Searching with Non-binding Asking Prices

Irisa Zhou [*]

University of Toronto

May 21, 2026

Abstract

Asking prices in the housing market are not binding, yet they shape buyer behavior and determine transaction outcomes. This paper examines empirically and theoretically the role of asking prices. Using detailed transaction data from Toronto, Canada, I study how sellers trade off days on market and expected sale price in response to anticipated market tightness in the segment of the market they are in. Sellers can either list low to attract competition and sell quickly, or list higher and wait longer in the market. I identified seller types via Natural Language Processing in listing texts, investigated which seller types benefit and suffer the most, and validated these findings with a directed search model involving financially heterogeneous buyers. Lower asking prices directly attract more buyers and raise the probability of a bidding war when buyer heterogeneity is high; they are costly when tightness is low or underpricing fails to coordinate a pool. The model predicts that the sold-to-ask ratio peaks where competition is strong enough to reward underpricing and auction failure is limited. This theory implies that relaxing buyer financial constraints can raise sale prices, shift surplus to sellers, and reduce bidding-war frequency.

Keywords: Asking Price, Ask-Sale Gap, Heterogeneous Agents Directed Search, Endogenous Housing Search, Financially Constrained Buyers

JEL Codes: D83, R00, R21, R31

---

[*]Department of Economics, University of Toronto. Email: irisa.zhou@mail.utoronto.ca. Comments and suggestions from the discussion with Ronald Wolthoff, William Strange, Mark Rempel, Min Fang, Robert McMillan, Han Lu, participants at the 2024 Canadian Economics Association Conference, Urban Working Group, Macroeconomic Brown bag, 4060 Graduate Research Group, 5060 Discussion Group, Public Economics Working Group and IO Structural Group were very constructive. I am also thankful to all the helpful discussions with Sijie Lin and Guangbin Hong, for their support and encouragement in writing this paper. All errors are my own.

1 Introduction

Real estate transactions are among the most important financial decisions households make, yet the first price observed in the transaction–the asking price–is not a binding price. In fact, the asking price is observed to be higher, lower, or equal to the sale price all the time. This real estate transaction is typically initiated by a seller (he) posting an asking price for the property he holds and one of the potential set of interested buyers (she) will bid and purchase the property with a sale price. The gap between these two prices is economically meaningful because it reflects seller strategy and buyer competition, and more importantly, allow us to understand how a non-binding price can be used as a strategic instrument to direct the transaction of a good.

When sale price is higher than the asking price, the market typically refer to this situation as a "Bidding War". Notably, most news concern the bidding wars as a result of frenzied buyers in the real estate market (**wahi˙bidding˙wars**), and this phenomenon has been long described as a problem in buying a house for buyers in these markets. However, people tend to overlook that asking prices have almost no commitment in the transaction, and is solely set by the sellers. Conversations with experienced real estate agent also reflected that "listing on the low side" could be used to "drive [buyer] emotion" to reach the end goal of selling above market value (**torontorealtyblog˙list˙low˙strategy**). In fact, interestingly, in the City of Toronto in 2021, 55.6% of houses have a higher sale price than the asking price. This paper studies when sellers use low asking prices to coordinate buyer attention, when they instead price near or above expected transaction value, and why both strategies can coexist in equilibrium.

The prevalence of sale prices above asking prices is unusual relative to much of the earlier housing literature. Most literature suggests extremely low occurances of higher sale price to asking price in the real estate market. For example, **carrillo2012empirical** states there are only 1.5% of sold over ask, **Horowitz** claims that "Houses are routinely sold at prices below, and are virtually never sold at prices above, their list prices", and **han2016role** reports 10% to 30% at most for houses sold between 2006 and 2009, inclusive. On the contrary, I found that in the City of Toronto, the average sold-to-asking ratio is above 1 in almost all neighbourhoods starting from 2016, and the average sold-to-asking ratio in 2021 is 1.06, i.e. the sold price is on average higher than the asking price (**report**). The empirical object is therefore not simply the existence of a few bidding wars, but the systematic use of the asking price as a strategic instrument across market segments.

Moreover, asking prices have a differential effect across income and the price of houses. Figure **??** plotted the average sold-to-ask ratio across sold prices. The sold-to-ask ratio first increases and then decreases with the sold price. The peak of the plot is around $1.4 million of the sold price. The central explanation I develop via theory and empirical support is that sellers price strategically in response to market thickness, buyer heterogeneity, and the risk that an attempted auction fails.The key object is therefore the seller's regime choice: whether to use a low asking price to coordinate many buyers into an auction-like outcome, or to use a higher asking price and wait for a buyer with a sufficiently high private valuation.

I ask three questions in this paper. First, how does a non-binding asking price affect the matching and trade process in the housing market? Second, which market conditions make underasking profitable rather than risky? Third, how do policies that relax buyer financial constraints change bidding-war probabilities, sale prices, and surplus allocation?

I answer these questions in two steps. First, in Section **??**, I use a simple directed-search model to discipline the mechanism: a non-binding asking price can still matter because it changes which buyers find it worthwhile to visit, how many buyers arrive, and how costly it is for sellers to attract a larger pool. The model is intentionally simple; its purpose is to organize the trade-off, not to carry the full empirical argument. Second, I use Toronto transaction and listing data to document where this trade-off appears in the market and which sellers are most exposed to auction failure.

Since this paper mainly focuses on how sellers use asking price to strategically price, the organizing decision problem is naturally seller-side centered. Each seller enters the market with two relevant state variables: the type of property being sold and the thickness of the market segment the seller faces. These inputs shape the expected buyer pool and therefore the seller's asking-price strategy. A seller can underask in an attempt to coordinate buyer attention and initiate an auction-like outcome, or ask near or above the expected transaction value and wait for a buyer with a sufficiently high private valuation to negotiate.

Figure **??** guides the thought process of the hypothesis for the paper. The model explains why the asking price can map seller-side conditions into buyer-side search behavior. The empirical analysis then studies the same decision tree: when auction entry is more likely, which listings can make underpricing effective, and what penalty sellers face when a bidding war does not materialize.

The empirical section, Section **??**, is the main evidence for this mechanism. Using the universe of Toronto residential sales in 2021, I show first that the sold-to-ask ratio follows a pronounced inverse-V pattern over the price distribution: sale-over-ask outcomes are largest in the middle of the market, not at the bottom or the luxury tail. This pattern is concentrated among fast-selling listings and is stronger in thicker neighbourhood-month markets, where the probability of selling above ask is higher. These facts show that asking prices are not simply noisy value signals; they are associated with distinct transaction regimes.

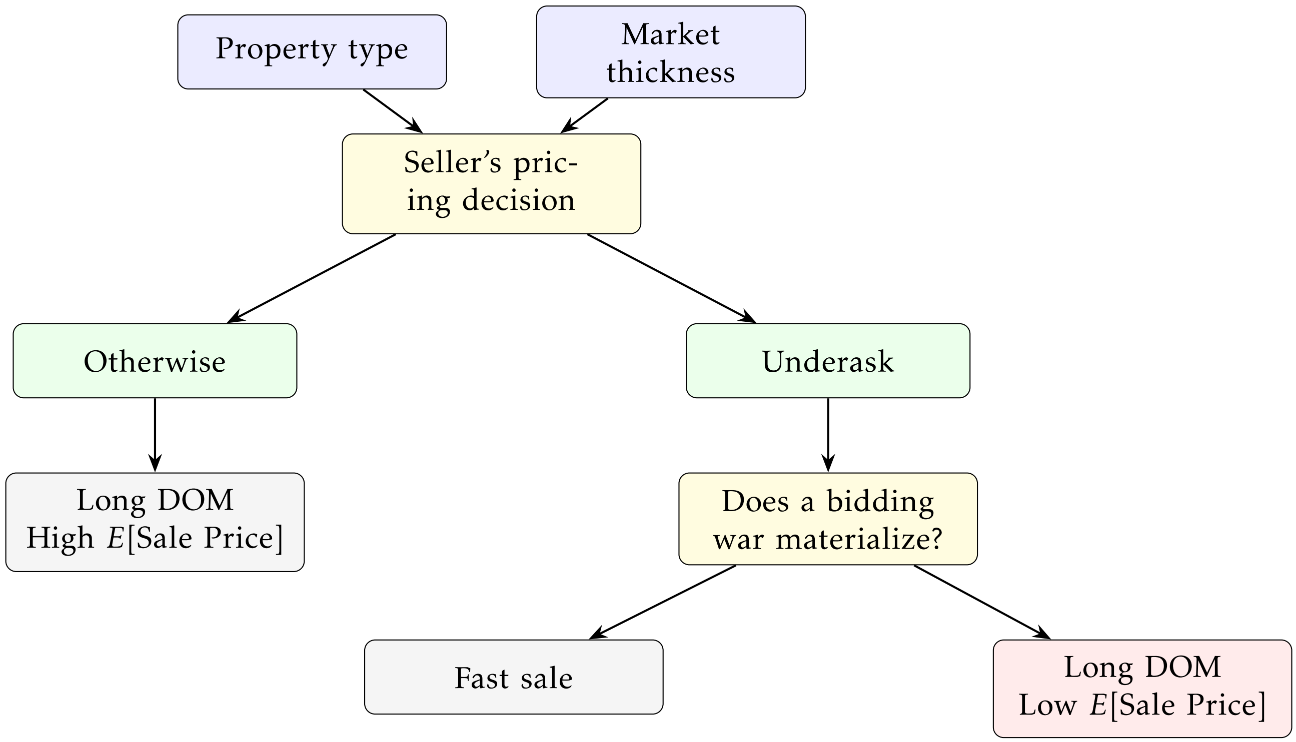


Figure 1: Seller Asking-Price Decision and Outcomes

Note: The figure summarizes the strategic decision studied in the paper. Property type and market thickness shape the seller's expected buyer pool. Sellers then choose whether to underask and attempt to generate an auction-like outcome, or to ask near or above market value and wait in a negotiation regime. Underasking can lead to a successful bidding war, but failed underasking leaves the property on the market longer and is associated with a sale-price penalty.

The key takeaway is that underpricing is valuable only when it successfully coordinates the right buyer pool. A simple empirical decomposition shows that the mid-market has the highest expected return to underpricing because both auction entry and auction payoffs are high there. Text-based listing classifications then show that underpricing works better when paired with property descriptions that make the listing attractive to buyers. Generic underpriced listings are more likely to fail, and failed auctioneers face a substantial sale-price penalty relative to successful auctioneers and patient sellers. The empirical contribution is therefore to show not only that sellers underask, but when underasking is likely to work and what sellers lose when it does not.

The paper proceeds as follows. The remainder of this section discusses related literature and the contribution of the paper. Section **??** develops the directed-search model. Section **??** brings the theory to the Toronto data by documenting the inverse-V sold-to-ask pattern, the role of days on market and market thickness, the regime payoffs decomposition, seller-type heterogeneity, and the penalty from failed underpricing. Section **??** concludes.

Related Literature and Contribution. The literature on directed search in housing and the role of asking prices is large. The review paper by **han2015microstructure** extensively summarizes the literature on asking prices and search models in housing in Sections 4 and 5. Most search models in housing consider the asking price as either a ceiling or a floor, mainly for simplification and for ease of identifying analytical solutions in equilibrium[1].

I contribute to the literature mainly in two ways. The first is to add to the literature on non-binding asking prices or posted prices: I show that non-binding asking prices with partial commitment can play an important role in directing search in multiple ways when heterogeneous buyers compete for a good. Among the potential roles of an asking price in the housing market, I present a potential explanation and mechanism for asking prices to serve as strategic instruments and non-binding commitment devices. This provides a unification of the different roles of asking prices in the literature with empirical evidence. Second, I add to the literature on directed search in the housing market by presenting a framework of heterogeneous buyers whose decisions are interdependent. Relative to the existing literature, the contribution is also empirical: I connect the seller's pricing decision to market thickness, property type, text-based listing signals, days on market, and auction-failure penalties in one framework. Below, I compare and discuss my contribution to the current literature.

The most related work on the role of asking price in real estate transactions is the paper by **han2016role**. They build a simple search framework with homogeneous buyers and sellers and a discrete number of buyer arrivals. They illustrate how their framework can incorporate sale prices above, below, or equal to the asking price. The main effect of the asking price is that it drives differences in the expected number of buyers that arrive at one seller. Sellers trade off more buyers against higher search costs. Buyers with high or low match-specific valuations arrive with different probabilities, and only buyers with high values can generate a positive surplus for sellers. Their model runs into two main problems. First, since they use a discrete number of buyers, it is hard to obtain analytical results for analysis or to calibrate the model to fit the empirics. As a result, they did not discuss potential counterfactual and policy issues related to the asking price in a structural framework. Moreover, with a simple extension to heterogeneous sellers, it can be shown that their framework is not enough to generate sorting across markets that resembles what I observe in the data: the ratio between the asking price and the sold price would be monotonic across market segments with different seller reservation values. My paper (1) provides alternative ways for the asking price to affect the market such that sellers have different incentives to

[1] See **lester2017competing**, **carrillo2012empirical**, and **arnold1999search** for asking prices as price ceilings; and **han2021effects** and **albrecht2016directed** for asking prices as price floors

price high or low in different segments, and (2) uses continuous notation to make calibration possible. I also draw on the argument that asking prices signal house quality (for example, see **wang2011**) in my framework.

The most related work for understanding how market characteristics affect sellers' choices through asking prices is **guren2018house**. In his framework, sellers can choose whether to follow the signal by posting an asking price that conforms with other similar properties, given that they face concave demand. With concave demand and all sales occurring at the asking price, the price sellers choose to post also determines how competitive their market is. His paper justifies the existence of concave demand by explaining that when sellers deviate from a prevalent price, raising the price slightly yields a large loss in the number of buyers, while lowering the price slightly yields only a small increase. I incorporate some of this idea into providing the rationale for attracting different buyers when sellers post different asking prices. One caveat of his model is that the asking price has to be a full commitment; without that, the effect of the asking price on price momentum would be lost. I present a framework in which the asking price is not a full commitment, yet sellers still have incentives to price lower or higher, i.e. not to follow the signal, in different settings.

In the directed search literature, this paper draws largely from **lang2005**, which studies heterogeneous buyers' search when sellers have preferences over buyers. Under the same payoff, sellers prefer one type of buyer over the other in their setup. It can be shown from their paper that if both buyers give sellers the same payoff, only a separating equilibrium exists. I add both the role of asking prices and interdependent buyer payoffs and queue lengths to their model. Therefore, the payoff of one type of buyer, observed by the sellers, is not strictly dominating the payoff of the other type. In this case, a pooling equilibrium could be supported. This paper also relates to segmented housing search [2] in the sense that buyers are constrained to a certain number of markets. Buyers trade off the probability of winning with the expected payoff if they win. Conditional on a positive probability of winning, buyers receive higher payoffs in markets with higher sale prices but have a lower chance of winning, and vice versa.

Alternative explanations include the atypicality of houses (for example, **jud1990atypicality**), information asymmetry that is heterogeneous across sale prices (e.g., descriptively discussed in **green2008imperfect**), different outcomes when considering stock-flow dynamics in the market (e.g., **smith2022short**), and neighbourhood norms in pricing (e.g., **pryce2011bidding**). The empirical section addresses these alternatives by showing that the inverse-V pattern is robust across geography and property types, is concentrated among fast-selling listings, is predicted by market thickness, and carries a measurable penalty when underpricing fails.

[2] See **Piazzesi2020** for neighbourhood segmentation and **WILLIAMS2018107** for income segmentation

2 A Simple Theoretical Model

This section develops the theoretical framework of the paper. I postulate the asking price to be a tool seller use to attract some pool of buyers with heterogeneous financial abilities. The key difference here from previous literature is that the asking price is not binding in any way. The key trade-off in this simple theoretical model is that seller faces some per-buyer visit cost and wants to maximize the sale price by initiating bidding war. I then use the theoretical results to guide the empirical results. I assume sellers compete in identical yet completely segmented market. Therefore, it suffices to only look at one specific market and derive some cross-market results in the calibration section.

2.1 Environment

Players. A static economy consists of an exogenously large measure of risk-neutral buyers and sellers[3]. Each seller possesses a single unit of an indivisible house, for which each buyer has a unit demand. There is no new construction in this economy.

Timing of Search. In the first stage, each seller j posts an asking price a_j publicly available to all buyers. In the second stage, each buyer, observing the asking price, pays a cost c_i to select one a_j at which the buyer wishes to attempt to match. I will follow the search literature and assume that coordination between buyers is not feasible and restrict buyers to symmetric strategies (**lester2017competing**). Multiple buyers can be matched with the same seller. After arriving at the seller with the asking price a_j, the match-specific value is realized. All market information is revealed to buyers and sellers at the bidding stage. Buyers bid after learning all the information and sellers award the house to the highest bidder or randomly at one of the highest bidders. The bidding stage is effectively an open first-price auction. I refer to all buyers and sellers choosing a particular a_j as a submarket.

Preferences. Buyers are heterogeneous in the maximum amount they can pay, which I will denote as their financial constraint from now on. For simplicity, suppose there are two types of buyers, unconstrained (U) and constrained (C). Type U buyers can pay as much as they want, and type C buyers can only pay up to a maximum amount - ω. In the full model, buyers have a distribution of $\omega_i \sim \mathcal{F}_\omega$, where $\mathcal{F}_\omega$ has support on $[0, \infty)$. All buyers have the same match-specific distribution $\mathcal{F}_U(x)$ with support in $[x_L, x_H]$, which will only be realized after arriving at the house. To simplify, suppose there are two potential realizations for

[3]Risk neutrality of sellers is assumed for simplicity of notation. Imposing risk aversion on sellers will make the analysis intractable, but the main comparative statistics and theorems will still hold

the buyers: high value x_H, and low value x_L. I will assume $x_L \leq \omega \leq x_H$ according to the definition of type U and C. The probability of a high match-specific value is p_H, which is the same for both types of buyers. I will discuss the cases where the probability differs for both buyers in equilibrium results. All sellers within a submarket have the same valuation of the house, and I will normalize it to be the lowest match-specific values of the buyers, x_L [4]. See Table **??** for a complete breakdown. At the time of bidding, all information regarding types and associated number of buyers and seller's reservation value will be perfectly known to all players in the market, but not before.

Table 1: Buyer Type and Maximum Amount to Bid

Drawn Value	Maximum Type U can Pay	Max Type C can Pay
x_H	x_H	ω
x_L	x_L	x_L

Meeting and Matching. Consider a submarket a_j with a measure of b_j^c constrained buyers, b_j^u unconstrained buyers and s_j^s sellers, then the queue length of each buyer is $\lambda_j^c \equiv b_j^c/s_j^s$ and $\lambda_j^u \equiv b_j^u/s_j^s$, respectively, in this submarket. A meeting between a buyer and a seller guarantees a match between the buyer and the seller. A buyer can be matched with at most one seller, but each seller can be matched with multiple buyers. The probability of a seller be matched with n^c constrained buyers and n^u unconstrained buyers, $n^c, n^u \in \mathbb{Z}+$, is given by $P_n^c(\lambda_j^c)$ and $P_n^u(\lambda_j^u)$. The probabilities are both assumed to be continuously differentiable. Later described in detail, these queue lengths are endogenous variables determined by the equilibrium behaviour of sellers and both types of buyers.

Sorting across Markets. Seller's reservation values and buyers' financial constraints are exogenously given under distribution $\mathcal{F}_s$ and $\mathcal{F}_\omega$, respectively. Sellers choose an a_j to join a submarket, and buyers choose one submarket to be matched and bid.

Inspection and Search Cost. Buyers and sellers both need to pay a search cost when matched. After buyers arrive at a seller, I assume that they must pay an inspection cost of k_i, $i \in c, u$ depending on the type of buyer. Following the literature (e.g. **carrillo2012empirical**), these inspection costs can include any travel costs, opportunity costs, and all non-pecuniary costs for buyers to tour this property, which is type-dependent. Each seller incurs a search cost of s for each buyer that arrives and matches with this seller. Therefore, sellers have

[4] This normalization is effectively equivalent to sellers only restrict their attention to those that could at least afford the reservation value. Altering the reservation value of the sellers will not change the results qualitatively

the incentive to limit the number of buyers that arrive. Seller search costs are an important difference with **han2016role**. The per-buyer search cost incurred by sellers ensures that sellers do not price at $1 for all houses. Therefore, the per-buyer seller search cost poses a disincentive for sellers to put a low asking price.

Bidding and Gains from Trade. At the time of bidding, each buyer knows perfectly the maximum value other buyers are willing to bid to this seller. The highest bid buyer (or one of the highest bidders) will win the house. Given that the highest bid will win the house, each buyer will bid up to the maximum value of the second highest valued buyer, or her own constraint, or her value. See Table **??** for the complete and detailed description of each case, the associated best response bid for the winner, and associated payoffs. If buyers and sellers do not trade, then their payoff is normalized to zero. The net payoff for each buyer is buyer payoff less search cost c_i, and the net profit for each seller is $\pi - (n^c + n^u)s$. If there is only one high-value buyer that arrives, bargaining can happen with buyer bargaining power θ: $\theta x_H + (1-\theta)x_L$[5].

