condition stated formally, with the truncated-normal facts behind it, in Online Appendix OA.B.

The same risk penalty governs liquidity. A meeting converts to a sale only if the buyer's private valuation PV^B clears the seller's reservation value $P^S_{h,t}$. Sellers anchor their reservation on recent comparable prices and adjust it slowly (Genesove and Mayer, 2001), so $P^S_{h,t}$ does not fall one-for-one with the buyer's risk discount. The per-meeting acceptance probability is therefore[16]

$$q(\sigma_P^2) = \Pr\big[PV^B \geq P^S_{h,t}\big], \qquad \frac{\partial q}{\partial \sigma_P^2} < 0 : \tag{13}$$

the risk discount shifts the buyer's valuation down relative to the anchored reservation, so fewer meetings clear. The sign is exact when valuations shift down at unchanged dispersion; when disagreement also widens the valuation distribution, extra mass reaches the accepting tail, and the acceptance rate still falls so long as the discount moves the threshold by more than the tail widens (automatic when at least half of meetings clear, as at the calibration, and a mild restriction otherwise); in either case the seller's partially optimal reservation, which absorbs part of the discount and rises with the option value dispersion creates, works in the same direction (Online Appendix OA.A). With buyer arrivals at rate λ, expected time on market is $1/(\lambda q)$ and the probability of sale within any fixed listing window rises with q, so harder-to-value units sell more slowly, sell less often, and must attract more meetings, more search and attention, to clear. This is the paper's central prediction.[17]

[15]The displayed sign is exact for the dispersion channel acting alone; with both channels active it holds under the dominance condition of Equation (OA.9), stated formally in Online Appendix OA.B.

[16]The displayed sign is exact when valuations shift down at unchanged dispersion; with both channels active it holds under the dominance condition of Equation (OA.3): automatic when at least half of meetings clear (the calibration), a mild restriction otherwise.

[17]The meeting rate λ is held independent of σ_P^2, so the prediction works through the acceptance margin q. This is testable: page views proxy arrivals into the browse stage, and their per-day gradient is flat-to-weakly-negative, never positive, in uncertainty, ruling out a thin-market arrival story, while the conversion margin moves (Proposition OA.1; the duration-mechanics diagnostic, Online Appendix OA.F.11); the contact and showing decline is the endogenous visit gate inside q (Equation (18), Proposition OA.8), not the arrival

6.2 Where the channels come from

The setup takes two objects as given: buyers perceive resale risk $\kappa\sigma_P^2$, and their beliefs disperse with valuation difficulty, $\sigma_b^{2\prime} > 0$. Neither is free. Both follow from one Gaussian information primitive: how precisely public comparables pin down a unit's idiosyncratic value (the signal-extraction problem of Quan and Quigley, 1991). Let the public signal about the unit carry precision τ_{pub} and each buyer's private assessment carry precision τ_η. Measured uncertainty is the residual variance given public information alone, $u \equiv 1/\tau_{\text{pub}}$, the population counterpart of the squared hedonic residual. Bayesian updating then delivers both channels at once (Proposition 3, Online Appendix OA.B.4):

$$\underbrace{\widetilde{\sigma}^2 \;=\; \frac{u}{1+\tau_\eta u} \;=\; \kappa(u)\, u}_{\text{perceived risk: the discount channel}} \qquad\qquad \underbrace{\sigma_b^2 \;=\; \tau_\eta\, \widetilde{\sigma}^4}_{\text{belief dispersion: the disagreement channel}} \qquad , \tag{14}$$

both strictly increasing in u. The information-channel scaling $\kappa(u) = 1/(1+\tau_\eta u) \in (0,1)$ is no longer a parameter: it is the share of measured dispersion that private inspection cannot resolve. And disagreement is the Bayesian echo of weak public information: where comparables say little, each buyer's posterior leans on her own signal, which is exactly where opinions differ. One primitive, two moments: risk aversion prices the first, search and bargaining transmit the second, and no preference heterogeneity is needed to generate either. The comparative statics survive endogenous attention (buyers choosing τ_η at convex cost) provided acquired precision responds weakly enough to difficulty: an explicit elasticity bound (binding in the right tail, not at the mean) derived in Online Appendix OA.B.4; the tested gradients are conditional-mean objects in the body of the distribution, where the modest spell-adjusted attention response respects it.

Proposition 3 (One information primitive, two channels)**.** *Bayesian updating delivers, for every buyer i:* (i) *perceived residual risk $\widetilde{\sigma}^2 = 1/(\tau_{\text{pub}} + \tau_\eta) = \kappa(u)\, u$ with $\kappa(u) = 1/(1+\tau_\eta u) \in (0,1)$, strictly increasing in u: the risk the discount channel prices is the share of measured dispersion that private inspection cannot resolve;* (ii) *cross-buyer belief dispersion $\sigma_b^2 = \tau_\eta/(\tau_{\text{pub}}+\tau_\eta)^2 = \tau_\eta\, \widetilde{\sigma}^4$, strictly increasing in u: disagreement is a result, not an assumption, because the posterior weight on the private signal is largest exactly where the public signal is least precise;* (iii) *with endogenously chosen precision (strictly convex cost, supermodular value), acquired attention is nondecreasing in u, delivering Equation (16), and (i)–(ii) survive provided the attention elasticity satisfies $\epsilon < 1/(\tau_\eta u)$.*

The proof is six lines of Gaussian projection algebra (Online Appendix OA.B.4, which also derives the elasticity bounds).

rate. Endogenizing arrivals cuts both ways (price-directed search would raise λ at a discounted listing, dispersion-directed search need not), so I do not sign the bias from λ. The liquidity result does not require it: by Proposition 4(b) a mean-preserving spread strictly lowers q with no anchoring and no risk discount.

The seller side needs no behavioral anchor for the dispersion channel. Suppose the reservation is not anchored at all but solves the optimal stopping problem: flow carrying cost c, discount rate r, offers $\mathcal{N}(\mu, s^2)$. Wider dispersion raises the option value of waiting, and the optimal reservation climbs by more than the accepting tail thickens whenever

$$c + r\mu \; > \; 0, \tag{15}$$

so the per-meeting acceptance rate strictly falls and time on market strictly rises with disagreement: for an asset with positive value and any carrying cost, always (Proposition 4, Online Appendix OA.A). The companion result disciplines what anchoring is for: an optimal seller passes a level shock like the risk discount into her reservation at rate $\lambda(1-\alpha)(1-\Phi)/[r+\lambda(1-\alpha)(1-\Phi)]$, which approaches one as $r \to 0$: into price, hardly into quantity. Anchoring therefore scales how much of the discount shows up in sale probability and time on market (the calibrated $\nu = 0.78$; a fully optimal seller delivers under a third of the measured time-on-market gradient, four-fifths of it through the dispersion-holdout channel); anchoring scales the magnitudes without supplying the mechanism. The division of labor is clean: the discount needs anchoring to slow trade, disagreement slows it on its own. The anchoring the framework does use is the model's reference-dependence channel, and it is disciplined by evidence rather than assumed: reference-dependent seller behavior is among the best-documented regularities in housing (Genesove and Mayer, 2001, 1994; Andersen et al., 2022), and the anchor weight this sample's own relisting behavior implies for gain-domain sellers (≈ 0.20, Online Appendix OA.A.8) matches the calibrated $1 - \nu = 0.22$; matching the estimated gradients requires $\nu = 0.78$, a seller who adjusts three-quarters of the way toward the optimal reservation (Online Appendix OA.A.7), and the within-spell concession path lands at the same partial adjustment (Online Appendix OA.F.12).

Proposition 4 (Optimal holdout: disagreement delays trade without anchoring)**.** *Let S^* solve the stopping problem $rS^* + c = \lambda(1-\alpha)\,\mathbb{E}[(V - S^*)^+]$ with offers $V \sim \mathcal{N}(\mu, s^2)$, and let q be the per-meeting acceptance probability. Then:* (a) *the reservation rises with dispersion, $\partial S^*/\partial s > 0$;* (b) *if $c + r\mu > 0$, a mean-preserving spread strictly lowers the acceptance rate and strictly lengthens expected time on market, with no anchoring and no risk discount anywhere in the argument;* (c) *a level shock passes into the reservation at rate $\lambda(1-\alpha)(1-\Phi)/[\,r + \lambda(1-\alpha)(1-\Phi)\,] \in (0,1)$, which approaches one as $r \to 0$: an optimal seller absorbs the risk discount almost fully into price, hardly into quantity.*

The proof is three lines: the inequality in (b) collapses through the stopping condition itself to $c + r\mu > 0$ (Online Appendix OA.A.3).

The same acceptance rate fixes where the lost sale goes. Listings also end for reasons unrelated to the sale itself (plans change, the household stays), and the estimated extension carries this as an exit hazard δ_w that does not vary with valuation difficulty (Online Appendix OA.A.9).

Corollary 1 (Abandonment is failed clearing, not accelerated exit)**.** *Let sale arrive at hazard $\lambda q(u)$ and exit unsold at hazard $\delta_w > 0$, with δ_w independent of u. Then* (a) *the probability a listing ever sells is $\lambda q/(\lambda q + \delta_w)$, increasing in q, so it falls with valuation difficulty exactly as q does, and the entire shortfall accrues to permanent withdrawal, the complementary terminal state;* (b) *for competing exponential risks the terminal time is distributed $\mathrm{Exp}(\lambda q + \delta_w)$ independently of which risk realizes, so conditional on ending unsold, a harder-to-value listing goes dark* later*, not sooner. Abandonment concentrates among hard-to-value homes because the sale never arrives, not because their owners pull them faster.*

Both implications are borne out, and neither was targeted: the withdrawal share of the lost sale and its accounting identity are Table 7, and conditional on not selling, harder-to-value listings delist later ($t = 14.7$, Section 5.1.2); the difficulty-invariance of δ_w is a tested restriction, not a convenience. At the estimated values the implied probability of ever selling is 0.755 against 0.775 in the data (Online Appendix OA.A.9).

A third primitive completes the identification logic: a unit can also be slow because its segment is thin: buyers rarely arrive. Jiang, Kotova and Zhang (2024*b*) show at the market level that duration alone cannot identify what moved; it is the joint behavior of prices, durations, and dispersion that does. The stopping model contains thinness as the meeting rate λ, and the three primitives leave mutually exclusive signatures (Proposition OA.1, Online Appendix OA.A.4): a thin market is slow because meetings are scarce while the optimal seller settles (per-meeting acceptance rises), whereas a hard-to-value unit is slow because meetings fail, and only disagreement moves per-exposure attention. The data sit on the uncertainty rows: the per-day arrival proxy is flat, conversion falls, and attention rises, a slow sale of the failing-meetings kind, not the scarce-meetings kind.

6.3 Attention: monitoring and costly inspection

In the model, buyers resolve their own valuation of a hard-to-value unit by paying attention to it: monitoring the listing, saving it, following its price. How much such evaluation a unit elicits scales with how much buyers disagree about it: when valuations are tightly clustered there is little to learn, but when they are dispersed, locating where one's own value falls is worth more attention. Let per-exposure committed attention A, the committed evaluation a unit of buyer exposure generates, increase in the dispersion of valuations,

$$A(\sigma_P^2) = c_A \ \mathrm{Var}_V(\sigma_P^2), \qquad \frac{\partial A}{\partial \sigma_P^2} > 0, \qquad c_A > 0, \tag{16}$$

so that per-exposure committed attention is a fingerprint of the dispersion channel, not the discount: a uniform downward shift in valuations gives buyers nothing more to investigate, whereas a wider spread does. The risk discount, by contrast, depresses liquidity by lowering the per-meeting acceptance rate q rather than by changing how much buyers investigate.

Two attention objects must therefore be kept apart. What the platform records per listing is a stock, the followers a listing accumulates before it clears, and the stock compounds the evaluation each unit of exposure draws with how long the listing is exposed:

$$\mathcal{A}(\sigma_P^2) = \frac{A(\sigma_P^2)}{q(\sigma_P^2)}, \qquad \frac{\partial \mathcal{A}}{\partial \sigma_P^2} > 0 \text{ through both channels:} \tag{17}$$

dispersion raises the committed evaluation per meeting (A), and the risk discount, by cutting the acceptance rate q, stretches the spell over which attention accrues, so the follower stock rises even for a unit nobody disagrees about. The stock is thus a liquidity outcome, the search a listing must accumulate to clear; the fingerprint specific to disagreement is the spell-adjusted component, the attention a listing draws per unit of exposure, and the empirical work reads the two accordingly (Section 5.2.1; Online Appendix OA.F.11). The same dispersion also raises the option value of waiting, so a seller facing more disagreement optimally holds out longer for a high draw (a theorem under an optimal reservation, not a tendency; Proposition 4), reinforcing the longer time on market.

Equation (16) is the reduced form of an information-acquisition problem. A buyer unsure where her own valuation of a hard-to-value unit falls can pay to evaluate it remotely (monitoring, re-reading, saving the listing) and optimally acquires more information the more dispersed her prior, because the value of resolving uncertainty grows with its scale (Gabaix, 2019). Under a standard information cost, optimal acquired precision, and hence monitoring effort, increases in the prior variance Var_V, which delivers Equation (16); Online Appendix OA.B.4 derives this jointly with the disagreement response from the same Gaussian structure. Per-exposure committed attention is thus the observable footprint of disagreement: the effort buyers expend to resolve it.

A costly physical inspection is a different object. A showing requires the buyer's expected gain from visiting to clear a real cost, so it inherits the threshold logic of the acceptance rate q rather than the information-demand logic of Equation (16): the risk discount shades the mean of valuations, fewer buyers clear the visiting threshold, and showings per unit of exposure fall with uncertainty even as monitoring rises. The model therefore predicts a split in the engagement funnel: uncertainty raises what is cheap (following) and cuts what is costly (visiting, transacting), which the recovered showing counts confirm (column (5) of Table 10; Online Appendix OA.F.12). The acceptance margin accordingly factors into a visit stage and a close stage,

$$q \;=\; p_{\text{visit}} \;\times\; q_{\text{buy}|\text{visit}}, \tag{18}$$

and selection works the two factors in opposite directions: only buyers whose interim assessment clears the visit gate come at all, so a pure discount makes the surviving visits convert worse (more showings to clear), while disagreement selects the visiting pool from a wider upper tail and makes them convert better (fewer), so the sign of the showings-to-clear gradient signs the race between the channels (Proposition OA.8). The recovered counts fall; once the inspections accumulated over the longer spell are netted out, the implied per-visit conversion moves mildly against the harder-to-value home (Online Appendix OA.F.12). Online Appendix OA.B.5 states the funnel formally, as nested thresholds on the buyer's posterior whose stage-by-stage conversion rates are the structural margins, the listing-level counterpart of the stage congestion Badarinza, Balasubramaniam and Ramadorai (2024) estimate for the housing matching function.

6.4 Mapping regression coefficients to the model

The reduced-form coefficients have a direct structural reading. The cross-sectional slope of the pricing wedge on uncertainty identifies the risk discount $(1-\alpha)\rho^2\bar{\gamma}/2$ of Equation (11), up to the scaling between the squared hedonic residual and σ_P^2. The slopes of the liquidity outcomes on uncertainty identify the acceptance channel of Equation (13): the sale-probability and time-on-market coefficients measure how strongly the buyer's risk discount, set against the anchored seller reservation, lowers the per-meeting clearing rate $q(\sigma_P^2)$.

The two-channel evidence maps the same way. The slope of the transaction-price mean on uncertainty identifies the risk discount of Equation (11); the slope of the list-to-deal wedge spread identifies the dispersion channel of Equation (12) (the raw price variance, being mechanically tied to the measure, is not used empirically); and the slope of the follower stock identifies the accumulation margin of Equation (17), a composite of the two channels whose spell-adjusted component isolates the information-demand margin of Equation (16). The model delivers the signs of these slopes and the cross-equation restriction of Proposition 1, which Sections 5.1.1–5.2.1 confirm. Recovering the relative magnitude of the two channels would require structurally estimating the risk-aversion and belief-dispersion parameters $(\bar{\gamma}, \sigma_b^2)$ from these slopes; the cross-sectional design does not sharply identify that decomposition, so I test the model's qualitative predictions and leave its structural estimation to future work. The price slope itself identifies the discount only up to the bargaining weight and the composite risk-pricing index $(1-\alpha)\bar{\gamma}\kappa$: I read it as a risk discount under $\alpha \perp \sigma_P^2$ within community and month, and do not separate it from any covariance of effective bargaining power with valuation difficulty.

7 Robustness and External Validity

The illiquidity results in Section 5 are robust. I summarize the most informative checks here; full specifications are in Online Appendix OA.G.

Tail-robustness. Because the squared hedonic residual has a heavy upper tail (house-price dispersion measures generically develop power-law tails; Ohnishi, Mizuno and Watanabe, 2020), the central concern is that a handful of extreme listings drive the estimates. They do not. Column (2) of Table 6 re-estimates each outcome after dropping the top one percent of the uncertainty distribution and re-ranking within cell; the coefficients are essentially unchanged (probability of sale -0.0066, log days on market $+0.052$, log followers $+0.042$), and dropping the top five percent leaves the same pattern. This is the step that matters: a return-premium reading of the same measure does not survive it.

Right-censoring of recent listings. Because the data are a single snapshot, listings posted late in the window may not yet have had time to transact; they are recorded as unsold with truncated durations, which attenuates the sale and time-on-market estimates toward zero. Column (3) of Table 6 addresses this by restricting to the mature 2016–2019 posting cohorts, which had ample time to clear before the data end; every coefficient holds or strengthens (probability of sale -0.0081, log days on market $+0.0534$, log followers $+0.0428$). The bias is sharpest on the sale margin: replacing the raw sale indicator with a censoring-proof outcome (sold within 365 days of posting, defined identically for every cohort) more than doubles the estimated effect, from -0.0081 to -0.0168 ($t = -7.0$). Right-censoring therefore biases the baseline toward zero; correcting for it makes the illiquidity result stronger, not weaker.

Selection into transaction. The probability-of-sale outcome is the extensive margin itself (it is defined over all listings, sold and unsold), so it is not subject to a conditioning-on-sale bias. Follower attention is observed for every listing; days on market, by contrast, is observed only for sold units and so conditions on sale. Reassuringly, the extensive margin that any such selection would operate through, the probability of sale, is itself negative and robust, and restricting to the matched-transaction and follower-to-viewing-ratio (FVR at most one) samples leaves the sign and significance of all three outcomes intact (Online Appendix OA.G.1).

Measurement and clustering. The results hold under distribution-free (rank) and level measures of uncertainty, under alternative hedonic specifications, and under coarser and two-way clustering of the standard errors (Online Appendix OA.G.2). Restricting the hedonic to progressively denser communities, where the within-community comparable set is larger and listing-level variation is more disciplined, leaves the illiquidity relationship intact (Online Appendix OA.G.3). A measure of valuation difficulty containing no choice of the seller at all (the scarcity of comparable transactions by other units in the year before the listing) reproduces the sale-probability and time-on-market gradients ($+2.3$ percentage points and -9 percent from fewest to most comparables, surviving size-decile fixed effects) and is essentially orthogonal to the baseline uncertainty rank, so the two capture distinct components of valuation difficulty (Online Appendix OA.D.8).

Sample period, data recording, and policy environment. The window ends in the managed market of 2020–21, when local governments actively supported the primary market: countercyclical land purchases, sales-permit restrictions, and new-home price floors (Chang, Wang and Xiong, 2025). These are city-level, primary-market interventions: within community and month they are absorbed by the fixed effects, and they do not mechanically bind household-negotiated resale outcomes. Consistent with this, the mature 2016–2019 posting cohorts, which largely predate the managed period, deliver the same illiquidity gradients (Table 6, column 3), and the cold-market amplification survives excluding 2020–21 altogether (Table 16, column 3). At the other end of the window, eight percent of listings are a platform backlog posted before January 2015; dropping them leaves the gradients unchanged or slightly stronger (probability of sale -0.0061, log days on market $+0.062$, log followers $+0.040$). Two recording conventions matter for measurement and do not drive the results: sixteen percent of sales record the posting and deal on the same day (retroactively posted listings; excluding them leaves the time-on-market gradient at $+0.052$, $t = 7.4$), and the platform stores a price field for unsold listings that tracks the asking price, which is why deal-side objects are defined only on the matched-transaction subsample throughout.

External validity across provinces. I observe the full listing-to-transaction record only for Beijing; for four additional provinces (Guangdong, Zhejiang, Shandong, and Henan), spanning some thirty-six cities, the data are transaction-only, so the probability of sale cannot be formed, but time on market, follower attention, seller revisions, and the list-to-deal wedge can.[18] Because community codes are reused across cities, I make the comparison collision-proof: I define a community by its city $\times$ community pair, re-estimate the hedonic with those fixed effects, recompute $\hat{u}^2$, and rank within the collision-proof cell. Bare community-code cells give nearly identical estimates (within-province collisions are only two to four percent of communities), so these collisions do not drive the result.

Figure 12 plots the within-cell uncertainty gradients across the five markets (the full estimates are in Online Appendix Table OA.12). The mechanism, and the disagreement channel in particular, travels well. Committed attention rises with uncertainty in all four provinces; seller revisions rise in all four, more steeply than in Beijing; and the negotiated wedge widens in all four. Time on market lengthens in three of the four (Zhejiang, Shandong, and Henan), the exception being a flat Guangdong slope. The one margin I cannot test out of

[18]The policy regime traveled with the replication: the 2016–17 tightening was a national wave, with purchase restrictions adopted in more than twenty cities from September 2016 and purchase and resale restrictions strengthened across many cities in March–May 2017 (Zhu et al., 2023), so the province samples span the same cycle the Beijing design absorbs within cell.

sample is the probability of sale, which requires the unsold listings present only in the Beijing data. I therefore keep Beijing, where I observe every margin and community identifiers are unambiguous, as the primary setting, and read the province evidence as showing that the channel generalizes: the replication strengthens external validity, not identification, which remains within-cell and associational.

What is setting-specific, and what should travel. The quantity margins are not a China-specific finding: atypical homes take longer to sell in U.S. data (Haurin, 1988), and U.S. time on market moves with price dispersion at the market level (Jiang, Kotova and Zhang, 2024*b*). Where settings can genuinely differ is the no-compensation conclusion, and the difference is an institution, not a contradiction. The burden of valuation uncertainty has three outlets: the sale price, the speed and certainty of sale, and, where a thick rental market lets owners monetize a discounted purchase, the rental yield. In Beijing, rental yields sat below the one-year risk-free rate (Liu and Xiong, 2018) and the rental and sale markets are comparatively segmented, so the yield outlet is shut and the burden surfaces at the point of sale; in Germany, Amaral (2024) finds part of it surfacing as yield compensation. The cross-setting prediction is therefore sharp and testable: the thicker the rental market, the more of the uncertainty burden should migrate from the liquidity margins toward yields; U.S. listing platforms, which record the same engagement funnel, are the natural place to test it.

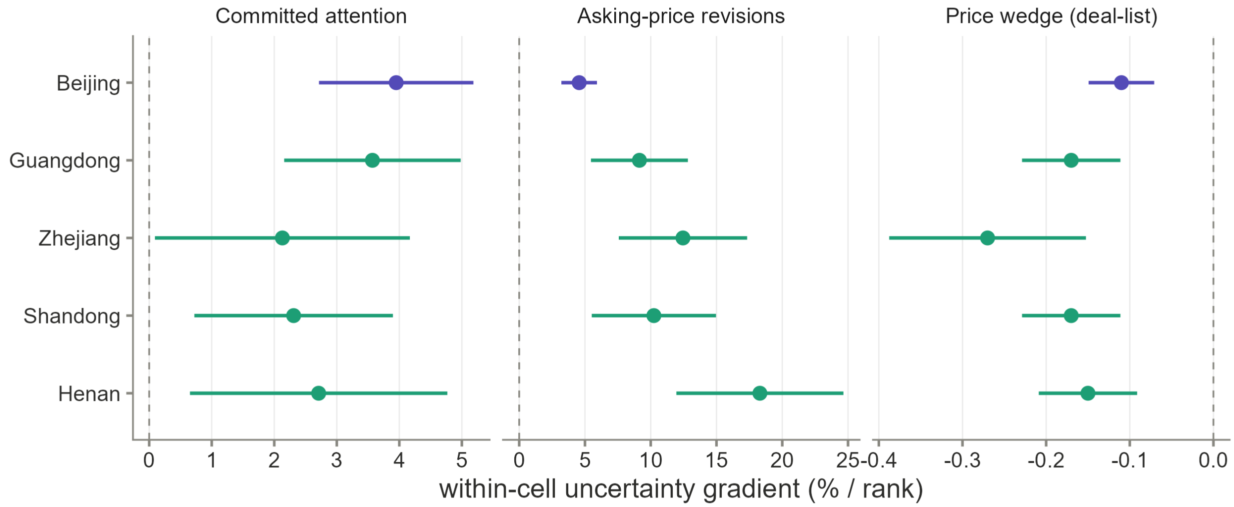


Figure 12: The Mechanism Replicates Across Markets

Notes. Within-(community × month) uncertainty gradients (point estimate and 95% confidence interval) for three channels in Beijing (purple) and four further provinces (teal): committed attention (log followers), asking-price revisions, and the list-to-deal price wedge. Each estimate is a separate regression with community, year–month, floor, and construction-type fixed effects and standard errors clustered by community; for each province a community is the city × community pair, the hedonic is re-estimated, and $\hat{u}^2$ is recomputed within the collision-proof cell. The full estimates, including the Beijing-only sale-probability margin, are in Online Appendix Table OA.12.

8 Conclusion

Where a home is hard to value, the price mechanism stops clearing the market, and the market's adjustment shifts from prices to quantities. Using listing and transaction records for roughly 586,000 Beijing apartments, this paper shows that idiosyncratic valuation uncertainty (the dispersion of apartment-specific pricing residuals within a community and month) governs how a home trades, and does so through the seller's own demand curve: in the cross-section, where comparables are noisy, the asking price barely moves the probability of sale, the dying lever of Figure 1. A seller who cannot discount the home into a trade often does not sell it at all: most of the sale that valuation difficulty costs is a withdrawal without a later observed sale, not a longer wait. The home that does sell goes at a discount its owner never recoups. Comparing two units in the same community and month that differ only in how hard they are to value, the harder-to-value one is less likely to sell, spends longer on the market when it does, and must draw more committed buyer attention to clear.

The illiquidity is economically meaningful and robust: on the full listing universe, a move from the least to the most difficult unit in a community and month lowers the probability of sale by about 1.3 percentage points on the realized measure and 2.4 on the predetermined index, and lengthens time on market by six to nine percent, and every estimate survives trimming the heavy upper tail of the squared-residual proxy. A search model with risk-averse, heterogeneous buyers reproduces the pattern through two channels: uncertainty lowers the mean of buyers' valuations (a risk discount that makes fewer meetings clear and the home sell for less) and widens their dispersion, which lengthens search and draws more attention. That the price falls while attention rises, a pattern neither channel produces alone, shows both are at work; and the framework is estimated, not only calibrated: an extension with a withdrawal margin and a costly inspection gate fits five liquidity and funnel gradients with three parameters, with the fit's one tension confined to a single funnel margin (the approximate diagonal-weighting over-identification diagnostics, $J = 7.15$, are reported in Online Appendix OA.A.9), and recovers the same one-percent risk discount the calibration finds. The effect is also stronger when trading is thin (the uncertainty–time-on-market slope is steeper in low-turnover months), though this state-dependence is specific to the turnover margin, and the cross-sectional design leads me to read it as suggestive rather than as a propagation channel.

These results locate the cost of holding an undiversified, hard-to-value home in liquidity, not only in price: such homes are slower and less certain to sell, a friction borne by every owner who must eventually transact. The friction scales with how hard the home is to value: about a third of a percent of value for the median-to-high move, but most of a percent for the hardest-to-value homes and larger still in thin, low-turnover markets, falling on owners of hard-to-value homes when liquidity is already scarce (Section 5.5). The finding also bears on how idiosyncratic risk is priced in housing. A recent literature reads valuation uncertainty as a return premium (Amaral, 2024), yet idiosyncratic housing risk is not compensated in appreciation (Peng and Thibodeau, 2017) and, in equities, such premia are largely a tail artifact (Bali, Cakici and Whitelaw, 2011); I find the listing-implied premium does not survive tail-robust measurement, and realized resale returns show no robust premium (any positive tilt arises through selection on which sales complete, the model's own prediction), while the illiquidity is unambiguous: the second moment of housing value is borne at the point of sale, in the price, speed, and certainty of trade, rather than compensated ex ante in returns to the holder. Both sides of the market adjust to it: buyers in what they monitor and decline to visit, sellers in what they revise and how long they wait. Because a forward-looking buyer pays less partly in anticipation of facing the same friction at her own exit, the burden travels with the home: it is borne by each owner in turn, at the moment they sell.

Beyond housing, the framework describes any market for listed, idiosyncratic, unshortable assets, and the scissors is a portable diagnostic. Any decentralized listing market that records prices and engagement (used cars, art, collectibles, jobs) can ask the same question: when an item's transaction price falls while the attention it draws rises, no single demand or quality shift can be the explanation: the pattern requires a force that lowers the mean of buyers' valuations alongside one that widens their dispersion, and it needs only the data listing platforms already collect. A causal effect of valuation uncertainty itself asks for a shock to a home's opacity, and Beijing supplies two candidates: the staggered multi-school-zoning reform, which turned a home's school assignment from a certainty into a lottery, and the 2021 reference-price disclosure. Section 5.3 takes the in-sample step on the first, with the Xicheng announcement lowering the probability of sale by 8.5 percentage points against flat pre-trends, an estimate in the extreme tail of placebo assignments though formally limited by the single treated district; extending that design through the post-2021 data and the disclosure, in the spirit of the credit-shock identification of Greenwald and Guren (2025), is the natural next step. A second design would hold the city fixed and vary the clearing technology: Beijing's judicial foreclosure auctions (*sifa paimai*) sell apartments through open bidding with visible bids, so the same city's housing stock can reveal whether aggregated bidders pay the optimist's premium that bilateral search denies.

References

Adelino, Manuel, Antoinette Schoar, and Felipe Severino. 2018. "Perception of house price risk and homeownership." National Bureau of Economic Research Report.

Akerlof, George A. 1970. "The Market for "Lemons": Quality Uncertainty and the Market Mechanism." *The Quarterly Journal of Economics*, 84(3): 488–500.

Allen, Jason, Robert Clark, and Jean-François Houde. 2019. "Search frictions and market power in negotiated-price markets." *Journal of Political Economy*, 127(4): 1550–1598.

Amaral, Francisco. 2024. "Price Uncertainty and Returns to Housing." Working Paper Report.

Amihud, Yakov. 2002. "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects." *Journal of Financial Markets*, 5(1): 31–56.

Amihud, Yakov, and Haim Mendelson. 1986. "Asset Pricing and the Bid-Ask Spread." *Journal of Financial Economics*, 17(2): 223–249.

Andersen, Steffen, Cristian Badarinza, Lu Liu, Julie Marx, and Tarun Ramadorai. 2022. "Reference dependence in the housing market." *American Economic Review*, 112(10): 3398–3440.

Anenberg, Elliot. 2016. "Information Frictions and Housing Market Dynamics." *International Economic Review*, 57(4): 1449–1479.

Ang, Andrew, Robert J Hodrick, Yuhang Xing, and Xiaoyan Zhang. 2006. "The cross-section of volatility and expected returns." *The Journal of Finance*, 61(1): 259–299.

Anundsen, André K., Plamen Nenov, Erling Røed Larsen, and Dag Einar Sommervoll. 2022. "Pricing and Incentives in the Housing Market." Working paper.

Badarinza, Cristian, Vimal Balasubramaniam, and Tarun Ramadorai. 2024. "In Search of the Matching Function in the Housing Market." SSRN Working Paper No. 4594519.

Badidi, Samia. 2026. "Forecast Dispersion and the Price Impact of Macroeconomic News." Working paper.

Bailey, Michael, Ruiqing Cao, Theresa Kuchler, and Johannes Stroebel. 2018. "The Economic Effects of Social Networks: Evidence from the Housing Market." *The Review of Financial Studies*, 31(7): 2224–2271.

Bajari, Patrick, and C. Lanier Benkard. 2005. "Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach." *Journal of Political Economy*, 113(6): 1239–1276.

Baker, Malcolm, and Jeffrey Wurgler. 2006. "Investor Sentiment and the Cross-Section of Stock Returns." *The Journal of Finance*, 61(4): 1645–1680.

Bali, Turan G, Nusret Cakici, and Robert F Whitelaw. 2011. "Maxing out: Stocks as lotteries and the cross-section of expected returns." *Journal of Financial Economics*, 99(2): 427–446.

Barber, Brad M., and Terrance Odean. 2008. "All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors." *The Review of Financial Studies*, 21(2): 785–818.

Barber, Brad M, Xing Huang, Terrance Odean, and Christopher Schwarz. 2022. "Attention-induced trading and returns: Evidence from Robinhood users." *The Journal of Finance*, 77(6): 3141–3190.

Bayer, Patrick, Christopher Geissler, Kyle Mangum, and James W Roberts. 2020. "Speculators and middlemen: The strategy and performance of investors in the housing market."

The Review of Financial Studies, 33(11): 5212–5247.

Ben-Shahar, Danny, and Roni Golan. 2022. "Price Dispersion and Time-on-Market in the Housing Market." *Journal of Housing Economics*, 58: 101875.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru. 2025. "Why Is Intermediating Houses So Difficult? Evidence from iBuyers." NBER Working Paper No. 28252 (revised June 2025); accepted, *Journal of Political Economy*.

Buraschi, Andrea, Fabio Trojani, and Andrea Vedolin. 2014. "Economic uncertainty, disagreement, and credit markets." *Management Science*, 60(5): 1281–1296.

Campbell, John Y, Martin Lettau, Burton G Malkiel, and Yexiao Xu. 2001. "Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk." *The Journal of Finance*, 56(1): 1–43.

Carlin, Bruce I., Francis A. Longstaff, and Kyle Matoba. 2014. "Disagreement and Asset Prices." *Journal of Financial Economics*, 114(2): 226–238.

Carrillo, Paul E. 2013. "To Sell or Not to Sell: Measuring the Heat of the Housing Market." *Real Estate Economics*, 41(2).

Chambers, David, Christophe Spaenjers, and Eva Steiner. 2021. "The Rate of Return on Real Estate: Long-Run Micro-Level Evidence." *The Review of Financial Studies*, 34(8): 3572–3607.

Chang, Jeffery (Jinfan), Yuheng Wang, and Wei Xiong. 2025. "Price and Volume Divergence in China's Real Estate Markets: The Role of Local Governments." Working paper, CUHK Shenzhen and Princeton University.

Chatterjee, Kalyan, and William Samuelson. 1983. "Bargaining under Incomplete Information." *Operations Research*, 31(5): 835–851.

Chetty, Raj. 2009. "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods." *Annual Review of Economics*, 1: 451–488.

Chetty, Raj, László Sándor, and Adam Szeidl. 2017. "The effect of housing on portfolio choice." *The Journal of Finance*, 72(3): 1171–1212.

Chinco, Alex, and Christopher Mayer. 2016. "Misinformed speculators and mispricing in the housing market." *The Review of Financial Studies*, 29(2): 486–522.

Chodorow-Reich, Gabriel, Adam M. Guren, and Timothy J. McQuade. 2024. "The 2000s Housing Cycle with 2020 Hindsight: A Neo-Kindlebergerian View." *Review of Economic Studies*. Forthcoming.

Clapp, John M., and Carmelo Giaccotto. 1992. "Estimating Price Indices for Residential Property: A Comparison of Repeat Sales and Assessed Value Methods." *Journal of the American Statistical Association*, 87(418): 300–306.

Cocco, Joao F. 2005. "Portfolio choice in the presence of housing." *The Review of Financial Studies*, 18(2): 535–567.

Constantinides, George M. 1986. "Capital Market Equilibrium with Transaction Costs." *Journal of Political Economy*, 94(4): 842–862.

Dachis, Ben, Gilles Duranton, and Matthew A. Turner. 2012. "The Effects of Land Transfer Taxes on Real Estate Markets: Evidence from a Natural Experiment in Toronto." *Journal of Economic Geography*, 12(2): 327–354.

Da, Zhi, Joseph Engelberg, and Pengjie Gao. 2011. "In Search of Attention." *The Journal of Finance*, 66(5): 1461–1499.