Partial Commitment of Asking Prices. If the asking price is above the bargaining outcome governed by θ, then the final sale price will be the bargaining outcome $\theta x_H + (1-\theta)x_L$. If the asking price is below, then sellers have to commit to the asking price, and the final sale price will be the asking price. However, sellers are free to choose any asking price they want. The motivation for the partial commitment of asking prices for the seller is simple: if there is very limited interest in the house - only one buyer(she) is willing to offer to buy the house - then the seller(he) is willing to transact with her as long as it is better than the reservation value. If the asking price is too high, then this buyer, learning that she is the only person giving an offer, is not willing to pay higher than what she can bargain for. If the asking price is low, this buyer would be better off putting a leave-it-or-take-it offer to the seller exactly at the asking price. This seller is still better off transacting if the asking price is greater than his reservation value, so the final sale price will be the asking price.

[5]This bidding process is similar to an open English auction with the key difference of the partial commitment of the asking price. In a classical English auction, there is no role of an asking price when it comes to the bidding process

Table 2: Outcomes under Different Situations for Each Seller

Situations	U	C	Winner	Sale Price	Buyer Gains from Trade	Seller Gains from Trade
1	1 High	no high	U	a	$x_H - a$	$a - x_L$
2	1 High	$\geq$ 1 High	U	$max\{a, \omega\}$	$x_H - max\{a, \omega\}$	$max\{a, \omega\} - x_L$
3	$\geq$ 1 High	Anything	U	x_H	0	$x_H - x_L$
4	no high or 0	1 High	C	$min\{a, \omega\}$	$x_H - min\{a, \omega\}$	$min\{a, \omega\} - x_L$
5	no high or 0	> 1 High	C	ω	$x_H - \omega$	$\omega - x_L$
6	no high	no high	{U,C}	x_L	0	0

Note: This table presents the different situations in the model. The second and the third column displays the number and match-specific value of each type of buyers that are matched with the seller. For example, "1 High" under column "U" represents one unconstrained type buyer matched with this seller that draws a high match-specific value. The third column indicates which type of buyer will win the bid, followed by the sale price the winner pays to the seller. The last two columns present the gains from trade without considering any search costs for the buyer and the seller, respectively.

2.2 The Decentralized Equilibrium

High Valued Buyers. Since sellers' reservation values are the same as the maximum willingness to pay for a buyer who draws a low value after visiting the house, sellers only earn a positive payoff from a high-value buyer. For cleaner notation, denote λ^c and λ^u for expected queue lengths for constrained high-value buyers and unconstrained high-value buyers, respectively. I will focus only on buyers drawing a high match-specific value from now on.

Unconstrained Buyers' Strategies and Payoffs Note that if an unconstrained buyer draws a low match-specific value x_L, regardless of winning the house, the payoff for the unconstrained buyer is zero. In that case, the U type buyer's net payoff is $-k_u$. The probability of being matched with a high match-specific value is p_H, which will be the only case to receive any positive payoff. Let $f(\lambda_j^c, a_j)$ be the ex-ante expected payoff for an unconstrained buyer that draws high match-specific value in submarket j. Therefore, the ex-ante expected payoff of an unconstrained buyer in submarket j is $U_j(\lambda^u, \lambda^c, a) = p_H e^{-\lambda^u} f(\lambda_j^c, a_j) - k_u$.

Table **??** shows that under different situations, the gains from trade for buyer and seller are different. Denote $q(\lambda^i)$ as the probability of type i buyer winning with an expected queue length of λ^i for type i, $i \in L, H$. This is equivalently also the matching probability of a high-value type i buyer.

$$q(\lambda^i) \equiv e^{-\lambda^i} + \sum_{n=1}^{\infty} \frac{1}{n+1} \frac{e^{-\lambda^i} (\lambda^u)^n}{n!} \tag{1}$$

Since the payoff is different under different situations outlined in table **??**, the expected payoff of a type U buyer with high value is further broken down to:

$$q(\lambda^u) = \underbrace{e^{-\lambda^c} e^{-\lambda^u}}_{\text{situation 1}} + \underbrace{(1 - e^{-\lambda^c}) e^{-\lambda^u}}_{\text{situation 2}} + \underbrace{\sum_{n=1}^{\infty} \frac{1}{n+1} \frac{e^{-\lambda^u} (\lambda^u)^n}{n!}}_{\text{situation 3}} \tag{2}$$

Re-writing the ex-ante expected payoff U is: $U_j(\lambda^u, \lambda^c, a) = p_H e^{-\lambda^u} q(\lambda^u) U^p(a_j, \omega) - k_u$, where $U^p(a_j, \omega)$ is the realized expected payoff. After some simplification, we will arrive at:

$$U_j = \begin{cases} p_H e^{-\lambda_j^u} \left(x_H - e^{-\lambda_j^c} a_j - (1 - e^{-\lambda_j^c}) \omega \right) - k_u & a_j \leq \omega \\ p_H e^{-\lambda_j^u} (x_H - a_j) - k_u & o.w. \end{cases} \tag{3}$$

Define $U^* \equiv \max_{\text{submarket} j} U_j^H$ the maximum expected payoff an unconstrained buyer can receive ex-ante across all submarkets. No unconstrained buyer would be willing to visit a submarket that yields an expected payoff of less than U^*.

Constrained Buyers' Strategies and Payoffs The constrained buyers solve a very similar problem to the unconstrained buyer. Similarly defined as the unconstrained buyers, $g(\lambda^u, \lambda^c, a)$ represents the ex-ante expected payoff of a constrained buyer with high match-specific value. Then, $C = p_H e^{-\lambda^u} g(\lambda^u, \lambda^c, a) - k_c$. Breaking down the winning probability into situations defined in Table **??** will yield the following:

$$e^{-\lambda^u} q(\lambda^c) = e^{-\lambda^u} \left(\underbrace{e^{-\lambda^c}}_{\text{situation 4}} + \underbrace{\sum_{n=1}^{\infty} \frac{1}{n+1} \frac{e^{-\lambda^c} (\lambda^c)^n}{n!}}_{\text{situation 5}} \right) \tag{4}$$

Note here that an extra term $e^{-\lambda^u}$ comes before the probability of an additional high-value constrained buyer because if there is a high-value unconstrained buyer, the constrained buyer would never win the house. Situation 6 is left out because it yields 0 payoffs. After some simplification:

$$C_j = \begin{cases} p_H e^{-\lambda_j^u} \frac{1-e^{-\lambda_j^c}}{\lambda_j^c} \left(x_H - \frac{\lambda_j^c e^{-\lambda_j^c}}{1-e^{-\lambda_j^c}} a - \frac{1-e^{-\lambda_j^c} \lambda_j^c e^{-\lambda_j^c}}{1-e^{-\lambda_j^c}} \omega \right) - k_c & a \leq \omega \\ p_H e^{-\lambda_j^u} \frac{1-e^{-\lambda_j^c}}{\lambda_j^c} (x_H - \omega) - k_c & o.w. \end{cases} \tag{5}$$

Similar to the unconstrained buyers, constrained buyers will also only go to submarkets that yield at least the maximum expected payoff $C^* \equiv \max_{\text{submarket} j} p_H C_j^H$.

2.2.1 Seller

All sellers are homogeneous; each seller can choose an asking price a_j to enter into submarket j and receives an ex-ante expected profit π_j. The asking price is bounded implicitly by $[x_L, x_H]$ since any asking price below x_L or above x_H will be unimportant in pricing in the sense that no house will be traded at the asking price. In addition, the asking price must be at least as good as the bargaining, for which the process is mentioned in the setup. Therefore, $a_j \leq \theta x_H + (1-\theta) x_L$. If not, the asking price also becomes unimportant in pricing because no trade will be made at the asking price. The probability of a seller meeting at least one buyer in submarket j is $1 - e^{-\lambda_j^c / p_H} e^{-\lambda_j^u / p_H}$. Denote $H(a_j, \lambda_j^u, \lambda_j^c)$ as the expected sale price for sellers in submarket a_j,The expected sale price of sellers conditioning on meeting a buyer is:

$$\mathbb{E}\Big[H(a_j,\lambda_j^u,\lambda_j^c)|\text{meet}\Big] = \underbrace{p_H\lambda_j^u e^{-p_H\lambda_j^u}(e^{-p_H\lambda_j^c})a_j}_{\text{situation 1}} + \underbrace{p_H\lambda_j^u e^{-p_H\lambda_j^u}(1-e^{-p_H\lambda_j^c})max\{a_j,\omega\}}_{\text{situation 2}}$$
$$+ \underbrace{(1-e^{-p_H\lambda_j^u}-p_H\lambda_j^u e^{-p_H\lambda_j^u})x_H}_{\text{situation 3}} + \underbrace{e^{-p_H\lambda_j^u}(p_H\lambda_j^c e^{-p_H\lambda_j^c})min\{a_j,\omega\}}_{\text{situation 4}} \tag{6}$$
$$+ \underbrace{e^{-p_H\lambda_j^u}(1-e^{-p_H\lambda_j^c}-p_H\lambda_j^c e^{-p_H\lambda_j^c})\omega}_{\text{situation 5}} - (\lambda_j^c+\lambda_j^u)s_j$$

With some simplification, the sellers sale price is:

$$\mathbb{E}\Big[H_j|\text{meet}\Big] + (\lambda_j^c+\lambda_j^u)s_j = \begin{cases} e^{-\lambda_j^u-\lambda_j^c}(\lambda_j^u+\lambda_j^c)a_j + (1-e^{\lambda_j^u}-\lambda_j^u e^{\lambda_j^u})x_H & \\ \quad + e^{\lambda_j^u}\Big(1-e^{-\lambda_j^c}-\lambda_j^c e^{-\lambda_j^c}+\lambda_j^u(1-e^{-\lambda_j^c})\Big)\omega & a_j \le \omega \\ \lambda_j^u e^{-\lambda_j^u} a + (1-e^{-\lambda_j^u}-\lambda_j^u e^{-\lambda_j^u})x_H & \\ \quad + e^{-\lambda_j^u}(1-e^{-\lambda_j^c})\omega & o.w. \end{cases} \tag{7}$$

Therefore, the total expected profit for sellers in submarket j are:

$$\mathbb{E}\Big[\pi_j\Big] = \mathbb{P}(\text{meet})\,\mathbb{E}\Big[H(a_j,\lambda_j^u,\lambda_j^c)|\text{meet}\Big] + x_L \tag{8}$$

Sellers choose an asking price to maximize the expected profit in (**??**), given the best response queue lengths from the two types of buyers.

2.2.2 Within-Market Results

For this section, I will discuss the results and equilibrium within a submarket j. Therefore, I will drop all the subscript j for simpler notation. The proofs of the propositions could be found in appendix **??**.

Characteristics of the Buyers' Problems. Note that given free-entry of buyers, the optimal U^* and C^* must equal to 0. In addition, note that an unconstrained buyer always has a higher trading payoff than constrained buyers since for every situation a constrained buyer could be traded, an unconstrained buyer would receive a higher payoff was an unconstrained buyer be present. Formally,

Corollary 1. *$f(\lambda^u,\lambda^c,a) > g(\lambda^u,\lambda^c,a)$ for all parameters and choice variables. Therefore, $k_u > k_c$.*

Corollary **??** provides the same result as in **WILLIAMS2018107**, for which he assumed the relationship by stating that these unconstrained buyers have higher incomes or wealth, which corresponds to higher opportunity costs.

Comparative Statistics The choice of asking price will subsequently simultaneously determine the expected queue lengths of both buyers. The relationship between the asking price and the queue length is determined by the optimality condition of both types of buyers. The following two propositions state the comparative statistics of this relationship for constrained and unconstrained buyers.

Proposition 2. *The expected queue length of constrained buyers is decreasing if the asking price is posted below the financial constraint of the constrained buyers, and increasing otherwise. Mathematically,*

$$\frac{d\lambda^c}{da}\begin{cases} < 0 & a \leq \omega \\ > 0 & o.w. \end{cases}$$

and,

$$\frac{d^2\lambda^c}{da^2} < 0 \ \forall a \in [x_L, x_H]$$

Proposition 3. *The expected queue length of unconstrained buyers is decreasing. Mathematically,*

$$\frac{d\lambda^u}{da} < 0 \ \forall a \in [x_L, x_H]$$

The more interesting result is that once asking price is greater than the maximum a constrained buyer can bid, the queue length increases with the asking price. This is because asking price can be potentially a sale price with its partial commitment feature, and conditional on asking price being the bargaining result, the higher asking price is, the fewer unconstrained buyers will arrive because the expected payoff will decrease. On the other hand, because constrained buyer never pays the asking price - they will always pay the constraint ω. In this case, it is always a constant payoff if winning. But at the same time, a constrained buyer will never win the house if there is an unconstrained buyer with high value present and only win with probabilities because she can be one of the many buyers who is willing to bid ω. When there are less unconstrained buyer in expectation, there will also be less high-value unconstrained buyer in expectation, and thus the probability of winning the house becomes higher for a constrained buyer. This corresponds to how a buyer would bid a low offer hoping that there will be no other buyers that arrive with higher bid and the seller is still willing to sell. Appendix Figure **??** reports the numerical comparative statics for both buyer queues.

Similarly as buyers, I will abstract the j from the subscript to discuss within-market equilibrium in this section.

Proposition 4 (Some Properties of the sellers' payoff.)**.** *1. The profit function is continuous*

2. The profit function is differentiable other than at $a = \omega$

Proposition 5 (Profit function is concave when $a > \omega$)**.** *For $a \in [\omega, x_H]$, there exists a set of parameters such that the decreasing slope of λ^u is larger in magnitude than that of λ^c. In that case, sellers' profit function is concave.*

The proof consists of two parts. First I can show that, given a set of weak conditions, $\mathbb{E}\left[\pi_j | a = \omega\right] > \mathbb{E}\left[\pi_j | a = x_H\right]$. Then, as long as the decreasing rate of constraint type buyer queue lengths is smaller, there exists $\tilde{x} \in (\omega, x_H)$ such that $\mathbb{E}\left[\pi_j | a = \omega\right] < \mathbb{E}\left[\pi_j | a = \tilde{x}\right]$. This condition ensures that optimal asking price could be above ω. Similarly, there also exists a set of parameters such that the profit function is concave when $a < \omega$.

Equilibrium Definition. I will follow the equilibrium definition in **lester2017competing**.

Definition 1. *An equilibrium is a distribution $G(a, y, z)$ and market utilities C^* and U^* such that every pair in the support of G is a solution to the seller's problem, given C^* and U^*. If the solution is unique, then G is degenerate.*

Proposition 6 (Comparative Statistics of the Optimal Asking Price)**.** *Optimal asking prices weakly increases as:*

1. *financial constraint gap decreases*

2. *search costs increase*

3. *proportion of high value buyers, p_H, increases*

The intuition of the proof is to utilize Topkis' theorem and proof by cases. In particular, if an optimal asking price is strictly smaller or greater than the financial constraint, then by supermodularity of the functions, $a^*(\omega)$ is non-decreasing. If the optimal asking price exactly coincides with the financial constraint, then if I increase ω by ϵ, $\omega + \epsilon$ remains the optimal asking price. Thus the change is one-to-one. Appendix Figures **??** and **??** show the corresponding numerical comparative statics for ω and seller search costs.

The economic intuition of this result is that if buyers are more heterogeneous in terms of financial constraint (ω is small), sellers are better off lowering the asking price to attract both types of buyers to create competition. With competition, the sellers could potentially sell at a high sale price, creating a high sold-over-ask ratio. However, if buyers are more homogeneous (ω is large), sellers do not have the incentives to pay the per-buyer search cost to increase, especially, the number of unconstrained buyers since the sellers will receive a similar payoff for both types of buyers. Depending on the search cost, the asking price could be the same or higher than the sold price.

Equilibrium Results Discussion. In general, when given a small high-value probability, moderate financial constraint, and low seller search cost, the seller will be willing to price their asking price lower than the constraint ω. This result is saying that depending on the market structure, sellers sometimes are willing to price lower to attract more buyers, even if that means some buyers that have less expected profit for the seller will arrive. This is only beneficial to sellers if the sellers are facing small per matched-buyer search costs; otherwise, the cost to induce a "bidding war" is not profitable. The following theorem inherits the result from Proposition **??**.

Theorem 7 (Sale-over-Ask Ratio)**.** *1. increases with the heterogeneity of buyers*

2. decreases with search costs

3. first increases then decreases with proportions of high type buyers

The economic intuition to Theorem **??** is that the more similar the two types of buyers are, the less incentive for sellers to ask lower. Asking a lower price results in more buyers, especially more unconstrained buyers, but also at an increasing cost. If the two types of buyers are similar in their payoff to the seller, which is equivalent to low heterogeneity in financial constraints, the sellers are almost equally better off transacting with a high value constrained buyer to a high value unconstrained buyer. Similarly, when search costs are lower, with the same amount of cost, the seller can afford to attract a larger pool of buyers, which increases the probability of getting a "bidding war" scenario. Interestingly, with the increase in proportions of high type buyers, the sale-over-ask ratio first increases as the optimal asking price increases. But once beyond a certain probability threshold, the sale price reaches the maximum possible while asking price continues to increase to attracts a shorter queue length. Imagine the extreme cases where this probability of high value is 1, the seller only need to have 2 high value unconstrained buyer for achieving the maximum possible sale price.

Theorem **??** has important implications for targeted policies to buyers. When buyers are homogeneous in financial ability, i.e., no buyer is constrained, there will not be any bidding wars. However, the expected sale price is higher: the previously positive surplus from the constrained buyer winning the house will all be re-distributed to the seller. Therefore, any policy that targets at the buyers only to expand their financial capability will eventually leads to extra surplus going to the seller, inducing a higher sale price. Appendix Figures **??**, **??**, and **??** collect the numerical comparative statics for the three parts of the theorem.

2.2.3 Discussion of the Within-Market Results

Convergence to han2016role. It is apparent that if buyers are homogeneous, which is also equivalent to no financial constraints, the results converge to what **han2016role** has with the extension of Poisson arrival of buyers. In that case, the optimal asking price will be a weighted average between buyer and seller search costs.

$$a^* = \frac{k}{k+s} x_L + \frac{s}{k+s} x_H \tag{9}$$

Extension to Heterogeneous Sellers Cannot Provide Theoretical Support. If I only treat sellers as heterogeneous and all buyers to be homogeneous, the equilibrium results will never fit the empirical data with reasonable parameter choices. That would be another extension from the continuous version of **han2016role**. In that case, homogeneous buyers sort between different sellers, varied by seller reservation value. Since the optimal asking price for each seller is a weighted average, the asking price will be a monotonic transformation of the sellers' reservation value, which almost translates into a monotonic relationship of the sold-to-ask ratio. However, if sellers' search costs are very close to zero and the difference between x_H and x_L is of magnitudes of 1,000 times, the sold-over-ask ratio could be non-monotonic.

Blind v.s. Open Bidding. In reality, most bidding in the real estate market in Toronto is blind bidding without learning about competitors and competing offers. In a blind bidding setting, the upper bound of each buyer's bid is still their maximum willingness to pay. In addition, such a setting motivates the constrained buyers to participate in the market even knowing that they are less competitive because of the constraints. Specifically, given that all buyers do not know if there exist competing buyers, all buyers could behave opportunistically such that they put in a low bid in hopes to win the house with a high valuation given that everything is blinded. However, there will no longer be a benchmark winning bid that is less than the maximum willingness to pay of the buyer but will win the bid, i.e. the second highest bid, since no other bid information will be revealed. Therefore, the results in this paper are a lower bound of the effects. The calibration results below will show that even the lower bound produced by this paper could replicate the empirical trend, which implies that the main mechanism is not within the bidding setting, but the ones mentioned in this paper.

Information Friction. An alternative explanation or an additional explanation to the heterogeneity of buyers could be systematic information friction persistent in the market. In the real estate market, the reservation value of the sellers is not revealed to anyone - some may even argue that sellers themselves might not know the exact reservation value since most sellers in the real estate market are naive sellers - they are not trained merchants but regular households. If information is similarly frictional for the least-valued and most-valued houses, and differently for the middle-valued houses, then the same shape of the sold-to-ask ratio could be generated. This is not of particular concern to this paper because of two reasons. Firstly, one could also view the match-specific valuations by the buyers as an incorporation of the information frictions each buyer receives. In this case, the heterogeneity of the buyers could also include some noise that captures this friction. Then, the results presented in this paper underestimate the sold-over-ask ratio. Secondly, it is difficult to construct and compare the level of information asymmetry across sale prices. Therefore, it is difficult to provide support for any characterization of information asymmetry.

Valuation Differences v.s. Financial Constraints. Another alternative explanation could be because of buyer valuation differences instead of financially constrained buyers. However, the main difference will be the gains from trade for the buyers, which subsequently determines the entrance of the buyers. If the pool of buyers is because of valuation differences instead of financial constraints, there will be less or no buyers' gains from trade for the constrained buyers and thus decreases significantly the incentive to enter and all competition mechanisms by the asking price will be missing. Also, financially constrained buyers are empirically observed, which is shown in the left panel of Figure **??**.