Deng, Yongheng, Congyan Han, Teng Li, and Yonglin Wang. 2024. "The effectiveness

and consequences of the government's interventions for Hong Kong's residential housing markets." *Real Estate Economics*, 52: 324–365.
de Silva, Tim, Eric C. So, and Kevin Smith. 2026. "Losing Is Optional: Retail Option Trading and Expected Announcement Volatility." *Review of Finance*, 30(2): 489–535.
Diether, Karl B., Christopher J. Malloy, and Anna Scherbina. 2002. "Differences of Opinion and the Cross Section of Stock Returns." *The Journal of Finance*, 57(5): 2113–2141.
Di Maggio, Marco, Francesco Franzoni, Shimon Kogan, and Ran Xing. 2023. "Avoiding Idiosyncratic Volatility: Flow Sensitivity to Individual Stock Returns." NBER Working Paper No. 31360.
Duffie, Darrell, Nicolae Gârleanu, and Lasse Heje Pedersen. 2005. "Over-the-Counter Markets." *Econometrica*, 73(6): 1815–1847.
Gabaix, Xavier. 2019. "Behavioral inattention." *Handbook of behavioral economics: Applications and foundations 1* Vol. 2, 261–343. Elsevier.
Gao, Zhenyu, Michael Sockin, and Wei Xiong. 2021. "Learning about the Neighborhood." *The Review of Financial Studies*, 34(9): 4323–4372.
Gargano, Antonio, and Marco Giacoletti. 2021. "Cooling Auction Fever: Underquoting Laws in the Housing Market." Working Paper No. 03/21.
Gargano, Antonio, Marco Giacoletti, and Elvis Jarnecic. 2023. "Local Experiences, Search, and Spillovers in the Housing Market." *The Journal of Finance*, 78(2): 1015–1053.
Garnache, Cloé, Martin Schmalz, and Dag Einar Sommervoll. "Sticky Asset Prices." SSRN Working Paper No. 5424215.
Gatzlaff, Dean H., and Donald R. Haurin. 1997. "Sample Selection Bias and Repeat-Sales Index Estimates." *Journal of Real Estate Finance and Economics*, 14(1–2): 33–50.
Genesove, David, and Christopher J Mayer. 1994. "Equity and time to sale in the real estate market." *American Economic Review*, 87(3).
Genesove, David, and Christopher Mayer. 2001. "Loss aversion and seller behavior: Evidence from the housing market." *The Quarterly Journal of Economics*, 116(4): 1233–1260.
Genesove, David, and Lu Han. 2012. "Search and Matching in the Housing Market." *Journal of Urban Economics*, 72(1): 31–45.
Giacoletti, Marco. 2021. "Idiosyncratic Risk in Housing Markets." *The Review of Financial Studies*, 34(8): 3695–3741.
Giacoletti, Marco, and Christopher A. Parsons. 2022. "Reference Points Spillovers: Micro-Level Evidence from Real Estate." SSRN Working Paper.
Gilbukh, Sonia, and Paul Goldsmith-Pinkham. 2024. "Heterogeneous Real Estate Agents and the Housing Cycle." *Review of Financial Studies*, hhae048.
Glaeser, Edward, Wei Huang, Yueran Ma, and Andrei Shleifer. 2017. "A Real Estate Boom with Chinese Characteristics." *Journal of Economic Perspectives*, 31(1): 93–116.
Glower, Michel, Donald R. Haurin, and Patric H. Hendershott. 1998. "Selling Time and Selling Price: The Influence of Seller Motivation." *Real Estate Economics*, 26(4): 719–740.
Goetzmann, William N, Christophe Spaenjers, and Stijn Van Nieuwerburgh. 2021. "Real and private-value assets." *The Review of Financial Studies*, 34(8): 3497–3526.
Greenwald, Daniel L., and Adam M. Guren. 2025. "Do Credit Conditions Move House Prices?" *Working Paper*.
Guerrieri, Veronica, and Robert Shimer. 2014. "Dynamic Adverse Selection: A Theory of Illiquidity, Fire Sales, and Flight to Quality." *American Economic Review*, 104(7): 1875–

1908.
Guren, Adam M. 2018. "House Price Momentum and Strategic Complementarity." *Journal of Political Economy*, 126(3): 1172–1218.
Guren, Adam M., and Timothy J. McQuade. 2020. "How Do Foreclosures Exacerbate Housing Downturns?" *The Review of Economic Studies*, 87(3): 1331–1364.
Halket, Jonathan, Lars Nesheim, and Florian Oswald. 2020. "The Housing Stock, Housing Prices, and User Costs: The Roles of Location, Structure and Unobserved Quality." Working paper.
Han, Lu. 2013. "Understanding the puzzling risk-return relationship for housing." *The Review of Financial Studies*, 26(4): 877–928.
Han, Lu, and William C. Strange. 2015. "The Microstructure of Housing Markets: Search, Bargaining, and Brokerage." In *Handbook of Regional and Urban Economics*. Vol. 5, , ed. Gilles Duranton, J. Vernon Henderson and William C. Strange, 813–886. Elsevier.
Harrison, J Michael, and David M Kreps. 1978. "Speculative investor behavior in a stock market with heterogeneous expectations." *The Quarterly Journal of Economics*, 92(2): 323–336.
Haurin, Donald. 1988. "The Duration of Marketing Time of Residential Housing." *Real Estate Economics*, 16(4): 396–410.
Head, Allen, Huw Lloyd-Ellis, and Hongfei Sun. 2014. "Search, Liquidity, and the Dynamics of House Prices and Construction." *The American Economic Review*, 104(4): 1172–1210.
Himmelberg, Charles, Christopher Mayer, and Todd Sinai. 2005. "Assessing High House Prices: Bubbles, Fundamentals and Misperceptions." *Journal of Economic Perspectives*, 19(4): 67–92.
Hong, Harrison, and Jeremy C. Stein. 2007. "Disagreement and the Stock Market." *Journal of Economic Perspectives*, 21(2): 109–128.
Hu, Grace Xing, Jun Pan, and Jiang Wang. 2013. "Noise as information for illiquidity." *The Journal of Finance*, 68(6): 2341–2382.
Jiang, Erica Xuewei, and Anthony Lee Zhang. 2025. "Collateral value uncertainty and mortgage credit provision." *Journal of Financial Economics*, 169: 104054.
Jiang, Erica Xuewei, Nadia Kotova, and Anthony Lee Zhang. 2024*a*. "Liquidity in residential real estate markets." *Available at SSRN 3386353*.
Jiang, Erica Xuewei, Nadia Kotova, and Anthony Lee Zhang. 2024*b*. "Liquidity in Residential Real Estate Markets." SSRN Working Paper No. 3386353.
Jurado, Kyle, Sydney C. Ludvigson, and Serena Ng. 2015. "Measuring Uncertainty." *American Economic Review*, 105(3): 1177–1216.
Krainer, John. 2001. "A Theory of Liquidity in Residential Real Estate Markets." *Journal of Urban Economics*, 49(1): 32–53.
Kurlat, Pablo, and Johannes Stroebel. 2015. "Testing for Information Asymmetries in Real Estate Markets." *The Review of Financial Studies*, 28(8): 2429–2461.
Liu, Chang, and Wei Xiong. 2018. "China's Real Estate Market." NBER Working Paper No. 25297.
Li, Zigang, Stijn Van Nieuwerburgh, and Renxuan Wang. 2026. "Understanding Rationality and Disagreement in House Price Expectations." *The Review of Financial Studies*, 39: 297–342.
Longstaff, Francis A. 1995. "How Much Can Marketability Affect Security Values?" *The*

Journal of Finance, 50(5): 1767–1774.
Malkhozov, Aytek, Philippe Mueller, Andrea Vedolin, and Gyuri Venter. 2017. "International illiquidity." *FRB International Finance Discussion Paper*, , (1201).
Milgrom, Paul R., and Robert J. Weber. 1982. "A Theory of Auctions and Competitive Bidding." *Econometrica*, 50(5): 1089–1122.
Miller, Edward M. 1977. "Risk, uncertainty, and divergence of opinion." *The Journal of Finance*, 32(4): 1151–1168.
Mullainathan, Sendhil, and Jann Spiess. 2017. "Machine Learning: An Applied Econometric Approach." *Journal of Economic Perspectives*, 31(2): 87–106.
Myerson, Roger B., and Mark A. Satterthwaite. 1983. "Efficient Mechanisms for Bilateral Trading." *Journal of Economic Theory*, 29(2): 265–281.
Nash, John F. 1950. "The Bargaining Problem." *Econometrica*, 18(2): 155–162.
Nathanson, Charles G., and Eric Zwick. 2018. "Arrested Development: Theory and Evidence of Supply-Side Speculation in the Housing Market." *The Journal of Finance*, 73(6): 2587–2633.
Ngai, L. Rachel, and Silvana Tenreyro. 2014. "Hot and Cold Seasons in the Housing Market." *American Economic Review*, 104(12): 3991–4026.
Novy-Marx, Robert. 2009. "Hot and Cold Markets." *Real Estate Economics*, 37(1): 1–22.
Ohnishi, Takaaki, Takayuki Mizuno, and Tsutomu Watanabe. 2020. "House Price Dispersion in Boom–Bust Cycles: Evidence from Tokyo." *The Japanese Economic Review*, 71(4): 511–539.
Peng, Liang, and Thomas G Thibodeau. 2017. "Idiosyncratic risk of house prices: Evidence from 26 million home sales." *Real Estate Economics*, 45(2): 340–375.
Piazzesi, Monika, and Martin Schneider. 2009. "Momentum Traders in the Housing Market: Survey Evidence and a Search Model." *American Economic Review: Papers & Proceedings*, 99(2): 406–411.
Piazzesi, Monika, Martin Schneider, and Johannes Stroebel. 2020. "Segmented Housing Search." *American Economic Review*, 110(3): 720–759.
Quan, Daniel C., and John M. Quigley. 1991. "Price Formation and the Appraisal Function in Real Estate Markets." *The Journal of Real Estate Finance and Economics*, 4(2): 127–146.
Sagi, Jacob S. 2021. "Asset-level risk and return in real estate investments." *The Review of Financial Studies*, 34(8): 3647–3694.
Sampford, M. R. 1953. "Some Inequalities on Mill's Ratio and Related Functions." *The Annals of Mathematical Statistics*, 24(1): 130–132.
Scheinkman, Jose A, and Wei Xiong. 2003. "Overconfidence and speculative bubbles." *Journal of Political Economy*, 111(6): 1183–1220.
Sinai, Todd, and Nicholas S. Souleles. 2005. "Owner-Occupied Housing as a Hedge Against Rent Risk." *The Quarterly Journal of Economics*, 120(2): 763–789.
Stambaugh, Robert F., Jianfeng Yu, and Yu Yuan. 2015. "Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle." *The Journal of Finance*, 70(5): 1903–1948.
Stein, Jeremy C. 1995. "Prices and Trading Volume in the Housing Market: A Model with Down-Payment Effects." *Quarterly Journal of Economics*, 110(2): 379–406.
Taylor, Curtis R. 1999. "Time-on-the-Market as a Sign of Quality." *Review of Economic Studies*, 66(3): 555–578.
van der Beck, Philippe, Lorenzo Bretscher, and Julie Zhiyu Fu. 2026. "A Bound on Price

Impact and Disagreement." Harvard Business School Working Paper 26-053.
Wheaton, William C. 1990. "Vacancy, search, and prices in a housing market matching model." *Journal of Political Economy*, 98(6): 1270–1292.
Xiong, Wei. 2013. "Bubbles, crises, and heterogeneous beliefs." National Bureau of Economic Research Report.
Xu, Mandi, Hefan Zheng, and Jing Wu. 2024. "Are Online-Only Real Estate Marketplaces Viable? Evidence from China." *Journal of Real Estate Finance and Economics*, 68(4): 583–623.
Zhang, X Frank. 2006. "Information uncertainty and stock returns." *The Journal of Finance*, 61(1): 105–137.
Zhu, Enwei, Jing Wu, Hongyu Liu, and Keyang Li. 2023. "A Sentiment Index of the Housing Market in China: Text Mining of Narratives on Social Media." *Journal of Real Estate Finance and Economics*, 66: 77–118.
Zhu, Enwei, Jing Wu, Hongyu Liu, and Xindian Li. 2022. "Within-City Spatial Distribution, Heterogeneity and Diffusion of House Price: Evidence from a Spatiotemporal Index for Beijing." *Real Estate Economics*, 50(3): 621–655.

A Proof of the Linearized Pricing-Wedge Equation

This appendix derives the linearized relationship between listing-implied expected returns and valuation uncertainty stated in Proposition OA.4.

Step 1: First-order expansion of $\operatorname{asinh}(\cdot)$**.** Recall the definition of the one-month listing-implied expected return:

$$\mathrm{ExpRet}^{(1m)}_{h,t} \equiv \operatorname{asinh}\big(\widetilde{P}^{\text{list}}_{h,t}\big) - \operatorname{asinh}\big(\bar{P}^{\text{deal}}_{c(h),t-1}\big). \tag{19}$$

Let $X_{h,t} \equiv \widetilde{P}^{\text{list}}_{h,t}$ and $B_{c,t-1} \equiv \bar{P}^{\text{deal}}_{c(h),t-1}$. Fix typical values $X_0 > 0$ and $B_0 > 0$ and take a first-order Taylor expansion:

$$\operatorname{asinh}(X_{h,t}) = \operatorname{asinh}(X_0) + \frac{1}{\sqrt{1+X_0^2}}\,(X_{h,t} - X_0) + o\big(|X_{h,t} - X_0|\big), \tag{20}$$

$$\operatorname{asinh}(B_{c,t-1}) = \operatorname{asinh}(B_0) + \frac{1}{\sqrt{1+B_0^2}}\,(B_{c,t-1} - B_0) + o\big(|B_{c,t-1} - B_0|\big). \tag{21}$$

Subtracting (21) from (20) and collecting constant terms into ζ_t yields

$$\mathrm{ExpRet}^{(1m)}_{h,t} \approx \zeta_t + \frac{1}{\sqrt{1+X_0^2}}\,X_{h,t} - \frac{1}{\sqrt{1+B_0^2}}\,B_{c,t-1}. \tag{22}$$

Step 2: Uncertainty enters through the pricing equation. In the model, the quality-adjusted listing price is a monotone transform of the equilibrium expected price after netting out observables:

$$\widetilde{P}^{\text{list}}_{h,t} = g\big(\mathbb{E}[P^*_{h,t}]\big), \qquad g'(\cdot) > 0.$$

(Under the partial-anchoring listing rule of Online Appendix OA.B, Equation (OA.7), the slope derived below is scaled by the adjustment weight $\theta \in (0,1)$; I set $\theta = 1$ here to reduce clutter, which affects magnitudes, not signs.) From Equation (11)—written with the information-channel scaling κ made explicit (Online Appendix OA.B; the main text normalizes $\kappa = 1$)—the equilibrium price has an affine structure in σ_P^2:

$$\mathbb{E}[P^*_{h,t}] = \underbrace{P^{(0)}_{h,t}}_{\text{terms independent of } \sigma_P^2} - (1-\alpha)\,\frac{\bar{\gamma}\,\kappa}{2}\,\rho^2\sigma_P^2. \tag{23}$$

Linearizing $g(\cdot)$ around a typical value P_0:

$$\widetilde{P}^{\text{list}}_{h,t} \approx g(P_0) + g'(P_0)\big(P^{(0)}_{h,t} - P_0\big) - g'(P_0)(1-\alpha)\frac{\bar{\gamma}\,\kappa}{2}\rho^2\sigma_P^2.$$

Step 3: Combine. Substituting into (22) and absorbing all terms not involving σ_P^2 into ζ_t and the control vector gives

$$\boxed{\mathrm{ExpRet}_{h,t}^{(1m)} \approx \zeta_t - \underbrace{\frac{g'(P_0)}{\sqrt{1+X_0^2}}\,(1-\alpha)\,\frac{\bar{\gamma}\,\kappa}{2}\,\rho^2}_{\Lambda}\;\sigma_P^2 + \text{(controls)} + \varepsilon_{h,t}.} \tag{24}$$

Since $g'(P_0) > 0$, $(1+X_0^2)^{-1/2} > 0$, $1-\alpha > 0$, $\bar{\gamma} > 0$, $\kappa > 0$, and $\rho > 0$, the slope satisfies $\Lambda > 0$, so listing-implied expected returns are decreasing in σ_P^2; Λ is a constant of the linearization, common across listings. □

B Generated-Regressor Bootstrap

Both the uncertainty measure and its within-cell rank are estimated objects, so the headline regressions carry first-stage noise that a textbook analytic standard error ignores. The direct check rebuilds everything inside the resampling loop: each of three hundred draws resamples the $7{,}231$ communities with replacement (with fresh community identifiers, so duplicated communities enter the fixed effects and the clustering as distinct units), re-estimates the full hedonic, recomputes the squared residual and its within community$\times$month rank, and re-runs the headline regressions on their published conventions: the three liquidity outcomes, the price-side uncertainty slopes, and the demand-curve lever interaction. Table 22 reports the comparison. The bootstrap standard errors sit within eleven percent of the analytic cluster-robust values on every margin (ratios 0.98 to 1.11), every headline confidence interval excludes zero, and the overpricing-side lever, which the paper reports as attenuating only modestly, is the one interval that includes it. That the rank is invariant to the scale of first-stage error is therefore not an assertion; on these draws it is an outcome.

Table 22: Generated-Regressor Bootstrap: The Full Headline Suite

	Estimate	Analytic SE	Bootstrap SE	Ratio	Boot. 95% CI
P(sale) $\sim$ rank	-0.0061	(0.0020)	(0.0022)	1.08	$[-0.0104, -0.0019]$
log DOM $\sim$ rank	$+0.0511$	(0.0067)	(0.0067)	1.00	$[+0.0382, +0.0639]$
log followers $\sim$ rank	$+0.0395$	(0.0063)	(0.0062)	0.98	$[+0.0281, +0.0512]$
ExpRaw: uncertainty (β)	-0.1729	(0.0204)	(0.0213)	1.04	$[-0.2206, -0.1373]$
ExpRaw: unc. $\times$ post (δ)	-0.2400	(0.0884)	(0.0916)	1.04	$[-0.4102, -0.1039]$
Lever: discount slope $\times$ disp.	-0.1819	(0.0467)	(0.0517)	1.11	$[-0.2869, -0.0834]$
Lever: overpricing slope $\times$ disp.	-0.0075	(0.0525)	(0.0524)	1.00	$[-0.1253, +0.0864]$

300 community pairs-cluster draws (0 failed); hedonic + u^2 + within-cell rank rebuilt per draw.

Notes. Three hundred community pairs-cluster draws, none failed, with fixed per-draw seeds. Each draw re-derives the hedonic of Equation (1), the squared residual, and the within-cell percent rank before re-estimating each outcome. Point estimates and analytic standard errors are the published full-sample values with community clustering; liquidity outcomes follow the baseline conventions (FVR ≤ 1; days on market on sold units), and the lever interaction follows the demand-curve conventions (cells of at least eight listings, leave-one-out submarket dispersion). The bootstrap interval is the 2.5–97.5 percentile range of the draws.

Table 23 retains the earlier, deliberately conservative exercise, which bootstraps a pre-tightening subsample slope from an interacted specification; its larger ratio (1.71) is a property of that narrower estimand and sample, not of the headline estimates above.

Table 23: Generated-Regressor Bootstrap: Standard Error Comparison

	Estimate	Analytic SE	Bootstrap SE
Valuation uncertainty ($\hat{\beta}$)	-0.1259	0.0331	0.0564

Notes. Pairs cluster bootstrap with 200 iterations, resampling communities with replacement. Each iteration re-estimates both the first-stage hedonic (with community–month fixed effects) and the second-stage uncertainty regression. *Sample:* FVR-restricted subsample ($N = 249{,}378$; FVR ≤ 1), used rather than the full $N = 385{,}818$ headline sample because each iteration is costly; the bootstrap question is whether analytic SEs understate sampling uncertainty once first-stage noise is compounded, which is answered adequately on this subsample. The point estimate $\widehat{\beta} = -0.1259$ is the uncertainty slope for pre-December-2016 postings in a specification that interacts uncertainty with the post-tightening indicator, estimated on the subsample; it is not the headline -0.173. Analytic SEs are cluster-robust at the community level. The bootstrap-to-analytic SE ratio is 1.71 for $\widehat{\beta}$.

Online Appendix

When Price Stops Clearing:
Valuation Disagreement and Housing Illiquidity

Tai Lo Yeung

Notation

This table collects the paper's notation. Each symbol carries a single meaning throughout; where one letter necessarily serves two roles, a decoration or subscript distinguishes them—most importantly the discount rate r versus the uncertainty rank r^{unc}, the bargaining weight α versus the community fixed effect $\mu_{c(i)}$, the arrival rate λ versus the inverse Mills ratio $M(z)$, and the uncertainty primitive u versus the relative ask.

Symbol	Meaning
Indices and sets	
$i,\ t,\ c,\ h,\ y$	listing, month, community, house, calendar year
$m \in \{\text{list}, \text{deal}\}$	market segment: listing vs. transaction
$\tau \in \{E, S\}$	buyer type: end-user (E) vs. speculative / short-horizon (S)
$S_{c,t}$	set of active listings (the "shelf") in community c, month t
Buyers, valuations, bargaining	
$\gamma_E,\ \gamma_S$	absolute risk aversion by buyer type, $0 \le \gamma_S < \gamma_E$
$\bar{\gamma}$	average risk aversion of the buyer pool a listing meets
V	a buyer's private valuation of the house (the offer draw)
PV^B	buyer's present value of owning the house
CE_τ	certainty equivalent of buyer type τ
α	buyer's Nash bargaining weight (weights P^S in P^*)
ρ	per-period discount factor
$a_{h,t},\ \bar{a}$	fundamental (level) component of value
Seller reservation	
$P^S_{h,t}$	seller's (anchored) reservation value
S^*	fully optimal stopping reservation
$\bar{S}$	backward-looking anchor (market norm)
ν	anchoring weight, $P^S = (1-\nu)\bar{S} + \nu S^*$ (calibrated $\nu = 0.78$)
Prices and wedges	
$P,\ P^*$	price; realized transaction price
$\widetilde{P}^m$	quality-adjusted (hedonic-residualized) price, segment m
$\bar{P}^{\text{deal}}_{c,t-1}$	community-mean quality-adjusted deal price, prior month
$\text{ExpRaw}^{(1m)}$	listing-implied one-month pricing wedge (empirical)
$\text{ExpRet}^{(1m)}$	listing wedge in the model (theoretical counterpart)
Prem	within-community listing premium

Symbol	Meaning
Valuation uncertainty	
u	value left unexplained by public comparables (the information primitive)
$\hat{u}^m$	hedonic residual, segment m
$\hat{u}^2$	squared hedonic residual: the valuation-uncertainty measure
σ_P^2	idiosyncratic valuation uncertainty (model counterpart of $\hat{u}^2$)
$\widetilde{\sigma}^2$	perceived risk, $\widetilde{\sigma}^2 = \kappa(u)\, u$ (the discount channel)
σ_b^2	belief dispersion, $\sigma_b^2 = \tau_\eta\, \widetilde{\sigma}^4$ (the disagreement channel)
$\kappa(u)$	precision shrinkage $1/(1+\tau_\eta u) \in (0,1)$ (derived, not a parameter)
$\tau_\eta,\ \tau_{\text{pub}}$	private-signal (attention) and public-comparable precision
$g,\ \Delta g$	risk discount and its least-to-most-uncertain increment
$s,\ \Delta s$	valuation dispersion and its increment
Search and matching	
λ	buyer meeting (arrival) rate
q	per-meeting clearing (acceptance) probability
χ	cost of a physical visit (the inspection gate)
$\Lambda_{c,t}$	community–month buyer-contact aggregate
$w(\cdot)$	listing attractiveness weight in contact allocation
$z,\ z^*$	standardized selectivity (threshold)
$M(z)$	inverse Mills ratio, $\phi(z)/\big(1-\Phi(z)\big)$
$\delta(z)$	$M(z)\big(M(z)-z\big)$, the truncated-variance term
$\Phi,\ \phi$	standard normal CDF and density
δ_w	withdrawal (delisting) rate (estimated extension)
Platform attention (data)	
$\text{VIEWING}_{i,t}$	page views: broad, low-cost browsing
$\text{FOLLOWER}_{i,t}$	saved-listing followers: committed attention
$\text{VISITING}_{i,t}$	agent-accompanied physical showings (*daikan*)
$\text{FVR}_{i,t}$	follower-to-viewing ratio, FOLLOWER/VIEWING
$\text{Popular}_{i,t}$	indicator: followers above the period 75th percentile
Empirical specification	
$r_{i,t}^{\text{unc}}$	within community×month rank of valuation uncertainty
$\mu_{c(i)}$	community fixed effect
τ_t	year–month fixed effect

Symbol	Meaning
Ξ_i	unit-attribute controls / fixed effects
β	uncertainty gradient (the coefficient of interest)
TOM	time on market (days)
$\mathrm{SupplyShare}_{c,t}$	community's share of citywide listings in month t
Calibration and estimation	
r	discount rate ($r = c = 5\%$ in the calibration)
c	flow carrying cost
$A,\ \mathcal{A}$	committed-attention flow (A) and stock ($\mathcal{A} = A/q$)
c_A	attention-cost scaling, $A = c_A\,\mathrm{Var}_V$
J	minimum-distance over-identification statistic
λ^*	Miller-limit meeting-rate threshold
Reference dependence (relisting)	
Δ	community price appreciation since the seller's purchase
$\mathrm{LOSS},\ \mathrm{GAIN}$	nominal loss / gain at relisting, $\max\{0,-\Delta\}$, $\max\{0,\Delta\}$
V_{now}	current market value of the unit at relisting

OA.A Search-and-matching microfoundation

This appendix gives the search-and-matching foundation for the acceptance channel of Section 3 and for the state-dependence documented in Section 5.2.5. It is deliberately lean: it states the environment, derives the per-meeting acceptance probability and its comparative statics in valuation uncertainty, and shows why the uncertainty–illiquidity link steepens when the market cools.

OA.A.1 Environment

A seller lists one house h in community c in month t. Potential buyers arrive according to a Poisson process with meeting rate λ_t, the market-liquidity state—high in hot markets, low in cold ones. On meeting, a buyer draws a private valuation

$$PV^B = \bar{v}_{c,t} - \tfrac{\gamma}{2}\,\sigma_P^2 + \epsilon, \qquad \epsilon \sim F(\cdot\,;\sigma_P^2), \tag{OA.1}$$

where $\bar{v}_{c,t}$ is the common community–month value, γ is buyer risk aversion, σ_P^2 is the idiosyncratic valuation uncertainty of the house (empirically, the squared hedonic residual $\hat{u}^2$), and ϵ is a buyer-specific match component whose dispersion $\sigma_\epsilon(\sigma_P^2)$ is increasing in σ_P^2: harder-to-value homes generate more disagreement across buyers. The $-\frac{\gamma}{2}\sigma_P^2$ term is the risk discount of Equation (8), with the discounting and information-channel scalings $\rho^2\kappa$ absorbed into γ to reduce clutter.

OA.A.2 Seller reservation and acceptance

The seller holds a reservation value $P_{h,t}^S$ anchored on recent community comparables and adjusted sluggishly (Genesove and Mayer, 2001); it does not fully absorb the buyer's risk discount. Such stickiness is an equilibrium feature of listing markets, not merely a behavioral one: when each seller's sale hazard depends on her price relative to neighboring listings, deviating from the going price is punished asymmetrically, and list prices become strategically sticky (Guren, 2018). A meeting converts to a sale iff $PV^B \geq P_{h,t}^S$, so the per-meeting acceptance probability is

$$q(\sigma_P^2;\lambda_t) = \Pr\Big[\,\epsilon \geq P_{h,t}^S - \bar{v}_{c,t} + \tfrac{\gamma}{2}\sigma_P^2\,\Big]. \tag{OA.2}$$

Higher uncertainty raises the risk discount and hence the threshold the buyer's draw must clear; it also widens the dispersion of ϵ, which feeds mass back into the accepting tail when the threshold sits in the upper half of the draw distribution. Writing z for the standardized threshold, the acceptance rate falls, $\partial q/\partial\sigma_P^2 < 0$, whenever the discount moves the threshold by more than disagreement widens the tail,

$$\frac{\gamma}{2} \;>\; z\,\frac{\partial \sigma_\epsilon}{\partial \sigma_P^2}, \tag{OA.3}$$

which is automatic when $z \le 0$ (at least half of meetings clear)—the regime of the calibration in Section OA.A.7, whose baseline acceptance rate settles slightly above one-half—and is reinforced when the reservation itself rises with the option value that dispersion creates, as the calibrated anchoring weight implies and Proposition 4 makes exact. Under this maintained condition, fewer meetings convert. This is the baseline illiquidity result, Equation (13) in the text.

OA.A.3 Dispersion delays trade under an optimal reservation

The anchored reservation above is a quantitative device, not the source of the dispersion channel's delay. Suppose instead the seller is fully optimal: her reservation S^* solves the standard stopping problem with flow carrying cost $c \ge 0$, discount rate $r > 0$, meeting rate λ, Nash buyer weight α, and offers $V \sim \mathcal{N}(\mu, s^2)$,

$$rS^* + c \;=\; \lambda(1-\alpha)\,\mathbb{E}\big[(V - S^*)^+\big] \;=\; \lambda(1-\alpha)\,s\,h(z^*), \qquad z^* = \frac{S^* - \mu}{s}, \tag{OA.4}$$

where $h(z) = \phi(z) - z\big(1 - \Phi(z)\big)$ is the normal mean-excess function, $h'(z) = -\big(1 - \Phi(z)\big)$. This is exactly the reservation structure of the calibration in Section OA.A.7 at anchoring weight $\nu = 1$. The offer distribution is the model's valuation object of Equation (10): $\mu = \mu_V$ and $s^2 = \mathrm{Var}_V = \rho^4 \sigma_b^2(u)$, where $u \equiv \sigma_P^2$ is the idiosyncratic valuation uncertainty (empirical counterpart $\hat{u}^2$)—a single dispersion scale under several glyphs, collected here once for the stopping algebra that follows.

Proposition 4 in the main text states the result; in the notation here its displayed objects are

$$\frac{\partial S^*}{\partial s} = \frac{\lambda(1-\alpha)\,\phi(z^*)}{r + \lambda(1-\alpha)\big(1 - \Phi(z^*)\big)} \;>\; 0, \qquad \frac{\partial S^*}{\partial \mu} = \frac{\lambda(1-\alpha)\big(1 - \Phi(z^*)\big)}{r + \lambda(1-\alpha)\big(1 - \Phi(z^*)\big)} \;\in\; (0,1),$$

with $q = 1 - \Phi(z^*)$ and $z^* = (S^* - \mu)/s$ from Equation (OA.4).

Proof. Write $H(S; \mu, s) \equiv rS + c - \lambda(1-\alpha)\,s\,h\big((S-\mu)/s\big)$. Then $\partial H/\partial S = r + \lambda(1-\alpha)(1 - \Phi(z)) > 0$, $\partial H/\partial s = -\lambda(1-\alpha)\big[h(z) + z(1-\Phi(z))\big] = -\lambda(1-\alpha)\phi(z) < 0$, and $\partial H/\partial \mu = -\lambda(1-\alpha)(1-\Phi(z)) < 0$; the implicit function theorem gives (a) and (c). For (b), q falls if and only if z^* rises, i.e. $\partial S^*/\partial s > z^*$. Using (a), this is $\lambda(1-\alpha)\big[\phi(z^*) - z^*(1 - \Phi(z^*))\big] > rz^*$, i.e. $\lambda(1-\alpha)\,h(z^*) > rz^*$. Substituting the stopping condition $\lambda(1-\alpha)\,h(z^*) = (rS^* + c)/s$

and $z^* = (S^* - \mu)/s$, the inequality collapses to

$$\frac{rS^* + c}{s} > \frac{r(S^* - \mu)}{s} \quad \Longleftrightarrow \quad c + r\mu > 0,$$

which holds for any asset with positive expected value, $\mu > 0$ (hence for housing), and otherwise whenever the carrying cost covers the discounted holding loss, $c > -r\mu$. Then $\mathbb{E}[\mathrm{TOM}] = 1/(\lambda q)$ rises. □

The economics: dispersion raises the option value of waiting, and the optimal reservation climbs by more than the accepting tail thickens—the seller spends the entire option-value gain on selectivity, and $c + r\mu > 0$ guarantees her time is too valuable to do otherwise. Two readings matter for the paper. First, the dispersion channel's delay is not a behavioral artifact: it survives, indeed is sharpest under, full seller rationality. Second, part (c) assigns anchoring its true role: an optimal seller routes a valuation-level shock—the risk discount—into her reservation nearly one-for-one when r is small, so the discount slows trade only to the extent reservations are sticky. Anchoring is the dial on the discount's quantity bite; disagreement needs no dial. This is quantitatively visible in the calibration: at $\nu = 1$ under a third of the measured time-on-market gradient survives, four-fifths of it the optimal-holdout channel of part (b) and the remainder the sliver of the discount that even an optimal seller does not absorb; moving to the calibrated $\nu = 0.78$ multiplies the discount's quantity contribution roughly fifteen-fold while leaving the dispersion contribution essentially unchanged—anchoring is the discount's dial, as part (c) says.

OA.A.4 Uncertainty versus thinness: a signature table

A unit can also be slow to sell for a reason that has nothing to do with valuation difficulty: it can sit in a thin segment where buyers rarely arrive. Jiang, Kotova and Zhang (2024*b*) make the market-level version of this point—time on market alone is uninterpretable, because the same duration movement means opposite things depending on how prices and price dispersion move with it—and identify market-level liquidity supply and demand from the joint movement. The stopping model delivers the within-market analogue: the meeting rate λ is a primitive alongside the risk discount and disagreement, and the three leave mutually exclusive signatures.

Lemma OA.1 (Thickness comparative statics)**.** *In the stopping model of Equation* (OA.4)*: $\partial S^*/\partial\lambda > 0$, so z^* rises and the per-meeting acceptance probability q falls as the market thickens; expected time on market $1/(\lambda q)$ nonetheless strictly falls; and the conditional dispersion of accepted valuations $s^2\big(1 - \delta(z^*)\big)$ falls. Equivalently, in a thin segment (λ low) the optimal seller settles—acceptance is high—yet sales are slow and the accepted set is wide.*

Proof. $\partial H/\partial\lambda = -(1-\alpha)\,s\,h(z) < 0$ gives $\partial S^*/\partial\lambda = (1-\alpha)\,s\,h(z^*)/\big[r+\lambda(1-\alpha)(1-\Phi(z^*))\big] > 0$, hence z^* rises, $q = 1-\Phi(z^*)$ falls, and $\delta(z^*)$ rises by Lemma OA.2(i). For time on market, $d\ln\mathbb{E}[\mathrm{TOM}]/d\lambda = -1/\lambda + (\phi/q)\,\partial z^*/\partial\lambda$, which is negative if and only if $q\,[\,r+\lambda(1-\alpha)q\,] > \lambda(1-\alpha)\,\phi(z^*)\,h(z^*)$, implied by $q^2 > \phi h$. Writing $t = q/\phi$, the inequality $q^2 > \phi\,[\phi - z(1-\Phi)]$ is $t^2 + zt - 1 > 0$, i.e. the Mills ratio satisfies $\phi/q < \frac{1}{2}\big(z+\sqrt{z^2+4}\big)$—the classical bound used in the proof of Lemma OA.2 (Sampford, 1953)—so time on market falls in λ for every z. □

Proposition OA.1 (Signatures: thinness, discount, disagreement). *Under the maintained structure (partially optimal reservation, the conditions of Propositions 4 and 1), the three primitives move the observables as follows:*

	Mean price	*TOM*	*Acceptance q*	*Attention/exposure*	*Cond. disp.*
Thin segment ($\lambda\downarrow$)	$\downarrow$	$\uparrow$	$\uparrow$	—	$\uparrow$
Risk discount ($\bar\gamma\kappa\uparrow$)	$\downarrow$	$\uparrow$	$\downarrow$	—	*weakly* $\downarrow$
Disagreement ($\sigma_b^{2\prime}\uparrow$)	*weakly* $\uparrow$	$\uparrow$	$\downarrow$	$\uparrow$	$\uparrow$

All three slow trade, but through opposite acceptance margins: a thin market is slow because meetings are scarce and the optimal seller settles; a hard-to-value unit is slow because meetings fail. And only disagreement moves per-exposure committed attention. The joint pattern in the data—price down, time on market up, per-exposure attention up, negotiated dispersion up, with the per-day arrival proxy (casual views) flat—is produced by the two uncertainty channels together and by no combination that excludes them.

Proof. The λ column is Lemma OA.1 with per-exposure attention $A = c_A\,\mathrm{Var}_V$ independent of λ; the discount and disagreement columns collect Propositions OA.3, OA.6, OA.7, 4, and the truncated-moment decompositions (OA.8)–(OA.9). The conditional-dispersion column assumes $z \geq 0$ (at most half of meetings clear), and the disagreement row's $q\downarrow$ uses the option-value response of the reservation (Proposition 4, $c + r\mu > 0$); under a fixed anchored reservation with $z > 0$, pure dispersion would instead raise q. □

Two corroborations already in the paper line up with the table's seams. First, the duration-mechanics diagnostic (Table OA.33) shows per-day casual views—the arrival proxy—essentially flat in uncertainty: the time-on-market gradient is a conversion phenomenon, not an arrival phenomenon, which is the uncertainty rows and not the thinness row. Second, the choice-free comparable-scarcity margin (Online Appendix OA.D.8) bundles segment thinness with public information; under size-segment controls its liquidity gradients survive while its attention gradient does not—exactly the thinness row—whereas the residual-uncertainty rank keeps both its conversion and its attention gradients, as the uncertainty rows require. The two measures separate along the seam the table predicts.