2.2.4 Across Market Implication

The within-market results above yield sharp predictions about how seller's asking price strategy responds to buyer heterogeneity, search costs, and the composition of buyer types. This subsection maps those predictions to the empirical distribution of sale-over-ask ratios documented in the Toronto data by means of a calibration exercise. The key mechanism is that the degree of financial-constraint heterogeneity among buyers could vary systematically across price segments of the housing market.

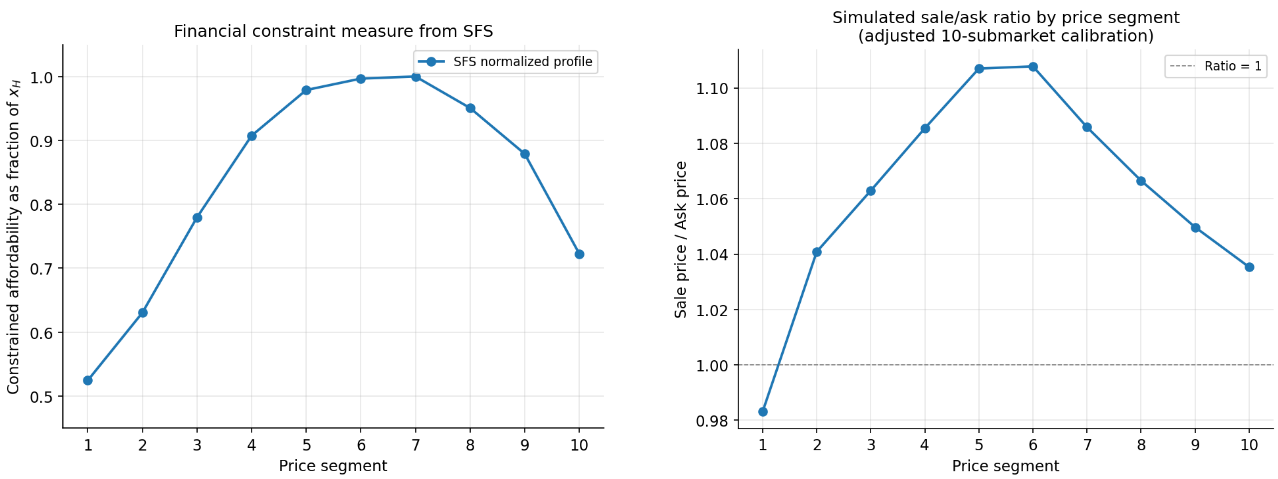
Suppose now for simplicity that there is no self-selection across different submarkets within the housing market (i.e. buyers could only choose whether to enter the submarket or not). Theorem **??** suggests that the heterogeneity of buyers along the distribution of houses could entirely drive the empirical observations. Starting from the lowest end (houses with the lowest price), the buyers are very similarly constrained. As one moves to higher valued houses, more buyers with deeper pockets are in the market, while the constrained buyers also wanted to take their chances to see if they could win the house. At the peak around \$ 1.4 million, the heterogeneity of buyers reaches the maximum. After the peak, very constrained buyers do not participate in competing because the probability of winning is effectively driven down to 0, resulting in a negative expected payoff. Moving towards the most luxurious houses, only those who are very unconstrained are still in the market, and the buyers become financially similarly unconstrained.

Datasets for Calibration. To calibrate wealth and the matches between buyers and sellers, I use the census table to obtain the after-tax income distribution and the public-use microdata file from the 2019 wave of the Survey of Financial Security to obtain the matches between buyers' income, wealth, and the price of the property they purchased. Both datasets are restricted to Toronto to correspond with the transaction data. Since the public-use microdata are extremely limited, there are only 136 observations of buyers who bought a property within a year of the survey. However, this provides a baseline for the matches between buyers and sellers. What is interesting is that there are two pricing strategies that are clearly separated by whether sellers engage in underasking behavior. I will be more precise on this point in the empirical section.

Preliminary Calibration results. Figure **??** summarizes the calibration inputs and fit. The left panel plots the financial-constraint profile from the Survey of Financial Security; the right panel reports the simulated sale-over-ask ratio across price segments. All units are in \$100,000. I suppose submarkets are independent, and the entire aggregate market is separated into 10 equally spaced price segments. The reason for not using sale-price deciles is that sale deciles are weighted by volume and therefore concentrated toward the lower end of the market. However, the full spectrum of sale prices needs to be considered, so I use a linear grid instead.

Within each submarket, the problem each seller solves is the same as in the model. Table **??** consolidates the calibrated parameter values. I set $x_L = 0.7 \cdot \text{price}$ and $x_H = 1.1 \cdot \text{price}$, so $x_H/x_L \simeq 1.57$ in each segment. Seller search cost is normalized to $s = 0.005$. Buyer search costs increase by segment, with $k_u/k_c = 1.3$ throughout. The financial constraint is expressed as ω/x_H. In the adjusted calibration, ω/x_H is lowest in the middle of the market, so the financial-constraint gap $x_H - \omega$ is largest where the model is intended to generate the highest sale-over-ask ratio. The SFS constraint profile in the left panel of Figure **??** motivates the use of a nonconstant affordability profile, while the right panel shows that the calibrated model reproduces the inverse-U pattern in the sale-over-ask ratio. Appendix **??** reports the additional calibration diagnostics and comparative statics.

Figure 2: Model Calibration: Financial Constraints and Simulated Sale-over-Ask Ratio



(a) Financial constraint measure (SFS)

(b) Simulated sale-over-ask ratio by price segment

Note: The left panel plots financial constraints from the public-use microdata file of the Survey of Financial Security. The y-axis shows the fraction of an unconstrained buyer's budget that a constrained buyer could afford in each market decile; construction details appear in Appendix **??**. The right panel plots the simulated sale-over-ask ratio from the adjusted calibration in Table **??**. The sale space is separated into 10 equally spaced submarkets. Assuming independence across submarkets, the model produces the inverse-U pattern when the financial-constraint gap is largest in the middle of the price distribution. Additional calibration diagnostics appear in Appendix **??**.

Model Implications. The model has three implications for the empirical analysis that follows. First, a non-binding asking price can matter because it changes buyer entry. A lower asking price raises the expected payoff from visiting and can attract a larger buyer pool, but sellers pay a cost for every arriving buyer. Underasking is therefore profitable only when the additional competition is valuable enough to compensate the seller for this search cost. Second, asking prices affect buyer composition, not only buyer volume. Lower asking prices can draw in more financially constrained buyers, while higher asking prices screen toward buyers with deeper pockets. The relevant object is therefore not simply how many buyers arrive, but whether the marginal buyers can generate competition for the property. Third, policies that relax buyer financial constraints need not make bidding wars more common. When buyers become more similar in financial ability, sellers have less reason to use low asking prices to attract heterogeneous bidders; the probability of sale above ask can fall even as the extra surplus is capitalized into higher sale prices. These implications motivate the empirical focus on market thickness, property type, seller advertising, days on market, and failed underpricing.

3 Empirical

3.1 Data

Data for Main Analysis. The main dataset used is the listed characteristics of all sold properties from Multiple Listing Services (MLS) in the city of Toronto from January 1st, 2021, to January 1st, 2022. The dataset contains all characteristics the seller reported in the actual online listings and publicly available characteristics of the property, including public schools and property tax. The dataset also includes both the asking price and the sold price, along with the accurate address. This rich data allows me to explore in detail the characteristics of sellers and their property listings.

See Table **??** for the summary statistics of the data I used for the main analysis. I took out outliers and parking spots. This is comparable to the published summary statistics by Toronto Regional Real Estate Board on their annual reports.

Table 3: Summary Statistics

Variable	Mean	Std. Dev.	Min	Max	Median
Sold Price	$ 1,065,682.8	$ 784,861.8	$ 210,000.0	$ 23,850,000.0	$ 840,000.0
Asking Price	$ 1,013,397.3	$ 797,788.8	$ 218,900.0	$ 29,800,000.0	$ 799,000.0
Bedroom	2.83	1.46	0	15	3
Bathroom	2.15	1.18	1	15	2
House Size (sqft)*	1,091.0	713.3	9	15,256	849.5
Days On Market	15.94	21.88	0	519	8
N			43,397		

Note: This table eliminates parking spots and extremely expensive houses. *House Size is reported in square feet and is available for 31,455 observations (72.5% of the sample).

Much recent literature on housing focuses on the Bay area, for example, **Piazzesi2020**. The major difference between my dataset and theirs is the inclusion of asking prices, detailed text information of listings, and full addresses. Without the asking price, it is impossible to obtain empirical support for the role of the asking price in real estate transactions. Relative to earlier literature, this dataset contains the universe of all sold prices in the city in 2021. Benchmark datasets lack the listing-side information necessary to study how non-binding prices direct search or inform seller types and their strategies; I establish the first empirical testing ground for these models by observing the complete universe of Toronto sales alongside their original posted prices. See Appendix **??** for limitations of this dataset.

Neighbourhood Dataset. An online housing platform publishes monthly neighbourhood-monthly level aggregate data that I will use to approximate the buyer-side information and inventory. The neighbourhood is very finely defined, which is at the size between the census tract and census tract and census subdivision. This neighbourhood notation is defined the same by the MLS, intended to group comparable houses together.

The competition measure is defined by the number of sales within a given month divided by the number of listings at the start of the month. If this ratio is greater than 100, that means that there are many houses that are sold after being listed within the same month. The investor measure is defined by the percentage of sales that become for lease within six months after being sold. Although not perfect, this measure approximates the ratio of investors in the market. In general, this will be a lower bound of the share of investors in the market.

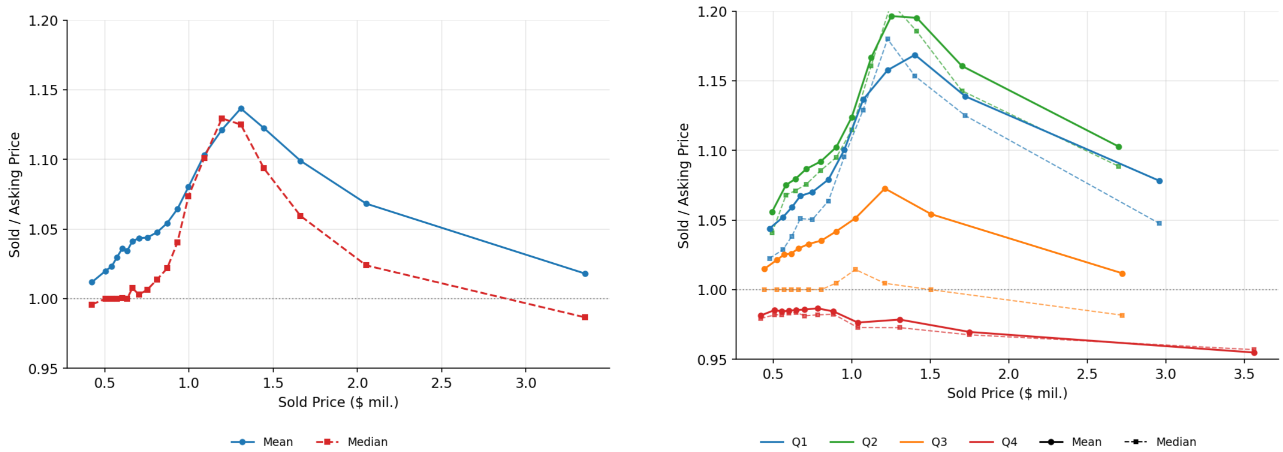
See Appendix **??** for detailed variable construction, including the derivation of the financial constraint measure from the Survey of Financial Security.

3.2 Stylized Facts

The paper is motivated by a stark empirical pattern between the relationship of sold-over-ask and sale price of properties transacted in the City of Toronto. The key implication of the model is that the asking price is strategically chosen by the seller in response to the expected market thickness and the ability to pay from the buyers. Therefore, the empirical analysis begins with three stylized facts from the data. Together, they document where the sold-to-ask ratio appears when it breaks down and which market conditions make sold above ask more likely. I then turn to a empirical data-driven decomposition that explains the hump shape as a mixture of two types of pricing strategies, or regimes, before using listing texts to examine heterogeneity in which sellers can successfully execute an underpricing strategy. The latter part provides evidence that sellers tradeoff time on market and expected sale price. Let me begin by describing the central fact.

Fact 1: The Sold-to-Ask Ratio Follows an Inverse-V Pattern. Figure **??** displays the aggregate relationship between the sale price and the sale-over-asking ratio in the entire City of Toronto. The left panel plots the sale-over-asking ratio across sold price bins, showing a clear inverse-V shape. This pattern implies that the asking price choice is carefully chosen

Figure 3: Underasking Behaviour of Sellers



(a) Sold-over-ask vs. sold price (aggregate)

(b) Sold-over-ask vs. sold price, by DOM quartile

Note: The left panel plots the mean and median sold-to-asking ratio across sold price bins for all 2021 City of Toronto transactions; a value above 1.0 indicates the house sold above the asking price. The right panel overlays the same sold-to-asking ratio for each of the four days on market quartiles: Q1 (≤ 5 days), Q2 (6–8 days), Q3 (9–18 days), Q4 (> 18 days). The reference line at 1.0 marks sold = ask. Results hold using log prices; see Appendix Figure **??**.

by sellers in different submarkets. The choice of asking prices is very similar to the sale price at the two ends and much lower than the sale price in the middle. The effect seems to be stronger for medium price than mean.

This non-linearity in sold-over-ask is robust across different neighbourhoods (Figure **??**) and property types (Figure **??**). This suggests that the hump is not a compositional artifact; it is a universal pattern across the city at different price points. The main takeaway from the model is that under-asking strategy depends critically on the seller's ability to attract buyers and on which buyers are available in that price range. This interpretation is further supported by looking more closely at property types: the hump is most pronounced for detached houses and less for appartments. This pattern is consistent with sellers pricing more strategically in competitive, heterogeneous submarkets, where detached houses are harder to substitute and inducing an auction more profitable. The first dimension that breaks the aggregate hump is days on market (DOM), which I introduce as Fact 2.

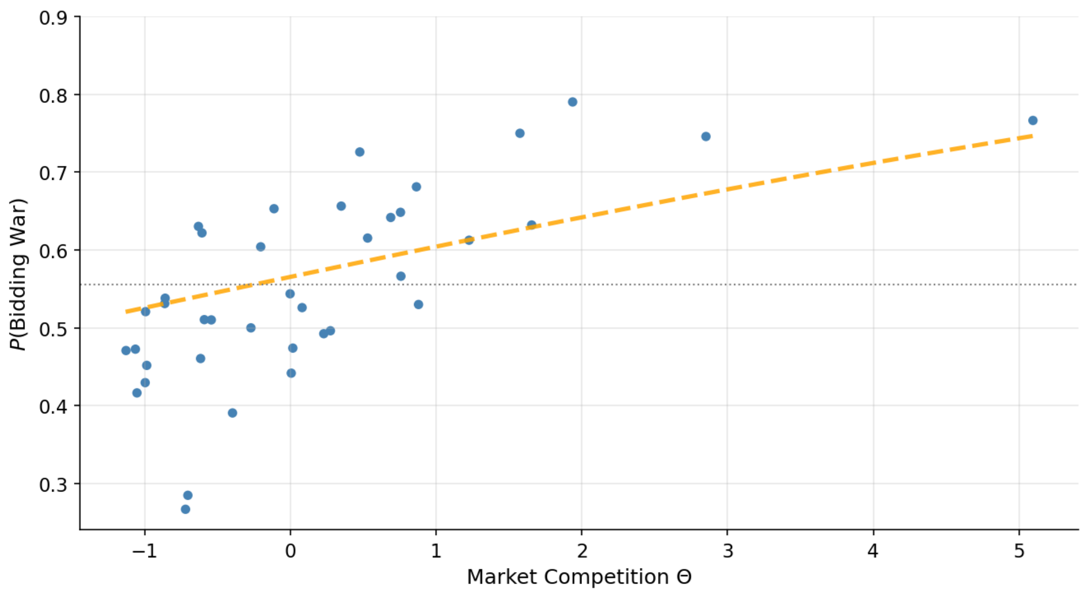
Fact 2: Sold Price over Asked Price Depends on Days on Market. The inverse-V shape weakens sharply with longer days on market. The right panel of Figure **??** overlays the sold-over-ask hump for each days on market (DOM) quartile. The hump is steeper for the faster-selling quartiles 1 and 2, and disappears almost entirely for the slowest quartile, quartile 4, with DOM up to 519 days, where the sold-over-ask ratio falls below 1. This implies that strategic underpricing only succeeds within those listings that sold fast. When a house sits on the market for an extended period, the role of the asking price as a tool to induce auction or bidding war disappears. Of course, from this figure alone, it is not possible to determine whether the days on market is a key object of the strategic choice of asking price or a consequence of it. Sellers could be either intentionally selecting into different pricing strategies, or everyone executes the same strategy, but only a few succeed. Or there could be a mixture of both. I will show that sellers intentionally choose one of these two strategies, and that expected profit balances with days on market (i.e., A higher expected sell price comes with longer time on market on average.). However, there must be key characteristics from the market that lead sellers to under-asking. And one of the biggest indicators is the market thickness.Fact 2: Sold Price over Asked Price Depends on Days on Market. What is interesting is that the significant inverse V-shape disappears with longer days on market. The right panel of Figure **??** overlays the sold-over-ask hump for each days on market (DOM) quartile. The hump is steeper for the faster-selling quartiles 1 and 2, and it disappears almost entirely for the slowest quartile, quartile 4, with DOM up to 519 days, where the sold-over-ask ratio falls below 1.

This implies that strategic underpricing succeeds only among listings that sell quickly. When a house sits on the market for an extended period, the asking price no longer appears to function as a tool for inducing an auction or bidding war. Of course, this figure alone cannot determine whether days on market is an object of the seller's pricing strategy or a consequence of it. Sellers may intentionally select into different pricing strategies, or everyone executes the same strategy, but only a few succeed. Or there could be a mixture of both.

will show below that sellers sort across these strategic regimes, and that expected sale price must be evaluated jointly with time on market (i.e., a higher expected sell price comes with longer time on market on average). Before decomposing those regimes directly, the next raw-data fact asks which market condition predicts whether a listing enters the auction regime at all. The central predictor is market thickness.

Fact 3: Auction Probability is Driven by Market Thickness. Unfortunately, it is not possible to directly observe the number of prospective buyers or the number of bidders for each property. However, disaggregate neighbourhood-month statistics on the number of sales in a neighbourhood by the end of the month divided by the number of active listings at the beginning of the month, known as the absorption rate, are a commonly used measure of market thickness in the real estate industry and serve as a good proxy for the number of buyers in the market[6]. For easier reference, I denote scenarios with a sold-over-ask ratio above 1 as the auction regime. I standardized the absorption rate and then plot the raw correlation against the probability of being in the auction regime. The results are plotted in Figure **??**. I then overlay the predicted probabilities from a logistic regression on the two variables, controlling for a full set of property characteristics, market conditions, and fixed effects. These two exercises indicate that there is a robust positive relationship between the level of market competition and the probability of entering into the auction regime.

Figure 4: Probability of Auction vs. Market Thickness



Note: Each point is a binned mean of observed auction-regime entry rates, or average shares of listings that sold above ask, against $\Theta_{m,n}$, the z-scored neighbourhood-month absorption rate. The dotted horizontal line marks the unconditional mean Pr(Bidding War) = 0.556. Predicted probabilities are from the logit Auction Regime$_i = \alpha + \beta_1 \Theta_{m,n} + \mathbf{X}'_{it}\gamma + \epsilon_i$, where $\mathbf{X}_{it}$ is the full control set that includes log asking price and its square, structural characteristics, amenities, maintenance fees, market-temperature, and property-type, season, and zone fixed effects. The coefficient of interest, $\hat{\beta}_1 = 0.158$ (bootstrap $z = 12.07$) confirms that market thickness is a robust, highly significant predictor of auction-mode entry. The raw slope from an OLS fit to the binned means is 0.069. Robustness checks, including a quadratic specification, are available in Appendix Figure **??**, and a seasonal correlation check between absorption rate and mean and median sold-to-ask ratio is available in Appendix Figure **??**.

3.3 Mechanism and Regime Decomposition

The three stylized facts establish the data features that any mechanism must explain: the sold-to-ask ratio rises and then falls with price, the pattern is concentrated among fast-

[6]Note that because of the timing mismatch in the construction of this index, it is common for the index to be well above 1 when there is substantial market activity during the month, which was often the case in the 2021 Toronto real estate market.

selling listings, and auction entry is more likely in thicker markets. I now move from documenting these facts to decomposing the hump into the two pricing regimes implied by empirical observations. Extremely luxurious properties are inherently different and therefore less likely to be sold above ask or enter an auction regime[7]. This means that market thickness and property type help determine whether sellers can enter the auction regime, but it remains to be shown that the mid-market is where underpricing has the highest expected payoff.