OA.A.5 Why the illiquidity is state-dependent

The reservation responds to market tightness, but only partially: in a cold market (low λ_t) the option value of waiting falls, yet an anchored seller cuts the reservation incompletely (Genesove and Mayer, 2001; Allen, Clark and Houde, 2019). Two forces then make the uncertainty penalty bite harder when λ_t is low. First, with anchoring incomplete the gap $P^S_{h,t} - \bar{v}_{c,t}$ stays high relative to the depressed pool of buyer valuations, so the risk discount moves a larger mass of buyers below the threshold. Second, expected time on market is $1/(\lambda_t q)$, convex in q, so a given uncertainty-induced fall in q translates into a larger increase in waiting when λ_t is small. Hence, with the anchored reservation held fixed at its level,

$$\frac{\partial^2\, \mathbb{E}[\mathrm{TOM}]}{\partial \sigma_P^2\, \partial(-\lambda_t)} > 0 : \tag{OA.5}$$

the uncertainty–time-on-market slope steepens, in levels, as the meeting rate falls. The statement describes the joint cold-market shock (a lower meeting rate against a sticky anchor), not a pure λ-derivative: under a fully optimal reservation the calibrated smooth model moves the log gradient the other way, which is exactly the non-match the calibration subsection reports. This matches the state-dependence in Section 5.2.5: the link is stronger when trading is thin. The cross-sectional design cannot establish that valuation uncertainty causally amplifies illiquidity over the cycle, so the prediction is read as a state-dependent regularity rather than a propagation channel.

OA.A.6 Outcomes

Expected time on market is $\mathbb{E}[\mathrm{TOM}] = 1/\big(\lambda_t\, q(\sigma_P^2; \lambda_t)\big)$, increasing in σ_P^2; the probability of sale within a listing window of length τ is $1 - e^{-\lambda_t q \tau}$, decreasing in σ_P^2; and the search a seller must attract to clear, proportional to $1/q$, is increasing in σ_P^2. These map to the three illiquidity outcomes in Table 6 and to Proposition OA.6. The same renewal logic signs the composition of the shelf: in steady state the active stock's density over the uncertainty rank is $f_{\text{stock}}(r) \propto f_{\text{in}}(r)/q(r)$, and because the within-cell rank is uniform in the inflow by construction, any over-representation of high-rank listings in the active stock is a regression-free trace of $1/q(r)$—an accounting check the data pass (Online Appendix OA.F.13).

The stock as the choice set: allocation and design invariance. The active stock is also the denominator of the buyer's problem. Let the buyers reaching cell c in month t allocate contacts across the shelf $S_{c,t}$ in proportion to an attractiveness weight,

$$\lambda_h \;=\; \Lambda_{c,t}\, \frac{w(r_h)}{\sum_{k \in S_{c,t}} w(r_k)}, \tag{OA.6}$$

with $w(\cdot) > 0$ inherited from the visit-gate value of Online Appendix OA.B.5. Two properties follow. First, design invariance: for any two units in the same cell, $\lambda_h/\lambda_{h'} = w(r_h)/w(r_{h'})$—the denominator cancels—so within-cell gradients identify own-attribute responses regardless of what else sits on the shelf, and composition moves only the cell aggregate that the community–month fixed effects absorb. Departures from proportionality (similarity-based substitution among close alternatives) would make the within-cell gradient vary with the competing mix; empirically it does not: active same-size competition is null on every margin (Online Appendix OA.D.8), and the uncertainty gradient is invariant to cell size (Online Appendix OA.D.2). Second, a congestion externality exists in principle—a hard-to-value entry contributes more listing-months to every denominator, in proportion to its duration—and the data now measure its two moments directly rather than merely bound them. The composition moment: within community, the active stock's mean rank rises when the cell cools (-0.021, $t = -6.7$ on the lagged cell sale rate)—the shelf hardens when the market cools, the sign the thin-market amplification implies—but a one-standard-deviation temperature swing moves the tilt by only 0.004 rank points, an order of magnitude below the cross-sectional gradients. The dilution moment: per-listing followers fall by 3.9 percent per log of shelf size within community ($t = -3.5$), so buyer entry absorbs roughly 96 percent of shelf growth (Online Appendix OA.F.13). The full equilibrium version, in which this entry margin amplifies thin-market downturns, is left to future work; the present data discipline it as quantitatively small. A fuller dynamic treatment with endogenous search intensity and seller learning—genuinely more involved, and not needed for the comparative statics the paper tests—is left to future work; the lean version here delivers those comparative statics.

OA.A.7 A calibrated comparative static in market thickness

How much of the paper's quantitative content can this lean structure carry? This subsection calibrates it. Buyers meet a listing at rate λ; a meeting draws $V \sim \mathcal{N}(1-g,\, s^2)$, where g is the risk discount and s the valuation dispersion, both increasing in uncertainty; the seller's reservation is anchored, $P^S = (1-\nu)\bar{S} + \nu S^*$, where S^* is the optimal-stopping value (flow carrying cost c, discount rate r, Nash weight α) and $\bar{S}$ the easy-unit, normal-market anchor, held at its normal-market value throughout the comparative static below: an anchor is a market norm, not a function of the meeting rate. I set $r = c = 5$ percent, $\alpha = 0.5$, baseline dispersion $s_0 = 4$ percent of value (the four-to-five percent buyer-valuation spread documented in the microstructure survey of Han and Strange, 2015), and the meeting rate pinned by a seventy-day mean spell at a one-half per-meeting acceptance rate. Three estimated moments then pin the three free parameters: the attention gradient pins the dispersion increment (Δs, through attention $\propto \mathrm{Var}_V$), and the price (-0.74 percent) and time-on-market ($+5.1$ percent) gradients jointly pin the risk-discount increment Δg and the anchoring weight ν.

Three results. First, matching the estimated gradients requires anchoring: the fit delivers $\nu = 0.78$—a seller who adjusts about three-quarters of the way to the optimal reservation—quantifying the anchored-reservation assumption of the framework (a fully optimal seller, $\nu = 1$, absorbs uncertainty in price and produces under a third of the measured time-on-market gradient). Because the three moments exactly identify the three parameters, their sampling uncertainty propagates directly to the calibration. I draw the three gradients from their clustered sampling distributions (cross-moment covariances set to zero, as the moments come from separate regressions) and pass each draw through the identical calibration map; $20{,}000$ draws give percentile intervals, with the delta-method standard errors as a cross-check.[19] The bootstrap puts ν at 0.78 with a 95 percent interval of $[0.65, 0.86]$ (delta method $[0.68, 0.88]$)—broadened slightly downward, still excluding the fully optimal seller in every draw—and Δg at 0.99 percent of value (SE 0.15). The robust statement about the optimist-pricing limit is the one Figure OA.1 can display directly: the price gradient stays negative across the entire plotted range, to 10^5 times Beijing's meeting rate, and every bootstrap draw keeps the threshold above that range—so the flip does not occur anywhere a housing market could plausibly sit. The implied threshold itself is astronomically larger ($\log_{10} \lambda^* = 33.7$; bootstrap one-sided floor ≈ 14), but its precise magnitude is a property of the Gaussian tail rather than a quantity to take literally: with fatter-tailed disagreement the threshold is only polynomial in $\Delta g/\Delta s$ and far smaller (Proposition OA.2). I therefore read the result qualitatively—optimist pricing is out of reach for any realistic housing market, by a margin the design cannot exhaust—and note only that the delta-method interval understated even the plotted range, which the bootstrap corrects. Second, Figure OA.1 traces the two margins of the burden as the meeting rate varies from a quarter of Beijing's to thirty-two times it: the price margin is nearly invariant (between -0.6 and -0.8 percentage points throughout), while the time margin shrinks with thickness—the trading technology moves the composition of the burden, not its sign. Third, the optimist-priced limit is quantitatively out of reach: at the dispersion increment the attention gradient implies, the price gradient remains negative for any meeting rate within five orders of magnitude of Beijing's. Miller's case is the $\lambda \to \infty$ limit of this model, and no housing market is anywhere near it. One non-match is reported for completeness: the smooth normal-hazard calibration moves the time-on-market gradient the wrong way as λ halves (the log gradient flattens by roughly a third), against the $+63$ percent steepening estimated in cold months—consistent with the paper's reading of the amplification as a thin-market nonlinearity rather than a smooth λ-derivative, and a further reason that result is labeled suggestive. The non-match also points at its own repair, through a documented mechanism rather than a free dial: replacing the

[19]The delta method is adequate for ν and Δg, which enter the map smoothly, but not for the threshold $\log_{10} \lambda^* = (\Delta g/\Delta s)^2/(2 \ln 10)$: a squared ratio of noisy gradients with a hard floor at zero is heavily right-skewed, so a symmetric interval understates the floor. None of the $20{,}000$ draws falls below the delta-method lower bound of $10^{4.7}$.

linear anchor with a loss-averse reference (Genesove and Mayer, 2001; Andersen et al., 2022) makes anchoring bind precisely when optimal reservations fall below the reference—in cold markets—routing more of the discount into the quantity margin exactly when the data say the slope steepens. The enrichment carries a testable implication: the cold-month steepening should concentrate among sellers in the loss domain, whom the quasi-repeat purchase prices identify (Section 5.1.4; the loss-domain asymmetry itself is measured directly in Online Appendix OA.A.8); the triple interaction is underpowered at sixteen thousand pairs, so the test awaits a larger repeat sample.

An asymptotic result turns the figure's numerical sweep into a closed form and explains why the optimist-priced limit is unreachable.

Proposition OA.2 (The optimist-priced limit is exponentially remote)**.** *In the stopping model of Equation (OA.4), as $\lambda \to \infty$ the optimal standardized selectivity and the expected transaction price satisfy*

$$z^* = \sqrt{2\ln\lambda}\,\big(1+o(1)\big), \qquad \mathbb{E}[P \mid \textit{sale}] = \mu + s\,\sqrt{2\ln\lambda}\,\big(1+o(1)\big) :$$

market thickness buys selectivity only at a $\sqrt{\ln}$ rate, because optimists are exponentially scarce in a normal tail. Comparing a harder-to-value unit $(\mu-\Delta g,\ s+\Delta s)$ with an easier one $(\mu,\ s)$, the price gradient turns positive only beyond

$$\ln\lambda^* \;=\; \frac{1}{2}\left(\frac{\Delta g}{\Delta s}\right)^2 \big(1+o(1)\big),$$

and the dropped higher-order term enters positively, so the leading term is a finite-λ lower bound on the threshold (confirmed numerically). At the calibrated increments ($\Delta g = 0.0099$, $\Delta s = 0.0008$), the leading term alone puts λ^* near 10^{33} times the Beijing meeting rate.

Proof sketch. For large z, the Mills ratio gives $h(z) = \phi(z)/z^2\,(1+o(1))$; substituting into Equation (OA.4) and taking logs gives $z^{*2} = 2\ln\lambda - 2\ln z^{*3} + O(1)$. A first pass yields $z^* = O(\sqrt{\ln\lambda})$; substituting that back, the correction term is $-2\ln z^{*3} = -3\ln z^{*2} = O(\ln\ln\lambda) = o(\ln\lambda)$, so $z^* = \sqrt{2\ln\lambda}\,(1+o(1))$. Because that dropped correction enters z^* positively, the leading expression is a lower bound on the threshold at finite λ—the sense in which it understates λ^*. Then $\mathbb{E}[P \mid \text{sale}] = S^* + (1-\alpha)\,s\,h(z^*)/\big(1-\Phi(z^*)\big)$, whose mean-excess term vanishes as $z^* \to \infty$, so the price tracks S^*. Through S^* the level shock $-\Delta g$ enters $\mathbb{E}[P \mid \text{sale}] = \mu + s\,z^*(1+o(1))$ with unit weight on μ (consistent with the $dS^*/d\mu \to 1$ pass-through of Proposition 4(c)), while the dispersion increment adds $\Delta s\,\sqrt{2\ln\lambda}$; the gradient flips where $\Delta s\sqrt{2\ln\lambda} = \Delta g$. □

Two readings. First, this is the precise sense in which the binding constraint on Miller (1977) here is search and not short sales: the disagreement premium requires meeting rates exponentially large in the squared discount-to-dispersion ratio, and no trading technology for houses is within thirty orders of magnitude of it. Second, the exponential form is the normal tail's: with Pareto-tailed disagreement the threshold would be polynomial in $\Delta g/\Delta s$, so optimist pricing re-emerges far sooner in fat-tailed markets—a discipline on where divergence-of-opinion premia should and should not be found, consistent with their documentation in speculative assets rather than ordinary housing.

Validation on untargeted moments. The fit uses three moments—the attention, price, and time-on-market gradients—and externally set primitives; every other moment in the paper is out of sample for the calibrated structure. Four families of them check it. First, the engagement funnel: the information block requires cheap monitoring to rise with disagreement and costly inspection to fall when the discount's shading dominates (Propositions 3(iii) and OA.8); at a fixed spell, followers rise ($+1.5$ percent, $t = 2.2$) while showings fall (-4.4 percent, $t = -7.2$; Table OA.36, Online Appendix OA.F.12)—opposite signs at the two costly stages, a split the fit never saw: the attention gradient pinned Δs, and nothing in the calibration used showings. Second, the stopping rule: the anchoring weight was pinned by two cross-sectional gradients, yet it implies a specific within-spell behavior—a seller who concedes as showings fail, but only part of the way—and the within-spell data land there: memorylessness is rejected in the updating direction with small magnitudes (-0.04 percent in the concession and -0.42 percent in the deal level per log showing; Online Appendix OA.F.12). Third, tenure: the burden the model prices is borne during the sale spell by whoever then owns the unit, so entry discounts should not be recouped at resale and returns should load on the resale spell rather than the purchase spell—both hold in the quasi-repeat pairs (Section 5.1.4, Table 9). Fourth, the congestion extension's two measured moments carry the signs it requires: the active stock tilts toward hard-to-value units when the local sale rate falls (-0.021, $t = -6.7$, lagged), and per-listing attention dilutes as the shelf grows (-3.9 percent per log listing; Online Appendix OA.F.13). One non-match is reported above—the smooth calibration understates the thin-market steepening—and these are sign-and-scale checks on a three-parameter calibration, not the moments of an estimated model; their value is that none of them entered the fit.

OA.A.8 The anchored seller in the data: reference dependence at relisting

The calibration above gives the seller a partial anchor: the reservation adjusts about three-quarters of the way toward the optimal value ($\nu = 0.78$), leaving roughly a fifth of the weight on the backward-looking reference point. That weight is a quantitative assumption, and the quasi-repeat-sale pairs let me hold it up against seller behavior directly. No repeat-sale data entered the calibration—it is pinned by three listing-level cross-sectional gradients alone—so these pairs furnish an untargeted check, a moment the fit never saw, not a re-estimation. I

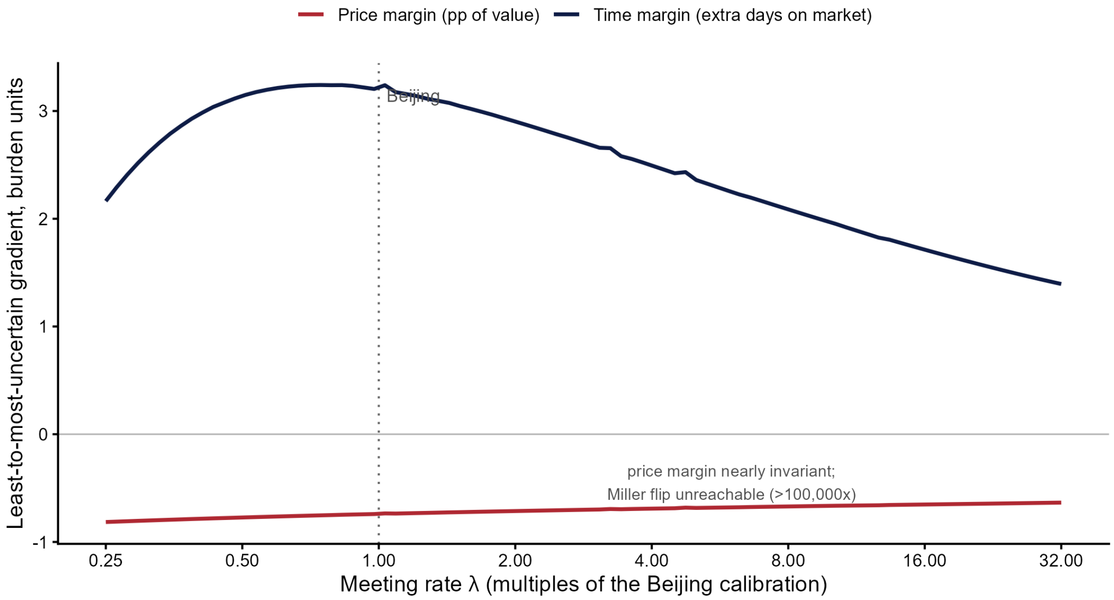


Figure OA.1: The Trading Technology Moves the Burden's Composition, Not Its Sign

Notes. Calibrated least-to-most-uncertain gradients as functions of the buyer meeting rate λ, in multiples of the Beijing calibration (dotted line; mean time on market of seventy days). The price margin (red) is the gradient of the expected transaction price in percentage points of value; the time margin (navy) is the gradient of expected time on market in days. Parameters: $r = c = 5$ percent, $\alpha = 0.5$, $s_0 = 4$ percent; Δs from the attention gradient; Δg and the anchoring weight $\nu = 0.78$ from the price and time-on-market gradients. The price gradient stays negative for all λ up to 10^5 times the Beijing rate.

take the clean-sequential pairs of Section 5.1.4—the same unit, or its identical twin, relisted after an earlier sale—and treat the relister's own earlier deal price as the reference point. Netting the community's appreciation between the two dates out of that purchase price gives the unit's current market value V_{now}, which differences out fixed quality and the local price path; the relist markup is the new asking price minus V_{now}. There are $19{,}878$ such pairs with a community–month price index at both dates.

Two facts emerge, and they pin the anchor from opposite directions. First, the relist ask is partially anchored: it loads positively on both the earlier purchase price and the current market value (both $t > 12$), so the seller neither marks fully to market nor freezes at what she paid. That pooled loading blends the gain and loss regimes, so the clean comparison to the calibration is the regime split, not the pooled rate. Second, the anchoring is asymmetric in the way Genesove and Mayer (2001) first documented. Table OA.2 splits the markup on the nominal gain and the nominal loss. In the gain domain—the community has risen above the purchase price—the seller passes about four-fifths of the gain into the ask (-0.20, an implied anchor weight near 0.20), of the same order as the calibrated $1 - \nu \approx 0.22$. In the loss domain the anchor takes over: she passes through only about a fifth of the market's decline and refuses to mark down the rest (an anchor weight near 0.81). Loss-domain relists also hold out on the other margins—they sit about 27 percent longer on the market ($t = 2.3$) and are 9 percentage points less likely to sell ($t = 2.9$), the Genesove–Mayer pattern reproduced at short horizon (Andersen et al., 2022). That shortfall takes the same form as the paper's central margin: it is overwhelmingly withdrawal rather than a slower sale—the underwater seller pulls the listing rather than mark the home down—a behavioral echo, on an axis orthogonal to within-cell valuation difficulty, of the friction-driven abandonment of Section 5.1.2.

The reading is twofold. Where reference dependence is mild—the rising market that covers most of the sample—the seller's anchor weight (≈ 0.20) is of the same order as the calibration's $1 - \nu$, so the model's partially anchored reservation is not a free device but a regularity visible in relisting. The two numbers are not the identical structural parameter: the calibration's ν weights a market-norm anchor $\bar{S}$ in the seller's reservation, whereas this coefficient weights the seller's own prior purchase price in the asking price. I therefore read the agreement as an order-of-magnitude consistency check on the anchoring assumption, not as a recovery of ν—it is informative precisely because nothing forced it. Where reference dependence is strong—the loss domain—anchoring blows up into loss aversion, the asymmetry the symmetric model deliberately abstracts from. That asymmetry is also the measured counterpart of the enrichment that Section OA.A.7 floats as the repair for its one non-match—a loss-averse reference that bites in cold markets, when optimal reservations fall below the purchase price: the asymmetry that repair posits is present in the data, though the triple interaction it would test (cold months $\times$ loss domain $\times$ uncertainty) stays underpowered at this sample. That the model can abstract from it is safe for the paper's claim: Proposition 4 shows the disagreement-induced delay survives a fully optimal, unanchored seller, so the illiquidity of hard-to-value homes does not rest on the anchor that this result confirms is present. Reference dependence governs the level of the ask along the gain/loss dimension; valuation uncertainty governs illiquidity along the orthogonal how-hard-to-value dimension—the foil of Section 5.4.3 sharpened into a measurement. Three caveats bound the evidence: the pairs identify the same unit or a near-identical twin rather than a guaranteed repeat sale; the current-value benchmark is built from sold comparables and inherits their selection; and the anchor's decay with staleness and its interaction with uncertainty are not robust across specifications, so I do not report them.

OA.A.9 An estimated extension: the visit gate and the withdrawal margin

The calibration above holds three moments exactly; this subsection asks the structure to earn more than it is given. Two additions make that possible. The first repairs a quantity-side inconsistency the lean model cannot escape: with a single exponential sale hazard, a mean

Table OA.2: Reference Dependence at Relisting: the Seller's Anchor in the Loss Domain

	(1) Relist markup over curr. value	(2) $\log$ days on market	(3) P(sale)
Nominal loss (mkt. below purchase)	$+0.810^{***}$	$+0.269^{**}$	-0.094^{***}
	(0.023)	(0.119)	(0.033)
Nominal gain (mkt. above purchase)	-0.199^{***}	$+0.104$	-0.019
	(0.023)	(0.068)	(0.022)
Community + relist-month FE (additive)	Yes	Yes	Yes
Observations	19,878	16,461	19,878

Notes. Clean-sequential quasi-repeat-sale pairs (the same unit or its identical twin, relisted after an earlier sale; Section 5.1.4). The reference point is the relister's own earlier deal price; V_{now} is that price plus the community's log price appreciation between the two dates, so it differences out fixed quality and the local price path. Column (1) regresses the relist markup, $\log(\text{ask}) - \log V_{\text{now}}$, on the nominal loss ($\max\{0, -\Delta\}$) and the nominal gain ($\max\{0, \Delta\}$), where Δ is community appreciation since purchase; a loss coefficient near $+0.8$ means the seller passes only about a fifth of the market's decline into the ask and anchors on the rest. Columns (2)–(3) repeat the split for $\log$ days on market (sold relists, real posting dates) and the probability of sale. Additive community and relist-month fixed effects throughout; standard errors clustered by community. $^{*}, ^{**}, ^{***}$: $p < 0.10, 0.05, 0.01$.

sold spell of 101.6 days[20] forces a hazard of 3.6 per year, which implies that essentially every listing sells within a year—whereas only 73 percent do. I therefore add a *withdrawal margin*: listings exit unsold at rate δ_w, so the total hazard is $H = \lambda q + \delta_w$, the mean sold spell is $1/H$, and the probability of selling within a year is $(\lambda q/H)(1 - e^{-H})$. The two level moments pin the two rates exactly: a sale hazard of 2.71 and a withdrawal hazard of 0.88 per year—a listed home is roughly three times as likely to sell as to withdraw in a given month. One implication is left untargeted and lands close: the implied share of listings that ever sell is 0.755, against 0.775 in the data. The second addition formalizes the funnel of Online Appendix OA.B.5 quantitatively: a buyer observes an interim signal $m \sim \mathcal{N}(\mu - g,\, \rho_m s^2)$ with $V = m + e$, $e \sim \mathcal{N}(0, (1 - \rho_m)s^2)$, pays an effective participation cost χ to visit, visits iff $(1 - \alpha)\, \mathbb{E}[(V - P^S)^+ \mid m] \geq \chi$, and buys iff $V \geq P^S$—so the sale probability per meeting is the product of the visit margin and the conversion margin, and the model now speaks to the showing counts directly. Funnel levels are absorbed by free scale constants; the model is asked only about gradients.

Estimation is by weighted minimum distance on five clean-window gradients: the deal price, time on market, the within-a-year sale probability, fixed-spell followers, and fixed-spell showings. Three parameters are estimated—the risk-discount increment Δg, the dispersion increment Δs, and the visit cost χ—so the system is over-identified by two. The anchoring

[20]The mean sold spell on the clean window (cohorts posted July 2016–December 2019, FVR ≤ 1). The calibration of Online Appendix OA.A.7 pins its meeting rate with the rounder seventy-day full-sample convention; the windows, and hence the levels, differ. The estimation disciplines its levels with the withdrawal margin, so the two exercises use the level information differently by design.

weight is fixed at the calibration's $\nu = 0.78$ (its delta-method interval is $[0.68, 0.88]$), and the signal share at $\rho_m = 0.5$ (χ and ρ_m govern gate selectivity nearly collinearly, so the data cannot separate them). The weighting matrix is the inverse of the moments' clustered variances, with cross-moment covariances set to zero; parameter variances are $(G'WG)^{-1}$.

Table OA.3 reports the estimates and Table OA.4 the fit. Three results. First, the risk discount is estimated at $\Delta g = 0.98$ percent of value (SE 0.17)—almost exactly the calibration's 0.99 percent, recovered here from a different architecture and moment set, which is reassuring about neither number being an artifact of its derivation. Second, the visit cost is precisely estimated at $\chi = 2.3$ percent of value. Read literally, that is far too large for the out-of-pocket cost of attending a showing. It instead measures a *participation threshold*: the expected bargaining surplus a buyer requires before initiating a physical inspection, bundling time, effort, and the seriousness the listing agent screens for. At that threshold the visit gate, not the acceptance margin, is the binding selection stage—which is what the funnel table shows in the data. Third, every targeted moment fits within 1.8 standard errors individually, while the joint over-identification distance is $J = 7.15$ ($p = 0.03$ against a χ^2_2 reference; the weighting matrix is diagonal while the bootstrap moment correlations reach 0.53, so the exact reference is a weighted-χ^2_1 mixture and the p-value is an approximation): with clustered precision this sharp, the test can detect that three parameters do not fully span five gradients. The binding tension is the pair it cannot relax simultaneously—the time-on-market gradient (under-predicted) against the showings gradient (over-predicted). Leave-one-moment-out re-estimation shows the rejection does not prop up the estimates and lays out the anatomy of identification: the risk discount Δg moves by at most 0.7 standard errors when any single moment is dropped (range 0.0097–0.0110), while Δs and χ each inherit their identification from one parent moment—the attention and showings gradients, respectively—and drift only when that parent is removed (χ by seven percent in level), leaving the other parameters in place. None of the paper's analytic claims depends on this extension passing: the scissors co-movement is reduced-form-identified, and the optimal-holdout and Miller-limit results are proved analytically. The over-identified fit is a disciplining stress-test I chose to expose, and a three-parameter system that rejects on a single funnel margin is more informative than an exactly-identified one that cannot reject.

The sixth row of Table OA.4 is untargeted, and what it measures deserves care, because the structural accounting corrects a reading the raw counts invite. Accumulated showings before a sale are inspections per unit of market time compounded over the spell, so their gradient is approximately the fixed-spell showings gradient plus the duration response—in the data, -4.4 percent plus a duration elasticity of 0.48 times $+5.3$ percent, which reproduces the measured -1.9 percent almost exactly. The model's counterpart—visit inflow times the expected sold spell—delivers -1.2 percent ($t = 0.96$): an accumulation identity the fitted model passes without being asked. The naive reading of the falling count as improving conversion per visit does not survive this accounting: dividing the sale hazard by the showing inflow, the implied conversion gradient is mildly negative (≈ -2.7 percent), the case in which the discount's shading outweighs the dispersion's selection at the final stage as well (Proposition OA.8(iii) signs exactly this race). The funnel's evidentiary weight therefore rests where the cheap–costly split put it—monitoring rises while inspection falls, which a one-sided discount cannot generate—and not on a conversion improvement.

Table OA.3: Estimated Extension: Parameters

	Δg	Δs	χ
Estimate	0.0098	0.00062	0.02290
SE	(0.0017)	(0.00027)	(0.00021)

Notes. Weighted minimum-distance estimates on the five targeted moments of Table OA.4. Δg and Δs are the least-to-most uncertain increments to the risk discount and valuation dispersion, in units of value ($s_0 = 0.04$); χ is the effective participation cost of a physical showing, in units of value. Externally set: $r = c = 0.05$, $\alpha = 0.5$, $\nu = 0.78$ (from the calibration of Online Appendix OA.A.7), $\rho_m = 0.5$. The sale and withdrawal hazards (2.71 and 0.88 per year) are pinned exactly by the mean sold spell (101.6 days) and the within-a-year sale probability (0.7345). Standard errors are delta-method, $(G'WG)^{-1}$, with cross-moment covariances set to zero.

Table OA.4: Estimated Extension: Moment Fit

Moment	Data (SE)	Model	t-distance
Deal price	-0.0074 (0.0015)	-0.0076	-0.14
Time on market (log, sold)	0.0534 (0.0074)	0.0424	-1.49
P(sale within a year)	-0.0168 (0.0024)	-0.0136	1.33
Followers (log, fixed spell)	0.0146 (0.0066)	0.0154	0.13
Showings (log, fixed spell)	-0.0437 (0.0061)	-0.0545	-1.77
Showings accumulated before sale (log)†	-0.0187 (0.0069)	-0.0121	0.96

Overidentification distance on the five targeted moments: $J = 7.15$ ($\sim \chi^2_2$, $p = 0.03$)
†Untargeted: the accumulation identity (inflow $\times$ expected sold spell); see text.

Notes. Data moments are clean-window within-cell rank gradients (cohorts posted July 2016–December 2019, FVR ≤ 1): the deal price and time on market from Tables 10 and 6, the censoring-proof within-a-year sale probability, and the fixed-spell follower and showing gradients from Table OA.36. Standard errors are clustered by community. The t-distance is (model$-$data)/SE. The final row is untargeted: the model's accumulated-showings gradient is visit inflow times the expected sold spell, an accumulation identity (see text).

OA.B Theoretical-framework appendix

This appendix collects the formal content of the framework in Section 3: six empirical-mapping predictions stated as propositions with proof sketches, the truncated-normal facts they use, the CARA–Normal certainty-equivalent and Nash-bargaining derivations, the information microfoundation that delivers both channels from one primitive (Section OA.B.4), full proofs, and the auxiliary algebra connecting the structural mapping of Section 6.4 to the pricing equation. The linearization of the asinh transform is in Appendix A and is cross-referenced where needed. Throughout, $\kappa > 0$ is the information-channel scaling of Equation (OA.12): measured uncertainty σ_P^2 enters the resale variance buyers perceive as $\kappa\,\sigma_P^2$, so larger κ means a larger perceived risk per unit of measured dispersion; the main text normalizes $\kappa = 1$.

OA.B.1 Propositions

The framework's three headline results are stated in the main text—Propositions 3 (the information primitive), 4 (optimal holdout), and 1 (the scissors). This subsection states the remaining empirical-mapping propositions with one-paragraph proof sketches; full proofs for all of them follow in Sections OA.B.7–OA.B.12. Each proposition maps directly to an empirical test in Section 5. The two listing propositions share one listing rule: the seller posts

$$P^{\text{list}}_{h,t} = \theta\, g\big(\mathbb{E}[P^*_{h,t}]\big) + (1-\theta)\, B_{c,t-1} + \eta_{h,t}, \qquad \theta \in (0,1), \tag{OA.7}$$

a partial adjustment between the model price (through a monotone transform g, $g' > 0$) and the lagged community benchmark $B_{c,t-1}$, with listing noise $\eta_{h,t}$ independent of σ_P^2. Interior θ is the listing analogue of the calibrated anchoring weight ν (Online Appendix OA.A.7): $\theta > 0$ makes the posted price load on the unit's expected transaction value, $\theta < 1$ keeps it anchored on comparables.

Proposition OA.3 (Uncertainty discount). *For any buyer with risk aversion $\gamma > 0$ and information-channel scaling $\kappa > 0$, and holding the buyer met fixed, the transaction price is strictly decreasing in valuation uncertainty:*

$$\frac{\partial P^*}{\partial \sigma_P^2} = -(1-\alpha)\,\frac{\gamma\,\kappa}{2}\,\rho^2 < 0.$$

Sketch. Differentiating the buyer's certainty equivalent in (8) with respect to σ_P^2 and substituting into the Nash solution (9) delivers the result; the slope is negative because $\alpha < 1$, $\gamma > 0$, $\kappa > 0$, and $\rho > 0$. The empirical counterpart is the negative coefficient on valuation uncertainty in the price regressions: within community–month, listings with higher ex-ante uncertainty trade at a discount in listing prices (Table OA.16) and deal prices (Table OA.15).

Proposition OA.4 (Listing-implied pricing wedge). *Define the listing-implied pricing wedge as*

$$\text{ExpRet}^{(1m)}_{h,t} \equiv \operatorname{asinh}\big(\widetilde{P}^{\text{list}}_{h,t}\big) - \operatorname{asinh}\big(\bar{P}^{\text{deal}}_{c(h),t-1}\big),$$

where $\widetilde{P}^{\text{list}}_{h,t}$ is the quality-adjusted listing price and $\bar{P}^{\text{deal}}_{c(h),t-1}$ is the lagged community deal benchmark. Under the listing rule (OA.7) with $\theta > 0$, $\text{ExpRet}^{(1m)}$ is decreasing in σ_P^2.

Sketch. Writing $\widetilde{P}^{\text{list}}_{h,t} = \mathbb{E}[P^*_{h,t}] + \eta_{h,t}$ and noting that the lagged benchmark is fixed at time t and that $\operatorname{asinh}(\cdot)$ is strictly increasing, the sign of $\partial\,\text{ExpRet}^{(1m)}/\partial\sigma_P^2$ follows Proposition OA.3. A formal linearization of the asinh transform around typical values is in Appendix A; the closed-form derivation under the full pricing equation is in Online Appendix OA.B. The empirical counterpart is the negative loading of the listing-implied pricing wedge on valuation uncertainty within community–month.

Proposition OA.5 (Listing–transaction wedge). *Under the listing rule* (OA.7) *with $\theta < 1$, and with the listing transform in price units ($g'(P_0) = 1$ at the linearization point), the expected listing–deal wedge $\mathcal{W}_h \equiv P^{\text{list}}_h - \mathbb{E}[P^*_h]$ is increasing in valuation uncertainty. (Empirically the wedge is reported with the opposite sign convention, as the negotiated outcome deal-minus-list, which falls in uncertainty under the dominance condition of Proposition 1; Section 5.2.1.)*

Sketch. The anchored component $(1-\theta)B_{c,t-1}$ of the posted price does not absorb the buyer's risk discount, while the expected transaction price carries it in full (Proposition OA.3); with the listing transform in price units ($g'(P_0) = 1$ locally), the wedge slope is $\partial\mathcal{W}_h/\partial\sigma_P^2 = (1-\theta)(1-\alpha)\frac{\bar{\gamma}\kappa}{2}\rho^2 > 0$, so the listing–deal wedge widens in valuation uncertainty. The realized wedge conditions on sale; under the dominance condition of Proposition 1 the conditional mean price also falls in σ_P^2, preserving the sign. The time-adjusted price analysis in Table OA.25 provides a complementary test, in that stale asking prices are penalized more in pricing wedges, consistent with tighter listing discipline.

Proposition OA.6 (Attention: accumulation and per-exposure evaluation). *Let buyers arrive at house h according to a Poisson process with meeting rate λ_t, and let $q(\sigma_P^2)$ be the per-meeting acceptance probability, decreasing in valuation uncertainty under the maintained condition* (OA.3) *(Online Appendix OA.A). Then:* (a) *the expected search a seller must attract before clearing, proportional to $1/q(\sigma_P^2)$, is increasing in σ_P^2—and rises with the risk discount even when buyers do not disagree;* (b) *per unit of exposure, committed evaluation $A = c_A\,\text{Var}_V$ rises in σ_P^2 if and only if disagreement responds, $\sigma_b^{2\prime} > 0$. The accumulated follower stock, $\mathcal{A} = A/q$ in Equation* (17)*, therefore rises through both channels; its spell-adjusted component isolates* (b).

Sketch. Since each meeting clears with probability $q(\sigma_P^2)$, the number of meetings before a sale is geometric with mean $1/q(\sigma_P^2)$, and $\partial q/\partial\sigma_P^2 < 0$ delivers part (a); its empirical counterpart is the follower stock a listing accumulates before clearing (Table 6). Part (b) is Equation (16): a downward shift in the mean of valuations gives a buyer nothing more to evaluate, a wider spread does; its empirical counterpart is the spell-adjusted attention gradient, with casual views flat per unit of exposure (Online Appendix OA.F.11). The level-count attention regressions are in Table OA.26.

Proposition OA.7 (Price dispersion)**.** *Under the valuation distribution in Equation* (10) *with normal disagreement, the variance of transaction prices conditional on sale is strictly increasing in σ_P^2 through the dispersion channel; with both channels active, and for $z \geq 0$ (at most half of meetings clear), $\partial \operatorname{Var}[P^* \mid \mathit{sale}]/\partial \sigma_P^2 > 0$ whenever the disagreement response outweighs the discount-induced tightening of the accepting tail,*

$$\sigma_V' \left[2\big(1 - \delta(z)\big) + z\, \delta'(z) \right] \;>\; \delta'(z)\, \frac{\rho^2 \bar{\gamma} \kappa}{2}$$

(notation of Lemma OA.2), a condition automatic as $\bar{\gamma}\kappa \to 0$. At the conservative calibration—whose dispersion increment is pinned by the attention stock—the two terms nearly cancel, so the observed widening of the negotiated spread (column 2 of Table 10) itself indicates a disagreement response above that attention-implied lower bound. Harder-to-value units then transact at a wider spread of prices within a community and month.

Sketch. The transaction price is the affine Nash map of the buyer valuation (Equation (9)), so its conditional variance is $(1-\alpha)^2 \operatorname{Var}[V \mid V \geq P^S]$, and Equation (OA.9) decomposes the response to σ_P^2 into the dispersion term, signed positive for $z \geq 0$ by Lemma OA.2(iii), and the discount's selection term, which the stated condition bounds. Because the squared price residual is by construction the uncertainty measure, the empirical counterpart is the centered dispersion of the list-to-deal wedge, which differences out the price level, in column (2) of Table 10.

Proposition 1 is stated in the main text. Its sketch is immediate from Equations (10), (16), (17), and the truncated-moment decompositions (OA.8)–(OA.9): at a fixed spell, the mean price loads on the risk-pricing index $\bar{\gamma}\kappa$ and the disagreement margins on the belief-dispersion response $\sigma_b^{2\prime}$, so each observable isolates one primitive; the full proof is in Section OA.B.12. The empirical counterpart is the contrast between column (1) and the spell-adjusted columns (2)–(3) of Table 10 (Online Appendix OA.F.11).

The search-and-matching structure underlying time on market and the acceptance margin is developed in Online Appendix OA.A; its empirical counterparts—sale probability, time on market, and the state-dependence of both—are tested in Section 5.

OA.B.2 Truncated-normal facts

The selection results above turn on how the mean and variance of a normal distribution truncated at the seller's reservation respond to its location and scale. The following facts are used throughout.

Lemma OA.2. *Let $M(z) = \phi(z)/\big(1 - \Phi(z)\big)$, the inverse Mills ratio, and $\delta(z) = M(z)\big(M(z) - z\big)$, so that for $V \sim \mathcal{N}(\mu, \sigma^2)$ and a threshold c, with $z = (c-\mu)/\sigma$, $\mathbb{E}[V \mid V \geq c] = \mu + \sigma M(z)$ and $\operatorname{Var}[V \mid V \geq c] = \sigma^2\big(1 - \delta(z)\big)$. Then: (i) $0 < \delta(z) < 1$ and $\delta'(z) > 0$ for all z; (ii)*

$M(z) - z\,\delta(z) > 0$ for all z, so the truncated mean is strictly increasing in σ at a fixed mean and threshold; (iii) $2\big(1-\delta(z)\big) + z\,\delta'(z) > 0$ for $z \geq 0$, so the truncated variance is strictly increasing in σ whenever the threshold does not lie below the mean valuation—the relevant case when at most half of meetings clear.

Proof. Part (i) is classical (Sampford, 1953). For (ii), use $M' = \delta$ and write $M - z\delta = M\,(1 - zM + z^2)$. For $z \leq 0$ the bracket is at least one. For $z > 0$, the Mills-ratio bound $M(z) < \frac{1}{2}\big(z + \sqrt{z^2+4}\big)$ (Sampford, 1953) gives $1 - zM + z^2 > \frac{1}{2}\big(2 + z^2 - z\sqrt{z^2+4}\big) > 0$, the last step because $(2+z^2)^2 - z^2(z^2+4) = 4 > 0$. For (iii), both terms are nonnegative when $z \geq 0$ by (i), and the first is strictly positive since $\delta < 1$. (Numerically the inequality in (iii) holds for all z; only $z \geq 0$ is used.) □

Differentiating the truncated moments along the model's two channels—the mean $\mu_V(\sigma_P^2)$ falling at rate $\rho^2\bar{\gamma}\kappa/2$ and the scale $\sigma_V(\sigma_P^2)$ rising at rate $\sigma_V' > 0$—gives the decompositions used in Propositions OA.7 and 1:

$$\frac{d\,\mathbb{E}[V \mid V \geq c]}{d\sigma_P^2} = -\frac{\rho^2\bar{\gamma}\kappa}{2}\big(1-\delta(z)\big) + \sigma_V'\,\big(M(z) - z\,\delta(z)\big), \tag{OA.8}$$

$$\frac{d\,\mathrm{Var}[V \mid V \geq c]}{d\sigma_P^2} = \sigma_V\Big\{\sigma_V'\,\big[\,2\big(1-\delta(z)\big) + z\,\delta'(z)\,\big] - \delta'(z)\,\frac{\rho^2\bar{\gamma}\kappa}{2}\Big\}. \tag{OA.9}$$

In each expression the first term is the dispersion channel, signed by the Lemma, and the second is the discount channel's selection effect. When the seller's reservation itself responds to σ_P^2—the partially optimal anchoring of the calibration—the discount rate $\rho^2\bar{\gamma}\kappa/2$ in both expressions is replaced by the net threshold-gap movement $d(P^S - \mu_V)/d\sigma_P^2$, which partial optimality compresses well below the raw discount; the qualitative statements are unchanged.

OA.B.3 CARA–Normal certainty equivalents and private valuation

Each buyer is risk averse with CARA utility over terminal wealth,

$$u_\tau(W) = -\exp(-\gamma_\tau W), \qquad \tau \in \{E, S\}.$$

For any normally distributed payoff $X \sim \mathcal{N}(\mu, \sigma^2)$,

$$\mathbb{E}[u_\tau(X)] = \mathbb{E}\big[-\exp(-\gamma_\tau X)\big] = -\exp\!\big(-\gamma_\tau\mu + \tfrac{1}{2}\gamma_\tau^2\sigma^2\big),$$

using the moment-generating function of the normal distribution. The certainty equivalent $\mathrm{CE}_\tau(X)$ solves $u_\tau(\mathrm{CE}_\tau(X)) = \mathbb{E}[u_\tau(X)]$, which gives the exact expression

$$\mathrm{CE}_\tau(X) = \mu - \frac{\gamma_\tau}{2}\sigma^2, \tag{OA.10}$$

identical to Equation (7) in the main text. This is an exact identity for the CARA–Normal pair, not an approximation. The buyer's total payoff from purchasing house h at price P_0 is the random variable

$$W_h = -P_0 + \rho\, r_{h,1} + \rho^2 P_{h,2},$$

where $r_{h,1} \sim \mathcal{N}(\mu_r, \sigma_r^2)$ and $P_{h,2} \sim \mathcal{N}(\mu_P, \sigma_P^2)$ are independent. Applying the certainty-equivalent identity (OA.10) to each period's payoff in turn and discounting period-by-period gives the buyer's private valuation,

$$PV_\tau^B = \rho\, \mathrm{CE}_\tau(r_{h,1}) + \rho^2\, \mathrm{CE}_\tau(P_{h,2}) = \rho\Big[\mu_r - \tfrac{\gamma_\tau}{2}\sigma_r^2\Big] + \rho^2\Big[\mu_P - \tfrac{\gamma_\tau}{2}\sigma_P^2\Big], \qquad \text{(OA.11)}$$

which is Equation (8) of the main text. Treating the two periods' payoffs as separate certainty-equivalent units (each priced at the buyer's marginal rate of risk substitution within its period and then discounted to the trade date) is standard in the housing literature (Amaral, 2024); the alternative of forming the certainty equivalent of the discounted sum gives the same sign of $\partial PV_\tau^B/\partial\sigma_P^2$ and therefore does not affect any of the propositions stated below. Under the information-channel parameterisation, the resale-payoff variance entering the certainty equivalent is $\widetilde{\sigma}_P^2 = \kappa\,\sigma_P^2$ rather than σ_P^2—with κ derived in Section OA.B.4 as the share of measured dispersion that private inspection cannot resolve—while the rental-income variance σ_r^2 remains unscaled because rental cash flows are observed directly. The private valuation under perceived uncertainty is therefore

$$PV_\tau^B = \rho\Big[\mu_r - \tfrac{\gamma_\tau}{2}\sigma_r^2\Big] + \rho^2\Big[\mu_P - \tfrac{\gamma_\tau}{2}\,\kappa\,\sigma_P^2\Big]. \qquad \text{(OA.12)}$$

OA.B.4 An information microfoundation: both channels from one primitive

The framework takes two objects as primitives: the perceived-risk scaling κ and the disagreement response $\sigma_b^{2\prime} > 0$. This subsection derives both—together with the attention demand of Equation (16)—from a single Gaussian signal-extraction structure, the housing analogue of Quan and Quigley (1991).

Environment. Within a community–month (the common factor is absorbed by the fixed effects), a unit's idiosyncratic value is $\xi_h \sim \mathcal{N}(0, \sigma_\xi^2)$. Public information—the comparable sales the platform displays—carries a signal $x_h = \xi_h + \varepsilon_h$ with precision τ_c, high when close comparables are dense and low when they are scarce; total public precision about ξ_h is $\tau_{\text{pub}} = \sigma_\xi^{-2} + \tau_c$. Measured valuation uncertainty is the residual variance of the unit's value given public information alone,

$$u \;\equiv\; \mathrm{Var}(\xi_h \mid x_h) \;=\; \frac{1}{\tau_{\text{pub}}},$$

the population counterpart of the squared hedonic residual. Buyer i supplements the public signal with a private assessment $s_i = \xi_h + \eta_i$, $\eta_i \sim \mathcal{N}(0, 1/\tau_\eta)$ i.i.d. across buyers (inspection, expertise, taste-free judgment). Resale value inherits the unit's idiosyncratic component—the buyer of a hard-to-value home resells it as a hard-to-value home—so beliefs about the resale price $P_{h,2}$ disperse with posterior beliefs about ξ_h.

Proposition 3 in the main text states the result: in the notation here, perceived residual risk is $\widetilde{\sigma}^2 \equiv \mathrm{Var}(\xi_h \mid x_h, s_i)$, belief dispersion is $\sigma_b^2 \equiv \mathrm{Var}_i\big(\mathbb{E}[\xi_h \mid x_h, s_i]\big) = \tau_\eta u^2/(1+\tau_\eta u)^2$ with posterior weight $w = \tau_\eta/(\tau_{\text{pub}} + \tau_\eta)$ on the private signal, the κ it microfounds is the scaling of Equation (OA.12), and part (iii)'s bound $\epsilon < 1/(\tau_\eta u)$ is the perceived-risk margin's, which implies the dispersion margin's weaker bound $2/(\tau_\eta u - 1)$ whenever $\tau_\eta u > 1$ (the dispersion margin is unrestricted when $\tau_\eta u \le 1$).

Proof. With Gaussian prior and signals, posterior precision is $\tau_{\text{pub}} + \tau_\eta$, giving (i); the posterior mean is $m_i = (1-w)\,\mathbb{E}[\xi_h \mid x_h] + w\, s_i$ with $w = \tau_\eta/(\tau_{\text{pub}} + \tau_\eta)$, and across buyers only s_i varies, with conditional variance $1/\tau_\eta$, so $\mathrm{Var}_i(m_i) = w^2/\tau_\eta = \tau_\eta/(\tau_{\text{pub}} + \tau_\eta)^2$, giving (ii); monotonicity in $u = 1/\tau_{\text{pub}}$ at fixed τ_η is immediate from both expressions. The endogenous-precision claim of (iii), and the form of Equation (16) it delivers, are established in the paragraph that follows. For the elasticity bounds, write $a = \tau_\eta^*(u)$ and $\epsilon = d\ln a/d\ln u$; logarithmic differentiation of $\widetilde{\sigma}^2 = u/(1+au)$ and $\sigma_b^2 = au^2/(1+au)^2$ gives

$$\frac{d\ln\widetilde{\sigma}^2}{d\ln u} = \frac{1-\epsilon\, au}{1+au}, \qquad \frac{d\ln\sigma_b^2}{d\ln u} = \frac{2+\epsilon\,(1-au)}{1+au}.$$

The first is positive iff $\epsilon < 1/(au)$; the second is positive for any $\epsilon \ge 0$ when $au \le 1$ and iff $\epsilon < 2/(au-1)$ when $au > 1$. Since $1/(au) < 2/(au-1)$ for $au > 1$, the single condition $\epsilon < 1/(\tau_\eta u)$ suffices for both. $\square$

Endogenous attention and the form of Equation (16). A buyer chooses monitoring effort $e \ge 0$ that delivers private precision $\tau_\eta(e)$, $\tau_\eta' > 0$, at strictly convex cost $C(e)$, $C'' > 0$. Resolving where her own valuation falls is worth more the wider the spread of valuations she might end up holding, so the gross value of effort is proportional to that dispersion, $G(e; \mathrm{Var}_V) = \mathrm{Var}_V \cdot v(e)$ with $v' > 0$: with no dispersion there is nothing to resolve. The buyer solves $\max_{e\ge 0}\ \mathrm{Var}_V\ v(e) - C(e)$. The cross-partial $\partial^2 G/\partial e\,\partial\,\mathrm{Var}_V = v'(e) > 0$ makes the objective supermodular in (e, Var_V), so by Topkis the optimum e^* is nondecreasing in Var_V; since $\mathrm{Var}_V = \rho^4\sigma_b^2(u)$ rises in u by (ii), acquired precision $\tau_\eta^* = \tau_\eta(e^*)$ is nondecreasing in u, which is (iii), and its elasticity $\epsilon = d\ln\tau_\eta^*/d\ln u$ inherits the bound above. The first-order condition $\mathrm{Var}_V\ v'(e^*) = C'(e^*)$ pins the level: with the leading specification $v(e) = e$ and $C(e) = e^2/2c_A$ it gives $e^* = c_A\ \mathrm{Var}_V$ exactly, and for general (v, C) a first-order expansion about $\mathrm{Var}_V = 0$ gives $e^* = c_A\ \mathrm{Var}_V + o(\mathrm{Var}_V)$ with $c_A = v'(0)/C''(0)$. Committed attention is this effort, $A = e^*$, which is Equation (16).

Four remarks. *First*, the main text's affine pricing form treats κ as a constant; here $\widetilde{\sigma}^2(u) = u/(1+\tau_\eta u)$ is increasing and concave, with the affine form its small-$\tau_\eta u$ approximation. Every signed prediction in the paper requires only $\widetilde{\sigma}^{2\prime}(u) > 0$ and $\sigma_b^{2\prime}(u) > 0$, which hold globally, so the propositions are unaffected. *Second*, the structure implies a curvature ordering: for small u the discount margin is linear in u while disagreement is quadratic ($\sigma_b^2 = \tau_\eta \widetilde{\sigma}^4$), and the ratio of dispersion to perceived risk, $\sigma_b^2/\widetilde{\sigma}^2 = \tau_\eta \widetilde{\sigma}^2$, rises with valuation difficulty—the disagreement channel should gain on the discount channel among the hardest-to-value units, a falsifiable shape restriction left for future work. *Third*, the primitive τ_{pub} is the precision of the comparables signal, so the choice-free comparable-scarcity count of Online Appendix OA.D.8 is its direct empirical counterpart: the horse race there—scarcity and the residual rank each predicting attention, essentially orthogonally—reads as two facets of the same primitive, the thinness of the public signal and the realized residual it leaves. *Fourth*, adding buyer-specific match utility on top of the common-value component ξ_h would widen Var_V further without altering any comparative static; the derivation deliberately generates disagreement about a common value, divergence of opinion in the sense of Miller (1977), with no taste heterogeneity required.

OA.B.5 The engagement funnel as nested thresholds

Badarinza, Balasubramaniam and Ramadorai (2024) estimate the housing matching function across observed stages—online views, contact requests, transactions—and find congestion at each stage, with match-quality variation carrying much of the load. The platform funnel here (browse $\to$ follow $\to$ physical showing $\to$ sale) has the same architecture, and the information block above delivers it as nested thresholds on a single latent object, the buyer's interim posterior m_i, with each stage's conversion rate—not its stock—as the structural margin.

Browse $\to$ follow. Committed monitoring is nearly free and pays in information: by Equation (16) and Proposition 3(iii), the per-exposure follow rate rises with disagreement and responds to neither the risk discount nor the meeting rate.

Follow $\to$ visit. A showing costs $\chi > 0$ and is taken when the expected surplus from the ensuing bargain clears it: buyer i visits iff $(1-\alpha)\ \mathbb{E}\big[(V_i - P^S)^+ \,\big|\, m_i\big] \geq \chi$, an option-value object increasing in m_i and in residual dispersion. The discount shifts the distribution of m_i down, thinning the set of buyers for whom the visit is worth its cost; dispersion raises both the tail mass above the visit threshold and each visitor's option value. The conversion rate per follower therefore falls with uncertainty exactly when the discount's shading outweighs the dispersion's option value—the visiting analogue of the dominance condition in Proposition 1—which is the empirically realized case: showings per follower fall sharply with uncertainty (-0.046, $t = -8.8$; Online Appendix OA.F.12).

Visit $\to$ transact. The final conversion is the per-meeting acceptance probability q of Equation (OA.2), falling with uncertainty under the maintained condition (OA.3).

The two costly stages interact through selection, and the recovered showing counts let the data sign the interaction.

Proposition OA.8 (Selection at the visit gate)**.** *Let the interim posterior be normal across buyers, $m_i \sim \mathcal{N}(\mu_m, \sigma_m^2)$ with $\sigma_m^2 = \sigma_b^2$ (Proposition 3). Buyer i visits iff her interim posterior clears the gate, $m_i \geq v^*$, learns $V_i = m_i + e_i$ on inspection ($e_i \perp m_i$, $e_i \sim \mathcal{N}(0, \sigma_e^2)$), and buys iff $V_i \geq P^S$; let $C \equiv \Pr[V_i \geq P^S \mid m_i \geq v^*]$ be conversion per visit, and suppose the gate is selective ($v^* > \mu_m$). Then:* (i) *a pure discount (a downward shift of μ_m) lowers both the visit propensity and C, so showings accumulated before sale, $1/C$, rise;* (ii) *a pure disagreement increase (a rise in σ_m) raises C—the visiting pool is selected from a wider upper tail—so showings before sale fall;* (iii) *the observed sign of the showings-to-clear gradient therefore signs the race between the channels at the conversion stage.*

Proof. (i) The gate mass $\Pr[m_i \geq v^*]$ falls directly. The truncated pool $m_i \mid m_i \geq v^*$ is stochastically increasing in its location parameter (truncation of a log-concave location family), and $C = \mathbb{E}\big[1 - \Phi\big((P^S - m_i)/\sigma_e\big) \,\big|\, m_i \geq v^*\big]$ is an increasing functional of that pool, so C falls as μ_m falls. (ii) Write the visitors' excess as $(m_i - v^*) \mid m_i \geq v^* \overset{d}{=} \sigma_m \left(Z - z_v \mid Z \geq z_v\right)$ with $z_v = (v^* - \mu_m)/\sigma_m$. A rise in σ_m scales the excess up and, because $v^* > \mu_m$, lowers z_v; since $(Z - z \mid Z \geq z)$ is stochastically decreasing in z for the normal, both movements raise the visitors' valuation distribution in first-order stochastic dominance, so C rises. (iii) is immediate from (i)–(ii). □

Empirically, showings accumulated before sale fall with uncertainty (-2.5 percent on the within-cell rank; column (5) of Table 10; Online Appendix OA.F.12). That count, however, is an accumulation object—inspections per unit of market time compounded over the spell—and the fixed-spell decomposition reproduces it almost exactly (-4.4 percent inflow plus a 0.48 duration elasticity times the $+5.3$ percent spell response; Online Appendix OA.F.12), so it carries the gate margin and the spell rather than a separate conversion fingerprint. Dividing the sale hazard by the showing inflow, the implied conversion per visit moves mildly against harder-to-value homes: by part (iii), the discount's shading outweighs the dispersion's selection at the final stage as well, and the funnel's evidentiary weight rests on the cheap–costly split—monitoring up, inspection down—rather than on conversion. The gate parameters are not separately identified by the cross-section, so the proposition contributes signs, not magnitudes.

Two readings. First, the funnel's signed pattern—the cheap stage up, both costly conversions down—is the model's split (the inspection paragraph of Section 6.3) stated stage by stage, and it is the pattern the recovered showing counts confirm. Second, conversion rates are the structural objects and stocks are conversion times exposure (Equation (17)); reading stocks as if they were rates conflates the funnel with the spell—the congestion point of Badarinza, Balasubramaniam and Ramadorai (2024) at the market level, and the duration-mechanics diagnostic of Online Appendix OA.F.11 at the listing level. The two lessons are the same lesson.

OA.B.6 Nash bargaining algebra

Conditional on meeting, the seller and buyer solve the Nash bargaining problem

$$\max_{V_B, V_S} \; V_B^{\alpha} V_S^{1-\alpha} \quad \text{s.t.} \quad V_B + V_S = PV_\tau^B - P_{h,t}^S, \tag{OA.13}$$

where $P_{h,t}^S$ is the seller's reservation value and $\alpha \in (0,1)$ is the buyer's bargaining weight. The first-order condition with respect to V_B at the binding surplus constraint gives $V_B = \alpha\,(PV_\tau^B - P_{h,t}^S)$. The transaction price P^* satisfies $V_B = PV_\tau^B - P^*$, so $P^* = PV_\tau^B - V_B = (1-\alpha)PV_\tau^B + \alpha P_{h,t}^S$,

$$P^*(\tau) = (1-\alpha)\, PV_\tau^B + \alpha\, P_{h,t}^S. \tag{OA.14}$$

Substituting (OA.12) into (OA.14) gives the type-conditional transaction price,

$$P^*(\tau) = (1-\alpha)\Big\{\rho\Big[\mu_r - \tfrac{\gamma_\tau}{2}\sigma_r^2\Big] + \rho^2\Big[\mu_P - \tfrac{\gamma_\tau}{2}\,\kappa\,\sigma_P^2\Big]\Big\} + \alpha\, P_{h,t}^S. \tag{OA.15}$$

Taking expectations over which buyer the seller meets and writing $\bar{\gamma}$ for the average risk aversion of the buyers a listing meets yields the expected transaction price,

$$\mathbb{E}[P_{h,t}^*] = (1-\alpha)\Big\{\rho\mu_r + \rho^2\mu_P - \frac{\bar{\gamma}}{2}\big(\rho\sigma_r^2 + \rho^2\kappa\,\sigma_P^2\big)\Big\} + \alpha\, P_{h,t}^S, \tag{OA.16}$$

which is Equation (11) of the main text with κ normalized to one and the σ_P^2-independent terms collected in $a_{h,t}$.

OA.B.7 Proof of Proposition OA.3

Proof of Proposition OA.3. In (OA.15), only the term $-(1-\alpha)\rho^2(\gamma_\tau\kappa/2)\sigma_P^2$ depends on σ_P^2. Differentiating yields

$$\frac{\partial P^*}{\partial \sigma_P^2} = -(1-\alpha)\,\frac{\gamma_\tau\,\kappa}{2}\,\rho^2,$$

which is strictly negative since $\alpha < 1$, $\gamma_\tau > 0$, $\kappa > 0$, and $\rho > 0$. □

OA.B.8 Proof of Proposition OA.4

Proof of Proposition OA.4. Under the listing rule (OA.7), the quality-adjusted listing price is $\widetilde{P}^{\text{list}}_{h,t} = \theta\, g(\mathbb{E}[P^*_{h,t}]) + (1-\theta)B_{c,t-1} + \eta_{h,t}$, with $g' > 0$, $\theta > 0$, and $\eta_{h,t}$ independent of σ_P^2 at the listing date. Then

$$\frac{\partial \widetilde{P}^{\text{list}}_{h,t}}{\partial \sigma_P^2} = \theta\, g'\big(\mathbb{E}[P^*_{h,t}]\big)\, \frac{\partial\, \mathbb{E}[P^*_{h,t}]}{\partial \sigma_P^2} < 0$$

by Proposition OA.3. Since the lagged benchmark $\bar{P}^{\text{deal}}_{c(h),t-1}$ is fixed at time t and $\operatorname{asinh}(\cdot)$ is strictly increasing, $\partial\, \mathrm{ExpRet}^{(1m)}/\partial\sigma_P^2$ has the same (negative) sign. A formal linearization of the asinh transform around typical values is provided in Appendix A, which derives the affine representation (24) of the listing-implied pricing wedge in terms of σ_P^2 and the composite slope $\Lambda \propto \bar{\gamma}\,\kappa$. □

OA.B.9 Proof of Proposition OA.5

Proof of Proposition OA.5. Define the wedge as $\mathcal{W}_h \equiv P^{\text{list}}_h - \mathbb{E}[P^*_h]$. Under the listing rule (OA.7), with the listing transform in price units ($g'(P_0) = 1$ at the linearization point) and the expected transaction price given by Equation (OA.16),

$$\frac{\partial \mathcal{W}_h}{\partial \sigma_P^2} = \big(\theta\, g'(P_0) - 1\big)\, \frac{\partial\, \mathbb{E}[P^*_h]}{\partial \sigma_P^2} = (1-\theta)\, \frac{1}{2}(1-\alpha)\rho^2\, \bar{\gamma}\, \kappa \;>\; 0 \quad \text{for } \theta < 1,$$

so the listing–deal wedge widens in valuation uncertainty: the anchored component of the asking price omits a risk discount that the transaction price applies in full, and the gap is larger precisely for harder-to-value units. The realized wedge conditions on sale. By Equation (OA.8) the conditional mean price falls in σ_P^2 under the dominance condition of Proposition 1; but the listed price falls as well, so signing the realized deal-minus-list wedge requires the stronger sufficient condition that the anchored component's omitted discount exceed the selection offset, $(\rho^2\bar{\gamma}\kappa/2)\,(1-\delta-\theta) > \sigma'_V\,(M - z\delta)$. The condition holds at the calibration; away from it, the wedge's sign is an empirical statement rather than an implication of dominance alone. □

OA.B.10 Proof of Proposition OA.6

Proof of Proposition OA.6. (a) By Equation (OA.2) and the maintained condition (OA.3), the per-meeting acceptance probability satisfies $\partial q/\partial\sigma_P^2 < 0$. Since meetings arrive as a Poisson process and clear independently with probability $q(\sigma_P^2)$, the expected number of meetings—and hence the accumulated search—before a sale is $1/q(\sigma_P^2)$, which is increasing in σ_P^2:

$$\frac{\partial}{\partial \sigma_P^2}\, \frac{1}{q(\sigma_P^2)} = -\frac{1}{q(\sigma_P^2)^2}\, \frac{\partial q}{\partial \sigma_P^2} > 0.$$

Because the risk discount alone lowers q, the accumulation rises even at $\sigma_b^{2\prime} = 0$. *(b)* $A = c_A \operatorname{Var}_V = c_A \rho^4 \sigma_b^2(\sigma_P^2)$, so $\partial A/\partial\sigma_P^2 = c_A\rho^4\sigma_b^{2\prime}$, strictly positive if and only if $\sigma_b^{2\prime} > 0$. Combining, $\mathcal{A} = A/q$ rises in σ_P^2 through both terms. The empirical counterparts are the follower stock accumulated before clearing (Table 6) and its spell-adjusted component (Online Appendix OA.F.11). $\square$

OA.B.11 Proof of Proposition OA.7

Proof of Proposition OA.7. Trade occurs iff $V \geq P^S$, and the transaction price is the affine Nash map $P^* = (1-\alpha)V + \alpha P^S$ with P^S fixed within the listing, so

$$\operatorname{Var}\big[P^* \mid \text{sale}\big] = (1-\alpha)^2 \operatorname{Var}\big[V \mid V \geq P^S\big] = (1-\alpha)^2\, \sigma_V^2 \big(1 - \delta(z)\big), \qquad z = \frac{P^S - \mu_V}{\sigma_V}.$$

Differentiating along the two channels gives Equation (OA.9). Under the dispersion channel alone ($\bar\gamma\kappa = 0$), the derivative reduces to $\sigma_V\sigma_V'\,[\,2(1-\delta(z)) + z\,\delta'(z)\,]$, strictly positive for $z \geq 0$ by Lemma OA.2(iii) (and numerically for all z). With both channels active, the derivative is positive exactly when the condition stated in the proposition holds, where the discount term applies to the net threshold-gap movement $d(P^S - \mu_V)/d\sigma_P^2$ once the reservation responds—partial optimality compresses it well below the raw discount. At the attention-pinned calibrated increments the two terms nearly cancel (the dispersion of accepted valuations moves -0.3 percent), so the clearly positive empirical gradient of the centered wedge dispersion (Table 10, column 2) reads as a disagreement response stronger than the attention stock alone implies, not as a calibration confirmation. $\square$

OA.B.12 Proof of Proposition 1

Proof of Proposition 1. Only if. Suppose $\sigma_b^{2\prime} = 0$. Then Var_V does not respond to σ_P^2, so per-exposure attention $A = c_A \operatorname{Var}_V$ is flat (Equation (16)) and, with $\sigma_V' = 0$, the conditional variance response in Equation (OA.9) is $-\sigma_V\delta'(z)\,\rho^2\bar\gamma\kappa/2 \leq 0$: at a fixed spell, neither disagreement margin rises. (The stock $\mathcal{A} = A/q$ still rises through q, which is why the proposition conditions on spell-adjusted margins.) So rising spell-adjusted disagreement requires $\sigma_b^{2\prime} > 0$. Suppose instead $\bar\gamma\kappa = 0$. Then μ_V is constant and, by Equation (OA.8), the transacted mean responds by $\sigma_V'\,(M(z) - z\,\delta(z)) > 0$ by Lemma OA.2(ii): pure dispersion strictly raises the conditional mean, so a falling transacted mean requires $\bar\gamma\kappa > 0$.

If. With $\bar\gamma\kappa > 0$ and $\sigma_b^{2\prime} > 0$, per-exposure attention rises (Equation (16)) and the negotiated dispersion rises under the condition of Proposition OA.7; the transacted mean falls exactly when the discount term in Equation (OA.8) dominates the selection term, the displayed condition, which the calibrated parameters satisfy. Under these conditions the scissors is observed. $\square$

OA.B.13 Auxiliary steps in the structural mapping

The structural mapping in Section 6.4 follows from the affine representation of the listing-implied pricing wedge derived in Appendix A. Equation (24) there gives

$$\mathrm{ExpRet}^{(1m)}_{h,t} \approx \zeta_t - \Lambda\,\sigma_P^2 + (\text{controls}) + \varepsilon_{h,t}, \qquad \Lambda = \frac{g'(P_0)}{\sqrt{1+X_0^2}}\,(1-\alpha)\,\frac{\bar{\gamma}\,\kappa}{2}\,\rho^2.$$

The within-community–month coefficient on $\mathrm{Unc}_{h,t}$ in the baseline cross-sectional specification identifies $-\Lambda$. Under the unit convention $g'(P_0) = \sqrt{1+X_0^2}$—equivalently, measuring the listing price in the units of the asinh benchmark, so that the two linearization constants cancel—the slope reads

$$\beta = -(1-\alpha)\,\frac{\bar{\gamma}\,\kappa}{2}\,\rho^2. \tag{OA.17}$$

Given the bargaining weight α and the discount factor ρ, the composite risk-pricing index $\Gamma \equiv \bar{\gamma}\,\kappa$—the average met buyer's risk aversion scaled by the information channel—is then recovered as

$$\Gamma = \bar{\gamma}\,\kappa = \frac{2|\beta|}{(1-\alpha)\rho^2}, \tag{OA.18}$$

up to that unit convention and the scaling between the squared hedonic residual and σ_P^2. The recovery is a unit normalization, not a point identification of preferences; the paper uses it only as a magnitude diagnostic (Online Appendix OA.D.1), and the magnitudes in Table OA.21 are equivalent to applying the reduced-form slope β directly.

OA.C Data appendix

This appendix collects supporting material for the Data section of the main paper (Section 2.2): the per-district summary statistics, the detailed interpretation of the hedonic uncertainty residual, the forecast regressions establishing forward-looking content of the listing-time uncertainty measure, the comparison with the ex-post dispersion measure of Amaral (2024), and the construction of platform-attention and other auxiliary variables.

OA.C.1 Listing characteristics by uncertainty quartile

Table OA.5 reports listing characteristics by quartile of valuation uncertainty for the full Beijing 2015–2021 sample. Moving from the lowest to the highest uncertainty quartile, mean floor space rises from about $82\,\mathrm{m}^2$ to about $87\,\mathrm{m}^2$ (roughly 6% larger), and mean days on market rises from 71 to 85 days (roughly 19% longer). Listing and deal prices per square meter are slightly lower in the high-uncertainty quartile (about 5% lower per m^2) than in the low-uncertainty quartile. Build year is essentially flat across quartiles, as is the

Table OA.5: Listing Characteristics by Valuation-Uncertainty Quartile

Characteristic	Q1 (low unc.)	Q2	Q3	Q4 (high unc.)
Observations	96,455	96,455	96,454	96,454
Mean valuation uncertainty	2e-04	0.0015	0.0052	0.0452
Size (m^2)	81.82	82.26	83.12	86.73
Build year	2,000.3	2,000.3	2,000.3	2,001.5
Built after 2000 (%)	59.7	60.0	61.0	67.2
Listing price (RMB/m^2)	57,576	57,333	57,116	54,615
Deal price (RMB/m^2)	56,365	56,105	55,904	53,417
Days on market	71.2	71.2	73.2	84.8
Followers	37.3	38.2	38.3	39.5
Total views	1,927	1,973	2,033	2,134
Price revisions	0.94	0.95	0.94	0.93
Share in Chaoyang (%)	33.3	32.6	32.2	30.0
Share in Haidian (%)	12.1	12.0	11.8	10.4

Notes. Listings are sorted into quartiles by `price_unc` (the squared hedonic residual). Q1 is the lowest-uncertainty quartile, Q4 the highest. Means are computed within each quartile across the full Beijing 2015–2021 sample.

average number of price revisions per listing. The district composition is broadly similar across quartiles: the Chaoyang share moves from 33% in Q1 to 30% in Q4, and the Haidian share from 12% to 10%. Taken together, the pattern is consistent with the high-uncertainty bin containing units that are harder to value—larger, slightly atypical, slower to sell—rather than units that are systematically of lower quality or located in different parts of the city.

OA.C.2 Per-district summary statistics

Table OA.6 reports the per-district sub-panels of the summary statistics referenced in Section 2.2. The Beijing-aggregate panel appears in the main text; the panels here decompose this aggregate into the three largest districts (Chaoyang, Haidian, and Fengtai), which together account for 58% of the Beijing sample.

OA.C.3 Hedonic residual: interpretation and bias

Economic content. The hedonic controls for community, year–month, size, construction year, floor category, construction type, facing, layout, and building type. After conditioning on these observables, two apartments that are "identical on paper" may still trade at very different prices because of attributes the listing's structured data do not capture: interior renovation quality, specific view orientation (park-facing vs. street-facing), noise exposure, exact floor number within a band, feng shui considerations, or seller-specific circumstances. $\text{Unc}_{i,t}$ measures how much the listing data fail to explain about this unit's price.

Table OA.6: **Summary statistics by district (Chaoyang, Haidian, Fengtai)**

	N	Mean	SD	P25	Median	P75
Chaoyang						
Listing Price (ten thousand RMB)	180 991	509	371	300	418	605
Deal Price (ten thousand RMB)	180 991	498	356	295	410	595
Size (m^2)	180 991	88.2	45.9	58.0	75.3	105.0
Construction year	180 991	2000	9.19	1994	2003	2007
Valuation Uncertainty	180 991	0.0124	0.1170	0.0006	0.0027	0.0089
Number of communities	1 477					
Haidian						
Listing Price (ten thousand RMB)	79 695	613	408	385	520	710
Deal Price (ten thousand RMB)	79 695	597	388	380	508	694
Size (m^2)	79 695	82.6	44.2	56.3	68.8	95.6
Construction year	79 695	1997	9.12	1990	1998	2003
Valuation Uncertainty	79 695	0.0130	0.2140	0.0006	0.0027	0.0088
Number of communities	1 201					
Fengtai						
Listing Price (ten thousand RMB)	78 027	390	210	256	345	475
Deal Price (ten thousand RMB)	78 027	382	203	252	339	466
Size (m^2)	78 027	79.2	32.8	56.9	71.2	94.9
Construction year	78 027	2000	7.90	1995	2002	2006
Valuation Uncertainty	78 027	0.0100	0.0734	0.0006	0.0025	0.0079
Number of communities	847					

Notes. The table reports summary statistics by district for the sample period January 2015 to January 2021. Valuation uncertainty is the squared residual from the hedonic specification described in Section 2.4.2. All prices are nominal and reported in ten-thousand RMB.

What the hedonic residual does and does not measure. A squared hedonic residual can capture several distinct economic objects. Four candidates are relevant here: (a) true idiosyncratic risk, the ex-ante uncertainty about how the market will value unobserved apartment characteristics; (b) informational opacity, the difficulty of pricing an apartment relative to its within-community comparables; (c) seller-specific mispricing, deviations from fundamentals due to anchoring, illiquidity, or urgency; or (d) omitted apartment attributes correlated with the outcome but excluded from the hedonic. Interpretation (d) is the hedonic econometrician's classic objection—residuals contain unobserved product quality, recoverable from the conditional price distribution (Bajari and Benkard, 2005); more broadly, residual-dispersion measures mix true pricing error with coefficient estimation error and the pricing of unobserved attributes (Peng and Thibodeau, 2017), the latter two of which the within-community design and the dense-cell community×month re-estimation below address. For the paper's main result, the cross-sectional uncertainty discount, the distinction among (a)–(c) is immaterial: all three are priced through the same risk channel in the model, and the within-cell slope captures the composite pricing of whatever the residual measures. Interpretation (d) is the most concerning, as omitted quality could generate a spurious correlation. Three features of the design mitigate this concern. Community fixed effects absorb permanent neighborhood quality. Floor and construction-type controls leave the discount unchanged. And results strengthen when the hedonic is re-estimated with community×month fixed effects in dense communities (Section OA.G.4). Community-level uncertainty also predicts the dispersion of future realized deal prices, supporting a forward-looking risk interpretation over a pure omitted-attribute explanation. The discount also holds under alternative operationalizations including percentile rank and community-mean uncertainty, though it is specific to the squared residual measure (the absolute-residual slope is weaker), consistent with the model's emphasis on variance-based risk pricing.