To confirm this, I estimate the expected gain from underpricing at each price point, measured by the decomposition-based expected sold-to-ask ratio, and test whether it peaks in the mid-market range. Since the previous analyses indicate that the market is segmented into distinct subpopulations, I use a Gaussian mixture model (GM model) to estimate the auction and negotiation regimes and their component means across the price distribution.

Decomposing the Inverse V-Shape: Auction vs. Negotiation Regimes. For each sold-price bin, I fit a two-component[8] Gaussian mixture model to the distribution of sold-to-ask ratios together with the estimated auction probability from the analysis in Figure **??**. Equation (**??**) formally introduces the expected sold-to-ask ratio for a property i in a given sold-price bin, where $\hat{\pi}_i$ is the fitted auction probability and $\hat{\mu}_{\text{Auction}}$, $\hat{\mu}_{\text{Negotiation}}$ are the upper and lower GM model component means for each price bin, respectively. Note the replication of the hump pattern in the expected sold-to-ask ratio is not mechanical. The auction probability is estimated from a logistic regression at the global level, controlling for a full set of property characteristics, market conditions, and fixed effects while Figure **??** plots the expected sold-to-ask ratio at the price-bin level. Thus, the closeness of the resemblance between the expected sold-to-ask ratio and the observed hump in Figure **??** indicates that the observed hump is indeed driven by an increasingly popular choice of underpricing in the mid-market range, as shown more specifically by the decomposition of the expected sold-to-ask ratio into its three components in the right panel of Figure **??**.

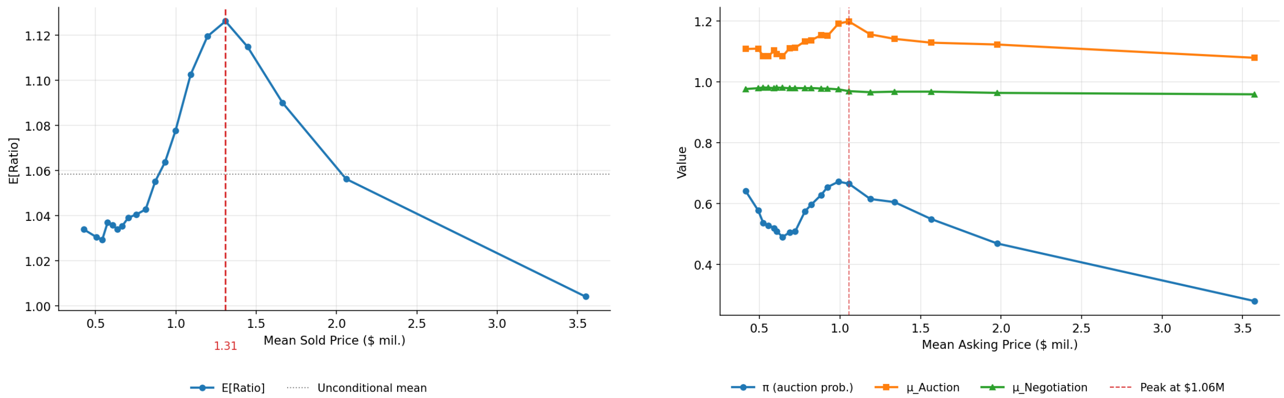
$$E[\text{Sold/Ask}]_i = \hat{\pi}_i \cdot \hat{\mu}_{\text{Auction}} + (1 - \hat{\pi}_i) \cdot \hat{\mu}_{\text{Negotiation}} \tag{10}$$

The regime decomposition shows that the hump reflects variation in the probability of entering the auction regime and in the payoff from doing so. To decompose even further behind what would contribute to this strategic underpricing, I break down the listing advertisement by using Natural Language Processing (NLP) techniques to capture listing quality

[7] For robustness, Appendix Figure **??** shows that the luxury tail (> $2.5M) has substantially higher hedonic residual variance than the rest of the market, confirming that it is structurally distinct.

[8] The actual empirical distributions within each sold-price bin are mostly bi-model, not affected by observable characteristics, like property types, days on market, geography. Theses are confirmed in Figures in Appendix **??**

Figure 5: Expected Sold-to-Ask Ratio and Its Components



(a) E[Sold/Ask] by sold-price bin

(b) Decomposition of E[Sold/Ask]

Note: E[Sold/Ask] from Equation (**??**) plotted across 20 equal-frequency sold-price bins. Expected returns to underpricing peak in the mid-market range, coinciding with the hump in Figure **??** with a maximum at 1.31 million sold price. The result is robust to using asking price as the binning variable; see Appendix Figure **??**.

and seller type. In particular, the key questions I aim to answer deeper is what are the listings that are more likely to succesfully execute the auction regime, and what are the tradeoff sellers face when they attempt to underprice.

This heterogeneity check is motivated by the same market-thickness mechanism. Underpricing works only if it brings enough high-willingness-to-pay buyers to the listing at roughly the same time. A more specific, high-quality advertisement can help coordinate that buyer pool by making the property worth visiting and by signaling that the listing is not merely cheap, but attractive. In other words, multiple buyers need to arrive roughly at the same time and have high willingness to pay for the property in order for the auction regime to work. Therefore, a more specific, high quality advertisement is needed to signal the quality of the property and attract the right pool of buyers.

Underpricing Only Works if Paired with High-Quality Advertising. High-quality advertising is categorized by the seller types that I have classified, detail in Appendix **??**. I estimate the following regression in Equation (**??**) to test the added value of underpricing for different seller types, relative to a baseline of generic listings without clear distinguishing features.

$$\frac{\text{Sold}}{\text{Ask}}_i = \sum_k \beta_k \mathbf{1}[\text{Type}_k] + \sum_k \delta_k \Big(\text{Underpricing}_i \times \mathbf{1}[\text{Type}_k]\Big) + \mathbf{X}_i'\gamma + \epsilon_i \qquad (11)$$

where $\text{Underpricing}_i = \hat{\epsilon}_{ask,i}$ is the asking-price residual from running an asking-price hedonic regression, and negative values indicate that the listing was priced below the market norm for comparable properties. Type_k is the seller-type indicator. Table **??** reports the estimated coefficients by seller type. Note that since Underpricing is negative for underpriced listings, a more negative $\hat{\delta}_k$ implies a larger positive return to underpricing for that type, relative to the generic listing baseline.

Table 4: **Underpricing Complementarity and Outcome Statistics by Seller Type**

	full sample		$1.1M–$1.5M subsample		
Seller type	$\hat{\delta}_k$	SE	Mean sold-to-ask ratio	P(sold<ask)	N
GenericListing	(baseline)		1.098	0.344	1,054
AttractiveAds	-0.0316^{***}	(0.0048)	1.130	0.263	2,310
LuxuryHome	-0.0291^{***}	(0.0052)	1.127	0.271	1,697
AttractingInvestors	-0.0188^{***}	(0.0055)	1.125	0.274	1,330
OwnHome	-0.0435^{***}	(0.0087)	1.128	0.258	620
FlipperRenovator	-0.0202	(0.0146)	1.138	0.253	170

Note: Full Sample: OLS estimates of Equation (**??**) on the full sample ($N = 43{,}273$). $\hat{\delta}_k$ is the interaction between seller type k and underpricing depth relative to GenericListing; a more negative $\hat{\delta}_k$ implies a larger positive return to underpricing (underpriced listings have $\text{Underpricing}_i < 0$). Standard errors are HC3 heteroskedasticity-consistent (**MACKINNON1985305**). Seller-type direct effects ($\hat{\beta}_k$) are in Appendix Table **??**. *$1.1M–$1.5M subsample*: sample restricted to the $1.1M–$1.5M sold-price range. Mean sold-to-ask ratio is the mean sold-to-ask ratio; P(sold<ask) is the share of listings that sold below the asking price (auction failure). $^{***}p < 0.001$; $^{**}p < 0.01$.

The finding suggests that the return to underpricing is highly heterogeneous across seller types, but all non-generic types have a significantly higher return to underpricing. This confirms the importance of having enough and the right buyers in order for underpricing to work. I next ask what are the consequences of failed underpricing. Specifically, I take a closer look at the $1.1M–$1.5M sold-price range, where the hump is most pronounced and the return to underpricing is highest. Table **??** reports the mean sold-to-ask ratio and the share of listings that sold below ask for each seller type within this range. Generic sellers fail at more than seven percentage points higher than all other types[9].

The next important question to answer, is thus, what are the consequences of failed underpricing? For the empirical equilibrium outcome to hold, i.e. some sellers underprice and some do not, there must exist a significant tradeoff. Otherwise, the dominant strategy would be to underprice for all sellers, and the signaling value of the asking price would disappear altogether since all sellers behave the same way.

In order to understand the tradeoff, I classify sellers into three groups based on their deviation of asking price relative to other asking prices of similar propertiesand their days

[9] See Appendix Figure **??** for the full distribution of the sale-to-ask ratios within this range across different seller types. It confirms that the distribution for generic listings has higher density at below 1.0 than any other type

on market. Specifically, I divide sellers into: (1) Failed Auctioneers, who underprice but do not attract the auction crowd (having a long DOM and a negative asking price residual); (2) Successful Auctioneers, who underprice and successfully trigger a bidding war (having a short DOM and a negative asking price residual); and (3) Patient Sellers, who deliberately price above market and wait for a private-value buyer (having a long DOM and a positive asking price residual)[10]. Table **??** shows that higher sale price takes more time to sell[11]. The penalty for failed auctioneers is approximately 10 log points below successful auctioneers, who themselves sell below the hedonic benchmar 81.5% of the time. Therefore, sellers that had a low asking price and stayed on market longer face a significant penalty.This inherent tradeoff between time on market and sale price is thus the key determinant of the choice of the sellers to enter into the auction regime or not[12].

Table 5: Mean Sale-Price Residual $\hat{\epsilon}_{sale}$ by Seller Group

Seller Group	N	Mean $\hat{\epsilon}_{sale}$	SE	% Below Hedonic	p-value vs. Failed
Failed Auctioneers	4,677	−0.230	0.002	98.7%	—
Successful Auctioneers	12,496	−0.132	0.001	81.5%	$< 10^{-200}$
Patient Sellers	5,640	+0.159	0.003	18.4%	$< 10^{-200}$

Note: Welch two-sample t-tests. Failed vs. Successful: $t = -35.77$; Failed vs. Patient: $t = -102.37$. All differences significant at $p < 10^{-200}$. Residual distributions and group means are shown in Appendix Figure **??**.

Summary. Taken together, the evidence paints a coherent picture of sellers strategically setting asking-price differently to suit their needs in Toronto's 2021 residential real estate market. The stylized facts show that the aggregate sold-to-ask ratio traces an inverse-V hump across the price distribution, that this hump is concentrated among fast-selling listings, and that auction entry rises with market thickness. The mechanism section then decomposes the hump into auction and negotiation regimes, showing that the expected return to underpricing is highest in the mid-market. Finally, the heterogeneity check shows that underpricing is most effective when paired with listing signals that can attract the right pool of buyers. Market thickness is therefore the mechanism, while days on market is the key trade-off separating the regimes: fast-selling listings use low asking prices to attract attention and liquidate quickly, while slow-selling listings either wait for private-value buyers

[10]See Figure **??** for distribution of listing types within each type of sellers

[11]See Appendix Figure **??** for the residual distributions and group means. The density distribution of patient sellers is much more skewed to the right than the other two groups, followed by successful auctioneers

[12]A natural concern is that these failed auctioneers are inferior by selection. I tested that by comparing the mentioning of explicit high-quality features, like "renovated," "granite," "stainless," "hardwood,", and implicit high-quality signals, like "hold offers," "review offers," "virtual tour", between the failed and successful autioneers and found no significant differences. See Appendix Figure **??** and Table **??** for details.

or suffer a penalty when underpricing fails to induce a bidding war.

4 Conclusion

This paper studies how sellers use non-binding asking prices as a strategic instrument in the housing market. The central claim is that the asking price is not simply a noisy signal of value or a starting point for negotiation. Even when it is not fully binding, it helps direct buyer search, shape the composition of prospective bidders, and determine whether a transaction takes place through an auction-like outcome or through a slower negotiation process.

The empirical analysis documents this mechanism in the Toronto residential real estate market. The sold-to-ask ratio follows a pronounced inverse-V pattern over the price distribution: the largest sale-over-ask outcomes occur in the middle of the market, not at the lowest or highest price points. This pattern is concentrated among fast-selling listings and is stronger in thicker neighbourhood-month markets, where the probability of selling above ask is higher. Decomposing the aggregate relationship into auction and negotiation regimes shows that the hump is generated by both the probability of entering an auction-like outcome and the payoff conditional on doing so. The expected return to underpricing is highest in the mid-market, where market thickness and buyer heterogeneity make a low asking price especially effective at coordinating competition.

The text-based evidence further clarifies why this strategy succeeds for some sellers and fails for others. Underpricing is most effective when paired with listing descriptions that make the property legible and attractive to the right group of buyers. Generic underpriced listings, by contrast, are more exposed to auction failure. The comparison among successful auctioneers, failed auctioneers, and patient sellers shows that the choice of asking price is closely tied to the time dimension of search. A low asking price can produce a fast sale and a high sold-to-ask ratio, but if the listing fails to attract a competitive buyer pool, the seller faces a substantial penalty. A high asking price may instead reflect a willingness to wait for a buyer with a sufficiently high private valuation. Days on market is therefore not only an outcome variable; it is part of the strategic trade-off sellers face.

The theoretical model isolates the static mechanism behind these facts. In a directed search environment with heterogeneous buyers and partial commitment through the asking price, sellers choose asking prices by balancing the benefit of attracting more buyers against the cost of managing a larger pool of visitors. Lower asking prices can increase competition and raise the probability of a bidding war, but only when the buyer pool contains enough willingness-to-pay heterogeneity and when seller search costs are not too high. The model therefore rationalizes why sale-over-ask ratios are highest in the middle of the market and why policies that relax buyer financial constraints need not reduce seller market power: if additional borrowing capacity makes buyers more similar, bidding wars may become less frequent while the extra surplus is still captured by sellers through higher sale prices.

The empirical evidence also points to a potential natural extension for the theory. The current model establishes why a non-binding asking price can direct search and change the expected buyer pool. A natural extension is to allow sellers to choose explicitly between pricing regimes over time. In such a model, the seller's decision would not be summarized by the level of the asking price alone. It would also include whether to pursue a fast auction strategy or a slower negotiation strategy, with the trade-off governed by time on market, holding costs, and the risk that an attempted auction fails to attract enough bidders.

This paper therefore provides a baseline theory and a set of empirical facts for a broader theory of seller regime choice. Market thickness helps determine whether auction entry is feasible, listing quality affects whether underpricing can successfully coordinate buyers, and days on market reveals the cost of choosing a strategy that fails. A dynamic extension would make these objects jointly endogenous, but the main message already emerges from the evidence and the model: asking prices matter because they organize search, and sellers use them to trade off attention, competition, time, and failure risk.

Appendices

A Proof

Proof to Proposition **??**. This proof is for constrained buyer's comparative statistics.

The first part of this proof is for when asking price is posted below the financial constraint, i.e. $a \le \omega$.

By definition above if $a \le \omega$,

$$f(\lambda^c, a) \equiv \left(x_H - e^{-\lambda^c} a - (1 - e^{-\lambda^c})\omega\right)$$
$$g(\lambda^c, a) \equiv \frac{1 - e^{-\lambda^c}}{\lambda^c}\left(x_H - \frac{\lambda^c e^{-\lambda^c}}{1 - e^{-\lambda^c}} a - \frac{1 - e^{-\lambda^c}\lambda^c e^{-\lambda^c}}{1 - e^{-\lambda^c}}\omega\right)$$

Then, because of free-entry, $U^* = C^* = 0$. Note also from Proposition **??** that $f > g$.

$$p_H e^{-\lambda^u} f(\lambda^c, a) = k_u \tag{12}$$
$$p_H e^{-\lambda^u} g(\lambda^c, a) = k_c \tag{13}$$

Take the ratio (**??**) / (**??**) will yield:

$$\frac{f(\lambda_j^c, a_j)}{g(\lambda_j^c, a_j)} = \frac{k_u}{k_c}$$

Take the partial derivative of $\frac{f}{g}$ w.r.t a and λ^c yield:

$$\frac{\partial \frac{f}{g}}{\partial a} = \frac{e^{-\lambda^c}(f - g)}{g^2} > 0 \text{ iff } f > g \tag{14}$$
$$\frac{\partial \frac{f}{g}}{\partial \lambda^c} = \frac{e^{-\lambda^c}(a - \omega)(f - g) + \frac{1 - e^{-\lambda^c} - \lambda^c e^{-\lambda^c}}{(\lambda^c)^2}(x_H - \omega) f}{g^2} > 0 \tag{15}$$

Note that (**??**) is positive independent of the sign of $(f - g)$, as long as f and g are both positive.

By total differentiation:

$$\frac{d\lambda^c}{da} = -\frac{\frac{\partial f/g}{\partial a}}{\frac{\partial f/g}{\partial \lambda^c}} < 0 \;\; \forall \; a \in [x_L, \omega]$$

Similarly for $a > \omega$, where

$$f(\lambda^c, a) = f(a) \equiv x_H - a$$
$$g(\lambda^c, a) = g(\lambda_j^c) \equiv \frac{1 - e^{-\lambda^c}}{\lambda^c}(x_H - \omega)$$

In this case,

$$\frac{\partial \frac{f}{g}}{\partial a} = \frac{-g}{g^2} < 0 \tag{16}$$
$$\frac{\partial \frac{f}{g}}{\partial \lambda^c} = \frac{-\frac{dg}{d\lambda^c} f}{g^2} > 0 \tag{17}$$

where

$$\frac{dg}{d\lambda^c} = -\frac{1 - e^{-\lambda^c} - \lambda^c e^{-\lambda^c}}{(\lambda^c)^2} < 0$$

By total differentiation:

$$\frac{d\lambda^c}{da} = -\frac{\frac{\partial f/g}{\partial a}}{\frac{\partial f/g}{\partial \lambda^c}} > 0 \;\; \forall\, a \in [\omega, x_H]$$

This completes the proof. □

Proof to Proposition **??**. This proof is for unconstrained buyers comparative statistics on queue lengths. Denote partial derivatives as a subscript, e.g. $\lambda_a^u \equiv \frac{\partial \lambda}{\partial a}$.

Total differentiate U and C w.r.t. a yields

$$\frac{dU}{da} = \frac{\partial U}{\partial \lambda^c}\frac{d\lambda^c}{da} + \frac{\partial U}{\partial \lambda^u}\frac{d\lambda^u}{da} + \frac{\partial U}{\partial a} \tag{18}$$
$$\frac{dC}{da} = \frac{\partial C}{\partial \lambda^c}\frac{d\lambda^c}{da} + \frac{\partial C}{\partial \lambda^u}\frac{d\lambda^u}{da} + \frac{\partial C}{\partial a} \tag{19}$$

At optimality, (**??**) is equal to $U^* = 0$ and (**??**) is equal to $C^* = 0$. Therefore, we can obtain two equations to simultaneously determine the two unknowns ($\frac{d\lambda^c}{da}$, $\frac{d\lambda^u}{da}$). In particular,

$$\frac{d\lambda^c}{da} = \frac{C_{\lambda^u} U_a - C_a U_{\lambda^u}}{C_{\lambda^c} U_{\lambda^u} - C_{\lambda^u} U_{\lambda^c}} \tag{20}$$
$$\frac{d\lambda^u}{da} = -\frac{U_a}{U_{\lambda^u}} - \frac{U_{\lambda^c}}{U_{\lambda^u}}\frac{d\lambda^c}{da} \tag{21}$$

For $a > \omega$, it is a direct result of partial derivatives and identities of probabilities that: $U_{\lambda^c} = 0$, $U_{\lambda^u} < 0$, $U_a < 0$. Therefore, (**??**) simplies to $\frac{d\lambda^u}{da} = -\frac{U_a}{U_{\lambda^u}}$.

$$\frac{d\lambda^u}{da} = -\frac{U_a}{U_{\lambda^u}} = -\frac{1}{p_H(x_H - a)} < 0$$

For $a \leq \omega$, the direct results of partial derivatives yield:

$$\begin{aligned} U_a &= -p_H e^{-\lambda^u - \lambda^c} \\ U_{\lambda^u} &= -(U + k_u) \\ U_{\lambda^c} &= -p_H e^{-\lambda^u - \lambda^c}(\omega - a) \end{aligned}$$

Substituting into (**??**) we get:

$$\frac{d\lambda^u}{da} = -\frac{e^{-\lambda^u - \lambda^c}}{U + k_u}\left(1 - (\omega - a)\frac{d\lambda^c}{da}\right) < 0 \tag{22}$$

This completes the proof.

□

Proof to Proposition **??**. Substitute a with ω for both the functions, and they both equal to $e^{-\lambda_j^u}(\lambda^u + 1 - e^{-\lambda^c})(a - x_L) + (1 - e^{-\lambda^u} - \lambda^u e^{-\lambda^u})(x_H - x_L)$. Thus, the function is continuous. Since every component of the functions on each piece is continuous and differentiable, the two functions are both differentiable in their domain. □

B Model Calibration and Comparative Statics

This appendix records the numerical checks behind the model calibration in Section **??**. The figures use the adjusted calibration in Table **??**. The single-market comparative statics use the middle price segment of the adjusted calibration, where the financial-constraint gap is largest and the model generates the peak in the sale-over-ask ratio.