The hedonic residual $\widehat{u}_{i,t}$ is mean-zero within each community (across all months) and within each year–month (across all communities) by OLS construction. A market-wide bubble that inflates all Beijing prices in a given month is absorbed by τ_t; a permanent community premium is absorbed by $\mu_{c(i)}$. The measure therefore cannot capture these systematic price components. A community-specific time-varying mispricing component (e.g., one community experiencing a larger bubble than another) is not absorbed by additive fixed effects; however, Section OA.G.4 shows that results strengthen when the hedonic is re-estimated with community×month fixed effects on communities with dense enough data, confirming that such residual variation does not drive the findings. Squaring the residual discards the sign and retains only the magnitude of the idiosyncratic pricing deviation, yielding a symmetric, non-negative measure of valuation dispersion.

Implementation. I estimate the hedonic specification (Equation (1) of the main paper) separately in the listing and deal samples. Inference in later sections uses uncertainty measures constructed out-of-sample by cross-fitting within time blocks to mitigate mechanical overfitting of residual variation.[21]

Incidental-parameters bias in the squared residual. The squared hedonic residual is a biased estimator of the true within-community variance, with the bias inversely proportional to the number of listings per community (or community–month cell under interaction

[21]Concretely, I split the sample into folds by calendar time, estimate the hedonic model on the complement, and predict residuals for the held-out fold. Results are robust to alternative fold definitions. The cross-sectional coefficient on uncertainty remains robust to using $|\widehat{u}^m_{i,t}|$ instead of $(\widehat{u}^m_{i,t})^2$, though the squared residual remains the theoretically preferred measure. Mullainathan and Spiess (2017) discuss the broader merits of cross-fitting and machine-learning approaches for prediction-oriented tasks in economics.

FE). Communities with few listings produce inflated uncertainty estimates. Because the design compares listings within the same community and month, any community-level bias is differenced out within the cell. It motivates the minimum-community-size robustness exercise in Section OA.G.3, where restricting to denser communities strengthens the result, consistent with small-community noise in the full sample attenuating the true uncertainty discount.

OA.C.4 Forward-looking content of valuation uncertainty

This subsection reports the forecast-regression machinery summarized in one sentence in Section 2.4.2 of the main paper.

A requirement for a risk interpretation is that uncertainty should forecast future dispersion in price outcomes, not merely reflect contemporaneous noise. Using a community–month series of median listing prices (asinh-transformed), define

$$p_{c,t} \equiv \mathrm{median}_{i\in(c,t)}\ \mathrm{asinh}\big(P^{\mathrm{list}}_{i,t}\big), \qquad \overline{\mathrm{Unc}}_{c,t} \equiv \frac{1}{N_{c,t}} \sum_{i\in(c,t)} \mathrm{Unc}^{\mathrm{list}}_{i,t}, \tag{OA.19}$$

and the forward- and backward-looking absolute movements

$$\mathrm{FutureAbs}_{c,t} \equiv \big|p_{c,t+1} - p_{c,t}\big|, \qquad \mathrm{PastAbs}_{c,t} \equiv \big|p_{c,t} - p_{c,t-1}\big|. \tag{OA.20}$$

I then estimate:

$$\mathrm{FutureAbs}_{c,t} = \beta\,\overline{\mathrm{Unc}}_{c,t} + \psi\,\mathrm{PastAbs}_{c,t} + \mu_c + \tau_t + \varepsilon_{c,t}, \tag{OA.21}$$

with community fixed effects μ_c, year–month fixed effects τ_t, and standard errors clustered by community. To mitigate thin-cell noise, the regression uses community–months with at least five listings. The estimates imply $\widehat{\beta} > 0$ and statistically significant: communities with higher valuation uncertainty today experience larger next-month absolute price movements even after controlling for recent volatility, indicating that uncertainty contains forward-looking information beyond simple persistence.

Finally, results are similar when prices are measured per m^2 using

$$p^{\mathrm{m}^2}_{c,t} \equiv \mathrm{median}_{i\in(c,t)}\ \mathrm{asinh}(P^{\mathrm{list}}_{i,t}/\mathrm{SIZE}_i),$$

indicating that the predictive content is not driven by time-varying size composition. Taken together, these diagnostics support viewing $\mathrm{Unc}^{\mathrm{list}}_{i,t}$ as a meaningful proxy for ex-ante idiosyncratic valuation risk faced by market participants at the time of listing.

OA.C.5 Comparison to ex-post hedonic dispersion (Giacoletti 2021)

The ex-ante measure defined in (1) is forward-looking: $\mathrm{Unc}^{\mathrm{list}}_{i,t}$ is constructed from the cross-sectional dispersion of listing residuals at posting. A natural benchmark is the ex-post dispersion measure of Giacoletti (2021), which uses the within-area standard deviation of residuals from a hedonic regression on realized transaction prices. The Giacoletti measure is backward-looking: it reflects the noise that was resolved in the settled sample. Replicating it on the Beijing data serves two purposes. First, it validates the micro panel by showing that a standard ex-post benchmark produces sensible cross-sectional dispersion levels. Second, it sharpens the economic interpretation of the forward-looking measure by contrasting the asset-pricing content of each series.

I estimate

$$\ln P^{\mathrm{deal}}_{i,t} = f(\mathrm{SIZE}_i, \mathrm{AGE}_{i,t}, \text{floor}, \text{building type}) + \alpha_{z(i),t} + \varepsilon_{i,t}, \tag{OA.22}$$

absorbing neighborhood×month fixed effects $\alpha_{z,t}$ (244 *chengqu*-level cells z), and recover $\widehat{\varepsilon}_{i,t}$. Within each (z,t) cell with at least five transactions, I compute $s_{z,t} \equiv \mathrm{sd}(\widehat{\varepsilon}_{i,t})$, and aggregate to the Beijing-wide monthly average $\widehat{\sigma}^{\mathrm{giac}}_t$, transaction-weighted across cells.

Result: same direction, very different magnitude. Figure OA.2 plots the two series over the estimation window (January 2015 through mid-2020; the sample is restricted to 2015 onward because earlier months have sparse transaction coverage at the *chengqu*-month level). The Giacoletti ex-post measure moves only modestly: its mean falls from 0.168 in the earlier window (2015:01–2016:11) to 0.154 in the later window (2016:12–2019:12), a decline of 8.1%. The forward-looking micro measure, in contrast, collapses from 0.0146 to 0.0097, a decline of 34.0% over the same horizon, roughly four times larger in relative terms. A mechanical break regression on standardized series, splitting at December 2016, yields a step coefficient of -1.163 (SE 0.136) for the Giacoletti series and -1.543 (SE 0.189) for the micro series: both series register a step larger than one standard deviation, but the raw collapse in the forward-looking measure is substantially sharper.

Does the ex-post measure replicate Giacoletti's predictive evidence in Beijing? A bare level contrast is not yet a replication of Amaral (2024). The central empirical content of Giacoletti (2021) is that within-area idiosyncratic dispersion forecasts future housing-market outcomes: high dispersion today predicts higher future returns (a priced idiosyncratic-risk premium in thin markets), and it also carries information about future trading volume and liquidity. I replicate this exercise on the Beijing monthly panel. For each horizon $h \in \{1, 3, 6, 12\}$, I run predictive regressions of four aggregate outcomes (city-wide log price growth, log volume growth, mean days-on-market, and the mean listing-to-deal wedge)

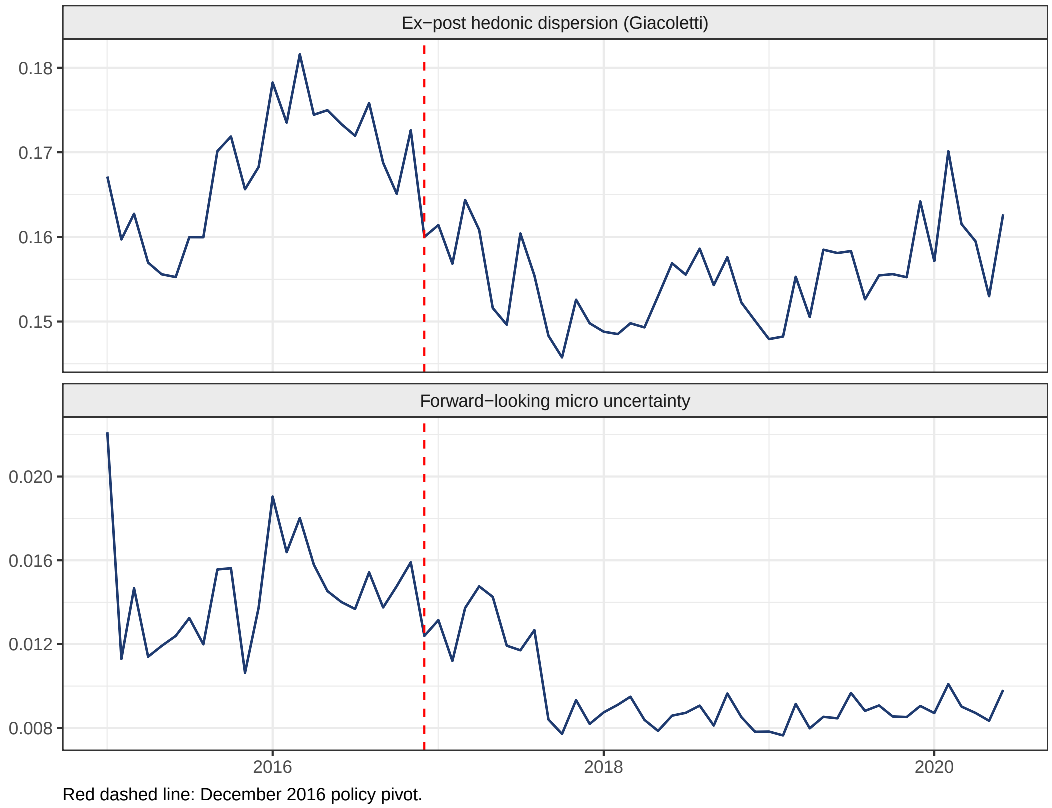


Figure OA.2: Ex-post Hedonic Dispersion (Giacoletti, 2021) vs. Forward-Looking Micro Uncertainty

Notes. Top panel: ex-post Giacoletti dispersion, $\widehat{\sigma}_t^{\text{giac}}$, the transaction-weighted mean across 244 chengqu-level cells of the within-cell standard deviation of residuals from hedonic regression (OA.22) on realized deal prices per m^2. Bottom panel: forward-looking micro uncertainty, the monthly mean of $\text{Unc}_{i,t}^{\text{list}}$ defined in Section 2.4.2. Red dashed line: December 2016. Sample restricted to January 2015 onward; both series use identical control sets. The ex-post measure falls by 8.1% from the earlier to the later window; the forward-looking measure falls by 34.0% over the same windows.

on a constant and one lagged dispersion measure, standardized so that coefficients are comparable across series:

$$y_{t+h} - y_t = a_h + b_h\,\widetilde{\sigma}_t + u_{t+h}, \qquad \widetilde{\sigma} \in \{\widehat{\sigma}^{\text{giac}}, \overline{\text{Unc}}^{\text{list}}\}, \tag{OA.23}$$

on the 2015:01–2019:12 monthly sample (60 observations), with Newey–West HAC standard errors (lag $= h+2$). A positive b_h on future price growth is the headline Giacoletti risk-premium prediction.

Table OA.7 delivers three findings. First, the Giacoletti prediction replicates strongly in Beijing. A one-standard-deviation increase in ex-post hedonic dispersion forecasts log price growth of $+1.9\%$ at one month, $+5.7\%$ at three months, $+11.1\%$ at six months, and $+16.7\%$ at twelve months, each significant at the 1% level under HAC inference. The univariate R^2

Table OA.7: **Predictive Regressions: Ex-post Dispersion vs. Forward-looking Micro**

		Giacoletti ex-post			**Forward-looking micro**		
Outcome	h	Coef	(HAC SE)	R^2	Coef	(HAC SE)	R^2
Price growth (log)	1	$+0.019^{***}$	(0.005)	0.34	$+0.010^{**}$	(0.004)	0.18
	3	$+0.057^{***}$	(0.012)	0.52	$+0.033^{***}$	(0.012)	0.32
	6	$+0.111^{***}$	(0.023)	0.63	$+0.059^{**}$	(0.025)	0.32
	12	$+0.167^{***}$	(0.035)	0.52	$+0.100^{***}$	(0.036)	0.34
Volume growth (log)	1	-0.025	(0.035)	0.00	-0.005	(0.033)	0.00
	3	$+0.037$	(0.089)	0.00	$+0.112^{*}$	(0.063)	0.07
	6	$+0.001$	(0.130)	0.00	$+0.092$	(0.080)	0.04
	12	-0.041	(0.156)	0.00	$+0.171$	(0.114)	0.09
Days-on-market	1	$+2.00$	(1.63)	0.02	$+1.09$	(1.11)	0.01
	3	$+2.52$	(3.29)	0.01	$+3.22$	(2.56)	0.02
	6	$+9.08^{*}$	(5.39)	0.05	$+8.92^{*}$	(5.01)	0.10
	12	$+27.11^{***}$	(10.29)	0.19	$+22.73^{**}$	(10.37)	0.25
Listing–deal wedge	1	$+0.001^{*}$	(0.001)	0.04	$+0.001$	(0.000)	0.03
	3	$+0.003^{**}$	(0.001)	0.08	$+0.002$	(0.001)	0.04
	6	$+0.007^{***}$	(0.002)	0.23	$+0.003$	(0.002)	0.10
	12	$+0.015^{***}$	(0.004)	0.49	$+0.007$	(0.004)	0.18
Observations per regression		$N = 60 - h$			$N = 60 - h$		

Notes. Each row reports a single time-series regression of the h-month-ahead change in the row outcome on the lagged and standardized predictor (Giacoletti ex-post dispersion or forward-looking micro uncertainty), estimated on monthly Beijing city-wide data from January 2015 through December 2019 (60 monthly observations, minus h for each horizon). Newey–West HAC standard errors in parentheses with lag $h + 2$. R^2 is the univariate regression R^2. Predictors are standardized to unit variance, so coefficients are the expected outcome change per one-standard-deviation increase in the predictor. Days-on-market coefficients are in days; the wedge is log listing minus log deal price. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

rises from 0.34 at $h = 1$ to 0.63 at $h = 6$, substantially higher than any single variable used for housing-return forecasting in the prior Beijing literature. This is a quantitatively strong replication of the priced idiosyncratic-risk premium that Giacoletti (2021) documents, and establishes that the ex-post dispersion measure genuinely captures a state variable with asset-pricing content, not a statistical artifact of the hedonic residual.

Second, the Giacoletti measure also forecasts the listing-to-deal wedge at horizons of three months and beyond, with R^2 rising from 0.08 at $h = 3$ to 0.49 at $h = 12$. This is an additional margin of validation: the same measure that predicts future returns also predicts the fraction of posted list prices that will be shaved in bargaining, consistent with the interpretation that cross-sectional dispersion in realized prices contains information about the intensity of idiosyncratic search frictions.

Third, the forward-looking micro measure is uniformly a weaker predictor of future outcomes than the Giacoletti measure. It forecasts price growth with the same sign pattern but lower coefficients and R^2 (e.g., $R^2 = 0.32$ at $h = 6$ vs. the Giacoletti measure's 0.63); it does not significantly forecast the listing-to-deal wedge at any horizon; and it tracks volume and days-on-market no better than the ex-post series. The predictive evidence, on its own, would favor the Giacoletti measure as the better proxy for priced housing risk in Beijing.

Why the forward-looking measure is still the right variable. The predictive evidence establishes what the Giacoletti measure is good for: summarizing the pricing of idiosyncratic risk as an asset-pricing object. It does not answer the question this paper asks, which is how a seller's perceived valuation uncertainty at the moment of listing maps into the unit's liquidity and price. Three observations pin down the complementarity.

First, the forward-looking measure is built from listing-stage residuals at posting, so it varies at the individual listing level within a community and month; the Giacoletti measure is a *chengqu*-month summary statistic of settled-sample dispersion and has essentially no within-cell variation. The identifying comparison in this paper is cross-sectional across listings within a community and month, so it requires the listing-specific measure rather than an area-level aggregate.

Second, the forward-looking measure is, by construction, observable before bargaining resolves: it is built from listing-stage residuals at the moment a unit is posted, while the Giacoletti measure is built from the settled transaction sample and therefore only accumulates information after search, negotiation, and match quality have already shaped which listings transact. A measure that reflects priced risk can still be the wrong object if the economic question is about the valuation uncertainty confronting a seller at the stage at which they commit to a list price.

Third, the Giacoletti measure does forecast the listing-to-deal wedge, but it does so at long horizons and as a summary statistic of cross-sectional dispersion. It is not a proxy for the individual seller's perceived valuation uncertainty on a specific listing. The identifying comparison in this paper is cross-sectional, within a month, across listings within a community; for that exercise the variable required is listing-specific uncertainty available at posting, which is exactly what $\mathrm{Unc}_{i,t}^{\mathrm{list}}$ provides.

The bottom line: the two measures capture different economic objects and answer different questions. The Giacoletti measure is an asset-pricing state variable, the object that should forecast returns, and it does. The forward-looking micro measure is a listing-level uncertainty proxy, the object that governs an individual unit's liquidity and pricing at the moment of listing, and it is exactly the variable the cross-sectional design requires. The paper's focus on $\mathrm{Unc}_{i,t}^{\mathrm{list}}$ for the identification exercise is motivated by the second role, and the Giacoletti replication validates the underlying micro panel rather than substituting for it.

OA.C.6 Platform-attention construction and other variables

FVR data cleaning and sample restriction. In the raw platform data, VIEWING occasionally contains non-numeric placeholders or is missing, which can generate mechanically extreme ratios. To mitigate these issues, I convert VIEWING to numeric when possible and apply a conservative filter restricting to $\mathrm{FVR}_{i,t} \leq 1$ in the main analysis. This trims observations likely driven by miscoding or missing view counts rather than economically meaningful conversion behavior. Section 7 verifies that the main results are robust to alternative trimming rules and to the unrestricted sample.

Listing premium.

Listing-implied pricing wedge and realized capital gain. The listing-implied pricing wedge (one-month horizon) is

$$\mathrm{ExpRaw}_{i,t}^{(1m)} = \operatorname{asinh}\big(\widetilde{P}_{i,t}^{\mathrm{list}}\big) - \operatorname{asinh}\big(\overline{P}_{c(i),t-1}^{\mathrm{deal}}\big), \tag{OA.24}$$

where $\widetilde{P}_{i,t}^{\mathrm{list}}$ is the quality-adjusted listing price of Section 2.4.1 and $\overline{P}_{c,t-1}^{\mathrm{deal}}$ is the community-mean quality-adjusted deal price in the prior month. This is the empirical counterpart of $\mathrm{ExpRet}_{h,t}^{(1m)}$ in Proposition OA.4; asinh accommodates right-skewness while preserving sign and the approximate log-difference interpretation for typical values.[22] For matched listings with a realized transaction, the realized raw capital gain is defined analogously with $\widetilde{P}_{i,t}^{\mathrm{deal}}$ in place of $\widetilde{P}_{i,t}^{\mathrm{list}}$ (Online Appendix OA.C.3). Figure OA.3 plots monthly averages of the wedge and the realized gain (computed with unadjusted prices); the two series co-move strongly at lower frequency over the sample (correlation 0.92), consistent with the listing-implied measure capturing components of subsequent price movement and mean-reversion to transaction benchmarks rather than pure noise.

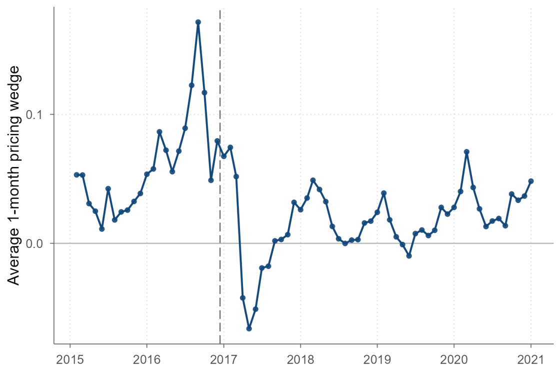
Within-community listing premium. To capture within-community relative over- or under-pricing at listing time, I define a listing premium as the deviation of a listing's raw price from the community–month mean raw listing price:

$$\mathrm{Prem}_{i,t} \equiv \operatorname{asinh}\big(\widetilde{P}_{i,t}^{\mathrm{list}}\big) - \operatorname{asinh}\Big(\overline{P}_{c(i),t}^{\mathrm{list}}\Big), \qquad \overline{P}_{c,t}^{\mathrm{list}} = \frac{1}{N_{c,t}} \sum_{j \in c,t} \widetilde{P}_{j,t}^{\mathrm{list}}. \tag{OA.25}$$

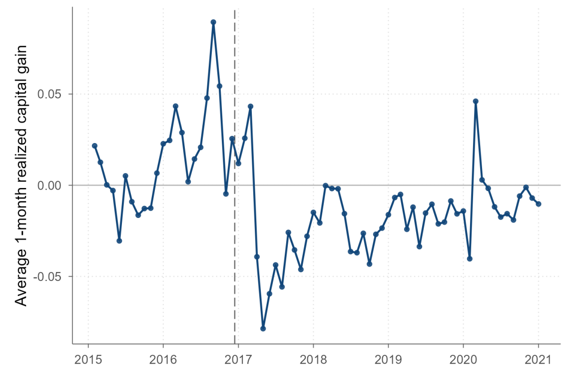
This object isolates relative pricing within a tight local comparison set.

Supply share. To measure local inventory pressure, I construct a community–month supply share. Let $N_{c,t}$ be the number of listings posted in community c in month t, and let

[22]For x large relative to 1, $\operatorname{asinh}(x) \approx \ln(2x)$, so the difference in asinh approximates a log ratio; asinh remains well-defined at zero.



(a) **1-Month Pricing Wedge**



(b) **1-Month Realized Capital Gain**

Figure OA.3: Short-horizon pricing wedges and realized capital gains over time

Notes. Panel (a) plots the monthly average listing-implied one-month pricing wedge, defined as the difference between a property's listing price and its community's average deal price one month earlier (asinh difference of per-m^2 prices, by listing month). Panel (b) plots the monthly average realized one-month capital gain for matched transactions, defined as the difference between the property's deal price and its community's average deal price one month earlier (same construction, by deal month). Both panels use unadjusted prices; the lagged community mean requires at least three deals. The regression object $\text{ExpRaw}^{(1m)}$ in Equation (OA.24) applies the same construction to quality-adjusted prices. The vertical dashed line marks the date of the 2016 Central Economic Work Conference.

$N_t = \sum_c N_{c,t}$ be the citywide total. Define:

$$\text{SupplyShare}_{c,t} \equiv \frac{N_{c,t}}{N_t}. \tag{OA.26}$$

This measure captures the fraction of citywide listing supply originating from a given community in a given month. Higher $\text{SupplyShare}_{c,t}$ reflects a larger local inventory footprint and, holding aggregate conditions fixed, potentially stronger seller competition and more informative local comparables. In the model, an increase in supply corresponds to a thicker sellers' market, which can improve the efficiency of matching and tighten the listing–deal wedge (Wheaton, 1990; Sagi, 2021).

OA.D Identification appendix

OA.D.1 Mapping reduced-form coefficients to structural parameters

Specification (3) is the direct empirical counterpart of the regression in Section 6.4. From the model (Equation OA.17):

$$\beta = -(1-\alpha)\,\frac{\bar{\gamma}\,\kappa}{2}\,\rho^2 \;<\; 0, \tag{OA.27}$$

where ρ is the discount factor, $\bar{\gamma}$ is the average risk aversion of the buyers a listing meets, and κ is the information-channel scaling. The cross-sectional slope β therefore identifies the composite risk-pricing index $\Gamma = \bar{\gamma}\,\kappa$ up to the bargaining weight α, the discount factor ρ, and the unit conventions of Online Appendix OA.B.13 (Equation (OA.18)). I report the coefficient and the implied Γ as a summary diagnostic of the mechanism's magnitude, not as a point estimate of preferences.

OA.D.2 Extended discussion of threats to identification

Heavy tail of the uncertainty distribution. Because the squared-residual measure is heavy-tailed, a level-based specification could load on its upper tail rather than on the bulk of the cross-section. I address this by reporting the discount under alternative operationalizations of uncertainty in Online Appendix OA.E.5. The cross-sectional discount on uncertainty is negative under all of them; a distribution-free percentile-rank specification, invariant to the residual variance scale by construction, recovers the same negative within-cell slope (Table OA.17), confirming that the discount is not an artifact of the tail.

Thin cells and mechanical dispersion. Because the hedonic absorbs community–month fixed effects, the effects are estimated more noisily in thin cells, which inflates the level of squared residuals there (an error of order σ^2/n). Two facts contain the concern. First, the inflation is common to every listing in the cell, so the within-cell rank—the regressor used throughout—is immune to it by construction. Second, the worry does not even appear in levels: across cells, mean squared residuals correlate positively with cell size ($+0.13$), and the spine is unchanged under cell-size-decile fixed effects (sale probability -0.0061 in both cases) with no interaction between the uncertainty rank and cell size (rank $\times$ log cell size: $t = -1.4$, 0.4, -0.1 across the three outcomes). For the state-dependence result, where thin months are the object of interest, noisier ranks in thin cells would attenuate the slope there—working against the measured steepening.

The hedonic is itself a specification. The residual depends on the hedonic that produces it. Re-deriving the measure under an independently re-estimated rich hedonic (size and its square, build-year bins, floor, facing, layout, construction and building type, total floors; rank correlation 0.66 with the shipped measure) reproduces and strengthens the spine: sale probability -0.0070 ($t = -3.5$), log days on market $+0.081$ ($t = 11.9$). A deliberately naive hedonic with no characteristics at all—the raw within-cell price deviation—still shows the days-on-market gradient ($+0.042$, $t = 6.2$) but attenuates the sale margin to insignificance (-0.0026, $t = -1.3$): stripping the controls mixes observable pricing into the residual and dilutes its idiosyncratic content, which is precisely the sense in which the measure captures deviation net of observables rather than the observables themselves (the predictable component carries the opposite sign; Online Appendix OA.D.5).

Confounding aggregate trends. The year–month fixed effects absorb all citywide aggregate shocks (e.g., interest rate changes, credit policy, macroeconomic conditions) that move listing-implied pricing wedges in parallel across listings. Identification therefore requires only that the within-month cross-sectional slope of pricing wedges on uncertainty be informative, not any aggregate time-series variation.

Listing composition. If which properties are listed varies with uncertainty, the cross-sectional relationship could reflect composition rather than pricing. The community fixed effects absorb permanent differences across communities, and the building controls (floor, construction type, CZ) absorb within-community sorting by observable characteristics. Section 7 further shows that results hold when restricting to the matched-transaction subsample, to the FVR ≤ 1 sample, and to specifications excluding the 2020–2021 (COVID) period.

OA.D.3 Platform coverage and selection

Platform coverage as a maintained assumption. Throughout the paper I assume that the online platform's coverage of Beijing listings is broadly stable over the sample window: that the set of communities and listings appearing in the data is not systematically shifting over time in a way that correlates with listing-level uncertainty. I do not have external data on platform market share by community or by month, so I cannot test this directly; I treat it as a maintained identifying assumption rather than a verified fact. The paragraph on platform selection below presents three indirect features that make large coverage shifts unlikely, but the identifying assumption itself is not testable with the available data.

Platform selection. A further consideration is selection into the platform. The data come from a single online listing platform, and differential market share across uncertainty levels could in principle generate a spurious cross-sectional correlation. Although I cannot observe platform market share directly, three features mitigate this concern. First, the platform was the dominant listing aggregator throughout the sample period, so market-share shifts at the listing level are unlikely to be large. Second, the analysis conditions on community and year-month fixed effects, absorbing any aggregate platform-level trends. Third, the within-community, within-month identification strategy requires that platform selection vary systematically with apartment-level uncertainty within narrow community-time cells, a pattern that would be difficult to generate from platform-level market-share changes alone.

OA.D.4 Generated regressor and pairs-cluster bootstrap

I do not observe a home's valuation uncertainty directly; I estimate it as the leftover from a price regression, so a home with a noisy price estimate can look artificially uncertain, and ignoring that first step would make the standard errors look tighter than they are. Formally, valuation uncertainty is a generated regressor: the squared residual from a first-stage hedonic estimated with community–month fixed effects, and the second-stage standard errors do not account for first-stage estimation error, which could lead to over-rejection. I take three steps to mitigate this concern. First, the hedonic is estimated using cross-fitting within time blocks (Section 2.4.2), which prevents in-sample overfitting from inflating the correlation between the uncertainty measure and the dependent variable. Second, in Section 7 I show that the uncertainty coefficient is robust to alternative hedonic specifications (reduced hedonic, deal-price hedonic), which alter the first-stage estimation error structure but produce similar second-stage estimates. Third, I implement a pairs cluster bootstrap that jointly re-estimates the hedonic and the second-stage regression in each iteration, sampling communities with replacement. Because the bootstrap is computationally expensive (each of 200 iterations re-fits the full first-stage hedonic), it is run on the FVR-restricted subsample ($N = 249{,}378$; FVR ≤ 1) rather than the full $N = 385{,}818$ headline sample. Table 23 reports $\widehat{\beta} = -0.1259$ on this subsample. I view this as a tight analog to the headline specification (the FVR-restricted subsample produces an uncertainty slope of roughly the same sign and statistical discipline as the full sample) rather than as a direct bootstrap of the headline number. The inference question the bootstrap is designed to answer, whether analytic standard errors understate the true sampling uncertainty once first-stage estimation error is compounded, is invariant to this choice of sample. Table 23 reports the results. The bootstrap standard error on $\widehat{\beta}$ is 0.056 (vs. the analytic 0.033, a ratio of 1.71). The bootstrap 95% confidence interval for $\widehat{\beta}$ on the FVR-restricted subsample is roughly $[-0.24, -0.02]$, excluding zero. Thus, while the analytic standard errors understate the true uncertainty by roughly 70%, the uncertainty coefficient remains significant at the 5% level under the bootstrap ($t \approx 2.2$) on the subsample where it is feasible to run. The headline liquidity gradients do not inherit this inflation, by construction rather than by simulation: the 1.71 inflation is specific to the level coefficient $\widehat{\beta}$: the rank transform that defines the headline regressor compresses first-stage estimation error, because a noisy hedonic residual seldom changes a unit's rank within its community and month even when it shifts the level of its squared residual.[23]

[23]The bootstrap resamples communities (the clustering unit) with replacement, re-estimates both the first-stage hedonic and the second-stage regression on each of 200 iterations, and collects the distribution of $\hat{\beta}$. This procedure accounts for both generated-regressor uncertainty and within-community correlation.

OA.D.5 Systematic versus idiosyncratic decomposition of valuation uncertainty

A standard concern with squared hedonic residuals as a valuation-uncertainty proxy is classical measurement error: the raw residual $\hat{u}_{i,t}^2$ equals the latent conditional variance $\sigma^2(x_i, t)$ plus an idiosyncratic noise component that attenuates outcome coefficients (Giacoletti, 2021). Following the two-stage construction of Giacoletti (2021), I project valuation uncertainty on a natural-spline basis in size, build year, and total floors, plus construction-type dummies and year fixed effects. The fitted value $\hat{g}(x_i, t)$ recovers the systematic component of squared residual variance predictable from observables. The orthogonal residual $\hat{u}_{i,t}^2 - \hat{g}(x_i, t)$ is the idiosyncratic component. The two pieces by construction span valuation uncertainty, and Table OA.8 reports the within-community within-month pricing-wedge slope on each.

The decomposition has three features the framework anticipates and the data confirm. First, the systematic component is small: it explains a tiny share of the variation in valuation uncertainty (correlation $\rho \approx 0.09$), so valuation uncertainty is dominated by the idiosyncratic component. Under the framework of Section 3, the idiosyncratic component is precisely the unhedgeable risk priced in equilibrium, and the level result on the raw proxy reflects this component. Second, the within-community within-month pricing-wedge slope on the idiosyncratic component is negative and significant at the 1% level, reproducing the headline result. Third, the slope on the systematic component is economically small and, where it is positive, is consistent with end-user buyers paying a quality or distinctiveness premium for units with extreme characteristics (very large, very old, atypical floor or construction type) rather than discounting them for risk. Either way, the systematic-vs.-idiosyncratic split confirms that the cross-sectional uncertainty discount in the main text reflects the idiosyncratic-noise channel and not a re-pricing of observable-characteristic uncertainty.

OA.D.6 Market-thickness instrumental variables

As an alternative identification check, I instrument valuation uncertainty with Amaral (2024)-style market-thickness measures $Z_i^m \equiv (x_i^m - \bar{x}_c^m)^2$ for property characteristic m: each unit's squared deviation from its community-level mean of m. The intuition is that units with extreme characteristics relative to their local comparable set generate harder-to-price residuals (relevance), conditional on the full set of fixed effects (Beijing-wide year–month, community). Table OA.9 reports OLS alongside three IV specifications: a full set of four characteristics (Z^{SIZE}, $Z^{\text{BUILD\_YEAR}}$, Z^{floors}, construction-type rarity), a narrow two-characteristic set, and a just-identified specification using only Z^{SIZE}. First-stage F-statistics are large across all three (113.5–313.2). The level coefficient on uncertainty is robust: the IV magnitude is similar to or larger than the OLS estimate (-0.22 to -0.40 across specifications versus OLS -0.17), consistent with attenuation bias correction. The Sargan overidentification

Table OA.8: Systematic vs. Idiosyncratic Decomposition of the Uncertainty Discount

	(1) Systematic alone	(2) Idiosyncratic alone	(3) Both	(4) Pre only Both	(5) Post only Both
z_{sys} (pre slope)	0.0015	—	0.0020	-0.0010	—
	(0.0026)	—	(0.0026)	(0.0025)	—
$z_{\text{sys}} \times \text{Post}$	0.0133***	—	0.0122***	—	—
	(0.0036)	—	(0.0029)	—	—
z_{sys} (post slope only)	—	—	—	—	0.0165***
	—	—	—	—	(0.0033)
z_{idio} (pre slope)	—	-0.0287***	-0.0287***	-0.0291***	—
	—	(0.0034)	(0.0034)	(0.0034)	—
$z_{\text{idio}} \times \text{Post}$	—	-0.0433***	-0.0430***	—	—
	—	(0.0159)	(0.0158)	—	—
z_{idio} (post slope only)	—	—	—	—	-0.0725***
	—	—	—	—	(0.0160)
Community FE	Y	Y	Y	Y	Y
Year–month FE	Y	Y	Y	Y	Y
Observations	385,284	385,284	385,284	134,645	250,035

Notes. The squared hedonic residual `price_unc` is decomposed into a systematic component $\hat{g}(x_i, t)$ (the fit from a stage-1 projection on a natural-spline basis in size, build year, total floors, with construction-type dummies and year fixed effects) and an idiosyncratic residual component. Both components are standardized. Columns (1)–(3) interact each component with `Post`. Columns (4) and (5) re-estimate the two components jointly on the pre- and post-pivot subsamples. The dependent variable in every column is the listing-implied one-month pricing wedge `exp_raw_1m`. Standard errors clustered by community in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

test rejects the full set and is borderline at the narrow set ($p = 0.026$), so the IV is best read as an alternative-measurement robustness check rather than a clean exclusion. The instruments thus corroborate the cross-sectional uncertainty discount under an alternative, characteristics-based measurement of how hard a unit is to price.