B.1 Calibration Diagnostics

This subsection documents how the ten-segment calibration is parameterized and whether the calibrated primitives generate equilibrium objects consistent with the within-market comparative statics proved in Section 2. The exercise is diagnostic rather than a separate estimation step: I choose parameter values so that (i) the simulated sale-over-ask ratio traces the empirical inverse-U across price segments, and (ii) the implied buyer queues and seller choices satisfy the qualitative predictions of Propositions **??**–**??** at interior parameter values.

Table B1: Adjusted Calibration Parameters by Price Segment

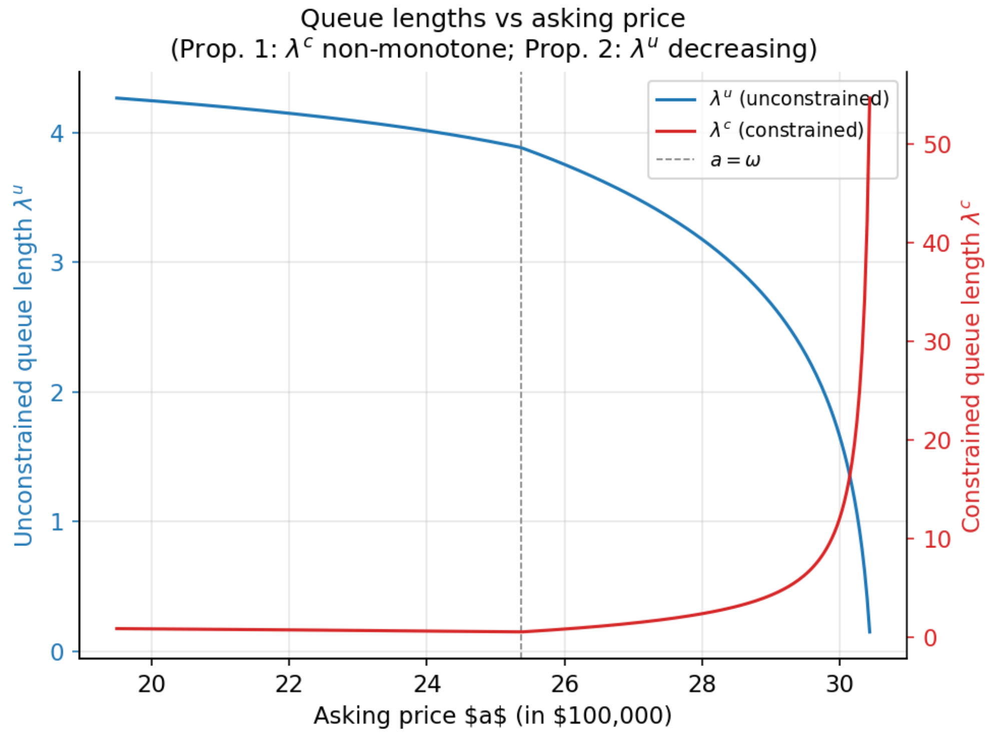
Segment	Price grid	x_L	x_H	ω/x_H	p_H	s	k_c	k_u	Simulated Sold/Ask
1	2.00	1.40	2.20	0.900	0.18	0.005	0.0080	0.0104	0.983
2	8.44	5.91	9.29	0.883	0.18	0.005	0.0093	0.0121	1.041
3	14.89	10.42	16.38	0.865	0.18	0.005	0.0109	0.0141	1.063
4	21.33	14.93	23.47	0.847	0.18	0.005	0.0127	0.0165	1.085
5	27.78	19.44	30.56	0.830	0.18	0.005	0.0148	0.0192	1.107
6	34.22	23.96	37.64	0.830	0.18	0.005	0.0172	0.0224	1.108
7	40.67	28.47	44.73	0.847	0.18	0.005	0.0201	0.0261	1.086
8	47.11	32.98	51.82	0.865	0.18	0.005	0.0234	0.0305	1.067
9	53.56	37.49	58.91	0.883	0.18	0.005	0.0273	0.0355	1.050
10	60.00	42.00	66.00	0.900	0.18	0.005	0.0318	0.0414	1.035

Note: All price and value units are in \$100,000. The buyer cost profile is $k_c = 0.008 \cdot \text{logspace}(0, 0.6, 10)$ and $k_u = 1.3k_c$.

Table **??** reports the adjusted calibration used throughout this appendix. Each row is one price segment on a linear grid from \$200,000 to \$6,000,000 (in \$100,000 units). The support $[x_L, x_H]$ scales with the segment's price grid point; buyer search costs (k_c, k_u) rise with price; and the financial constraint enters through the ratio ω/x_H. The key cross-market variation is in ω/x_H: it is lowest in segments 5–6, so the gap $x_H - \omega$ is largest in the mid-market where the model is meant to generate the peak sale-over-ask ratio. The rightmost column reports the simulated ratio at each segment's calibrated optimum. The simulated values rise through the middle of the distribution and fall in the tails, reproducing the inverse-U documented in the main text (Figure **??**).

Figure **??** checks the queue-length comparative statics at the middle segment (segment 5), where the financial-constraint gap is largest. The seller posts a range of counterfactual asking prices a and solves the free-entry conditions holding all other calibrated parameters fixed. The unconstrained queue λ^u (left axis) declines monotonically in a, as Proposition **??** predicts. The constrained queue λ^c (right axis) is non-monotone: it falls when $a \leq \omega$ because higher asking prices reduce constrained buyers' expected surplus, then rises when $a > \omega$ because fewer unconstrained high-value buyers arrive and constrained buyers face a higher chance of winning at bid ω. The vertical line at $a = \omega$ marks the kink in how constrained buyers trade off participation and bidding.

Figure B1: Buyer Queue Lengths and Asking Price



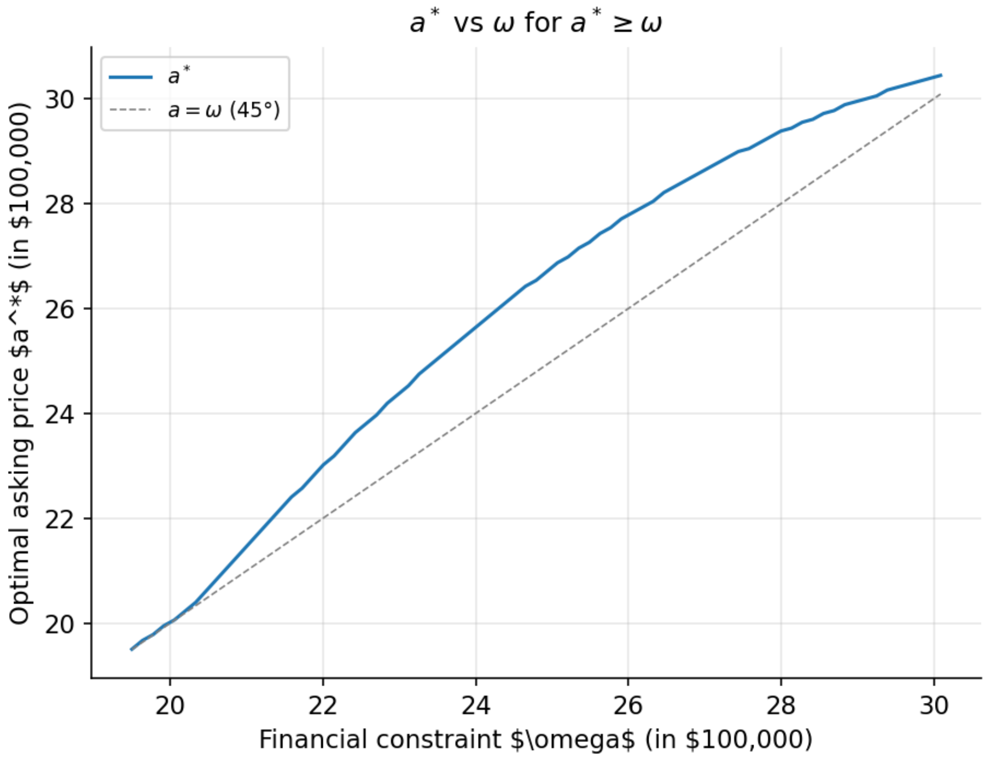
Note: The unconstrained queue length is plotted on the left axis and the constrained queue length on the right axis. The figure verifies Proposition **??**: constrained-buyer queues fall below ω and rise above ω. It also verifies Proposition **??**: unconstrained-buyer queues decline with the asking price.

B.2 Comparative Statics for Propositions and Theorem

The figures below report numerical comparative statics for Proposition **??** and Theorem **??**, again using the middle price segment (segment 5). For each counterfactual parameter value, I re-solve for the optimal asking price a^* and the implied sale-over-ask ratio and market moments. When the optimum lies below the financial constraint ($a^* < \omega$), the seller is effectively pricing to the constrained bid; those points are omitted from the line plots because the paper's bidding-war mechanism is active only when $a^* \geq \omega$.

Financial constraint ω. Figures **??** and **??** vary ω holding other parameters at their segment-5 values. A higher ω narrows the financial-constraint gap and reduces the seller's return to underpricing. Consistent with Proposition **??**, the optimal asking price is weakly increasing in ω. Consistent with the first part of Theorem **??**, the sale-over-ask ratio is highest at intermediate ω, where buyer heterogeneity is most useful for generating competition, and falls when buyers become either very similar or very dissimilar in effective bidding capacity.

Figure B2: Optimal Asking Price and Financial Constraint



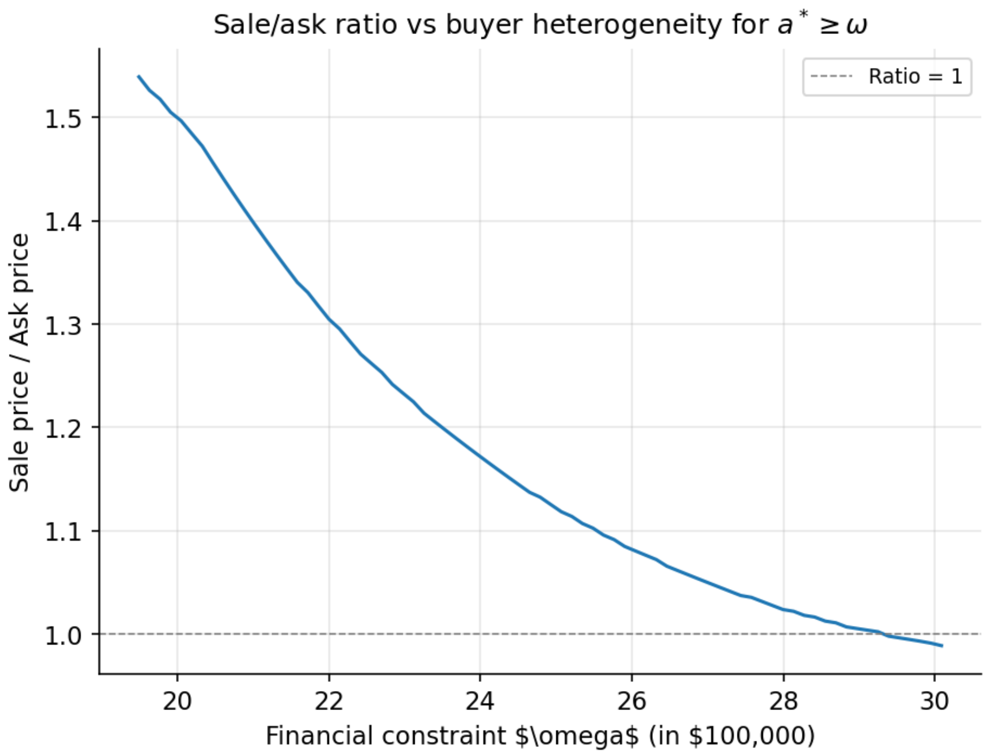
Note: The optimal asking price is non-decreasing in the financial constraint ω, consistent with Proposition **??**. The plotted comparative static keeps the region where the calibrated optimum satisfies $a^* \geq \omega$.

High-value probability p_H. Figure **??** varies the probability that a buyer draws the high match value. When p_H is low, sellers still benefit from a larger pool of buyers and moderate underpricing. As p_H rises, the seller can achieve the upper tail of the price support with fewer arrivals, so the optimal asking price rises and the sale-over-ask ratio eventually falls. The plotted curve is therefore hump-shaped, as in the third part of Theorem **??**.

Seller search cost s. Figure **??** traces how the optimum and sale-over-ask ratio respond to the per-buyer seller search cost. Higher s makes it more expensive to attract additional showings, so the seller posts a higher asking price and attracts a shorter queue. The sale-over-ask ratio declines with s, matching the second part of Theorem **??** and the search-cost comparative static in Proposition **??**.

Joint moments. Figure **??** summarizes how four equilibrium outcomes co-move as ω varies along the same comparative static: the sale-over-ask ratio, the probability of selling above ask, expected bidders, and a days-on-market proxy. The panels show that the mid-range

Figure B3: Sale-over-Ask Ratio and Financial Constraint

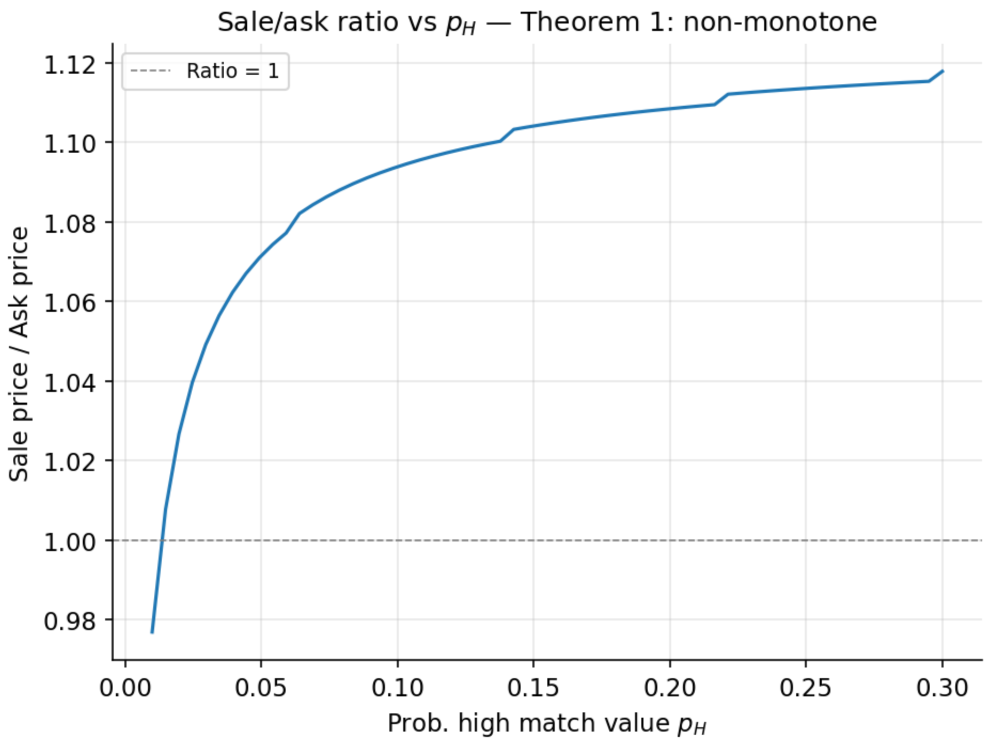


Note: The sale-over-ask ratio is highest when financial constraints create the most useful buyer heterogeneity. This is the numerical counterpart to the first part of Theorem **??**. The plotted comparative static keeps the region where the calibrated optimum satisfies $a^* \geq \omega$.

ω values that maximize the sale-over-ask ratio are also the values at which above-ask sales and buyer traffic are most likely, linking the theoretical bidding-war channel to observable market-thickness objects.

Parameter-space summary. Figure **??** maps the sale-over-ask ratio over a grid in (ω, p_H) at the segment-5 baseline. Cells are blanked out when the optimum satisfies $a^* < \omega$, i.e. when the seller does not find it optimal to price at or above the constrained bid. The heat map shows that sale-over-ask ratios above one arise in an interior region of the parameter space rather than at arbitrary (ω, p_H) pairs. The black dashed contour marks combinations with an interior optimum and expected competition ($\mathbb{E}[\text{bidders}] > 1$); it delineates where the model's mechanism is economically active.

Figure B4: Sale-over-Ask Ratio and High-Value Probability



Note: The sale-over-ask ratio is non-monotone in the probability of a high match value, consistent with the third part of Theorem **??**. Points where the optimum falls below the financial constraint are omitted from the plotted line.

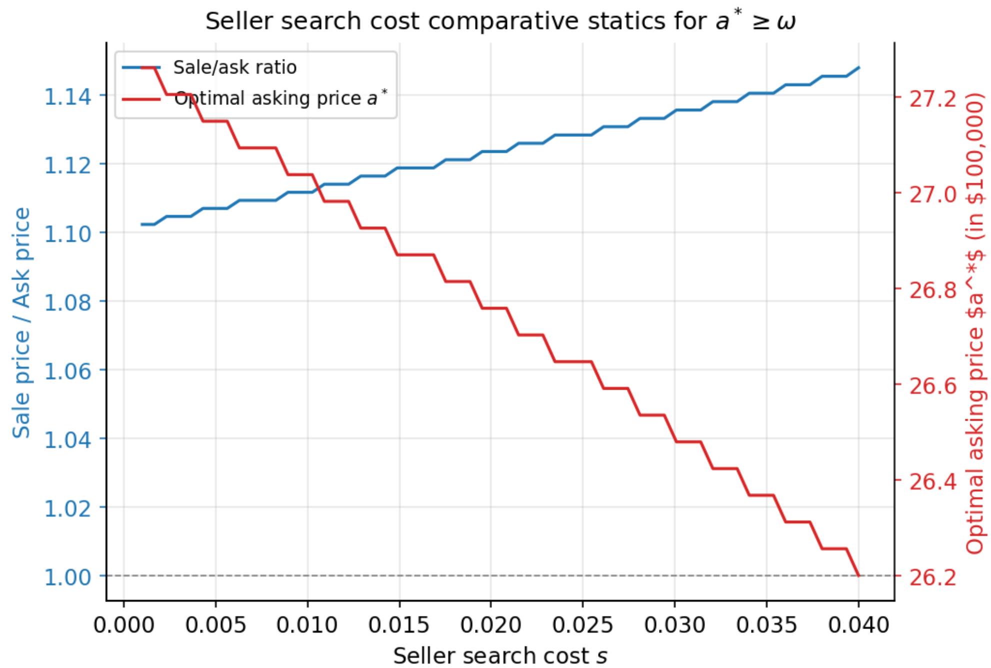
C Data Appendix

C.1 Main Dataset: Toronto MLS 2021

Source. Toronto MLS residential transactions, calendar year 2021, from a proprietary feed obtained through HouseSigma. The raw dataset contains 43,415 listings. After dropping 18 records with sold prices below \$200,000 or a sold-to-ask ratio ≥ 2.0, the working sample is **43,397 transactions**.

Coverage. The dataset spans four property types: Apartment (53 %), Detached (26 %), Townhouse (11 %), and Semi-Detached (10 %). Four geographic zones are defined by MLS region: Core/Downtown ($N = 11{,}697$), East Gate ($N = 10{,}984$), North Hub ($N = 9{,}662$), and West Gate ($N = 10{,}948$). The geographic zone classification follows MLS submarket boundaries, intended to group comparable properties together.

Figure B5: Seller Search Cost Comparative Statics



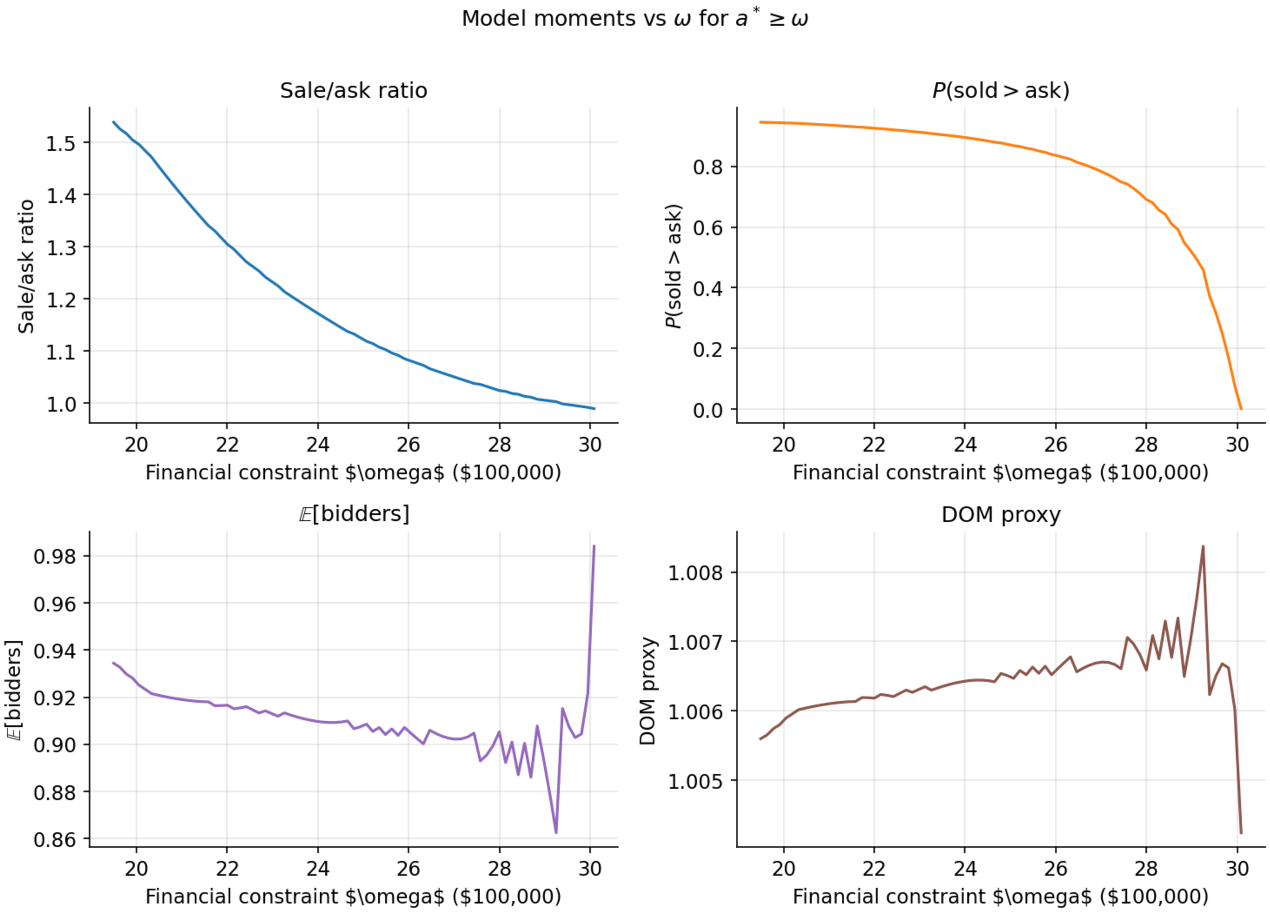
Note: Increasing seller search cost raises the optimal asking price and lowers the sale-over-ask ratio, matching Proposition **??** and the second part of Theorem **??**. Points where the optimum falls below the financial constraint are omitted from the plotted line.

Key variables. The primary outcome is Sold/Ask = sold_price/asking_price. Additional variables include: days on market (DOM), bedroom and bathroom counts, maintenance fee (binary presence and level), square footage, listing remarks text, and property-type and seasonal indicators. Square footage is available for 31,455 records (72 %), heavily over-represented by apartments, which systematically report it; detached houses frequently do not.

Market temperature variables. Monthly absorption rate (number of sales divided by number of listings at the start of the month, at the neighbourhood level) and investor share (percentage of sold properties subsequently listed for lease within six months) are merged from a separate HouseSigma aggregate dataset. These variables are available for 43,291 transactions (99.7 % of the working sample).

Sample selection. Summary statistics for the working sample are reported in Table **??** in the main text. The dataset used for the main analysis removes parking spots and extreme outliers. The resulting sample is comparable to the published summary statistics by the Toronto Regional Real Estate Board in their annual reports.

Figure B6: Model Moments over Financial Constraints



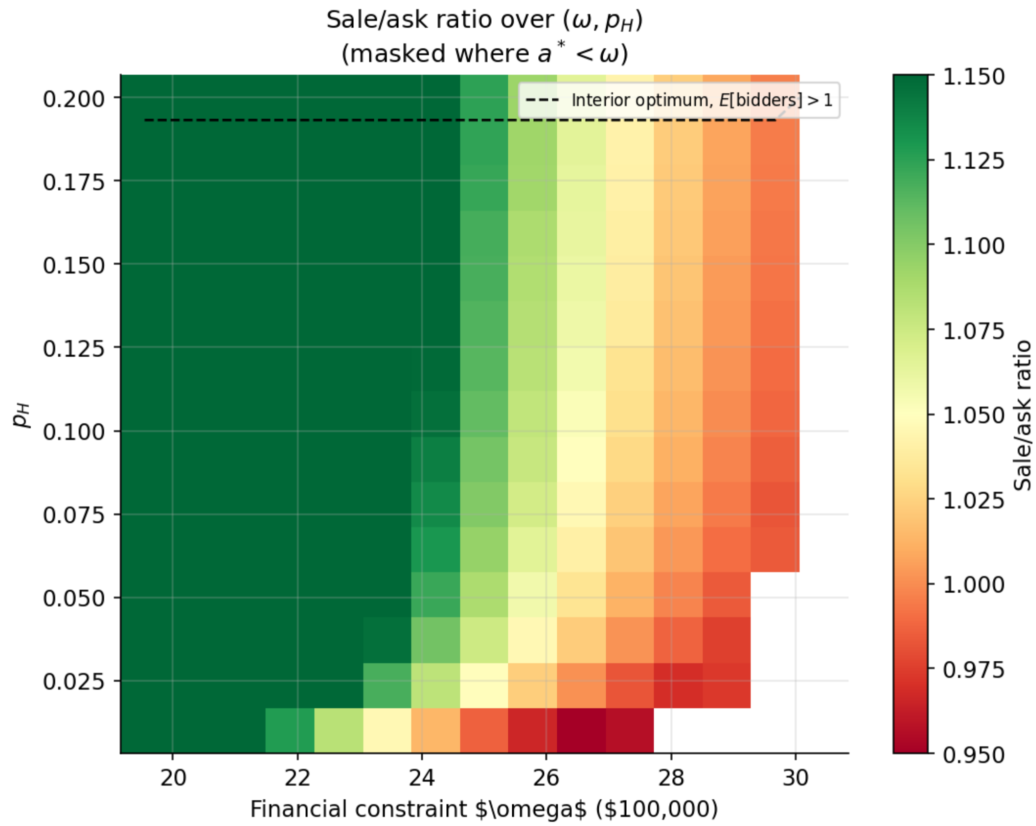
Note: The panels summarize how the sale-over-ask ratio, the probability of selling above ask, expected bidders, and the DOM proxy co-move as the financial constraint ω varies. The plotted comparative static keeps the region where the calibrated optimum satisfies $a^* \geq \omega$.

C.2 Dataset Limitations

There are several limitations to this dataset. First of all, there is no measure of the volume of houses, thus creating difficulty in obtaining vacancies and an accurate measure of queue length. This limitation can be partially solved by combining the monthly neighbourhood-level aggregate sale-to-list ratio of each sub-type of the house (for example, an apartment). Although this is an approximation to the sale-to-list ratio across prices, which is more ideal, this neighbourhood-level dataset is a good approximation since the price of the houses is highly correlated with the types of property.

Secondly, this dataset only included transactions in 2021. The housing market this year is experiencing shocks from COVID-19 and low-interest rates; moreover, the usual structure of visiting the house in person before purchasing has been interrupted. Therefore, it is harder

Figure B7: Feasibility Region in (ω, p_H) Space



Note: The heat map reports the sale-over-ask ratio over the (ω, p_H) parameter space at the optimal asking price. Cells are masked when the calibrated optimum satisfies $a^* < \omega$. The black dashed contour marks parameter pairs with an interior optimum and expected buyer competition ($\mathbb{E}[\text{bidders}] > 1$); inside this region the model can generate sale-over-ask ratios above one.

to distinguish the impact of the asking price on the housing market during this period from other shocks resulting from COVID-19.

Thirdly, there is no measure of buyers' characteristics. The data contained in this dataset only includes house characteristics, without any bidding information, exchanges of offers, etc. This problem is harder to solve because the housing market typically does not have much buyers' side information. **han2016role** uses a survey result and tests their model predictions, which are similar to my predictions. In this paper, I used a combination of the census income distribution table and results from the 2019 wave of the Survey of Financial Security both from Statistics Canada (**census**; **sfs**).

C.3 Secondary and Supplementary Datasets

Historical dataset (2000–2012). I use a supplementary dataset of detached houses and land sold in Toronto between 2000 and 2012 to assess the historical persistence of the mid-price hump pattern.

Neighbourhood dataset. An online housing platform publishes monthly neighbourhood-level aggregate data used to approximate buyer-side information and inventory. The neighbourhood definition follows MLS submarket boundaries at a granularity between the census tract and census subdivision. The competition measure (absorption rate) and investor measure (share of properties for-lease within six months) are constructed from this source. The absorption rate exceeding 100 indicates more properties were sold within the month than were available at the start of the month.

Survey of Financial Security (SFS). For calibrating buyer wealth and financial constraints, the 2019 wave of Statistics Canada's Survey of Financial Security is used together with the census after-tax income distribution table, both restricted to Toronto. The public-use microdata file contains 136 observations of buyers who purchased a property within a year of the survey, providing a baseline for matching buyers' income and wealth to transaction price deciles.

C.4 Financial Constraint Measure Construction

Definition. The financial constraint fraction is defined as the median net worth of households who purchased a house in Toronto in 2018–2019, as recorded in the 2019 SFS wave. All house prices are inflated to 2019 values using the Teranet–National Bank of Canada housing price index.[13]

Construction. SFS observations are merged to the sale data by decile of the sale price. Each SFS respondent is matched to one of ten sale price deciles. Due to insufficient observations in the 10th decile, values there are linearly extrapolated. The financial constraint for each decile is computed as the ratio of median net worth to the 10th percentile of net worth in the sample. This ratio is then normalised so that the maximum equals 1. The normalised constraint profile is plotted in the left panel of Figure **??** in the main text and is used directly in the calibration of Section **??**. Figure **??** in Appendix **??** reports the sale-price residual

[13] Data downloaded from `https://housepriceindex.ca/#maps=on_toronto`.

distributions and group means for failed auctioneers, successful auctioneers, and patient sellers.

D Seller-Type Classifier: Definitions and Methodology

This appendix documents the seller-type taxonomy introduced in Section **??**.

D.1 Design Principles

The classification is intended to capture the seller's apparent motivation, rather than the physical characteristics of the property alone[14]. A renovated house, for example, is not automatically treated as a flipper listing; it is classified that way only when the description uses renovation as the central justification for value. Similarly, a condominium with standard amenities is not assigned to a more strategic type unless the text gives evidence of investor targeting, urgency marketing, luxury positioning, or another seller narrative.

The categories are intentionally designed to not be mutually exclusive at the listing level. A single listing can contain signals for more than one type, so the classifier records a score for each category. To keep these scores interpretable, however, a given keyword or phrase is assigned to only one type in the signal dictionary. For example, the phrase "separate entrance" can be used to trigger both the Attracting Investors and the Own Home types, but it is assigned to only one of the more dominant types according to the sentence. This avoids mechanically counting the same piece of language as evidence for several different seller motives.

For each listing, the procedure reports both the continuous score for every type and the set of matched phrases that generated those scores. I manually inspected a random set of listings to ensure that the classifier is working as intended. The dominant seller type is then defined as the category with the highest score. This structure keeps the classification transparent: I can see both the assigned type, but also the language in the original listing that produced it.

D.2 Type Definitions

Type 1: Attracting Investors. This type captures listings in which the seller presents the property primarily as an income asset, redevelopment opportunity, or land-value play rather than as a personal home. The signals that support this type are financial framing, such as

[14] This analysis is also accompanied by a regression analysis on keywords.

the listing is written for a buyer who evaluates the property through expected cash flow or future development feasibility. Investor-oriented language includes references to rental income, income potential, existing tenants, explicit rent amounts, separate entrances, dual kitchens, multiple stoves, and basement-apartment potential. Development-oriented listings also receive this classification when they emphasize builders, developers, rebuilding, severance, lot dimensions near the beginning of the remarks, "as-is, where-is" sale language, approved plans, permits, zoning approval, or future value and appreciation potential.

Type 2: Own Home (Owner-Occupant Seller). This type identifies the narrow subset of listings where the seller appears to be an individual owner-occupant selling for personal life reasons. The narrative is not that the property is optimised as an asset, but that it has been someone's home and is being sold because of a life transition. The key type of seller I want to capture with this type are people that value their home much higher than an asset. This type of seller could be very different from an average seller, especially in terms of motivation to sell and expectation for selling price. Textual evidence includes personal possessive language such as "my home," "our home," or "we've lived here," references to a duration of occupancy or having raised a family in the home, and explicit selling reasons such as downsizing, relocation, and etc.

Type 3: Flipper Renovator (Non-Luxury). This type describes sellers who anchor the asking price to capital invested in renovation. The implicit narrative is that the seller has put substantial work into the property and that the renovation justifies the price. Renovation is the primary value signal, so this type is distinct from Luxury Home. Flipper Renovator is about return on renovation work, not a complete rebuild for luxury. Signals that support this type include explicit renovation dollar amounts, work permits or closed permits, recent-renovation timelines, and broad-scope renovation language such as "completely renovated" or "reno,". A listing fits this type also when it names several major systems that have been replaced or upgraded, such as a new roof, kitchen, bath, floors, windows, HVAC, electrical work, or addition. If luxury appliance brands, trophy amenities, or prestige-neighbourhood signals appear alongside renovation language, the listing is classified as Luxury Home rather than Flipper Renovator.

Type 4: Attractive Ads. This type captures listings whose main signal is marketing intensity. The seller or agent manufactures urgency, scarcity, and competitive tension in order to generate attention and multiple offers. The text is emotionally charged rather than purely informational. Signals include offer deadlines, urgency language such as "won't last," "act fast," or "don't miss out," rarity claims such as "rarely offered" or "seldom available," and must-see language. Pre-list inspection language, positive superlatives such as "coveted," "sought-after," "one of a kind," or "irreplaceable," unusually long narrative sentences, and high exclamation-mark density also support this classification.

Type 5: Luxury Home. This type identifies listings that position the property at the top of the market through premium finishes, brand names, architectural vocabulary, and other luxury signals. The property is presented as exceptional, not as something that needs urgency language to sell. Signals include explicitly high square footage, luxury appliance brands such as Sub-Zero, Wolf, Thermador, Miele, Viking, Fisher-Paykel, Gaggenau, La Cornue, or AGA, smart-home systems, glass elevators, heated floors or driveways, wine cellars, home theaters, indoor pools, outdoor kitchens, spas, and other trophy amenities. Luxury descriptors such as "high-end," "upscale," "prestigious," "exclusive," or "bespoke," resort-style language, gated-estate language, tennis or golf amenities, named architects, new construction paired with luxury signals, and custom premium finishes such as millwork, custom cabinetry, or coffered ceilings also place a listing in this type.

Type 6: Generic Market Listing. This type is reserved for listings with no distinctive strategic angle. The property is presented as a solid, standard unit whose description is functional and descriptive. It is most common among condos, stacked townhouses, and low-rise apartments in non-luxury neighbourhoods. Unlike the other types, Generic Market Listing is partly defined by absence: there is no renovation anchoring, no luxury positioning, no personal seller narrative, and no investor framing.

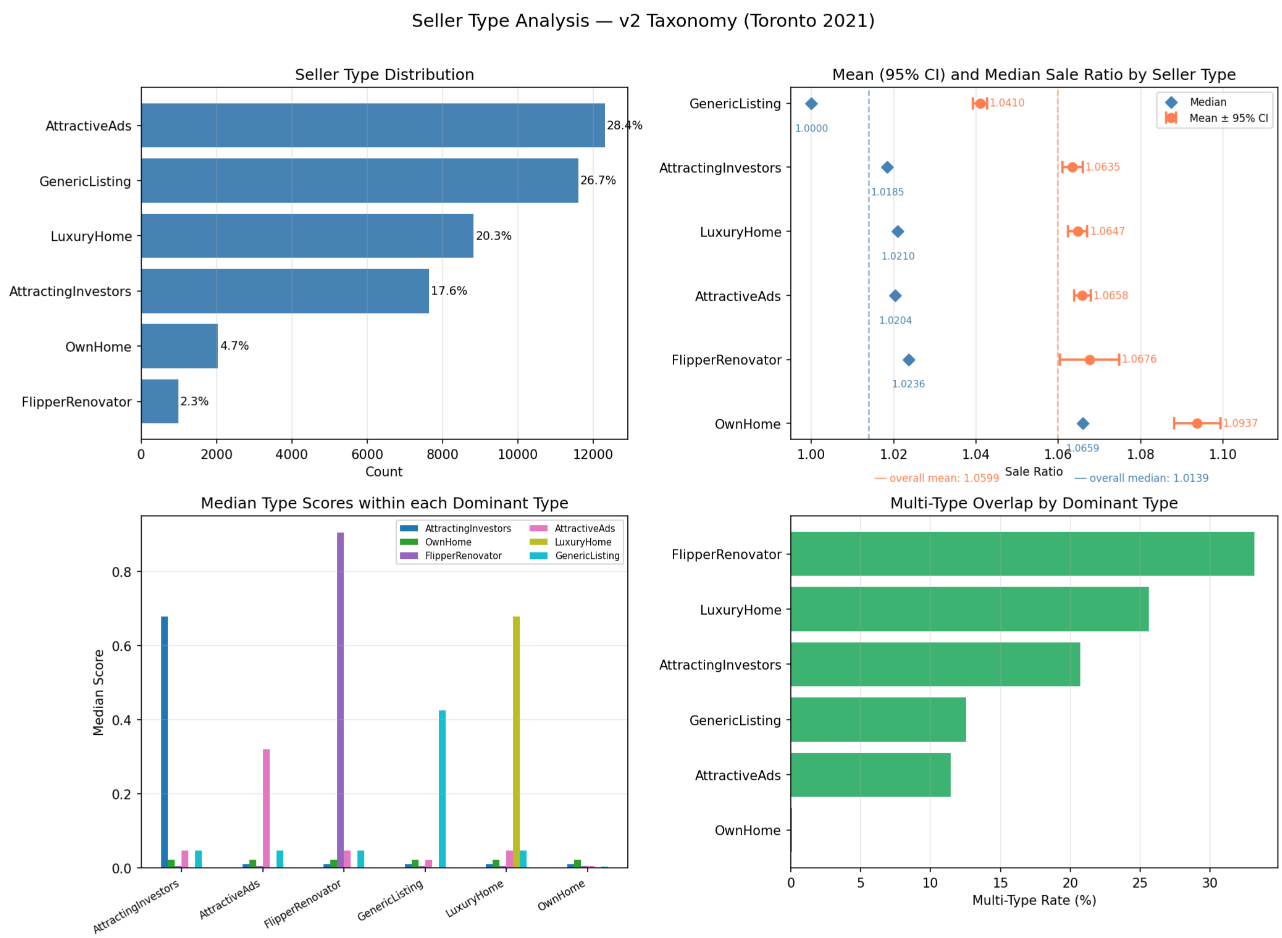
D.3 Summary Statistics

Figure **??** shows the distribution of dominant seller types across the 43,379 classified listings, along with mean sold-to-ask ratios by type.

D.4 Regression Results

To assess how much listing language explains sold-to-ask outcomes beyond standard property and market controls, I estimate four nested OLS models with sold-to-ask ratio as the dependent variable. **Model M0** includes only controls: bedrooms, bathrooms, log maintenance fee, log square footage, market-temperature z-scores, and property-type, season, and zone fixed effects. **Model M1** adds five seller-type dummies to M0 (baseline category: GenericListing). **Model M2** adds approximately 60 individual keyword indicators to M0.

Figure D1: Seller Type Distribution and Mean Sold-to-Ask Ratio



Note: Left panel shows the share of listings assigned to each dominant seller type. Right panel shows mean sold-to-ask ratio by type with 95% confidence intervals. GenericListing is the modal type and serves as the regression baseline.

Model M3 combines the seller-type dummies and keyword indicators from M1 and M2. Table **??** compares fit across these nested specifications. Figures **??**–**??** report coefficient estimates from the full M3 specification.