OA.D.7 Atemporal vs. diffusion variance: Sagi (2021) decomposition

Sagi (2021) shows that the within-property dispersion of housing returns decomposes into a diffusion component that scales linearly with the holding period and an atemporal component that does not. Under the random-walk-with-drift null, only the diffusion component is present, so a positive atemporal component is, in his interpretation, transactional or illiquidity risk: the wedge between the price a buyer pays and the value the asset would settle to in a frictionless trade. Sagi attributes this atemporal component to disagreement between buyers and sellers over an asset's value at the point of transaction. This is exactly the object the present paper measures: idiosyncratic valuation uncertainty is the dispersion in what counterparties think a given unit is worth, which is the kind of buyer–seller valuation dispersion Sagi shows generates a transactional, atemporal variance component rather than

Table OA.9: Market-Thickness Instrumental Variables for Valuation Uncertainty

	(1) OLS baseline	(2) 2SLS Z^m full set	(3) 2SLS narrow set	(4) 2SLS just-identified
Valuation uncertainty	-0.1724***	-0.2195***	-0.3978***	-0.3895***
	(0.0203)	(0.0741)	(0.1418)	(0.1425)
Uncertainty $\times$ Post	-0.2373***	0.1439	0.1495	0.1447
	(0.0873)	(0.2109)	(0.1548)	(0.1560)
Community FE	Y	Y	Y	Y
Year–month FE	Y	Y	Y	Y
First-stage F (min)	—	113.5	157.3	313.2
Sargan J stat	—	334.6	7.3	—
Sargan J p-value	—	<0.001	0.026	(just-id)
Wu-Hausman p-value	—	<0.001	<0.001	<0.001
Instruments	—	8	4	2
Observations	385,284	385,284	385,284	385,284

Notes. The endogenous variables are `price_unc` and `price_unc`$\times$`Post`. Instruments are Amaral (2024)-style market-thickness measures $Z_i^m \equiv (x_i^m - \bar{x}_c^m)^2$ for property characteristic m, plus their Post interactions, each standardized. Column (2) uses $m \in \{$SIZE, BUILD_YEAR, total_floors, construction-type rarity$\}$. Column (3) uses $m \in \{$SIZE, BUILD_YEAR$\}$. Column (4) uses only Z^{SIZE} and its Post interaction (exactly identified, no overid test). Standard errors clustered by community in parentheses.
$^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

a holding-period premium. Estimating the variance equation $\hat{e}^2_{i,k} = \sigma_0^2 + \sigma_1^2 k$ across horizons $k \in \{1, 3, 6, 12\}$ months, the atemporal variance σ_0^2 is positive and statistically significant ($\sigma_0^2 = 0.0197$ over the full sample, and 0.0296 in the pre-tightening subsample closest in regime to his window, Table OA.10). One construction difference bounds the comparison: returns here are benchmarked against contemporaneous community comparables rather than the same unit's own resale, so σ_0^2 also absorbs within-community pricing dispersion and is best read as an upper bound on the within-unit atemporal component, qualitatively in line with, rather than identical to, Sagi's within-property commercial-real-estate estimate of 0.0326. The decomposition therefore supplies an asset-pricing bridge for the paper's central reading: the second moment of housing value is borne as illiquidity—a transactional discount and longer time on market—not as a holding-period return premium.

Table OA.10: Sagi (2021) Atemporal vs. Diffusion Variance Decomposition of Buyer Returns

	(1) Full sample	(2) Pre-Dec 2016	(3) Post-Dec 2016
Mean equation: $r_{i,k} = \alpha_0 + \alpha_1 \cdot k$			
α_0 (intercept)	0.0216	0.0594	0.0050
α_1 (drift per month)	0.0052	0.0171	-0.0026
Variance equation: $\hat{e}^2_{i,k} = \sigma^2_0 + \sigma^2_1 \cdot k$			
σ^2_0 (atemporal)	0.0197***	0.0296***	0.0162***
	(0.0012)	(0.0032)	(0.0005)
σ^2_1 (diffusion per month)	0.0033***	0.0037***	0.0004***
	(0.0001)	(0.0001)	(0.0000)
Observations	1,066,255	381,653	684,602

Notes. Buyer ex-post realized returns are stacked over horizons $k \in \{1, 3, 6, 12\}$ months, where $r_{i,k} = \log(\text{community-mean deal price at } t+k) - \log(\text{deal price}_{i,t})$. The mean equation regresses $r_{i,k}$ on a constant and the horizon k in months. The variance equation regresses the squared residual from the mean equation on a constant and k. Under the random-walk-with-drift null, $\sigma^2_0 = 0$; a positive estimate indicates atemporal (transactional) risk. Standard errors clustered by community in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

OA.D.8 Choice-free measures: comparable scarcity and comparable dispersion

The baseline uncertainty measure is built from the seller's own posted price, so a natural objection is that its gradients reflect seller pricing behavior rather than valuation difficulty. This subsection re-estimates the illiquidity margins on two measures that contain no choice of seller i at all, both constructed from other units' past transactions. For listing i in community c posted in month t, the comparable set is the units sold in c during the trailing twelve months with size within $15\,\text{m}^2$ of i's. *Comparable scarcity* is the count of such comparables—few comparables make i hard to value in exactly the signal-extraction sense of Quan and Quigley (1991)—and *comparable dispersion* is the standard deviation of deal-price hedonic residuals among them (computed where at least five comparables exist). Both are ex ante (the window precedes the listing), both are leave-own-out by construction, and both enter as within-cell ranks, as in the baseline design.

Table OA.13 reports the spine. Comparable scarcity reproduces the liquidity gradients with no seller price in the measure: moving from the fewest to the most comparables in a cell raises the probability of sale by 2.3 percentage points and shortens time on market by about 9 percent ($t = 10.1$ and $t = -12.5$), and both margins survive size-decile fixed effects, so the scarcity gradients are not a relabelling of size popularity. The followers gradient does not survive the size controls: the attention margin belongs to the idiosyncratic-deviation component, not the comparability component—consistent with the two-channel reading, in which disagreement, and the attention it requires, attaches to a unit's own pricing deviation rather than to the thinness of its comparable set. The scarcity rank is essentially orthogonal to the baseline uncertainty rank (correlation -0.06), and both survive a joint horse race on the attention margin, so the two measures capture distinct components of valuation difficulty: how scarce a unit's comparables are, and how far its pricing sits from them.

Comparable dispersion, by contrast, produces no gradient. This is expected by construction rather than evidence against the mechanism: within a community–month cell, comparable sets overlap heavily, and a standard deviation computed on a median of six comparables is mostly sampling noise, so its within-cell rank carries little signal. The measurement lesson is that the comparability margin is identified by counts, which are precise, not by comparable dispersions, which are not.

Two caveats keep the reading calibrated. The comparable count bundles information scarcity with segment-level demand persistence—a size band that has been trading actively offers both more comparables and plausibly more demand—so I read the scarcity gradients as corroborating that the illiquidity of hard-to-value homes is not an artifact of the seller's own pricing, not as a pure information measure. The signature table of Online Appendix OA.A.4 sharpens this split: a thin segment moves the liquidity margins through scarce meetings with no response in per-exposure attention—exactly how the scarcity margin behaves once size segments are absorbed—while the residual rank moves the conversion and attention margins, as valuation uncertainty should. And because the baseline residual rank is orthogonal to scarcity, the paper's headline gradients are not explained by it; the two results are complements.

OA.D.9 Concurrent twins: observed disagreement and a pass-through check

The attribute key of Online Appendix OA.D.10 identifies $5{,}077$ pairs of listings of the same unit type whose spells overlap in time: two owners pricing the same asset at the same moment. Their asks disagree, and by a lot. The absolute log ask gap averages 5.3 points across all concurrent pairs; among the $1{,}091$ pairs posted in the same calendar month, where market drift cannot enter the comparison, it averages 4.2 points with a median of 2.7, and thirty percent of pairs sit more than five percent apart. Adjusting the earlier ask by the district index between the two posting months, or controlling for the days between postings, changes little. The gap is the seller-side counterpart of the four-to-five-percent buyer valuation dispersion that Han and Strange (2015) report as the canonical magnitude.

Two disciplines keep this evidence honest. First, the gap cannot be regressed on the pair's own uncertainty rank: within an attribute key the two asks share a predicted value, so the gap and the pair's squared residuals are close to the same object, and the strong positive slope that regression produces ($t = 7.2$) is mechanical. I therefore use regressors that contain no prices. The gap widens with the configuration's rarity in its community, the negative log share of the community's listings sharing the pair's coarse configuration ($+0.4$ log points per standard deviation, $t = 2.1$; $t = 2.4$ controlling for the posting gap), while the leave-one-out submarket dispersion, which is nearly absorbed by community fixed effects, carries no residual slope ($t = -0.8$). Second, within-pair price variation cannot estimate the demand curve: with pair fixed effects the higher-priced twin is *more* likely to sell within thirteen weeks ($+0.48$, $t = 3.0$), because even between twins the residual ask difference is dominated by twin-specific quality such as renovation. The pattern is itself informative, and it is why the paper claims only differences in the demand curve's slope across submarkets, never its level. Consistent with the flattening, the within-pair coefficient falls from $+0.76$ ($t = 3.4$) in transparent submarkets to $+0.20$ ($t = 0.8$) in opaque ones (interaction -0.56, $t = -1.7$), and the pass-through check reported next shows the drop is not a change in what the ask means.

The pass-through check regresses the transaction-price residual of sold listings on the listing markup, by submarket transparency, with community×month fixed effects ($N = 66{,}283$). The markup carries into the realized price at 0.98 in transparent submarkets and 0.99 in opaque ones. The ask is equally informative about value on both sides of the split; what differs, by a factor of fourteen on the discount side (Table 13), is whether that ask moves the probability of sale. Standard errors are clustered by community throughout.

Figure OA.4 plots the gap as a monthly series, dated by the second twin's posting month. Observed disagreement peaks in the early-2016 run-up, at roughly eleven log points after adjusting for community composition, and falls through the December 2016 policy turn and the March 2017 tightening to about five points from 2018 onward. The shape mirrors the aggregate decline of the residual-based measure in Figure 5, which also roughly halves over the sample: when prices move fastest, comparable sales age in weeks and sellers anchored on different vintages post far apart; once the tightening froze prices, the comparables regained precision. The correlation with the citywide thirteen-week sale rate is $+0.18$, so the gap is a hot-market object, while Online Appendix OA.G.7 shows the illiquidity per unit of disagreement steepens after the turn: less disagreement in the bust, but each unit costs more liquidity, the convexity the model delivers through $1/(\lambda q)$. The pre-2016 months rest on fifteen to nineteen pairs each, so the early band is wide, and the drift adjustment does its work precisely in the boom months; the figure shows both the raw and the adjusted series for that reason.

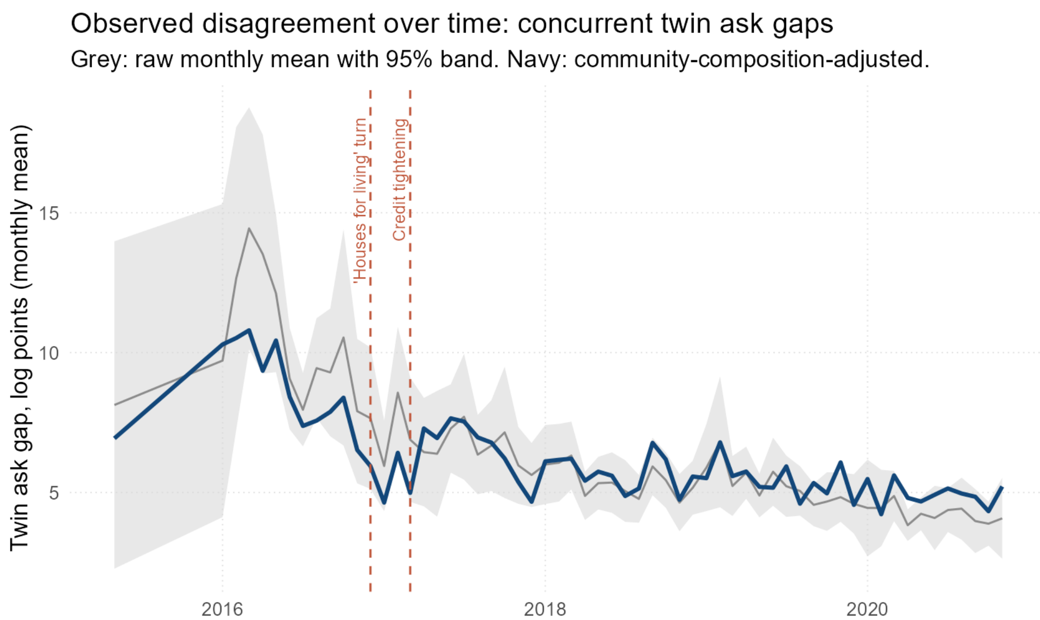


Figure OA.4: Observed Disagreement Over Time: Concurrent Twin Ask Gaps

Notes. Monthly mean absolute log ask gap between concurrent attribute-key twins, dated by the second twin's posting month and adjusted for district-level drift between the two posting dates; months with at least fifteen pairs. The grey line is the raw monthly mean with a pointwise 95 percent band; the navy line demeans by community and adds back the sample mean, so composition shifts across communities do not drive the shape. Dashed lines mark the December 2016 Central Economic Work Conference ("houses for living, not for speculation") and the March 2017 credit tightening.

OA.D.10 Within-unit evidence: repeat and twin listings

The data carry no unit identifiers, but the exact attribute combination—community, size, floor band, facing, layout, build year, and total floors—identifies listings of the same physical unit or of its identical twin in the same building. A fixed effect on that key absorbs every time-invariant unit attribute, observed or unobserved, so within-key estimates cannot be driven by fixed unit quality. The key-level panel contains $76{,}694$ listings over $31{,}094$ keys (FVR ≤ 1); a stricter "clean repeat" subsample keeps keys with exactly two listings in which the first sold before the second was posted ($12{,}522$–$15{,}640$ pairs, median sale-to-relist gap of nineteen months)—the closest available analogue of the same home listed twice, in two different valuation environments.

The gradients keep their signs within key, with the attenuation and loss of power one expects when most of the rank variation is absorbed: time on market $+0.032$ ($t = 1.7$), about sixty percent of the baseline slope, across all multi-listing keys; the probability of sale -0.013 ($t = -2.1$) among clean sequential repeats (identified off second spells, since first spells sold by construction); attention positive but insignificant. The within-key sale-probability point estimate is fragile—it attenuates to -0.002 ($t = -0.4$) in the broader all-keys estimator and toward zero as the key is drawn tighter—so I do not lean on it; the robust within-key statement is the demand-curve sign reversal below. I read these as calibrated support rather than headline estimates—their value is what they exclude: a story in which the illiquidity gradients reflect fixed unobserved quality of hard-to-value homes would not survive a design in which the unit's fixed attributes are differenced out entirely.

The same panel recovers the demand curve a seller faces—the sale-hazard response to her relative list price, the object Guren (2018) estimates and builds price stickiness on. Across units, the raw within-cell slope of the sale probability on the signed relative ask is flat to positive: a unit listed above its comparables is often a genuinely better unit, so quality signaling masks the price effect. Within identical-attribute keys, where fixed quality is differenced out, the demand curve emerges with its expected sign: a ten-percent-higher relative ask lowers the probability of sale by 0.71 percentage points (-0.071, $t = -2.4$). Two further descriptive facts follow the model's logic. The conditional-on-sale ask–TOM slope attenuates by about half as the uncertainty rank rises (interaction $+0.74$, $t = 2.7$, against a base of -1.44): the relative ask carries less information—and exerts less discipline—for hard-to-value units, which is the demand-curve counterpart of the model's dispersion channel ($\partial q/\partial\,\text{ask} \propto -1/s$, flatter where valuation dispersion s is high). I label these descriptive because the conditional-on-sale margin retains selection even within key; the clean statement is the sale-margin sign reversal between the across-unit and within-unit designs.

The measure also travels, and this is worth stating as a result in its own right: comparable scarcity is computable from bare transaction records—a community identifier, a size, and a date; no listings, no attention data, no hedonic model—which is the data environment most housing researchers face. Table OA.11 recomputes it in the four transaction-only provinces (thirty-six cities, collision-proof communities) and re-estimates the gradients. The liquidity and seller-side margins replicate everywhere: time on market falls with comparable abundance in all five markets (-0.09 to -0.13, every t above 6.6, surviving size-decile fixed effects in all five), sellers revise less where comparables are abundant in all five, and the negotiated wedge narrows in all five. The attention margin is mixed across provinces—consistent with the Beijing finding that scarcity's attention gradient is size-popularity rather than disagreement, which belongs to the residual component. A researcher with deeds or registry data alone can therefore construct a property-level measure of valuation difficulty whose liquidity content this paper validates against the full listing-and-attention apparatus.

Two further checks sharpen the reading. First, current competition is not the channel. Counting the other listings whose spells overlap the posting month in the same size band—the supply of close substitutes—and entering its within-cell rank jointly with the past-sales rank leaves the past-sales gradients unchanged (probability of sale $+0.0242$, $t = 9.2$; time on market -0.0958, $t = -11.3$) while the active-competitor rank carries no effect on any margin ($|t| < 0.6$, robust to the censoring cap on unsold spells). What predicts a sale is the information and demand revealed by closed transactions, not the contemporaneous stock of

Table OA.11: Comparable Scarcity Travels: A Transaction-Record Measure, Five Markets

	Beijing	**Guangdong**	**Zhejiang**	**Shandong**	**Henan**
Time on market (log)	-0.0905^{***}	-0.1272^{***}	-0.1018^{***}	-0.1334^{***}	-0.1103^{***}
	(0.0077)	(0.0101)	(0.0153)	(0.0138)	(0.0157)
adding size-decile FE	-0.0671^{***}	-0.0875^{***}	-0.0384^{**}	-0.1028^{***}	-0.0629^{***}
	(0.0077)	(0.0102)	(0.0153)	(0.0139)	(0.0153)
Committed attention (log foll.)	$+0.0180^{**}$	$+0.0319^{***}$	$+0.0523^{***}$	-0.0311^{***}	-0.0151
	(0.0073)	(0.0091)	(0.0138)	(0.0101)	(0.0133)
Asking-price revisions	-0.0457^{***}	-0.1336^{***}	-0.1280^{***}	-0.1002^{***}	-0.1323^{***}
	(0.0076)	(0.0227)	(0.0346)	(0.0280)	(0.0405)
Price wedge (deal−list)	$+0.0026^{***}$	$+0.0032^{***}$	$+0.0020^{**}$	$+0.0018^{***}$	$+0.0015^{***}$
	(0.0002)	(0.0003)	(0.0008)	(0.0003)	(0.0004)

Notes. Each cell is a separate regression of the row outcome on the within-community–month rank of comparable scarcity: the count of units sold in the same (collision-proof, city×complex) community within $15\,\mathrm{m}^2$ of the listing over the trailing twelve months—a measure computable from transaction records alone. High rank = many comparables = easy to value, so predicted signs are the opposite of the uncertainty gradients. Fixed effects: community, year–month, floor, construction type (the second row adds size-decile fixed effects); standard errors clustered by community; FVR ≤ 1. Province files are transaction-only, so time on market and the wedge condition on sale by construction. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table OA.12: The Mechanism Replicates Across Provinces (Collision-Proof)

	Beijing (headline)	**Guangdong** (replication)	**Zhejiang** (replication)	**Shandong** (replication)	**Henan** (replication)
Probability of sale	-0.0061^{***}	—	—	—	—
	(0.0020)				
Time on market (log)	$+0.0511^{***}$	$+0.0015$	$+0.0379^{***}$	$+0.0359^{***}$	$+0.0347^{***}$
	(0.0067)	(0.0081)	(0.0116)	(0.0104)	(0.0130)
Committed attention (log foll.)	$+0.0395^{***}$	$+0.0359^{***}$	$+0.0221^{**}$	$+0.0229^{***}$	$+0.0270^{***}$
	(0.0063)	(0.0072)	(0.0104)	(0.0080)	(0.0105)
Asking-price revisions	$+0.0456^{***}$	$+0.0914^{***}$	$+0.1237^{***}$	$+0.1024^{***}$	$+0.1833^{***}$
	(0.0069)	(0.0188)	(0.0249)	(0.0241)	(0.0324)
Price wedge (deal−list)	-0.0011^{***}	-0.0017^{***}	-0.0028^{***}	-0.0017^{***}	-0.0015^{***}
	(0.0002)	(0.0003)	(0.0006)	(0.0003)	(0.0003)
Cities	1	10	9	9	8

Notes. Full estimates behind Figure 12. Each cell is a separate regression of the row outcome on the within-(community × month) rank of valuation uncertainty, with community, year–month, floor, and construction-type fixed effects and standard errors clustered by community (follower-to-viewing ratio at most one). Beijing uses the headline data; for each province a community is the city × community pair, the hedonic is re-estimated with those fixed effects, and $\hat{u}^2$ is recomputed. Province files are transaction-only, so the probability of sale is estimable only for Beijing. Standard errors in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

substitutes. Second, a third choice-free measure—the disagreement among three valuation models trained only on the prior year's transactions (a community-fixed-effects hedonic, a pure comparables average, and a coarse district hedonic)—validates as a forecaster of valuation difficulty (the ex-post absolute pricing error rises strongly in its rank, $t = 16.6$) yet produces no illiquidity gradient on any margin and loses every horse race to the baseline rank. Model-computable valuation difficulty does not make a home hard to sell; what does is the scarcity of transacted comparables and the unit's own market-revealed deviation from them.

Table OA.13: Choice-Free Valuation Difficulty: Comparable Scarcity and Dispersion

	(1) Probability of sale	(2) Days on market $\log$, sold units	(3) Followers $\log(1+F)$
Comparable scarcity: rank of comp count	$+0.0234^{***}$	-0.0919^{***}	$+0.0210^{***}$
	(0.0023)	(0.0074)	(0.0070)
adding size-decile FE	$+0.0089^{***}$	-0.0619^{***}	$+0.0006$
	(0.0023)	(0.0074)	(0.0070)
Comparable dispersion: rank of comp-residual SD	-0.0025	$+0.0009$	-0.0071
	(0.0026)	(0.0086)	(0.0082)
Observations (scarcity rows)	268,560	207,134	268,560
Observations (dispersion row)	196,598	157,002	196,598

Notes. Each cell is a separate regression of the column outcome on the within-community–month rank of the row measure, with community, year–month, floor, and construction-type fixed effects (the "size-decile FE" rows add within-sample size-decile fixed effects); standard errors clustered by community (FVR ≤ 1 sample). Comparable scarcity is the count of units sold in the same community in the trailing twelve months with size within $15\,\mathrm{m}^2$ of the listing (high rank = many comparables = easy to value, so the predicted signs are the opposite of the baseline uncertainty gradients); comparable dispersion is the standard deviation of deal-price hedonic residuals among those comparables, defined where at least five exist. Both measures use only transactions of other units completed before the listing was posted. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

OA.E Additional empirical results

This appendix collects supplementary empirical results referenced from Section 5.

OA.E.1 Number of price revisions

The platform records the count of asking-price revisions per listing (mean 0.94, median 1, max 44). Read as a seller-side learning measure, the count admits two competing interpretations. If the seller iterates toward a market-clearing price by re-pricing in response to incoming attention and offers, then high-uncertainty listings should revise more as the seller searches for the right price. Alternatively, if a high-uncertainty asset has no clear comparable so the seller anchors on the initial post and treats revisions as costly admissions of mispricing, then

Table OA.14: Number of Price Revisions and the Uncertainty Discount

Dep. variable	(1) Revisions, levels $\log(1+\#\text{rev})$	(2) Revisions, rank $\log(1+\#\text{rev})$	(3) Wedge with revisions $\text{ExpRaw}^{(1m)}$
Valuation uncertainty (level)	-0.0544***	—	-0.1956***
	(0.0090)	—	(0.0228)
Uncertainty rank r^{unc}	—	0.0071***	—
	—	(0.0022)	—
$\log(1+\#\text{rev})$	—	—	-0.0151***
	—	—	(0.0006)
Community FE	Y	Y	Y
Year–month FE	Y	Y	Y
Observations	385,284	346,574	385,284

Notes. The number of price revisions per listing is the platform's revision count (mean 0.94, max 44). Column (1) regresses $\log(1+\#\text{rev})$ on valuation uncertainty in levels; column (2) uses the within-community–month percentile rank of uncertainty, the paper's tail-immune metric (the sign reversal between the two is the levels-versus-rank tail artifact discussed in the text); column (3) adds revisions as a control to the pricing-wedge regression of Section 5. Standard errors clustered by community. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

high-uncertainty listings should revise less. The two readings differ in their implications for whether a revision-based channel could account for the cross-sectional uncertainty discount documented in Section 5.

Empirically, the answer is measure-dependent in exactly the way the paper's levels-versus-rank comparisons lead one to expect. In levels—the specification of Table OA.14, Column (1)—high-uncertainty listings revise less: the coefficient on valuation uncertainty in a regression of $\log(1+\#\text{revisions})$ on uncertainty with community and year–month fixed effects is negative and significant (-0.054, $t=-6.0$). Under the tail-immune within-cell rank used throughout the paper, the sign reverses: revisions rise with the uncertainty rank ($+0.0071$, $t=3.2$ in the same sample and fixed effects, Column (2); $+0.036$, $t=5.3$ for the revision count in the mature-cohort specification of Table 10, Column (4)). As with the sale-probability margin (Section 5.4.3), the levels estimate is carried by the extreme upper tail of $\hat{u}^2$; the rank-based estimate is the economically relevant one, and it is the anchoring-versus-search prediction it supports: sellers of hard-to-value homes re-price more, consistent with iterating toward a market-clearing price under disagreement. More importantly for the identifying argument, the headline uncertainty coefficient in the pricing-wedge regression is essentially unchanged when revisions are added as a control (Column (2)), so the uncertainty discount is not driven by an unmodeled seller-revision channel under either sign convention.

OA.E.2 Listing-price levels

Table OA.15 reports the deal-price fixed-effects ladder discussed in Section 5.1.3, and Table OA.16 its listing-price counterpart. The saturated listing-price coefficient (-0.3652, s.e. 0.0247) is nearly identical to the deal-price estimate (-0.3667 on the matched-transaction subsample), so the level discount is set at listing and survives essentially unchanged into the negotiated price. The same tail-fragility documented in Section 5.1.3 applies throughout.

Table OA.15: Valuation Uncertainty and Deal Prices

	(1)	(2)	(3)
	Dependent variable: log **deal price per** m^2		
Valuation uncertainty	-0.4270^{***}	-0.3670^{***}	-0.3667^{***}
	(0.0328)	(0.0257)	(0.0256)
Community FE	Y	Y	Y
Year–Month FE	N	Y	Y
Floor FE	N	N	Y
Constr. Type FE	N	N	Y
Observations	407,995	407,995	407,995
Adj. R^2	0.7045	0.9269	0.9293
Within R^2	0.1091	0.2654	0.2712

Notes. The sample consists of Beijing apartment listings matched to transactions, January 2015 to January 2021. The dependent variable is the log deal price per m^2 in RMB. Valuation uncertainty is the squared hedonic residual from Equation (1). All specifications include community fixed effects; additional fixed effects are indicated. Standard errors are clustered by community and reported in parentheses. ***, **, * denote significance at the 1/5/10% levels.

Table OA.16: Valuation Uncertainty and Listing Prices

	(1)	(2)	(3)
	Dependent variable: log **listing price per** m^2		
Valuation uncertainty	-0.4249^{***}	-0.3651^{***}	-0.3652^{***}
	(0.0320)	(0.0247)	(0.0247)
Community FE	Y	Y	Y
Year–Month FE	N	Y	Y
Floor FE	N	N	Y
Constr. Type FE	N	N	Y
Observations	578,264	578,264	578,264
Adj. R^2	0.6751	0.9302	0.9326
Within R^2	0.0712	0.2068	0.2124

Notes. The sample consists of Beijing apartment listings, January 2015 to January 2021. The dependent variable is the log listing price per m^2 in RMB. Valuation uncertainty is the squared hedonic residual from Equation (1). All specifications include community fixed effects; additional fixed effects are indicated. Standard errors are clustered by community and reported in parentheses. ***, **, * denote significance at the 1/5/10% levels.

OA.E.3 Sign convention of the listing-implied pricing wedge

A natural question is why the slope coefficient is negative for both price levels and listing-implied returns, when standard asset-pricing logic predicts that riskier assets should command higher expected returns. The resolution lies in the construction of ExpRaw$^{(1m)}$: it equals $\operatorname{asinh}(\widetilde{P}^{\text{list}}_{i,t}) - \operatorname{asinh}(\bar{P}^{\text{deal}}_{c(i),t-1})$, where the lagged community benchmark is common to all listings in the same community–month. Because ExpRaw is derived directly from the listing price, it is a seller-side markup over the community benchmark, not the buyer's forward-looking return. When a property carries high valuation uncertainty, the seller must list at a discount to attract risk-averse buyers (negative slope on price levels); this discount feeds mechanically into a lower listing-implied "return" (negative slope on ExpRaw). The buyer's compensation for bearing risk is the mirror image: purchasing at the discount generates a higher forward-looking expected gain. An analogy to bond markets is instructive. A risky bond has both a lower price and a lower price-to-last-period-price ratio (same sign), yet a higher yield = coupon/price (opposite sign). ExpRaw is a price-relative measure; the buyer's risk premium corresponds to the yield, which is not directly observed at the listing stage.

OA.E.4 Shared-hedonic mechanical contamination check

Both the dependent variable (ExpRaw) and the key regressor (Unc) are derived from the same hedonic regression (1), raising a potential concern about mechanical correlation. Three features of the construction mitigate this concern. First, the two objects use different components of the hedonic output: $\text{Unc}_{i,t} = \widehat{u}^2_{i,t}$ is the squared full residual (after removing observables and all fixed effects), while the quality-adjusted price that enters ExpRaw retains the community and time fixed effects, stripping only the observable-characteristics component $\mathbf{X}'_i\widehat{\boldsymbol{\theta}}$. The two measures therefore vary along different dimensions: Unc captures within-community valuation dispersion, whereas ExpRaw captures the listing's price positioning relative to the community-month benchmark.

Second, and most relevant to identification, any mechanical relationship between the residual $\widehat{u}$ and its square $\widehat{u}^2$ is a distributional property of the residual itself. The OLS coefficient from regressing $\widehat{u}$ on $\widehat{u}^2$ is proportional to $\mathbb{E}[\widehat{u}^3]/\,\mathbb{E}[\widehat{u}^4]$, the skewness-to-kurtosis ratio of the residual distribution, which is approximately zero under symmetry. Such mechanical contamination would be common to all listings within a community and month and is differenced out by the within-cell comparison; it cannot generate the cross-sectional uncertainty discount, which is the paper's central object.

Third, the pairs cluster bootstrap reported in Appendix B re-estimates the hedonic in every bootstrap draw and reconstructs Unc from it before running the second-stage regression; the outcome is held fixed across draws, so the exercise propagates the regressor's estimation error. The uncertainty slope remains significant at the 5% level with a bootstrap standard error 71% larger than the analytic one (a 95% confidence interval of roughly $[-0.24, -0.02]$), supporting the view that inference is not overturned by the generated-regressor construction.

OA.E.5 Within-R^2 and functional-form robustness

A note on explanatory power: the within-R^2 of the pricing-wedge regressions is modest (0.07–0.08), indicating that valuation uncertainty explains a small share of within-community variation in listing-implied returns. This is not surprising. Listing-level pricing wedges are noisy objects driven primarily by idiosyncratic seller pricing decisions, and the uncertainty channel operates at the margin. The economic significance lies in the cross-sectional pricing of uncertainty rather than in the overall explanatory power of the regression.

Table OA.17: Functional-Form Robustness: Alternative Uncertainty Measures

	(1) $\hat{u}^2$	(2) $\lvert\hat{u}\rvert$	(3) Pctile rank	(4) Pctile rank
Uncertainty measure	-0.1690^{***}	-0.3807^{***}	-0.0096^{***}	-0.0095^{***}
	(0.0204)	(0.0284)	(0.0024)	(0.0024)
Community FE	Y	Y	Y	Y
Year–Month FE	Y	Y	Y	Y
Floor FE	Y	Y	Y	N
Constr. Type FE	Y	Y	Y	N
Observations	346,574	346,574	346,574	346,574
Within R^2	0.0798	0.0552	0.0003	0.0003

Notes. Dependent variable: $\text{ExpRaw}^{(1m)}$ in all columns. Column (1) uses the baseline squared hedonic residual $\hat{u}^2$. Column (2) uses the absolute residual $|\hat{u}|$. Columns (3)–(4) use the within-community-month percentile rank of $|\hat{u}|$, which is distribution-free and invariant to the scale of the residual variance. Column (3) includes the full FE set; column (4) uses only community and year–month FE. The uncertainty coefficient is negative and significant under every operationalization. Standard errors clustered by community in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

Beyond the choice of functional form, the squared-residual measure is heavy-tailed by construction: its ninety-ninth percentile is about 0.10 while its maximum exceeds 40. Table OA.18 traces how the uncertainty discount responds to this tail. Under smooth, monotone transforms that compress the tail while retaining every observation—$\text{asinh}(\hat{u}^2)$ and $\log(1 + \hat{u}^2)$—the discount is preserved in direction and remains significant at the one-percent level. The exact magnitude on the untransformed variance measure, by contrast, is sensitive to the tail: capping or trimming the upper one percent of $\hat{u}^2$ changes the point estimate, so the level-based measure loads on the extreme tail in magnitude, even though its direction does not. I read the evidence as a robust negative price of variance whose exact multiple is not separately identified from the heavy tail of the proxy; the paper's quantitative claims accordingly rest on the direction of the discount and on its corroboration in realized transaction prices and in the hedonic-model-free return measure, and on the within-cell

Table OA.18: Tail and Functional-Form Robustness of the Regime-Shift Interaction

	(1) Baseline $\hat{u}^2$	(2) $\mathrm{asinh}(\hat{u}^2)$	(3) $\log(1+\hat{u}^2)$	(4) Winsorize [1,99]	(5) Drop top 1%
Uncertainty × Post (δ)	-0.240^{***}	-0.181^{***}	-0.197^{***}	0.246	0.309^{***}
	(0.088)	(0.041)	(0.053)	(0.177)	(0.114)
Pre-period slope (β)	-0.173^{***}	-0.569^{***}	-0.705^{***}	-0.708^{***}	0.056
	(0.020)	(0.032)	(0.040)	(0.145)	(0.077)
Post/pre slope ratio	2.39	1.32	1.28	0.65	6.57
Observations	385,284	385,284	385,284	385,284	381,426

Notes. Dependent variable: ExpRaw$^{(1m)}$. All columns use the saturated specification (community, year–month, floor, and construction-type fixed effects); standard errors clustered by community in parentheses. The uncertainty regressor is the squared hedonic residual $\hat{u}^2$, which is heavy-tailed (99th percentile ≈ 0.10, maximum ≈ 41). Columns (2)–(3) apply smooth, monotone, tail-accommodating transforms that retain all observations; columns (4)–(5) cap or drop the upper tail. The negative steepening is preserved in direction under the smooth transforms but is sensitive to capping or trimming the extreme tail, indicating that the *magnitude* of the level-based interaction (notably the baseline post/pre ratio) loads on the upper tail of the variance measure, while its *direction* does not. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

rank measure used throughout the main text, rather than on the point estimate from the untransformed variance.

OA.E.6 Deal-price-based uncertainty cross-check

As a direct test of the generated-regressor concern, I construct an entirely independent uncertainty measure from deal prices: for each community–month cell, I compute the within-community variance of realized transaction prices per m^2 (lagged one month to avoid look-ahead bias). This "deal-price uncertainty" is constructed solely from the transaction market and shares no common component with the listing-side hedonic used to derive both $\mathrm{Unc}_{i,t}$ and ExpRaw$^{(1m)}$. Table OA.19 reports the results. Column (1) relates the pricing wedge to deal-price variance, and column (2) to the deal-price standard deviation; both deal-side measures enter significantly, confirming that an independently measured source of valuation dispersion is priced into listing behavior. Column (3) includes both deal-side and listing-side uncertainty simultaneously; both enter significantly, indicating that the two measures capture distinct components of valuation risk.

OA.E.7 Time on market and sale probability

Table OA.20 links listing-time valuation uncertainty to market frictions. Column (1) uses days on market as the dependent variable. The coefficient on valuation uncertainty is positive and significant: higher-uncertainty listings remain on the market longer. Quantitatively, a one-unit increase in valuation uncertainty is associated with approximately 3.9 additional

Table OA.19: Deal-Price-Based Uncertainty and Pricing Wedges

	(1)	(2)	(3)
	Dependent variable: ExpRaw$^{(1m)}$		
Deal-price variance (lag)	0.1282^{***}		0.1360^{***}
	(0.0215)		(0.0227)
Deal-price SD (lag)		0.1465^{***}	
		(0.0232)	
Listing-side uncertainty			-0.1617^{***}
			(0.0284)
Community FE	Y	Y	Y
Year–Month FE	Y	Y	Y
Floor FE	Y	Y	Y
Constr. Type FE	Y	Y	Y
Observations	199,520	199,520	199,520
Adj. R^2	0.064	0.064	0.170
Within R^2	0.016	0.015	0.127

Notes. Deal-price uncertainty is constructed as the within-community variance (column 1) or standard deviation (column 2) of realized deal prices per m^2, lagged one month. This measure is entirely independent of the listing-side hedonic used to construct $\mathrm{Unc}_{i,t}$ and ExpRaw$^{(1m)}$. Column (3) includes both deal-side and listing-side uncertainty simultaneously. Standard errors clustered by community. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

days on market (s.e. 1.45), conditional on community and month fixed effects and property controls. This is consistent with bargaining frictions prolonging the time to sale when the property's value is harder to benchmark (Genesove and Mayer, 1994; Sagi, 2021).

Column (2) uses an indicator for whether a listing is ever matched to a realized transaction. Measured in levels, the coefficient on valuation uncertainty is positive and highly significant (0.038, s.e. 0.006), implying that higher-uncertainty listings are more likely to transact eventually. This level estimate, however, is driven by the heavy upper tail of $\hat{u}^2$: under the tail-immune within-community–month rank used in the headline specification (Table 6) the sign reverses to negative. The level coefficient should therefore be read with caution; the rank estimate is the one I rely on.