D.5 Output Schema

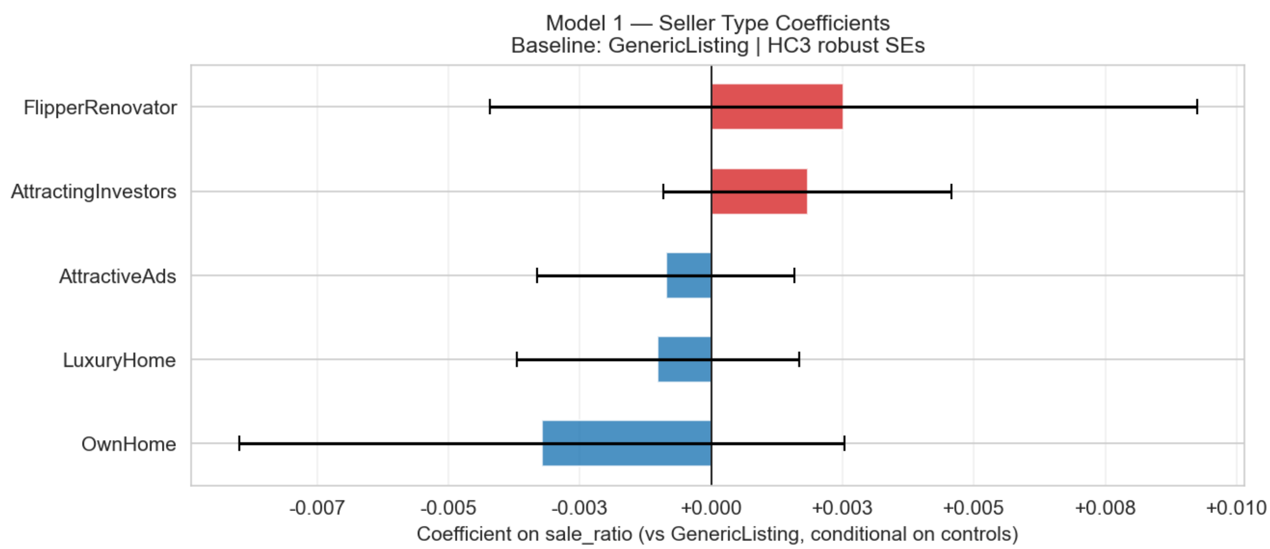
The classifier writes one row per listing. Table **??** documents each output column: the six continuous type scores (0–1), the dominant-type assignment, flags for multi-type listings, the matched keyword strings used for transparency, and the sold-to-ask ratio used in the regressions above.

Table D1: Regression Model Comparison — Dependent Variable: Sold-to-Ask Ratio

Model	Additional regressors	R^2	ΔR^2 vs M0
M0 (baseline)	Controls only	0.1896	—
M1	Seller-type dummies (5)	0.1897	+0.0001
M2	Keyword indicators (≈60)	0.1909	+0.0013
M3	Seller-type + keyword indicators	0.1910	+0.0014

Note: $N = 43{,}273$. HC3 robust SEs. Baseline: GenericListing, Spring, first zone alphabetically, Apartment. All five seller-type coefficients in M1 are statistically indistinguishable from zero (range: -0.0032 to $+0.0025$; all $p > 0.18$). The incremental R^2 from seller-type and keyword indicators combined is 0.0014.

Figure D2: Seller-Type Coefficients from OLS Regression (Full Specification)

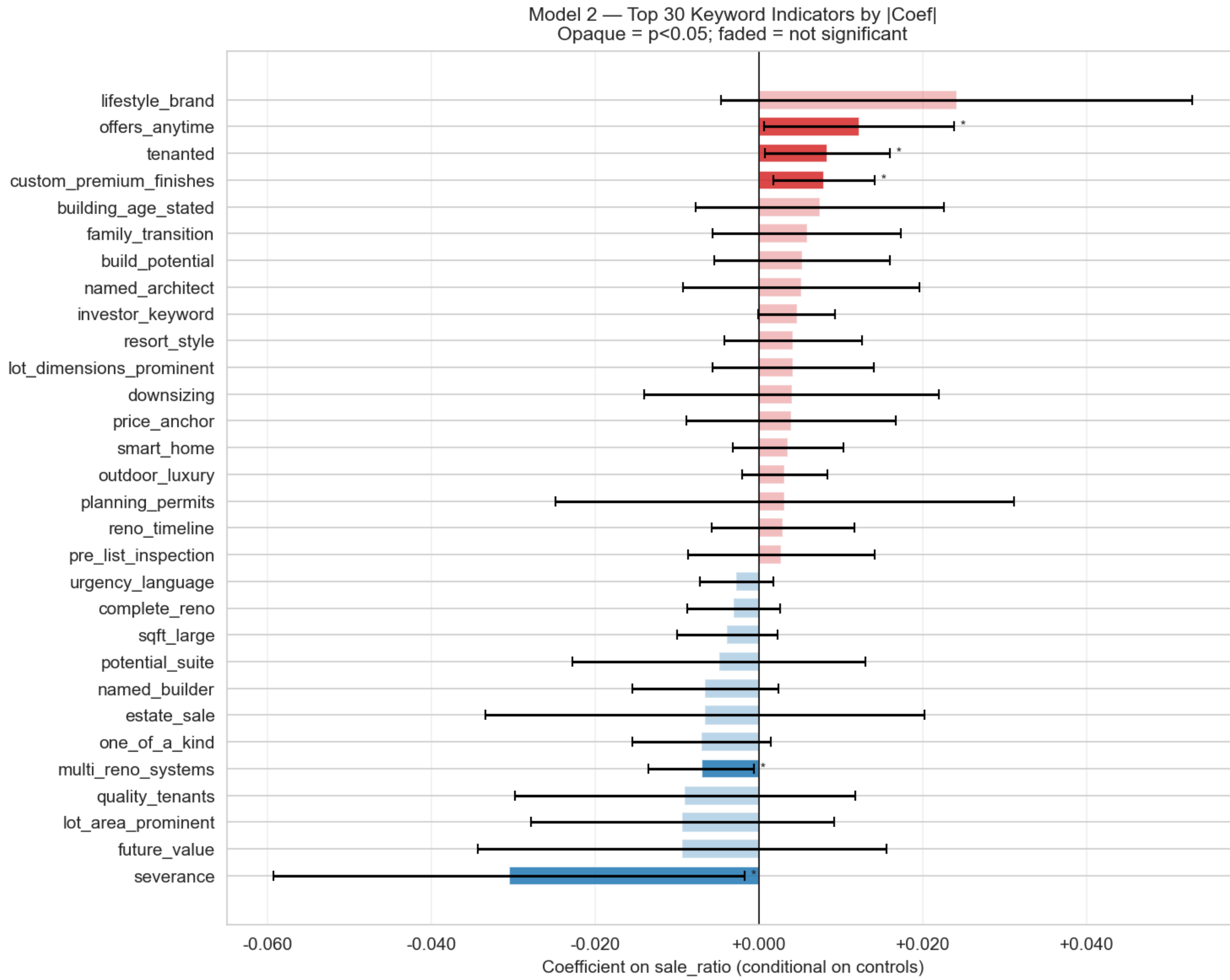


Note: Point estimates and 95% confidence intervals for seller-type indicators in the full regression specification. Baseline: GenericListing. Outcome: sold-to-ask ratio. HC3 robust standard errors. $N = 43{,}379$.

D.6 Additional Empirical Figures

Uniqueness in the Luxury Tail. The luxury market is largely heterogeneous, making it structurally unsuitable for sellers to induce a bidding war. The number of buyers for such properties is small and can arrive at large intervals, and the idiosyncratic preferences of buyers play a much larger role in determining willingness to pay, which makes aligning asking price to sale price more difficult. To show this large variation and heterogeneity, the average variance of residuals from the hedonic regression is computed across different sold-price bins. The most luxury bin has a median sold price of \$3.6M with a variance of 0.159, more than three times the variance in any other bin. Figure **??** plots the resulting variance profile.

Figure D3: Keyword Coefficients from OLS Regression (Full Specification)



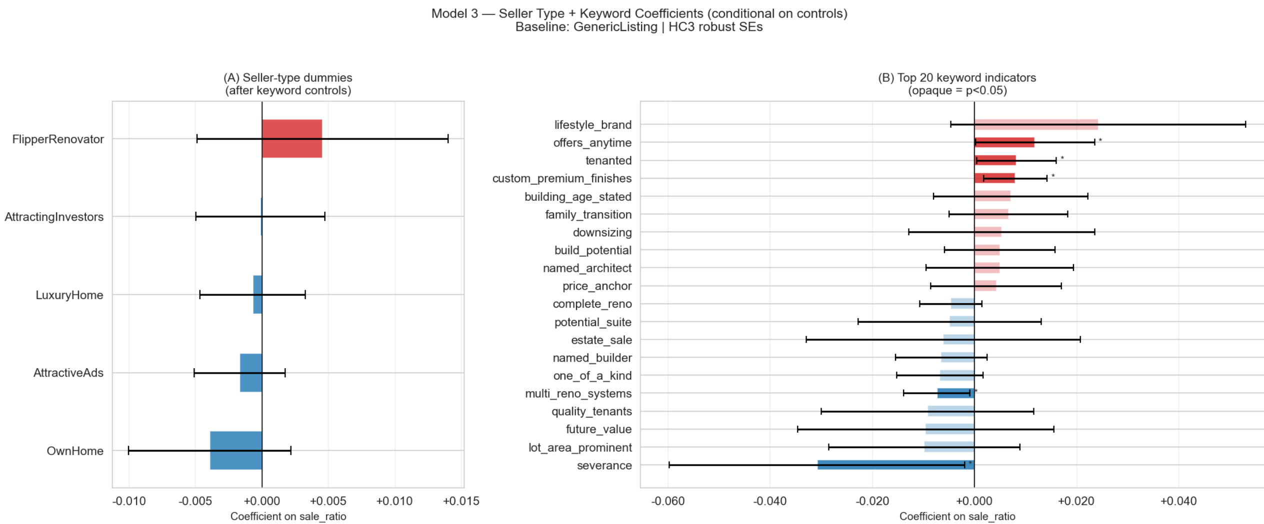
Note: Estimated coefficients on individual keyword presence indicators from the full regression. Outcome: sold-to-ask ratio. HC3 robust standard errors. $N = 43{,}379$.

D.7 Bimodality of the Sold-to-Ask Distribution

The aggregate sold-to-ask hump conceals a sharper structural feature: within mid-price bins, the distribution of Sold/Ask is not unimodal but bimodal, with one mass concentrated below 1.0 (the Negotiation Regime) and a second concentrated above 1.0 (the Auction Regime). This appendix documents the bimodality with four complementary tests.

Test 1 — Hartigan's Dip Test. Figure **??** applies Hartigan's dip test to the sold-to-ask distribution within each of 20 equal-frequency log-price bins. The left panel plots the dip statistic by bin; all but one bin are flagged as statistically bimodal ($p < 0.05$, shown in red). The right panel overlays the dip statistic against the median sold-to-ask ratio, showing that bimodality strength peaks in the same mid-price range where the aggregate hump is tallest,

Figure D4: Combined Coefficient Panel — Model 3



Note: Combined coefficient plot for the full specification (Model 3), showing both seller-type and control variable estimates. Outcome: sold-to-ask ratio. HC3 robust standard errors. $N = 43{,}379$.

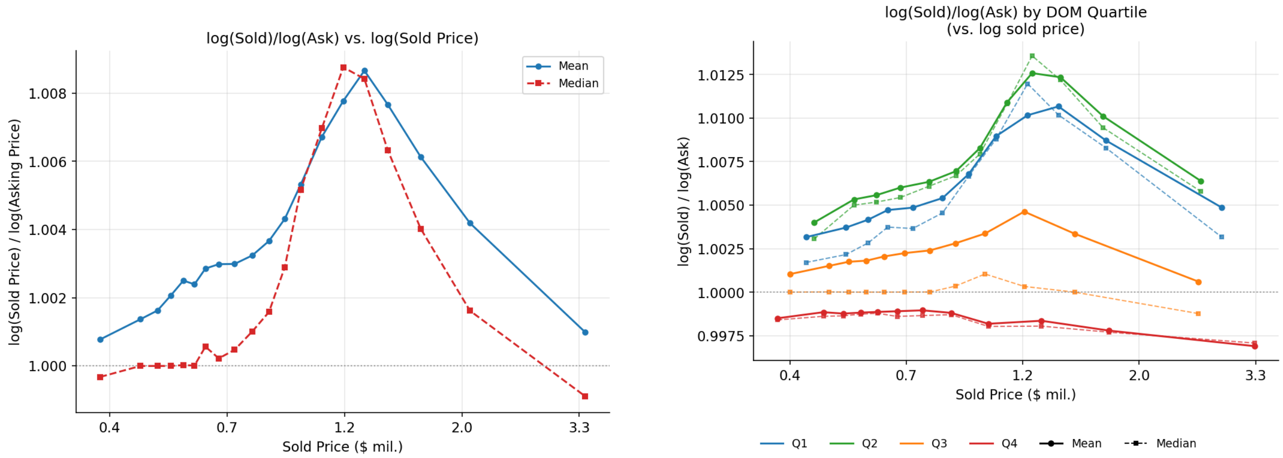
Table D2: Classifier Output Schema

Column	Type	Description
`type_investor_score`	float 0–1	Attracting Investors likelihood
`type_ownhome_score`	float 0–1	Own Home (owner-occupant selling for personal reasons)
`type_flipper_score`	float 0–1	Flipper Renovator (non-luxury)
`type_attractive_score`	float 0–1	Attractive Ads / urgency marketing
`type_luxury_score`	float 0–1	Luxury Home
`type_generic_score`	float 0–1	Generic Market Listing
`dominant_type`	string	Name of highest-scoring type
`dominant_score`	float 0–1	Score of dominant type
`multi_type`	bool	True if 2+ types score above 0.40
`[type]_keywords`	string	Matched keywords for each type (pipe-separated)
`sale_ratio`	float	`sold_price / asking_price`

and attenuates in the cheap and luxury tails.

Test 2 — Property-Type Decomposition. A natural alternative explanation is that bimodality is a compositional artifact: Freehold (Detached/Semi-Detached) listings and Apartments co-exist in mid-price bins and have systematically different sold-to-ask ratios, mechanically generating two modes in the pooled distribution. Figure **??** shows the within-bin eta-squared (η^2) from a one-way ANOVA of the sold-to-ask ratio on property type. The bimodal bins (red) have uniformly low η^2 — property type explains at most 10% of within-

Figure D5: Underasking Behaviour of Sellers — Log-Price Version

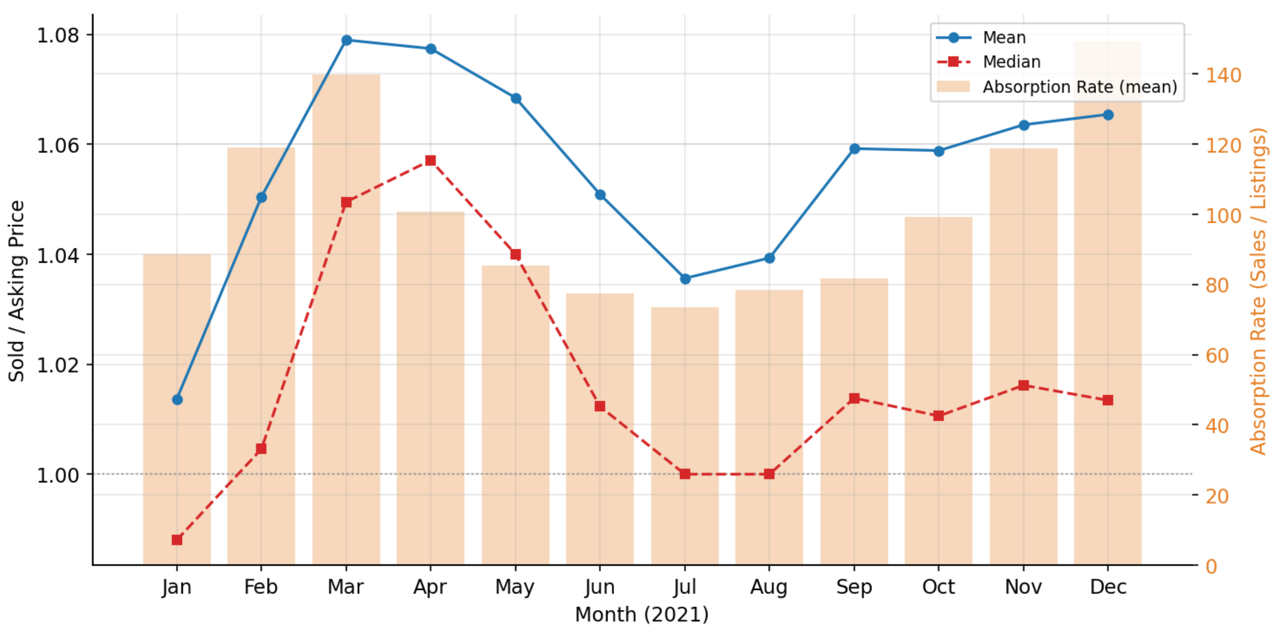


(a) log(Sold)/log(Ask) vs. log sold price (aggregate)

(b) log(Sold)/log(Ask) vs. log sold price, by DOM quartile

Note: Log-price counterpart to Figure **??** in the main text. The left panel plots log(Sold)/log(Ask) across log sold price bins for all 2021 Toronto transactions. The right panel overlays the same log sold-to-asking measure for each of the four DOM quartiles (Q1 fastest, Q4 slowest). The inverse-V shape and the DOM gradient are both preserved under the log transformation.

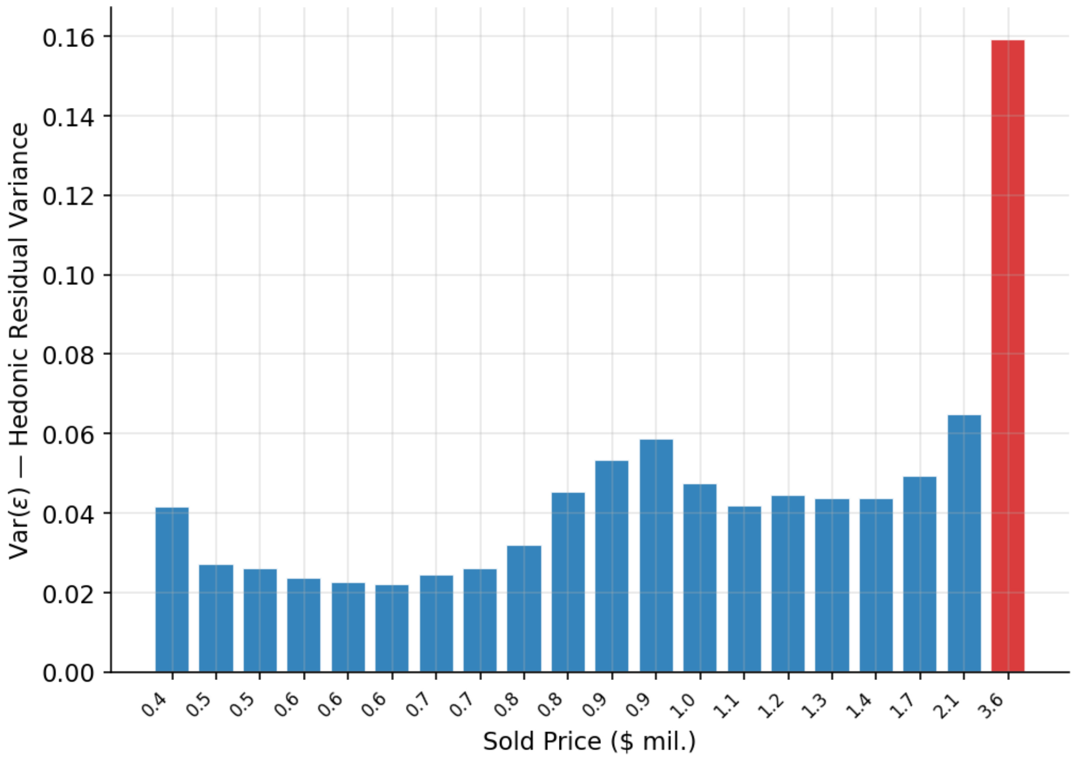
Figure D6: Sold-over-Ask and Absorption Rate by Month



Note: Monthly mean (solid) and median (dashed) sold-to-asking ratio with mean neighbourhood-level absorption rate (right axis, orange bars); reference line at 1.0 marks sold = ask.

bin variance and is not concentrated in the most bimodal bins. Bimodality is therefore not a property-type artifact.