Column (3) examines the probability of selling within 90 days. The same tail-sensitivity appears, and more sharply: the level coefficient is positive ($+0.036$, s.e. 0.008), but the tail-immune rank coefficient is negative and highly significant (-0.026, $t = -10.8$). The rank result is the economically coherent one—harder-to-value units take longer to sell (Column 1, which is positive under both measures) and are correspondingly less likely to clear within a fixed window—and it matches the acceptance-margin mechanism of Section 3, in which the buyer's risk discount lowers the per-meeting clearing rate and lengthens the search required before a sale. The contrast between the positive level coefficients and the negative rank coefficients in this table is a direct illustration of why valuation uncertainty enters throughout the paper as its within-cell rank rather than its level.

Table OA.20: Valuation Uncertainty, Time on Market, and Sale Probability

	(1) **Time on market** Days on market	(2) **Sale indicator** $\mathbb{1}\{\text{sold}\}$	(3) **Sale $\leq$ 90 days** $\mathbb{1}\{\text{sold} \cap \text{DOM} \leq 90\}$
Valuation uncertainty	3.929^{**}	0.0383^{***}	0.0358^{***}
	(1.447)	(0.0062)	(0.0082)
Community FE	Y	Y	Y
Year–Month FE	Y	Y	Y
Floor FE	Y	Y	Y
Construction Type FE	Y	Y	Y
Observations	578,264	578,264	351,565
R^2	0.14381	0.19531	0.07524
Within R^2	2.44×10^{-5}	0.00022	0.00010

Notes. The sample consists of Beijing apartment listings, January 2015 to January 2021. The dependent variable in Column (1) is days on market, the number of days a listing remains active; this column is estimated on cohorts posted from July 2016 onward, the window with real posting dates (earlier posting dates are backfilled; Figure 4). Column (2) is an indicator equal to one if the listing is matched to a realized transaction (non-missing deal date). Column (3) is an indicator equal to one if the listing sold and the time between posting and deal dates is at most 90 days; the sample for Column (3) is restricted to listings with valid posting and deal dates. All specifications include community and year–month fixed effects and additional property controls (floor and construction type). Standard errors are clustered by community. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

OA.E.8 Cross-market benchmark to Amaral (2024)

The magnitudes in Table OA.21 can be compared with Amaral (2024), who estimate uncertainty discounts for German apartments using repeat-sales identification. Their preferred specification implies an IQR discount of approximately 2–4% of the purchase price. For a comparable German apartment (median price ≈ €250,000), this translates to €5,000–10,000. My Beijing estimate at the IQR (about 6,200 RMB) corresponds to approximately 0.15% of the median purchase price (≈ 4.2 million RMB), an order of magnitude smaller in percentage terms. Two factors reconcile the gap. First, Amaral (2024) identify off repeat-sale price dispersion, which captures a broader notion of uncertainty than the within-community hedonic residual used here; their estimates therefore embed both idiosyncratic and model-specification uncertainty. Second, the Beijing market features tighter listing regulation and higher liquidity than the German market, both of which compress equilibrium risk premia. A third difference concerns how the discount is read. Amaral (2024) reads it as compensated because higher-uncertainty units carry higher rental yields. Total-return constructions built from rent-to-price ratios are measurement-sensitive—matched property-level accounting of income, costs, and prices cuts consensus long-run residential returns roughly in half (Chambers, Spaenjers and Steiner, 2021)—and in Beijing rental yields sat below the one-year risk-free rate over the sample period (Liu and Xiong, 2018), so a yield-compensation channel has little room to operate: a higher yield on a discounted unit is largely the discount re-expressed (Himmelberg, Mayer and Sinai, 2005).

Table OA.21: Implied uncertainty discount for a 76 m^2 apartment

σ_P^2 percentile	Discount (RMB)
P25 (0.0006)	400
Median (0.0028)	2,000
P75 (0.0093)	6,600
P95 (≈ 0.04)	28,300
IQR gap (P75–P25)	6,200

Notes. The discount is computed as $(P_0 - P_0 \exp(\text{slope} \times \sigma_P^2)) \times 76$, where $P_0 = 55{,}000$ RMB/m^2 and slope $= -(1-\alpha)(\Gamma/2)\rho^2 = \beta$, the cross-sectional wedge gradient itself: the structural relabeling through $\Gamma = \bar{\gamma}\,\kappa$ of Equation (OA.18) reconstructs exactly the reduced-form slope, so these are reduced-form magnitudes—conservative relative to the larger but tail-fragile level slope of Table OA.15.

The search-and-matching microfoundation, the systematic/idiosyncratic and Sagi (2021) variance decompositions, and a discussion of extensions (including the relationship to Amaral (2024) and CARA vs. CRRA) are developed in the Online Appendix.

OA.E.9 Attention proxies: popularity and follower-to-viewing ratio

The model predicts that variables capturing demand-side attention should be associated with pricing wedges (Proposition OA.6). I use two attention proxies: a binary popularity indicator based on follower counts above the 75th percentile and the follower-to-viewing ratio (FVR), which captures the composition of attention (committed versus casual) rather than its level.

Popularity. Within community and month, the coefficient on Popular is positive and significant (Table OA.22): popular listings—those that accumulate more follower attention—exhibit higher listing-implied one-month pricing wedges, consistent with harder-to-clear units having to draw more committed attention before they sell.

Follower-to-viewing ratio. Restricting to FVR ≤ 1 to remove mechanically extreme ratios, the within-community–month FVR slope on pricing wedges is negative and significant ($p < 0.01$; Table OA.23). When listing-time uncertainty is the dependent variable (Column 4), the FVR slope is also negative (-0.023, $p < 0.01$): conditional on community and month, listings with a higher committed-to-casual attention ratio are, if anything, slightly easier to value. The follower-to-viewing ratio therefore captures the composition of attention rather than the level, and is informative about how committed the demand for a given listing is.

Table OA.22: Popularity and Listing-Implied Pricing Wedges

Dependent variable: Listing-Implied 1-Month Pricing Wedge (ExpRaw$^{(1m)}$)			
	(1)	(2)	(3)
Popular	0.003671**	0.009355***	0.009575***
	(0.001198)	(0.001114)	(0.001099)
Community FE	Y	Y	Y
Year–Month FE	N	Y	Y
Floor FE	N	N	Y
Constr. Type FE	N	N	Y
Observations	385,818	385,818	385,818
Adj. R^2	0.015488	0.028447	0.028577
Within R^2	0.008090	0.007657	0.007443

Notes. The sample consists of Beijing apartment listings, January 2015 to January 2021. Popular equals one if the listing's follower count exceeds the within-subperiod (earlier vs. later, split at 2016-12-14) 75th percentile. All specifications include community fixed effects. Columns (2)–(3) add year–month fixed effects; Column (3) further adds floor and construction-type fixed effects. Standard errors are clustered by community. ***, **, * denote significance at the 1/5/10% levels.

Table OA.23: Follower–Viewing Ratio (FVR), Pricing Wedges, and Valuation Uncertainty

	(1)	(2)	(3)	(4)
Dependent var.	ExpRaw$^{(1m)}$	ExpRaw$^{(1m)}$	ExpRaw$^{(1m)}$	Valuation uncertainty
FVR	0.025731***	-0.027763^{***}	-0.027418^{***}	-0.023020^{***}
	(0.002950)	(0.003785)	(0.003789)	(0.003862)
Community FE	Y	Y	Y	Y
Year–Month FE	N	Y	Y	Y
Floor FE	N	N	Y	Y
Constr. Type FE	N	N	Y	Y
Observations	249,378	249,378	249,378	351,565
Adj. R^2	0.012088	0.036872	0.038007	0.006262
Within R^2	0.001258	5.82×10^{-4}	5.659×10^{-4}	4.339×10^{-4}

Notes. The sample consists of Beijing apartment listings, January 2015 to January 2021, restricted to FVR ≤ 1. FVR $=$ FOLLOWER/VIEWING. Columns (1)–(3) use listing-implied expected one-month returns ExpRaw$^{(1m)}$ as the dependent variable. Column (4) uses listing-time valuation uncertainty. All specifications include community fixed effects; additional fixed effects are indicated. Standard errors are clustered by community. ***, **, * denote significance at the 1/5/10% levels.

OA.F Mechanism evidence appendix

This appendix collects supporting material for the mechanism analysis in Section 5, including the proxy-to-mechanism map, the time-adjusted-price diagnostic, the FVR-uncertainty regression, supply-discipline and matching-function evidence, size and community heterogeneity, and behavioral validation of the attention proxy.

OA.F.1 Proxy-to-mechanism map

Table OA.24 maps each empirical proxy to the mechanism it captures, the identifying assumption required, and an important caveat. This table is intended as a roadmap for the analysis that follows.

Table OA.24: Mapping Proxies to Mechanisms

Proxy	Mechanism	Identifying assumption	Caveat
Valuation uncertainty (σ_P^2)	Risk pricing by the met buyer pool	Hedonic residual variation is orthogonal to unpriced amenities	Could proxy for information asymmetry
Within-cell uncertainty slope	Cross-sectional pricing of risk	No omitted within-cell confounder correlates with σ_P^2	Could proxy for information asymmetry
Listing–deal wedge	Seller pricing power / markup	Deal price reveals buyer WTP	Seller may concede strategically
FVR (attention)	Buyer search intensity	Platform views $\propto$ demand visits	Supply-side listing refreshes contaminate
FVR residual	Composition vs. mispricing	Orthogonal to listing premium after projection	Residual may capture other unobservables
Time on market (DOM)	Matching efficiency	DOM reflects demand-side frictions	Seller pricing strategy affects DOM
Community listing count	Local supply pressure	Within-community variation is exogenous	Endogenous entry
Community turnover	Local trading intensity	Turnover is predetermined within cell	Turnover also reflects liquidity
Time-adjusted price	Listing staleness / revision	TAP captures seller updating speed	Mechanically related to listing price
Matching tightness (buyers/listings)	Search market conditions	Community-month FVR ratio proxies tightness	Measurement error in buyer count

Notes. Each row maps an empirical proxy used in this section or in Section 5 to the economic mechanism it is intended to capture, the key identifying assumption under which it is informative, and a caveat or limitation. The evidence below is suggestive: the proxies are consistent with the model's predictions but cannot individually establish causation.

OA.F.2 Listing discipline and time-adjusted prices

To probe whether stale asking prices are penalized in the pricing wedge, I study the platform's time-adjusted asking price measure. This variable captures systematic adjustments in the posted asking price associated with listing timing and, in practice, correlates with staleness and revision behavior. Under tighter listing discipline, stale or slow-to-update asking prices should be penalized more in terms of short-horizon pricing wedges.

I estimate:

$$\text{ExpRaw}^{(1m)}_{i,t} = \beta\,\text{TAP}_{i,t} + \mu_{c(i)} + \tau_t + \mathbf{X}'_{i,t}\psi + \varepsilon_{i,t}, \tag{OA.28}$$

where $\text{TAP}_{i,t}$ is the time-adjusted asking price, $\mu_{c(i)}$ are community fixed effects, τ_t are year–month fixed effects, and $\mathbf{X}_{i,t}$ includes floor and construction-type fixed effects when indicated. Standard errors are clustered by community.

Table OA.25 shows that, within community and month, higher $\text{TAP}_{i,t}$ is associated with lower listing-implied pricing wedges: stale, over-asking listings are penalized in the short-horizon wedge, consistent with listing discipline and price discovery. A related regulatory channel operating through listing discipline is documented by Gargano and Giacoletti (2021), who exploit Australian underquoting laws to show that statutory interventions on listing strategies materially compress listing-to-deal wedges. This result also provides indirect support for Proposition OA.5: when sellers anchor on stale asking prices that the market does not validate, the listing–deal wedge widens, particularly for harder-to-value properties.

Table OA.25: Time-Adjusted Prices and Listing-Implied Pricing Wedges

	(1)	(2)	(3)
	Dependent variable: $\text{ExpRaw}^{(1m)}_{i,t}$		
Time-adjusted price	-0.005197^{***}	-0.005556^{***}	-0.005433^{***}
	(0.000201)	(0.000200)	(0.000196)
Community FE	Y	Y	Y
Year–Month FE	N	Y	Y
Floor FE	N	N	Y
Construction Type FE	N	N	Y
Observations	385,818	385,818	385,818

Notes. The dependent variable is $\text{ExpRaw}^{(1m)}_{i,t}$. All specifications include community fixed effects; additional fixed effects are indicated. Standard errors are clustered by community. ***, **, * denote significance at the 1/5/10% levels.

OA.F.3 Attention responses to valuation risk

The model predicts that harder-to-value listings must accumulate more committed attention before they clear (Proposition OA.6(a)). Table OA.26 relates the raw counts of platform attention to valuation uncertainty while controlling for time on market and a saturated set of fixed effects. In these level-count specifications, valuation uncertainty is not systematically related to either raw follower counts or raw page views once the fixed effects and listing duration are absorbed—consistent with the heavy upper tail of $\hat{u}^2$ that motivates the rank design throughout. The committed-attention margin emerges in the main-text Table 6, which measures uncertainty by its within-cell rank and the attention outcomes in logs: there, harder-to-value units accumulate significantly more followers before clearing. The accumulation bundles the model's two margins—per-exposure evaluation and the longer spell over which it accrues (Equation (17))—and the duration-mechanics diagnostic of Online Appendix OA.F.11 (Table OA.33) reports the split: a third of the follower gradient survives at a fixed spell while casual page views do not rise, the composition shift the dispersion channel predicts and spell length alone cannot produce.

Table OA.26: Platform Attention and Idiosyncratic Valuation Risk

Panel A. Dependent Variable: FOLLOWER	
	(1)
Valuation uncertainty	0.2182
	(0.2384)
Days on market	0.1014***
	(0.0025)
Community FE	Y
Year–Month FE	Y
Floor FE	Y
Constr. Type FE	Y
Observations	578,264
Adj. R^2	0.0948
Within R^2	0.0277

Panel B. Dependent Variable: VIEWING	
	(2)
Valuation uncertainty	−53.988
	(36.092)
Days on market	6.550***
	(0.108)
Community FE	Y
Year–Month FE	Y
Floor FE	Y
Constr. Type FE	Y
Observations	412,339
Adj. R^2	0.2348
Within R^2	0.0920

Notes. The dependent variables are raw platform attention counts: FOLLOWER (Panel A) and VIEWING (Panel B). Each column estimates $\text{Attention}_{i,t} = \beta\, \text{Unc}_{i,t} + \theta\, \text{DOM}_{i,t}$ with community, year–month, floor, and construction-type fixed effects. Standard errors clustered by community. ***, **, * denote significance at the 1/5/10% levels.

OA.F.4 Attention composition and uncertainty

If FVR captures the composition of platform engagement (the conversion of broad browsing into high-intent tracking), it may also covary with listing-time valuation uncertainty itself, providing a direct link between attention composition and the risk measure. Table OA.27 regresses valuation uncertainty on FVR, holding constant a saturated set of fixed effects and clustering standard errors by community.

Within community and month, higher-FVR listings exhibit lower valuation uncertainty (-0.023, $p < 0.01$), consistent with high-intent attention being concentrated in listings whose values are easier to benchmark (Chinco and Mayer, 2016; Bayer et al., 2020). This is the same cross-sectional association reported in column (4) of Table OA.23: the composition of attention (the committed-to-casual ratio) covaries with how hard a unit is to value, providing a direct link between attention composition and the risk measure.

Table OA.27: Follower–Viewing Ratio (FVR) and Listing-Time Valuation Uncertainty

Dependent variable: Valuation Uncertainty	
FVR	-0.0230^{***}
	(0.0039)
Community FE	Y
Year–Month FE	Y
Floor FE	Y
Constr. Type FE	Y
Observations	351,565
Adj. R^2	0.0063
Within R^2	0.00043

Notes. The dependent variable is listing-time valuation uncertainty. FVR = FOLLOWER/VIEWING. All specifications include community and year–month fixed effects, as well as floor and construction-type fixed effects. Standard errors clustered by community. ***, **, * denote significance at the 1/5/10% levels.

OA.F.5 Market tightness and a matching-function diagnostic

The preceding subsections established that attention is related to uncertainty within community and month and that the composition component of attention carries independent predictive content for pricing wedges. A remaining question is whether platform attention is an economically meaningful proxy for local demand pressure, that is whether variation in attention maps into actual transaction activity at the community level. To address this, I estimate a simple matching-function diagnostic that relates attention-based demand to inventory sell-through rates.

Construction of community–month supply, demand, and volume. Using the listing panel, I expand each listing into the set of months in which it is active: from its posting month through the month prior to its deal month (or only the posting month if the deal date is missing or the deal occurs in the same month). For each community c and month t, I construct:

$$\text{SUPPLY}_{c,t} \equiv \#\{\text{active listings in } (c,t)\}, \tag{OA.29}$$

$$\text{DEMAND}_{c,t} \equiv \sum_{i\in(c,t)} \log(\text{VIEWING}_i), \tag{OA.30}$$

$$\text{VOLUME}_{c,t} \equiv \#\{\text{completed deals in } (c,t)\}. \tag{OA.31}$$

I then form two ratios that mirror standard market-tightness objects in search-and-matching models (Wheaton, 1990; Head, Lloyd-Ellis and Sun, 2014):

$$\text{VS}_{c,t} \equiv \frac{\text{VOLUME}_{c,t}}{\text{SUPPLY}_{c,t}}, \qquad \text{DS}_{c,t} \equiv \frac{\text{DEMAND}_{c,t}}{\text{SUPPLY}_{c,t}}, \tag{OA.32}$$

where $\text{VS}_{c,t}$ is an inventory sell-through rate (transactions per active listing) and $\text{DS}_{c,t}$ is a demand-per-supply index based on platform attention.

A reduced-form matching regression. I estimate the following community–month specification:

$$\text{VS}_{c,t} = \beta\,\text{DS}_{c,t} + \mu_c + \tau_t + \varepsilon_{c,t}, \tag{OA.33}$$

where μ_c are community fixed effects and τ_t are month fixed effects. Standard errors are clustered by community.

The estimate of β is positive and highly significant (Table OA.28: 0.030, s.e. 0.0008): communities experiencing higher attention per unit of active supply convert inventory into transactions at a higher rate, even after absorbing permanent cross-community differences and aggregate monthly fluctuations. This result provides a micro-level validation of the platform data: attention is not merely noise but reflects economically meaningful variation in local demand pressure. It also supports the use of attention-based measures (FVR, Popular) to characterize the composition and intensity of demand-side participation, as in the preceding analysis.

OA.F.6 Supply pressure and price discovery

A final mechanism check asks whether local supply pressure disciplines listing prices relative to transaction anchors. In the model, an increase in local supply corresponds to a thicker sellers' market that improves matching efficiency and tightens the listing–deal wedge (Wheaton, 1990; Sagi, 2021), so higher supply should be associated with more compressed listing premia.

Table OA.28: Community–Month Matching: Sell-Through and Demand Pressure

	(1) **Dep. var.: $\mathrm{VS}_{c,t} = \mathrm{VOLUME}/\mathrm{SUPPLY}$**
$\mathrm{DS}_{c,t} = \mathrm{DEMAND}/\mathrm{SUPPLY}$	0.0301***
	(0.0008)
Community FE	Y
Month FE	Y
Observations	241,230
Adj. R^2	0.2246
Within R^2	0.0191

Notes. The unit of observation is community c by month t. $\mathrm{SUPPLY}_{c,t}$ is the stock of active listings in (c, t). $\mathrm{DEMAND}_{c,t}$ is the total $\log(\mathrm{VIEWING})$ across active listings in (c, t), and $\mathrm{VOLUME}_{c,t}$ is the number of completed deals. Standard errors clustered by community. ***, **, * denote significance at the 1/5/10% levels.

I proxy for community-level supply pressure using the supply share measure defined in Section OA.C.6 and estimate:

$$\mathrm{Prem}_{i,t} = \beta\,\mathrm{SupplyShare}_{c(i),t} + \mathrm{FE} + \varepsilon_{i,t}, \tag{OA.34}$$

where FE denotes the fixed effects included in each column.

Table OA.29 reports the results. The supply-share coefficient is small and only marginally significant once the fixed effects absorb aggregate and permanent cross-community variation, consistent with supply pressure exerting at most a modest disciplining effect on listing premia within community and month.

Summary of mechanism evidence. Together, the pieces of evidence in this section paint a coherent picture consistent with the illiquidity mechanism: (i) platform attention maps meaningfully into transaction activity at the community level, validating the data; (ii) committed follower attention rises with valuation uncertainty within community and month; and (iii) stale, over-asking listings are penalized in the pricing wedge, consistent with listing discipline. These patterns support reading platform attention as an economically meaningful proxy for demand-side search.

I provide three additional pieces of mechanism evidence below: size-based heterogeneity (Section OA.F.8), community-level attention heterogeneity in the uncertainty discount (Section OA.F.9), and behavioral correlates that validate FVR as an attention-intensity proxy (Section OA.F.10).

OA.F.7 Size-based transaction tax incentives

In addition to credit and eligibility rules, transaction taxes and fees vary discretely with unit size: in many cities and periods, units below the salient $90\,\mathrm{m}^2$ cutoff have been eligible for

Table OA.29: Supply Share and Listing Premium

	(1)	(2)	(3)
	Dependent variable: Listing premium (asinh)		
Supply share	0.0229	0.1670*	0.1482*
	(0.0285)	(0.0799)	(0.0775)
Community FE	Y	Y	Y
Year–Month FE	N	Y	Y
Floor FE	N	N	Y
Constr. Type FE	N	N	Y
Observations	351,565	351,565	351,565
Adj. R^2	-0.0203	-0.0206	-0.0193
Within R^2	5.13×10^{-6}	1.02×10^{-5}	1.01×10^{-5}

Notes. The dependent variable is the within-community listing premium. Supply share is the number of listings in community c in month t divided by the citywide total. Standard errors clustered by community. ***, **, * denote significance at the 1/5/10% levels.

preferential deed-tax treatment, and the size distribution shows a pronounced kink at the threshold, consistent with sorting into the policy-favored segment (Figure OA.5).[24] Because the tax wedge enters the buyer's effective purchase price, the segment below the cutoff plausibly concentrates short-horizon, entry-cost-sensitive participants—the model's type-S buyers—so SIZE $\leq 90\,\mathrm{m}^2$ is the natural segment for the buyer-composition heterogeneity examined next.

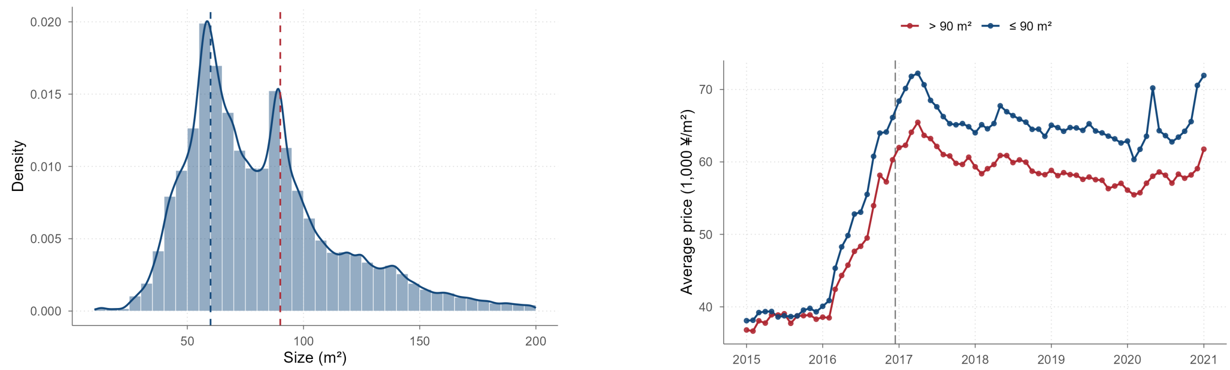
OA.F.8 Size heterogeneity and the tax-favored segment

The $90\,\mathrm{m}^2$ cutoff makes apartments below this threshold a distinct market segment. A natural question is whether the uncertainty discount is confined to this tax-favored niche or holds across the size distribution.

Table OA.30 splits the sample at $90\,\mathrm{m}^2$ and re-estimates the levels specification separately for small ($\leq 90\,\mathrm{m}^2$) and large ($> 90\,\mathrm{m}^2$) apartments. Both segments exhibit a negative and significant level discount of comparable magnitude (small: -0.157; large: -0.181), so no single niche drives the levels result.

The rank-based design sharpens the split into a composition test of the risk-discount channel: the model's discount slope is proportional to the average risk aversion of the segment's buyer pool, and Section OA.F.7 argues the tax-favored small segment concentrates short-horizon, entry-cost-sensitive buyers. The tilt is not an assertion—three direct margins document it. Within a community and month, sellers of small units are 6.0 percentage points more likely to be short-tenure owners (holding under five years by the listing's tax-status

[24]Institutional details (cutoffs, rates, and eligibility definitions) are implemented locally and can change over time. The empirical implication for this paper is that the $90\,\mathrm{m}^2$ cutoff is widely recognized and used as a rule-of-thumb for "policy-favored" small units, potentially concentrating marginal demand in that segment.



(a) **Distribution of Apartment Size** (b) **Average Listing Price per m^2 Over Time**

Figure OA.5: Size-Based Transaction Incentives and Listing Prices over Time

Notes. Panel (a) shows the distribution of apartment size for units no larger than $200\,\mathrm{m}^2$. The vertical lines at $60\,\mathrm{m}^2$ and $90\,\mathrm{m}^2$ mark two salient policy thresholds: the former is commonly used to classify economically affordable housing, while the latter is relevant for transaction-tax treatment. Panel (b) plots the average listing price per square meter over time, with listings split at the $90\,\mathrm{m}^2$ threshold. The vertical line marks the timing of the 2016 Central Economic Work Conference.

field, $t = 21$); investment-use property types are twice as concentrated in the small segment (4.2 versus 2.0 percent); and sale-to-relist gaps conditional on relisting are similar across segments (median fourteen months in both), so the tilt is in who owns the segment, not in relisting speed. A wealth-based intuition cuts the other way—lower-income buyers might be more risk averse per yuan—which makes the observed flattening more informative, not less: what the model's $\bar{\gamma}$ prices is the marginal buyer's exposure to this unit's resale risk over their holding horizon, and short-tenure, investment-type owners bear less of it whatever their wealth. Interacting the within-cell uncertainty rank with a small-unit indicator (mature 2016–2019 cohorts, the specification of Table 10) concentrates the gradients exactly where the model puts the risk-averse buyers: among large units the deal-price gradient is -0.0222 ($t = -9.0$) and shrinks to essentially zero among tax-favored small units (interaction $+0.0217$, $t = 6.9$); the time-on-market and sale-probability gradients flatten by a third to a half (interactions -0.0471, $t = -2.8$, and $+0.0111$, $t = 2.3$). The attention gradient does not differ by segment ($t = -0.6$)—disagreement is present in both segments; its pricing differs with who bears the risk. The heterogeneity is not a thin-segment artifact: controlling the uncertainty rank interacted with the comparable-count rank of Online Appendix OA.D.8 leaves the small-unit interaction unchanged ($+0.0223$, $t = 7.0$). (The contrast with the levels split mirrors the paper's running levels-versus-rank theme: the levels estimates are dominated by the same heavy tail in both segments, while the rank design isolates the cross-sectional gradient.) Two related splits are descriptive only: investment-use property types (2.7 percent of the sample) show a flatter sale-probability gradient but a steeper price discount,[25] and high-transaction-tax listings show slightly steeper gradients, consistent with entry costs amplifying the friction.

Table OA.30: Size Heterogeneity: Small vs. Large Apartments

	(1) Small ($\leq 90\,\mathrm{m}^2$)	(2) Large ($> 90\,\mathrm{m}^2$)
Valuation uncertainty	-0.1574^{***}	-0.1814^{***}
	(0.0450)	(0.0112)
Community FE	Y	Y
Year–Month FE	Y	Y
Observations	257,768	128,050
R^2	0.139	0.228
Within R^2	0.062	0.110

Notes. Dependent variable: listing-implied one-month pricing wedge (ExpRaw$^{(1m)}$). Column (1) restricts to apartments $\leq 90\,\mathrm{m}^2$; column (2) to apartments $> 90\,\mathrm{m}^2$. Standard errors clustered at the community level in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

OA.F.9 Community-level attention heterogeneity

A natural question is whether the uncertainty discount varies with the intensity of platform attention a community attracts. I proxy for community attention intensity using the community-level average follower-to-viewing ratio (FVR). Communities above the median FVR, where listings attracted disproportionately high follow-up engagement relative to casual browsing, are classified as "high-attention."

Table OA.31 reports the results. The cross-sectional uncertainty discount is present in both groups and is somewhat larger in high-attention communities (-0.210, $p < 0.001$) than in low-attention communities (-0.148, $p < 0.001$), consistent with attention being more tightly linked to pricing where engagement is more intense.

OA.F.10 Behavioral validation of the attention proxy

The preceding subsection uses FVR as a proxy for attention intensity. To validate this interpretation, I examine the cross-sectional relationship between FVR and two listing outcomes: days on market (DOM) and the listing–deal price wedge. If FVR captures the

[25] The investment-use cell (apartments on commercial land, hotel-style and office-titled units) is not a clean buyer-composition test: these are structurally different assets—40–50-year land tenure, commercial utility rates, no school-district rights, and tighter financing—so the steeper price discount plausibly reflects greater fundamental uncertainty about the asset class itself, priced by whoever buys it, rather than the buyer pool's risk aversion. I therefore lean on the within-asset-class size split, where the bundle is held fixed and only the buyer pool shifts.

Table OA.31: Attention Heterogeneity: Split by Community FVR

	(1) Low-attention	(2) High-attention
Valuation uncertainty	-0.1479^{***}	-0.2104^{***}
	(0.0246)	(0.0150)
Community FE	Y	Y
Year–Month FE	Y	Y
Observations	175,822	202,975
R^2	0.114	0.118
Within R^2	0.074	0.091

Notes. Dependent variable: listing-implied one-month pricing wedge (ExpRaw$^{(1m)}$). Communities are split at the median of the community-level average follower-to-viewing ratio (FVR). Standard errors clustered at the community level in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

intensity of platform engagement (including both casual browsers and serious potential buyers), higher-FVR listings should exhibit systematically different transaction dynamics.

Table OA.32 reports the results. Within community and month, FVR positively predicts log DOM ($+0.0070$, $p < 0.001$) and positively predicts the listing–deal wedge ($+0.0001$, $p < 0.001$). Higher-FVR listings take modestly longer to sell and trade at a slightly wider gap between the asking and deal price. This pattern is consistent with FVR measuring the breadth of attention: listings with many followers relative to views attract extensive browsing but may face longer price negotiation. The positive coefficients rule out a pure "informed trader" interpretation (in which high FVR would predict faster sales and tighter wedges) and support the broader "attention intensity" framing adopted throughout the paper.

Table OA.32: FVR Behavioral Validation (Pre-2016 Period)

	(1) log(DOM)	(2) Wedge
FVR	0.0070^{***}	0.0001^{***}
	(0.0005)	(1.16×10^{-5})
Valuation uncertainty	-0.0399^{*}	0.0006
	(0.0213)	(0.0008)
Community FE	Y	Y
Year–Month FE	Y	Y
Observations	59,430	99,383
R^2	0.650	0.128
Within R^2	0.009	0.001

Notes. Sample restricted to listings posted before December 2016. Column (1): dependent variable is log days on market. Column (2): dependent variable is the listing–deal price wedge (deal price minus listing price, divided by listing price). FVR is the follower-to-viewing ratio. Standard errors clustered at the community level in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

OA.F.11 Alternative explanations for the scissors: underpricing, bid shading, duration stigma, duration mechanics

This subsection tests the two rational alternatives to the two-channel reading of the price-down/attention-up co-movement (Section 5.2.1), plus a duration-stigma diagnostic and a duration-mechanics diagnostic. Table OA.34 reports the results; all specifications mirror Table 10 (mature 2016–2019 posting cohorts, FVR ≤ 1, within-cell uncertainty rank, community, year–month, floor, and construction-type fixed effects, community-clustered standard errors).

Underpricing for traffic. Anundsen et al. (2022) show a seller can manufacture attention with a low ask: in Norwegian auctions a one-percentage-point markdown raises viewers and bidders while cutting the sale price—attention up, price down, no role for valuation difficulty. The strong attention response to the relative ask estimated below is the engagement analogue of the demand curve facing a seller in Guren (2018). Two facts rule this out as the source of the scissors. First, uncertainty is not underpricing: the squared residual is symmetric in sign, so the top within-cell uncertainty decile is posted below comparables 50.3 percent of the time, and the within-cell rank is uncorrelated with the signed relative ask (-0.006). Second, column (1) adds the signed relative ask (the log listing premium over the hedonic benchmark) to the attention regression. The ask price matters exactly as the traffic-buying channel predicts—listings posted below comparables draw far more followers (-1.63, $t = -16$)—yet the uncertainty gradient is essentially unchanged ($+0.047$ baseline, $+0.044$ controlled). The underpricing channel is alive in the data and orthogonal to the uncertainty effect. (Because the uncertainty measure is the squared list residual, only the linear ask control is a clean test—any nonlinear-in-ask control would mechanically absorb the uncertainty level itself; the linear term is also the direction the traffic channel works through.)

Winner's-curse bid shading. Under common-value uncertainty, rational buyers shade their bids, and the shading deepens with the number of competitors (Milgrom and Weber, 1982). Column (2) interacts the uncertainty rank with the within-cell rank of realized follower attention: the discount is indeed concentrated among high-attention listings. I do not read this as evidence for shading, because realized attention is itself an outcome of disagreement—the model's dispersion channel predicts precisely that the listings requiring the most search are those with the most disagreement—so this moment is consistent with both readings and cannot separate them. The separation rests on margins shading does not produce: bid shading is silent on why sellers of hard-to-value homes revise their asks more often and why the negotiated wedge disperses (Table 10, columns 2 and 4)—both two-sided, seller-inclusive margins. And the setting itself does not clear by auction: Beijing resale matches a single buyer to a seller through sequential agent-accompanied showings and bilateral negotiation (Section 2.3), so there is no common-value auction in which a winner could be cursed.

Duration stigma. Taylor (1999) shows buyers rationally read long time-on-market as a bad-quality signal, so the discount could be spell-length repricing rather than valuation difficulty. Column (3) adds log days-on-market to the price regression: the stigma margin is present (longer spells close lower) but the uncertainty discount survives essentially intact (-0.0074 to -0.0066). Days-on-market is an outcome, so I report this as a diagnostic rather than a causal decomposition; its message is that the discount is not a relabelling of stale-listing dynamics. Note also that the uncertainty measure is fixed at listing, before any spell accumulates.

Duration mechanics. The remaining mechanical alternative is accumulation itself: followers, page views, and revision counts are stocks that grow over the listing spell, and the risk discount lengthens the spell (Equation (17)), so a pure discount raises every stock with no role for disagreement. Table OA.33 re-estimates the disagreement margins with and without log days-on-market (sold listings; the same diagnostic-not-causal caveat as column (3) applies, and the uncertainty measure is fixed at listing). The split is sharp. The follower gradient is $+0.054$ unconditionally and $+0.017$ ($t = 2.5$) at a fixed spell—about a third survives, and followers per day of exposure rise $+0.006$ ($t = 2.1$)—while casual page views do not rise once the spell is held fixed (-0.040, $t = -5.2$ on the analysis screen): per unit of exposure, hard-to-value homes draw more committers but no more browsers, a composition shift that spell length alone cannot produce, since exposure accumulates browsers and committers alike. Because the FVR ≤ 1 screen conditions on the view count, both attention margins are re-estimated without it: the spell-adjusted follower gradient survives—indeed strengthens ($+0.023$, $t = 3.6$)—while the views decline is screen-sensitive (-0.010, $t = -1.3$ unscreened), so the claim the paper relies on is the screen-robust one: at a fixed spell, committed attention rises and casual attention does not. Revision counts lose significance at a fixed spell ($+0.003$, $t = 0.5$): sellers of hard-to-value homes revise more because they list longer, not faster, so column (4) of Table 10 reads as an accumulation margin rather than independent evidence of seller-side disagreement intensity. The centered wedge dispersion is exactly unchanged ($+0.0013$, $t = 10.2$): negotiated-outcome disagreement is no spell artifact. The scissors accordingly rests on the fixed-spell margins—the wedge dispersion and the surviving attention response—with the follower stock read as what Proposition OA.6(a) says it is: the search a hard-to-value listing must accumulate before it clears.

A final price-formation fingerprint, again built without the price residual. Sellers who cannot pin down a home's value should price it more coarsely, and they do: net of the listing's price level, the trailing-zero coarseness of the asking price rises with the within-cell uncertainty rank ($+0.043$, $t = 11.5$), and the share of listings posted on a round 50-wan ($500{,}000$-yuan) grid climbs monotonically from 11.9 to 16.3 percent across uncertainty quintiles. I read coarse pricing as a signature of valuation difficulty rather than a liquidity

Table OA.33: Duration Mechanics: Which Disagreement Margins Survive a Fixed Spell

	(1) Followers $\log(1+F)$	(2) Views $\log(1+V)$	(3) Revisions count	(4) Wedge dispersion |wedge dev.|
Uncertainty rank r	$+0.0507^{***}$	$+0.0064$	$+0.0312^{***}$	$+0.0013^{***}$
	(0.0079)	(0.0096)	(0.0075)	(0.0001)
r, spell-adjusted	$+0.0160^{**}$	-0.0427^{***}	$+0.0043$	$+0.0013^{***}$
	(0.0067)	(0.0077)	(0.0066)	(0.0001)
$\log$ days on market	$+0.5650^{***}$	$+0.7981^{***}$	$+0.4389^{***}$	$+0.0006^{***}$
	(0.0027)	(0.0034)	(0.0040)	(0.0000)
Observations	169,967	169,967	169,967	169,967

Notes. Sold listings in the specification of Table 10 (mature 2016–2019 posting cohorts, follower-to-viewing ratio at most one, within-cell rank of valuation uncertainty $r^{\mathrm{unc}} \in [0,1]$, community, year–month, floor, and construction-type fixed effects; standard errors clustered by community; $N = 170{,}637$). For each outcome, the first row is the unconditional gradient and the second adds log days-on-market; the third row reports the log days-on-market loading. Days-on-market is an outcome, so spell-adjusted rows are accounting diagnostics, not causal decompositions. Column (4) is the absolute deviation of the list-to-deal wedge from its fixed-effects benchmark, as in column (2) of Table 10. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table OA.34: The Scissors Is Not Underpricing, Bid Shading, or Stigma

	(1) Attention $\log$ followers ask-controlled	(2) Price level $\log$ deal price attention interaction	(3) Price level $\log$ deal price TOM-controlled
Uncertainty rank r	$+0.0387^{***}$	$+0.0136^{***}$	-0.0066^{***}
	(0.0071)	(0.0018)	(0.0015)
Relative ask u	-1.8019^{***}		
	(0.1078)		
$r\times$ rank(followers)		-0.0417^{***}	
		(0.0020)	
$\log$ days on market			-0.0123^{***}
			(0.0004)
Observations	217,967	169,967	169,967

Notes. Specifications mirror Table 10: mature 2016–2019 posting cohorts, follower-to-viewing ratio at most one, within-cell rank of valuation uncertainty $r^{\mathrm{unc}} \in [0,1]$, with community, year–month, floor, and construction-type fixed effects; standard errors clustered by community. Column (1): log saved-listing followers on the uncertainty rank and the signed relative ask a (log listing price minus the hedonic benchmark). Column (2): log deal price per m^2 on the uncertainty rank interacted with the within-cell rank of follower attention (sold listings). Column (3): log deal price on the uncertainty rank and log days on market (sold listings; days-on-market is an outcome, so this column is a diagnostic, not a causal decomposition). $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table OA.35: **Disagreement versus Disamenity: the Committed-Attention Sign Split**

	(1) P(sale)	(2) $\log$ TOM	(3) $\log$ foll.
A. Configuration rarity vs. generic disamenity			
Configuration-rarity rank	-0.0167^{***}	$+0.1804^{***}$	$+0.0607^{***}$
	(0.0030)	(0.0099)	(0.0098)
Generic-disamenity indicator	-0.0145^{***}	$+0.0455^{***}$	-0.0261^{***}
	(0.0024)	(0.0078)	(0.0084)
Price-residual rank r (baseline)	-0.0028	$+0.0621^{***}$	$+0.0416^{***}$
	(0.0029)	(0.0096)	(0.0094)
Observations	351,066	257,534	351,066
B. Resale-restricted housing (institutional friction)			
Valuation-uncertainty rank r	-0.0090^{***}	$+0.1021^{***}$	$+0.0528^{***}$
	(0.0028)	(0.0095)	(0.0091)
Resale-restricted	-0.1572^{***}	$+0.0442^{***}$	-0.0500^{***}
	(0.0055)	(0.0162)	(0.0168)
Formerly public housing	-0.1214^{***}	-0.0083	-0.0553^{***}
	(0.0029)	(0.0092)	(0.0087)
Observations	351,122	257,553	351,122
Community $\times$ month FE	Yes	Yes	Yes

Notes. Within-community $\times$ month regressions of the sale probability, log days on market (sold units posted July 2016 onward), and log followers; standard errors clustered by community. Panel A enters the configuration-rarity rank (a price-free physical-rarity measure as in Table 19, here on the room $\times$ bath $\times$ facing $\times$ size-band configuration), a generic-disamenity indicator (bottom or basement floor, or at least five households per elevator), and the baseline price-residual rank r^{unc}. Panel B replaces the disamenity controls with two institutional frictions read from the ownership-type field—resale-restricted units (affordable, price-capped, and resettlement housing) and formerly-public housing—alongside the valuation-uncertainty rank. The discriminant is the follower (committed-attention) column: difficulty that is valuation disagreement raises committed attention, while difficulty that is generic disamenity or an institutional sale friction lowers it, even though both slow the sale. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

channel in its own right—round-priced listings also draw fewer followers and sell faster, the pattern of a motivated or anchored seller—so the clean content is the first-stage gradient from difficulty to coarse pricing.

OA.F.12 Showings versus followers: the costly stage of the funnel

The platform's raw transaction exports record a field the cleaned data had dropped: the number of showings (*daikan*)—times an agent physically took a buyer to the unit—one step costlier than following and one step short of an offer. I recover it by merging the raw export back to the matched-transaction sample on a tight composite key (community, posting date, deal date, deal price, size; ambiguous keys dropped), which matches $250{,}825$ sold listings with a numeric showing count.

The result splits the engagement funnel exactly where the model says it should. Within community and month, harder-to-value homes draw fewer showings (-0.026, $t = -4.0$ on the within-cell rank; -0.018, $t = -2.8$ after trimming the top one percent of $\hat{u}^2$; -0.0204, $t = -2.9$ under the uniform mature-cohort specification of Table 10, column 5), and conditional on their (higher) follower counts the shortfall deepens (-0.046, $t = -8.8$): per follower, hard-to-value homes convert markedly fewer monitors into visitors. The same sample reproduces the deal-price discount (-0.0058, $t = -4.1$), so this is not a sample artifact. Following is cheap evaluation and rises with disagreement (Equation (16)); a showing requires the buyer's expected gain to clear a real cost, so it inherits the threshold logic of the acceptance rate—the shaded mean thins the set of buyers for whom a visit is worthwhile, just as it thins the set for whom the purchase is. Uncertainty raises what is cheap and cuts what is costly: monitoring up; showings, sales, and prices down. This is the strongest single piece of evidence that the followers gradient is disagreement-driven evaluation rather than generic popularity—popular listings get both followers and showings; hard-to-value listings get only the followers. The falling accumulated count itself decomposes as arithmetic—fewer inspections per unit of market time compounded over a longer spell (-4.4 percent and a 0.48 duration elasticity times $+5.3$ percent reproduce the total)—so it adds the spell to the gate margin rather than measuring conversion; the per-visit conversion implied by the sale hazard and the showing inflow moves mildly against hard-to-value homes, the dominance case of Proposition OA.8(iii) (Online Appendix OA.B.5).

Table OA.36: The commitment funnel: cheap and costly attention, total and at a fixed spell

	(1) log foll.	(2) log foll.	(3) log show.	(4) log show.	(5) log $\frac{\text{show.}}{\text{foll.}}$	(6) log days
Uncertainty rank r	$+0.0427^{***}$	$+0.0146^{**}$	-0.0187^{***}	-0.0437^{***}	-0.0614^{***}	$+0.0520^{***}$
	(0.0079)	(0.0068)	(0.0070)	(0.0060)	(0.0067)	(0.0077)
log days on market		$+0.5409^{***}$		$+0.4794^{***}$		
Observations	158,089	158,089	158,089	158,089	158,089	158,089

Notes. One sample throughout: sold listings matched to a numeric showing count on the composite key, posted July 2016–December 2019 (real durations only; the platform's listing history begins mid-2016), FVR ≤ 1, days on market ≥ 1 ($N = 158{,}089$). All columns regress on the within-community-month uncertainty rank with community, year–month, floor, and construction-type fixed effects, clustering by community. Columns (2) and (4) add log days on market, so they compare listings at the same spell length; the duration elasticities (0.54 for followers, 0.48 for showings) are well below one, which is why ratio-per-day measures over-adjust. Trimming the top one percent of $\hat{u}^2$ and re-ranking leaves the fixed-spell gradients at $+0.016$ ($t = 2.4$) and -0.037 ($t = -6.2$) and the passage gradient at -0.056 ($t = -8.4$).

Table OA.36 puts the whole funnel on this one sample, total and at a fixed spell. Followers rise with uncertainty ($+4.3$ percent), and still rise when the comparison holds log days on market fixed ($+1.5$ percent, $t = 2.2$): the extra monitoring is not just a longer shelf life mechanically accumulating eyeballs. Showings fall (-1.9 percent), and fall nearly twice as fast at a fixed spell (-4.4 percent, $t = -7.2$): per unit of market time, harder-to-value homes are inspected strictly less, while they are monitored weakly more. The passage rate—showings per follower, a ratio in which the spell largely cancels—falls 6.1 percent ($t = -9.2$). Reading columns (2), (4), and (5) together with Proposition OA.8: disagreement widens the monitoring pool, and the visit gate then passes a thinner, more selected slice of it per arrival. Selection is visible at the two costly margins in opposite directions—monitoring up, inspection down—which is the pattern a pure risk discount cannot generate on any single stage.

The stopping rule in the data. The showing counts also let the data interrogate the seller's stopping rule. Under a fixed reservation, search is memoryless: the accepted price is the first draw to clear P^S, so conditional on the cell and the rank, the transaction price should not vary with how many showings failed before one converted. Under an updating seller—the partially anchored reservation $P^S = (1-\nu)\bar{S} + \nu S^*$ of Online Appendix OA.A.7—failed showings are news, and the reservation concedes. The data reject strict memorylessness, in the updating direction but with small magnitudes: given the spell length, each log unit of accumulated showings is associated with $+0.019$ more asking-price revisions ($t = 6.6$); the realized concession (log deal price minus log original ask) deepens by -0.04 percent per log showing ($t = -5.1$); and the deal-price level falls by -0.42 percent per log showing ($t = -8.5$). A seller who ignored the showing record entirely would show none of these gradients; a fully optimal Bayesian seller would show much larger ones. The small, precisely estimated concessions sit exactly where the calibrated anchoring weight $\nu = 0.78$ puts the seller: adjusting, but only part of the way. These are equilibrium associations on completed sales—showings, revisions, and prices are jointly determined over the spell—so they discipline the stopping rule's direction, not its structural magnitude.

Caveats: the showing count exists only for matched sold transactions (coverage stated above), and showings accumulate over the spell, so like time on market this margin conditions on sale; revisions count asking-price changes in either direction; the within-cell rank, trimming, and the same fixed-effect structure apply throughout.

OA.F.13 Information arrival and the sale hazard

If hard-to-value homes were merely waiting for public information, news that resolves valuation uncertainty should disproportionately help them sell. The platform makes comparable transactions visible, so each closing is an information event. This subsection builds a monthly spell panel—one row per listing per month at risk, spells capped at twelve months (1.4 million rows from $261{,}000$ listings posted from July 2016, the real-duration window)—and regresses an indicator for selling in month m on the (asinh) number of comparable deals—same community, size within $15\,\mathrm{m}^2$—that closed in month $m-1$, its interaction with the listing's posting-time uncertainty rank, community×calendar-month fixed effects, and months-since-posting fixed effects, clustering by community.

Three results. First, the headline replicates in hazard form under the stricter fixed effects: within a community and calendar month, harder-to-value listings are less likely to sell this month (-0.0052, $t = -6.0$, against a mean monthly hazard of 0.14). Second, comparable closings predict sales: a one-unit rise in asinh arrivals raises the hazard by 1.7 percentage points—twelve percent of the mean monthly hazard ($t = 17.7$)—while closings in distant size bands (30–$60\,\mathrm{m}^2$ away)—which carry segment demand but little valuation information about the unit—slightly lower it, consistent with cross-segment substitution. Read in levels, this is a listing-level estimate of the price-discovery externality that information-based models of housing posit (Gao, Sockin and Xiong, 2021): each transaction emits information that helps neighboring sellers of close substitutes convert their next meeting. A timing falsification sharpens this: next-month arrivals in the own band enter with the opposite sign (lag $+0.023$, $t = 25.6$; lead -0.019, $t = -17.9$). Smooth band-level demand persistence would make both positive; the sign reversal says closings are a flow event—last month's closings reveal information and arriving demand that help a listing clear, while next month's closings are the competitors that absorbed the buyers it did not convert. Third, the benefit of comparable information appears to accrue to the easy-to-value units: the arrival×uncertainty interaction is negative (-0.0029, $t = -2.7$) and the distant-band placebo interaction is zero ($t = -0.4$), consistent with idiosyncratic valuation difficulty that other units' prices cannot resolve—the search-mechanism reading, under which the friction on a hard-to-value home is not the wait for public information but the search for the particular buyer whose private valuation clears the bargain. In the count specification this interaction attenuates and loses significance when next-month arrivals are added (-0.0016, $t = -1.4$). A binary within-spell event-study form sharpens it. Building a monthly hazard panel and flagging whether any close comparable (same community and bedroom count, size within $15\,\mathrm{m}^2$, a different listing) resolves in the current spell-month, the resolution×uncertainty interaction is -0.031 ($t = -21.7$), while the future-comparable placebo interaction—a close comparable resolving in the next spell-month—is a clean null (-0.0001, $t = -0.1$, with size and bedroom controls), an order of magnitude smaller. The differential the count form could only call suggestive is robust in event-study form, and the future-comparable placebo rules out a symmetric pre-trend: comparable information that helps a typical unit clear does disproportionately little for the hard-to-value one. Why—because the hard-to-value unit has fewer comparables to begin with: holding community listing thickness fixed, a unit one step higher in the within-cell uncertainty rank sits in a physical cell (community × size band × bedrooms) with about 15 log points fewer contemporaneously-live close substitutes (-0.151, $t = -18.5$), so the thin substitute shelf that lengthens its sale is itself a consequence of idiosyncratic difficulty. I read the differential resolution response as robust in the event-study form; the level results—the hazard replication and the lead–lag asymmetry—remain the content the appendix leads with.

One construction choice deserves a check of its own: the paper ranks uncertainty within the posting cohort, but a long-lived listing competes with the current stock of listings, not with the cohort it entered with. Re-ranking within the active stock of the community in each calendar month makes the hazard gradient roughly twice as steep (-0.0152, $t = -15.1$, versus -0.0080 for the cohort rank in the same spell panel): the cohort-based headline is the conservative choice, not a flattering one.

The stock's composition provides a regression-free cross-check of the acceptance gradient. Because the within-cell rank is uniform in the inflow by construction, steady-state renewal logic makes the active stock's rank profile a direct trace of relative durations, $f_{\text{stock}}(r) \propto 1/q(r)$ (Online Appendix OA.A). Counting listing-months, the shelf over-represents the hardest-to-value decile relative to the easiest by a factor of 1.05 (and 1.07 among fully observed sold spells)—almost exactly the 1.05 the time-on-market regression implies. Selection also shows up as survivor drift: the surviving stock's mean posting rank rises from 0.500 at listing to 0.513 by the sixth month, sign-consistent with q falling in uncertainty though quantitatively small, so composition accounts for only a sliver of the hazard's duration profile. (Interior deciles of the rank pile mechanically at 0, $\frac{1}{2}$, and 1 in small cells, so the accounting uses the symmetric extremes.)

The composition also moves with the cycle, in the direction the thin-market amplification implies: when a community cools, its shelf hardens. Regressing the active stock's mean cohort rank on the cell's sale rate, with community and calendar-month fixed effects (cells with at least ten listing-months, weighted by size; the panel applies the FVR screen and the real-duration window, cohorts posted from July 2016), gives -0.029 ($t = -9.0$) contemporaneously and -0.021 ($t = -6.7$) on the previous month's sale rate—cold months leave behind a measurably harder-to-value stock. The magnitude is honest, not dramatic: a one-standard-deviation swing in the within-community sale rate (0.121, against a mean cell sale rate of 0.140) moves the tilt by about 0.004 rank points, an order of magnitude below the cross-sectional gradients, which is why community$\times$month fixed effects absorb it harmlessly in the main design. Finally, the allocation denominator in Equation (OA.6) is itself visible as mild attention dilution: within community, per-listing followers fall by 3.9 percent per log of active-shelf size ($t = -3.5$; 7.1 percent per day on real-duration sold spells, $t = -5.5$). Under $\lambda_h = \Lambda_{c,t}\, w(r_h)/\sum_{k \in S_{c,t}} w(r_k)$ this puts the buyer-entry elasticity near 0.96: entry almost, but not quite, scales one-for-one with the shelf. Hot cells attract listings and buyers together, biasing the dilution estimate toward zero, so the measured -3.9 percent is a lower bound on congestion; even so, it is small. Together the cyclical tilt and the dilution elasticity are the first two moments of the equilibrium feedback the framework leaves to future work, and both come back with the predicted sign and second-order size.

OA.G Robustness appendix

This appendix collects the robustness exercises summarized in Section 7, including sample-definition and FVR-cleaning checks, alternative clustering levels and uncertainty proxies, hedonic-precision checks (minimum community size and community×month fixed effects), and a double-sorted portfolio analysis.

OA.G.1 Sample definitions and FVR cleaning

The baseline results are reported on the FVR ≤ 1 sample to remove mechanical outliers arising from missing or miscoded view counts (Section 2.4.4). A concern is that this filter systematically excludes listings with specific attention profiles, potentially biasing the estimates. The illiquidity results are robust to the choice of sample: as reported in Section 7, restricting to the matched-transaction subsample and to the FVR ≤ 1 sample leaves the sign and significance of all three outcomes (probability of sale, time on market, and follower attention) intact, and the same holds under the full unrestricted sample, a winsorized-FVR sample, and a sample that trims extreme page-view realizations.

OA.G.2 Clustering levels and alternative uncertainty proxies

The baseline specification clusters standard errors at the community level. If uncertainty is spatially correlated across nearby communities, this may understate the true standard errors. I re-estimate the main specification clustering at three alternative levels: (i) the construction zone (CZ), which groups nearby communities; (ii) the district (REGION), a coarser administrative unit; and (iii) two-way clustering by community and year–month. The identifying variation is within-community cross-sectional, so clustering at coarser geographic levels provides a conservative adjustment for any residual spatial dependence. The point estimates are mechanically unchanged across clustering levels, and the standard errors are very close to the baseline; the illiquidity relationship remains significant under all of these clustering schemes (Section 7).

In addition, I verify that the results are not sensitive to the specific hedonic specification by replacing the baseline uncertainty proxy with alternatives: (i) uncertainty constructed from deal-market residuals ($\mathrm{Unc}^{\mathrm{deal}}_{i,t}$) rather than listing-market residuals, (ii) the absolute residual $|\widehat{u}_{i,t}|$ instead of the squared residual, and (iii) a hedonic specification that omits the year-interacted polynomials in size and construction year, relying solely on community–month fixed effects and discrete housing attributes. The squared residual is the theoretically correct measure for the risk-pricing interpretation: under the CARA–Normal model the risk premium is proportional to the variance σ_P^2, which the squared residual $\hat{u}_i^2$ estimates without bias under a Gaussian specification, whereas the absolute residual estimates σ_P rather than

σ_P^2 ($\mathbb{E}[|\hat{u}_i| \mid \sigma_P] = \sigma_P\sqrt{2/\pi}$). The results are stable across these measurement choices.

OA.G.3 Hedonic precision: minimum community size monotonicity

A potential concern is that communities with very few listings produce imprecisely estimated hedonic fixed effects, which would contaminate the squared residual and inflate the tails of the uncertainty distribution. To address this, I re-estimate the entire hedonic regression (1) on progressively restricted samples that require each community to have at least X total observations, then re-derive both $\text{Unc}_{i,t}$ and $\text{ExpRaw}^{(1m)}$ from scratch before re-running the baseline specification.

Table OA.37 reports results for $X \in \{30, 50, 100, 200\}$. The cross-sectional uncertainty discount $\widehat{\beta}$ is negative and significant across every threshold (-0.19 to -0.21) and, if anything, sharpens as the sample is restricted to denser communities, where the hedonic fixed effects are estimated more precisely. This pattern implies that small-community noise in the full sample attenuates the true discount: the better estimated the hedonic fixed effects, the stronger the result. This monotonic pattern directly addresses the omitted-quality concern: if the discount reflected changed valuation of omitted attributes, restricting to denser communities, where the hedonic absorbs more local variation and omitted-variable bias is compressed, should weaken the result, not strengthen it.

Table OA.37: Robustness: Minimum Community Size for Hedonic Estimation

Min. obs	Communities	N	$\hat{\beta}$	Within R^2
30	3,110	212,143	-0.213^{***}	0.0503
50	2,170	199,111	-0.196^{***}	0.0510
100	1,050	156,962	-0.191^{***}	0.0525
200	316	83,852	-0.189^{***}	0.0573
Baseline (no community size filter): $\hat{\beta} = -0.173$				

Notes. Each row re-estimates the full hedonic regression (1) on communities with at least the indicated number of total observations, then re-derives price uncertainty (squared residual) and the quality-adjusted listing price from the new hedonic. $\text{ExpRaw}^{(1m)}$ is reconstructed using the re-derived raw prices. The baseline specification (community, year–month, floor, and construction-type FE; community-clustered SEs) is then re-run on the restricted sample. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

OA.G.4 Community-times-month fixed effects

The baseline hedonic (1) uses additive community and year–month fixed effects. This absorbs time-invariant community heterogeneity and aggregate market trends, but does not absorb community-specific time-varying demand shocks: for instance, a community experiencing a localized boom in a particular quarter. As a stronger specification, I re-estimate the hedonic with community×month fixed effects ($\mu_{c(i),t}$ replacing $\mu_{c(i)} + \tau_t$), so that the residual is mean-zero within each community–month cell and captures only the apartment-specific deviation from contemporaneous within-community comparables.

The tradeoff is that many community–month cells contain only a handful of listings, leaving insufficient variation to estimate the interaction FE precisely. I therefore restrict to communities that average at least X observations per active month (and span at least 12 active months), and test $X \in \{3, 5, 8\}$. For each threshold I re-derive price uncertainty and quality-adjusted listing prices from the new hedonic and reconstruct ExpRaw$^{(1m)}$ from these updated inputs. The baseline additive-FE construction is retained for the dependent variable; only the uncertainty regressor is replaced with the interaction-FE version.

Re-deriving valuation uncertainty from the community$\times$month hedonic and re-running the cross-sectional specification, the uncertainty discount strengthens relative to the additive-FE baseline as the hedonic fixed effects are estimated on denser cells. At the most inclusive threshold ($X = 3$; 519 communities, 120,000 observations) the cross-sectional discount is $\widehat{\beta} = -0.217$ (vs. -0.173 in the additive-FE baseline); at $X = 5$ (109 communities) $\widehat{\beta} = -0.390^{***}$; and at $X = 8$ (25 communities, 14,669 observations) $\widehat{\beta} = -0.684^{***}$.

The pattern mirrors the minimum-community-size results in Table OA.37: better-estimated hedonic fixed effects produce a stronger, not weaker, uncertainty discount. This implies that the baseline additive specification attenuates the true signal by leaving community-specific time variation in the residual. Under community$\times$month FE, the squared residual is mean-zero within each community–month cell, supporting the view that the result is not driven by community-specific time-varying mispricing or rational bubbles.

The same promotion applies to the outcome regressions. Table OA.38 re-estimates every headline outcome (probability of sale, log days on market, log followers, the list–deal wedge, price revisions, and withdrawal) replacing the additive community and year–month effects with fully interacted community$\times$month fixed effects, each on its published sample and conventions. Every coefficient is unchanged to the third decimal, because the within-cell rank regressor already carries only within-cell information: rows in singleton cells, where a within-cell rank does not exist, drop from both specifications (10–27 percent of each sample), and the interacted effects then absorb nothing the additive effects and the rank had not.

Note. Online Appendix OA.G.7 examines the post-Dec-2016 timing transparently and non-causally.

OA.G.5 The predetermined index: time-split validation and ablation

Three audits back the index of Section 2.4.3. First, a strictly forward-looking construction trains the hedonic and the variance model only on listings posted through December 2017 and predicts listings posted from 2018 onward, so no future information of any kind enters. The level calibration degrades, as expected across a regime change (the out-of-sample R^2 in levels turns negative), but every rank-based object the paper uses survives: the index

Table OA.38: Interacted Community×Month Fixed Effects: The Full Outcome Suite

	Additive FE (published)			Community×month FE		
Outcome	Coef.	(SE)	N	Coef.	(SE)	N
Probability of sale	-0.0061^{***}	(0.0020)	267,911	-0.0060^{***}	(0.0020)	267,911
Days on market (log)	$+0.0511^{***}$	(0.0067)	193,815	$+0.0512^{***}$	(0.0067)	193,815
Followers (log)	$+0.0395^{***}$	(0.0063)	267,911	$+0.0393^{***}$	(0.0063)	267,911
List–deal wedge	$+0.0011^{***}$	(0.0002)	193,815	$+0.0011^{***}$	(0.0002)	193,815
Price revisions ($\log(1 + \#)$)	$+0.0071^{***}$	(0.0022)	346,574	$+0.0070^{***}$	(0.0022)	346,574
Withdrawal	$+0.0124^{***}$	(0.0016)	468,698	$+0.0123^{***}$	(0.0016)	468,698

Notes. Each row pairs the published additive-FE specification with its fully interacted community×month counterpart, estimated on the identical sample and conventions: sale, days on market, and followers under the baseline attention screen (FVR ≤ 1; days on market on sold units); the wedge on sold units; revisions on the published revision-regression sample; withdrawal under the three-fate conventions of Table 7 with no attention screen. The regressor is the within-cell rank of valuation uncertainty; standard errors are clustered by community. Listings in singleton community–month cells, where the rank is undefined, are excluded from both columns. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

rank predicts the absolute out-of-sample error ($t = 23.9$), the twin ask gap ($t = 8.5$), and the wedge dispersion ($t = 8.8$), and the sale and withdrawal gradients on the 2018-onward sample (-0.0249 and $+0.0239$) are about ninety percent of their cross-fitted counterparts on the identical window (Table OA.39). Second, an ablation rebuilds the index from feature subsets (Table OA.40): transaction-comparable features alone and configuration surprisal alone each deliver about ninety percent of the execution gradients, the two blocks are complementary in the within-cell variance they span, and the comparable-count and recency features are nearly redundant given similarity, dispersion, and surprisal. No single ingredient carries the result. Third, because the index itself carries estimated weights, a nested bootstrap rebuilds the entire pipeline inside each of two hundred community pairs-cluster draws (the trailing-comparables join, the configuration surprisal, the cross-fitted hedonic and variance model, and the frozen within-cell rank) and re-estimates the sale and withdrawal gradients on the full universe. The bootstrap standard errors are 0.0020 on both margins, within twelve percent of the analytic cluster-robust values (0.0018), with bootstrap means (-0.0245, $+0.0239$) on top of the point estimates (-0.0241, $+0.0236$): the index's generated-regressor uncertainty, like the residual measure's (Appendix B), is immaterial.

OA.G.6 Xicheng audits: permutation inference, per-day attention, intensity

Three audits of the multi-school event study (Section 5.3). Table OA.41 reports permutation inference: the treatment is reassigned to each of fourteen never-treated districts at the real date and to shifted announcement dates within the pre-reform sample (forward shifts are infeasible, as the data end in December 2020), and the real estimate is ranked in the

Table OA.39: Time-Split Validation of the Predetermined Index

2018+ sample	Time-split index		Cross-fold index	
	Coef.	(SE)	Coef.	(SE)
$\lvert$OOS hedonic error$\rvert$ $\sim$ rank	$+0.0141^{***}$	(0.0006)	—	
Twin $\lvert$log ask gap$\rvert$ $\sim$ pair index (z)	$+0.0057^{***}$	(0.0007)	—	
Cell SD(wedge) $\sim$ cell index (z)	$+0.0017^{***}$	(0.0002)	—	
P(sale) $\sim$ rank	-0.0249^{***}	(0.0026)	-0.0284^{***}	(0.0026)
Withdrawal $\sim$ rank	$+0.0239^{***}$	(0.0025)	$+0.0272^{***}$	(0.0026)

Trained through 2017-12; forward OOS R^2 (abs. error, 2018+) $= -0.032$.

Table OA.40: Feature Ablation of the Predetermined Index

Variant	Within-cell R^2 vs full	P(sale)		Withdrawal	
		Coef.	(SE)	Coef.	(SE)
(a) comp features only	0.647	-0.0229^{***}	(0.0018)	$+0.0220^{***}$	(0.0018)
(b) surprisal only	0.525	-0.0229^{***}	(0.0019)	$+0.0223^{***}$	(0.0019)
(c) full $-$ surprisal	0.647	-0.0229^{***}	(0.0018)	$+0.0220^{***}$	(0.0018)
(d) full $-$ count/recency	0.986	-0.0241^{***}	(0.0018)	$+0.0234^{***}$	(0.0018)
(e) full	1.000	-0.0250^{***}	(0.0018)	$+0.0243^{***}$	(0.0018)

resulting placebo distribution. The sale estimate sits third of seventeen valid assignments (pseudo-$p = 0.222$; the -8-quarter date placebo is degenerate, with no pre-period, and excluding it improves the rank to second of sixteen). The per-day follower rate, unlike the follower stock, is unremarkable in the placebo distribution. Table OA.42 splits Xicheng communities by terciles of their pre-reform ask premium over the district mean: the sale effect is largest in the middle tercile, not the top, so the design does not confirm the monotone school-premium intensity gradient a pure opacity mechanism would predict.

Table OA.41: Xicheng: Permutation Inference

Outcome	Real estimate	Placebo mean	Placebo SD	max $\lvert$placebo$\rvert$	Rank	Pseudo-p
P(sale $\leq$ 13 wk)	-0.0854^{***}	$+0.0057$	0.1039	0.2957	3/17	0.222
log(1+followers)	$+0.1925^{***}$	$+0.0767$	0.4099	1.4494	3/17	0.222
log followers per day	-0.0497^{*}	$+0.1756$	0.4182	1.7511	14/17	0.833

19 placebos: 14 district reassignments (real date) + 5 date shifts ($-8..-4$q, pre-reform sample); $+4..+8$q shifts infeasible (data end 2020-12). Pseudo-$p = (1+k)/(1+n)$.

OA.G.7 Timing: the gradients concentrate in the post-2016 market

The paper's identification is cross-sectional, but the sample spans a structural turn in Chinese housing policy: the 14 December 2016 Central Economic Work Conference, which elevated

Table OA.42: Xicheng: Pre-Reform Ask-Premium Intensity Terciles

Outcome	Bottom premium tercile Coef.	(SE)	Middle Coef.	(SE)	Top (school-premium) Coef.	(SE)	Top − bottom Diff.	(SE)
P(sale $\leq$ 13 wk)	-0.0838^{***}	(0.0182)	-0.1293^{***}	(0.0217)	-0.0489^{**}	(0.0199)	$+0.0349$	(0.0264)
log(1+followers)	$+0.1422^{***}$	(0.0521)	$+0.3137^{***}$	(0.0454)	$+0.1285^{**}$	(0.0528)	-0.0138	(0.0724)
log followers per day	-0.0672	(0.0462)	-0.1181^{***}	(0.0403)	$+0.0255$	(0.0633)	$+0.0926$	(0.0770)

Terciles of pre-reform (2018-05..2020-04) community mean log-ask premium over the Xicheng mean.

"houses are for living, not for speculation" to official doctrine and ushered in the 2017 tightening of Beijing purchase and credit rules. As a transparent—but explicitly non-causal—look at timing, Table OA.43 interacts the within-community–month uncertainty rank with an indicator for postings after that date, holding the community and year–month fixed effects fixed; the interaction is the change in the within-cell uncertainty gradient at the break.

The within-cell gradients are flat or negative before the break and emerge or steepen sharply after it: the uncertainty–time-on-market slope rises by $+0.124$ ($t = 6.2$; this row is estimated on the real-duration window, cohorts posted from July 2016, whose short pre-break stretch is the boom tail with a negative gradient of -0.050), committed attention by $+0.069$ ($t = 3.5$), and asking-price revisions by $+0.055$ ($t = 5.1$), while the sale probability falls by -0.011 ($t = 1.8$) and the deal-price discount by -0.009 ($t = 2.5$)—each from a pre-break gradient that is flat or of the opposite sign (the revisions pre-gradient is -0.013, marginal at ten percent). The one exception is the negotiation-wedge dispersion, which is positive throughout and does not move at the break. The pattern is consistent with the belief channel being a feature of the modern, speculation-constrained market, in which idiosyncratic valuation disagreement—rather than a rising tide that lifts every unit—governs how a hard-to-value home trades.

I do not read this as causal, for three reasons. The break is a single aggregate date, so it is confounded with the broader 2017 Beijing cooling and the concurrent purchase restrictions. The pre-period is short and thin ($38{,}347$ listings, 13% of the sample, concentrated in the noisier early data), so the flat pre-break gradients are imprecise rather than firmly zero. And this is the very timing the abandoned regime-difference design tried to exploit; the paper relies on the cross-sectional result and treats the 2021 reference-price reform as the cleaner causal follow-up. Table OA.43 is therefore suggestive corroboration of the mechanism, not a test of it.

OA.G.8 Double-sorted portfolio analysis

As a finance-style robustness exercise, I sort listings into quintiles of the within-community–month uncertainty rank (cells with at least five listings) and compute the average listing-implied pricing wedge per quintile. The high-minus-low (H−L) spread summarizes the cross-sectional pricing of uncertainty without imposing a functional form.

Table OA.43: Timing: Within-Community–Month Uncertainty Gradients Before and After 14 December 2016

	Pre-Dec-2016 gradient		Change after break ($\times$ Post)	
Outcome	Coef.	(SE)	Coef.	(SE)
Probability of sale	$+0.0025$	(0.0056)	-0.0108^{*}	(0.0060)
Days on market (log)	-0.0502^{***}	(0.0184)	$+0.1245^{***}$	(0.0200)
Followers (log)	-0.0157	(0.0186)	$+0.0692^{***}$	(0.0199)
Asking-price revisions	-0.0128^{*}	(0.0075)	$+0.0551^{***}$	(0.0108)
Negotiation-wedge dispersion	$+0.0014^{***}$	(0.0004)	-0.0001	(0.0004)
Deal price (log)	$+0.0007$	(0.0036)	-0.0091^{**}	(0.0036)
Observations	pre-Dec-2016 38,347		post 249,259	

Notes. Each row is a separate regression of the row outcome on the within-community–month uncertainty rank r^{unc}, an indicator for postings after 14 December 2016 (Post), and their interaction, with community, year–month, floor, and construction-type fixed effects and standard errors clustered by community (2015–2019 postings, follower-to-viewing ratio at most one). The first pair of columns reports the pre-break gradient (r); the second reports the change after the break ($r\times$Post). The Post main effect is absorbed by the year–month fixed effects. This is a descriptive timing decomposition, not a causal design (see text). $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Figure OA.6 plots the portfolio averages. The H$-$L spread is negative and significant (-0.0036, $t = -3.9$; the spread's standard error is per-listing, not community-clustered), and the shape is informative about where the discount lives: the mean wedge is flat to mildly rising through the fourth quintile and drops sharply in the highest-uncertainty quintile, so the discount is concentrated among the hardest-to-value listings in a cell rather than spread evenly across the distribution. The sort reproduces the cross-sectional uncertainty discount nonparametrically; its concentration in the top quintile is consistent with the two-channel reading, in which the risk discount and the search friction bind hardest where valuation difficulty is greatest.

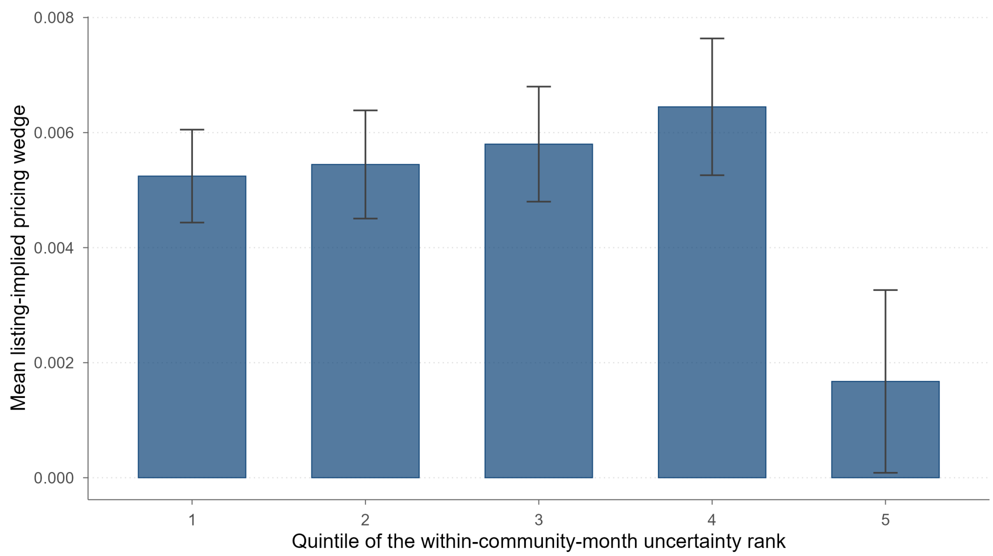


Figure OA.6: Portfolio Sort: The Pricing Wedge by Uncertainty Quintile

Notes. Average listing-implied pricing wedge ($\text{ExpRaw}^{(1m)}$) by quintile of the within-community–month rank of valuation uncertainty, computed in community–month cells with at least five listings. The high-minus-low (Q5−Q1) spread is -0.0036 (standard error 0.0009). Error bars are 95% confidence intervals.