Figure D7: Residual Variance from Hedonic Regression by Sold-Price Bin

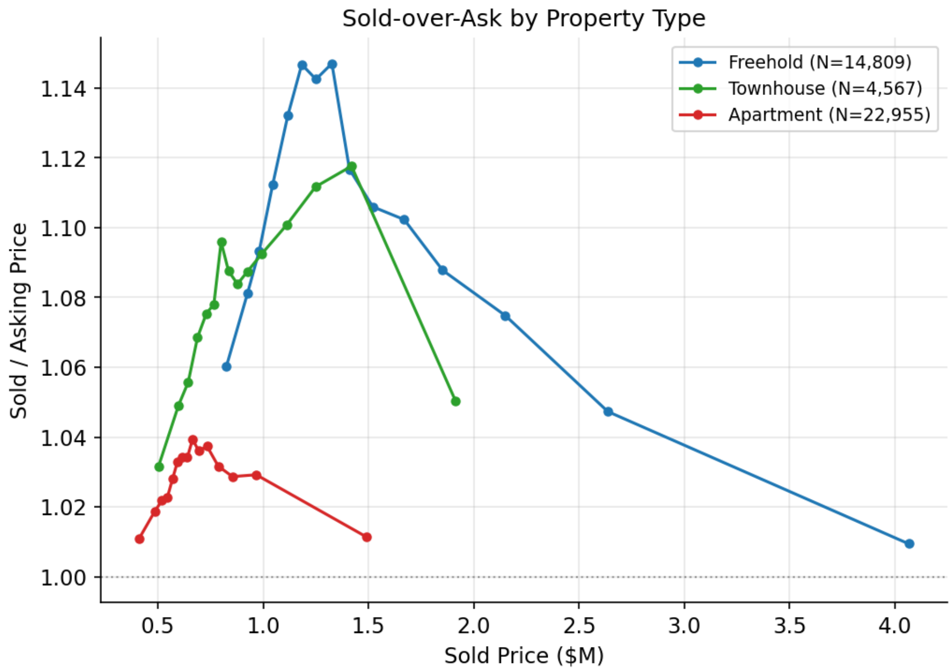


Note: Each bar is one of 20 equal-frequency sold-price bins (N = 43,273). Blue bars are bins with mean sold price below $2.5M; red above. The residuals $\hat{\varepsilon}_i$ are from the hedonic OLS $\log(\text{Sold}_i) = \alpha + \mathbf{X}_i'\beta + \varepsilon_i$, where $\mathbf{X}_i$ includes bedrooms, bathrooms, square footage (with missingness indicators), maintenance fee, eight amenity dummies, z-scored absorption rate and investor share, and property-type, season, and zone fixed effects. Variance rises sharply above $2.5M, consistent with idiosyncratic matching in the luxury segment.

Test 3 — GMM Two-Component Fits. Figure **??** fits a two-component Gaussian Mixture Model (GMM) to the sold-to-ask distribution in the six most bimodal bins. In every bin, one component centres below 1.0 (the Negotiation Regime, green) and a second centres above 1.0 (the Auction Regime, red), consistent with a discrete behavioural split rather than a continuous spread of outcomes.

Test 4 — Zone KDE. Figure **??** plots kernel density estimates of the sold-to-ask ratio separately for each of the four Toronto geographic zones within the same six peak bimodal bins. The bimodal shape persists within every zone (η^2 reported in each panel title), ruling out geographic sorting as the driver.

Figure D8: Sold-over-Ask by Property Type



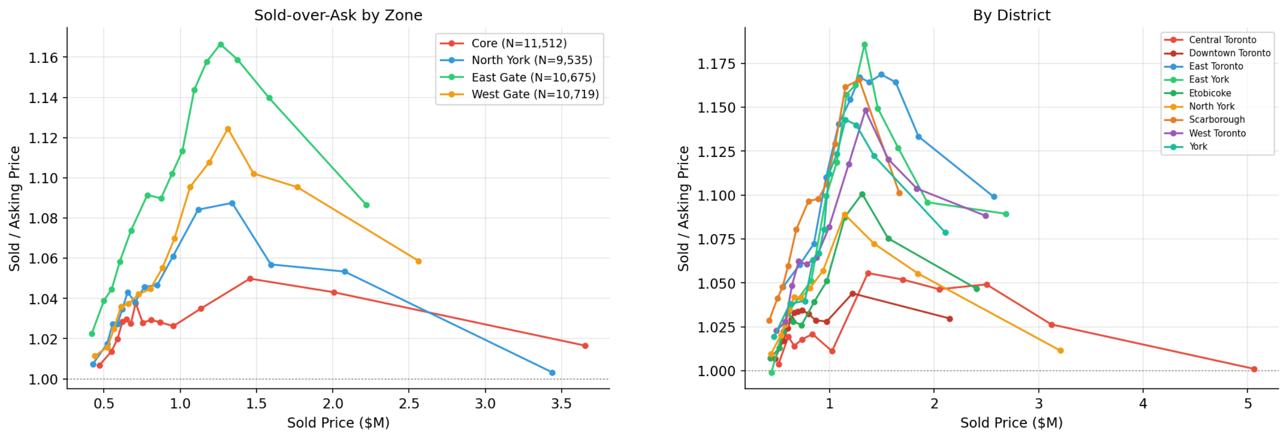
Note: Mean sold-to-asking ratio across sold price bins for three property groups: Freehold (Detached and Semi-Detached), Townhouse, and Apartment. The x-axis is the sold price in millions of dollars. The inverse-V shape is strongest for Freehold and attenuated for Apartments, consistent with the paper's prediction that strategic underpricing is more effective in heterogeneous, less-substitutable submarkets.

D.8 Coordination Failure vs. Signaling Error

The main text classifies sellers into failed auctioneers, successful auctioneers, and patient sellers and asks whether failed underpricers differ systematically in listing quality or coordination language. Figure **??** and Table **??** report seller-type composition and two text-based signal indices for each group. Failed and successful auctioneers use nearly identical coordination language: mean coordination-signal counts are 0.031 and 0.031 per listing (Welch $t = 0.020$, $p = 0.98$). Hard-quality signals are also statistically indistinguishable: 0.943 versus 0.963 ($t = -1.141$, $p = 0.25$). Although the overall seller-type mix differs across the three groups ($\chi^2 = 168$, $p < 10^{-34}$), the language deployed by failed and successful auctioneers is effectively the same.

The pattern supports interpreting auction failure as a coordination event rather than a quality shortfall. Sellers in the low-price trap made similar strategic choices and posted similar signals to sellers who succeeded; the key difference is whether enough buyers arrived in time to convert underpricing into a bidding war. When coordination fails, the low asking price can become a persistent anchor on the sale price rather than a temporary floor. Fig-

Figure D9: Sold-over-Ask by Zone and District

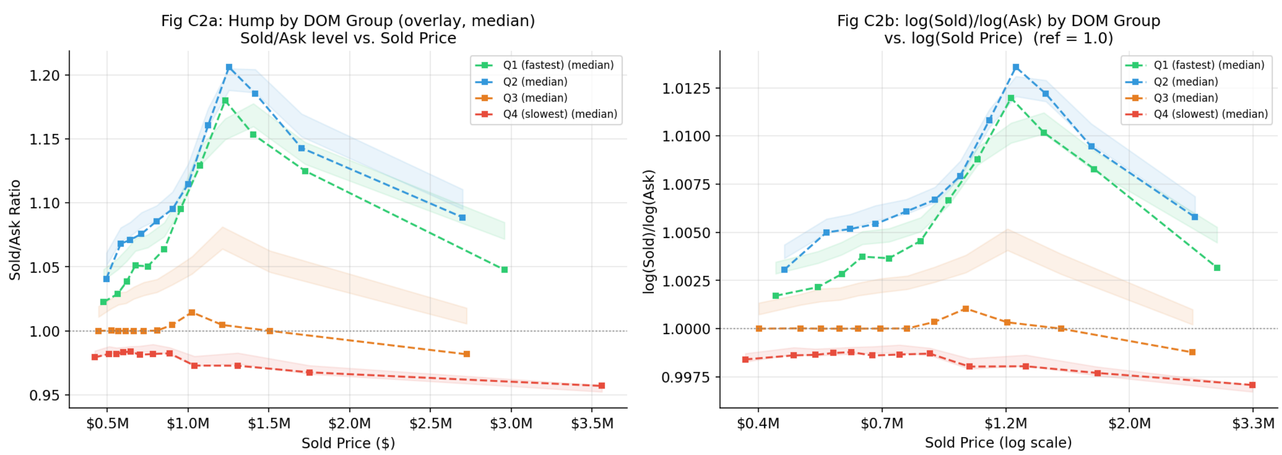


(a) Mean sold-to-asking ratio by zone

(b) Mean sold-to-asking ratio by district

Note: Left panel plots mean sold-to-asking ratio across sold price bins for four Toronto zones: Core (Downtown and Central Toronto), North York, East Gate (Scarborough, East York, East Toronto), and West Gate (Etobicoke, West Toronto, York). Right panel repeats the exercise for each of the nine individual MLS districts. The x-axis is the sold price in millions of dollars. The inverse-V shape persists across all zones and districts.

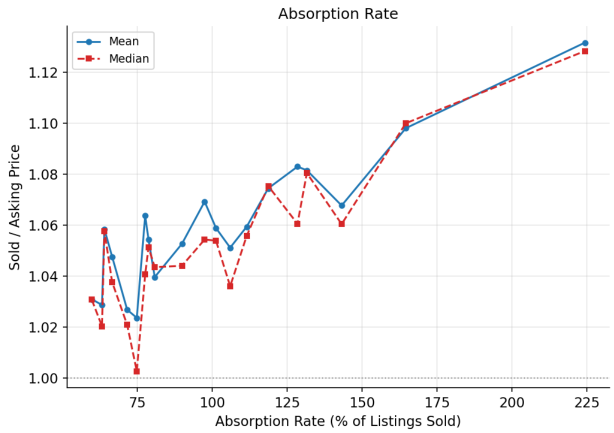
Figure D10: Underasking Behaviour by DOM Quartile — Overlay with Confidence Intervals



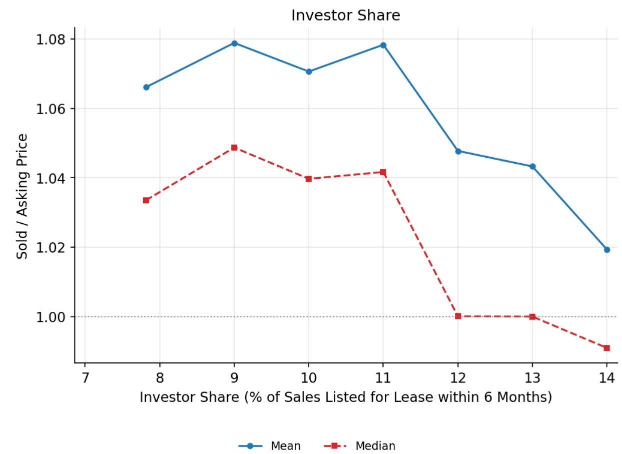
Note: Each panel overlays the median sold-to-asking ratio across sold price bins for each of the four DOM quartiles (Q1 fastest, Q4 slowest). The shaded bands are 95% confidence intervals around the mean (not the median), constructed as $\bar{x} \pm 1.96 \cdot \widehat{\text{SE}}$ within each bin. Because the plotted lines are medians while the bands are mean CIs, the bands do not necessarily enclose the point estimates; the two statistics can differ when the within-bin distribution is skewed. For a version showing only mean and median point estimates without bands, see Figure **??** in the main text.

ure **??** and Table **??** show that this conclusion is robust to an alternative, broader definition of slow-sale groups.

Figure D11: Sold-to-Ask Ratio, Absorption Rate, and Investor Share



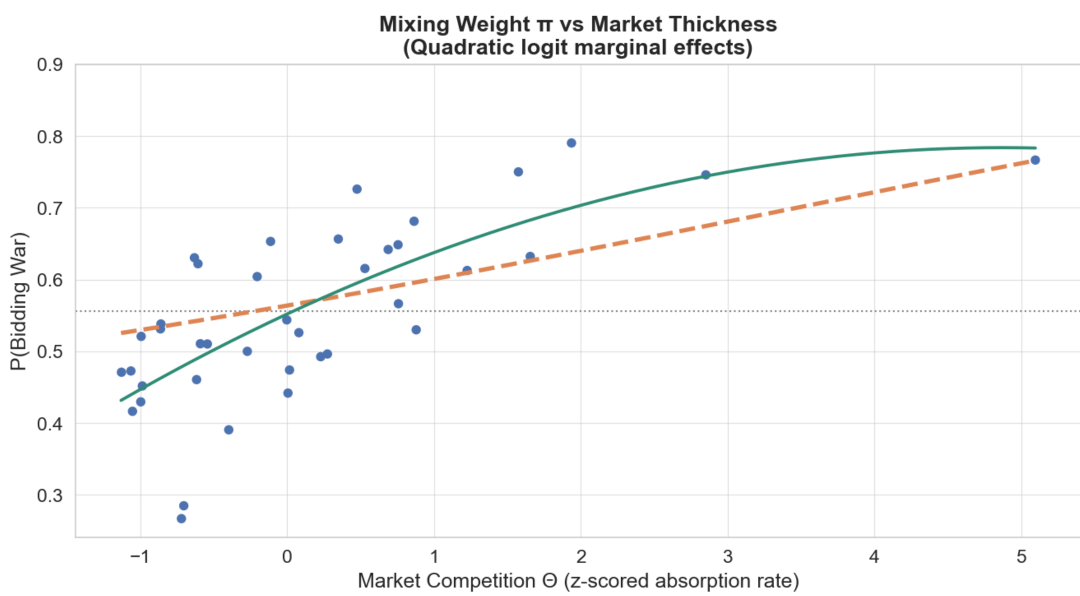
(a) Absorption rate



(b) Investor share

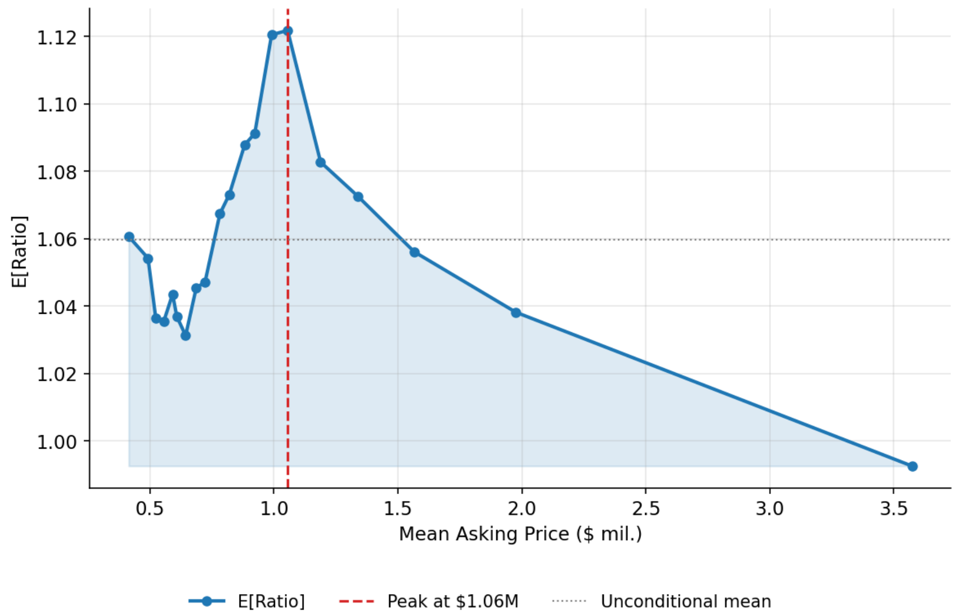
Note: Mean and median sold-to-asking ratio across 20 equal-count bins. The left panel bins by neighbourhood-month absorption rate (percentage of active listings sold); the right panel bins by investor share (percentage of sold properties subsequently listed for lease within six months). Reference line at 1.0 marks sold = ask.

Figure D12: Probability of Auction vs. Market Thickness (Quadratic)



Note: Each point is a binned mean of observed auction-regime entry rates against the z-scored neighbourhood-month absorption rate. The dotted horizontal line marks the unconditional mean Pr(Bidding War) = 0.556. The dashed curve (orange) plots predicted probabilities from a quadratic logit $\Pr(\text{Auction}_i) = \Lambda(\alpha + \beta_1 \Theta_{m,n} + \delta \Theta^2_{m,n} + \mathbf{X}_i' \gamma)$, estimated at covariate means; $\hat{\beta}_1 = 0.145$, $\hat{\delta} = 0.007$. The solid curve (teal) is a degree-2 polynomial fitted directly to the binned means.

Figure D13: Expected Sold-to-Ask Ratio by Asking-Price Bin (Robustness)



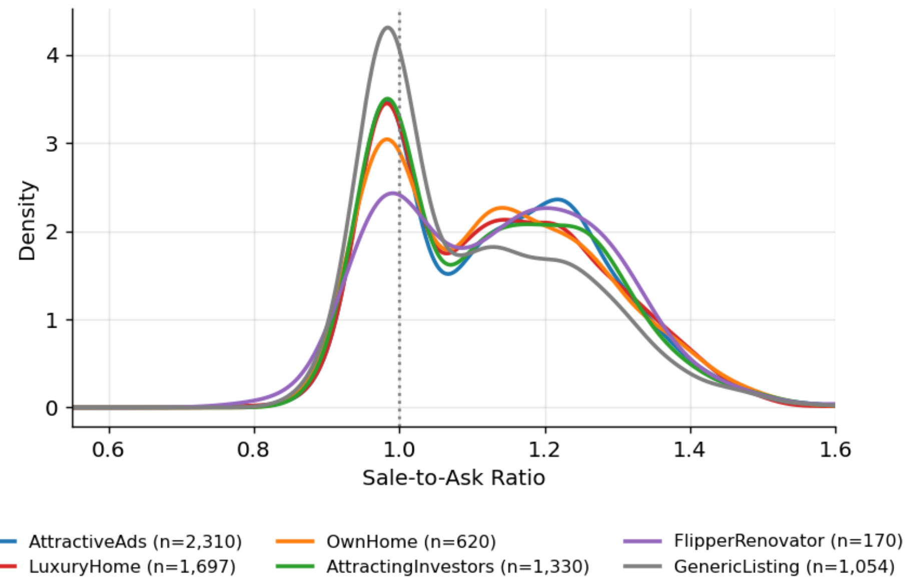
Note: E[Sold/Ask] from Equation (**??**) plotted across 20 equal-frequency asking-price bins (N = 43,273). The global maximum is 1.122 at $1.06M asking price, coinciding with the hump peak in Figure **??**. This replicates the main result of Figure **??** using asking price rather than sold price as the binning variable, confirming the peak is not an artefact of the price measure chosen.

Table D3: **Signal Intensity and Seller-Type Composition by Outcome Group**

	Successful Auctioneers	**Failed Auctioneers**	**Patient Sellers**
N	12,496	4,677	5,640
% Generic Listing	23.3%	31.7%	29.8%
Coordination signals	0.031	0.031	0.040
	(0.002)	(0.003)	(0.003)
Hard quality signals	0.963	0.943	0.980
	(0.009)	(0.015)	(0.014)

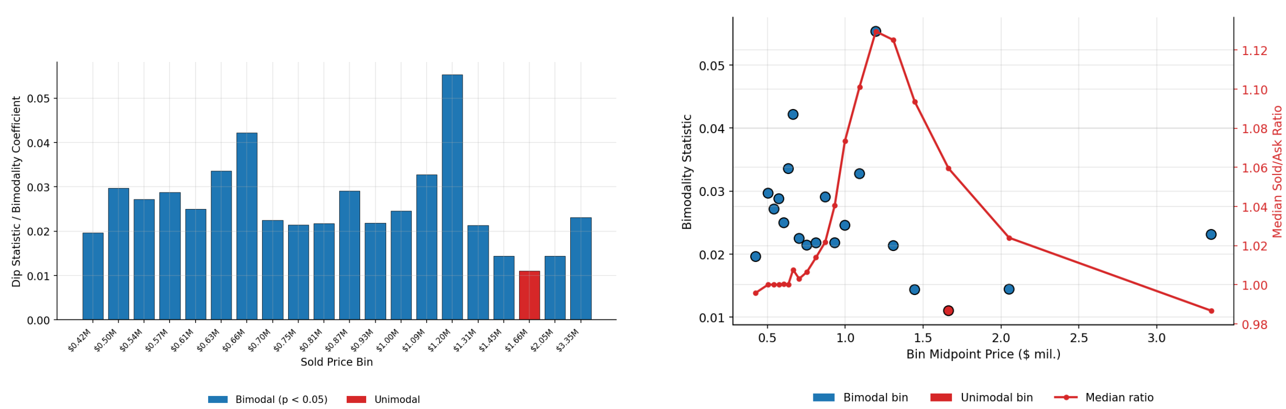
Note: Baseline outcome groups: failed and patient sellers require DOM in Q4; successful auctioneers require DOM in Q1/Q2, with $\widehat{\text{Resid}}_{\text{ask}} < 0$ for auctioneers and $\widehat{\text{Resid}}_{\text{ask}} > 0$ for patient sellers (left panel of Figure **??**). Coordination signals count mentions of offer-date language, hold-offers clauses, and virtual tours per listing. Hard quality signals count mentions of premium finishes (stainless steel, renovated, granite, hardwood, quartz, marble, pot lights, crown moulding). Failed vs. Successful Auctioneers are statistically indistinguishable on both signal dimensions: Welch $t = 0.020$ ($p = 0.98$) for coordination signals; $t = -1.141$ ($p = 0.25$) for hard quality signals.

Figure D14: Sale-to-Ask Distribution by Seller Type, \$1.1M–\$1.5M Range



Note: Kernel density estimates of the sale-to-ask ratio for each seller type, restricted to the \$1.1M–\$1.5M sold-price range (N = 7,181). The vertical dotted line marks sold = ask. GenericListing's distribution is shifted left and more diffuse; all other types are concentrated above 1.0. Summary statistics and failure rates by seller type are reported in Table **??**.

Figure D15: Bimodality Strength by Sold-Price Bin (Hartigan's Dip Test)



(a) Dip statistic per bin; blue bars are bimodal ($p <$ 0.05); the single red bar fails to reject unimodality.

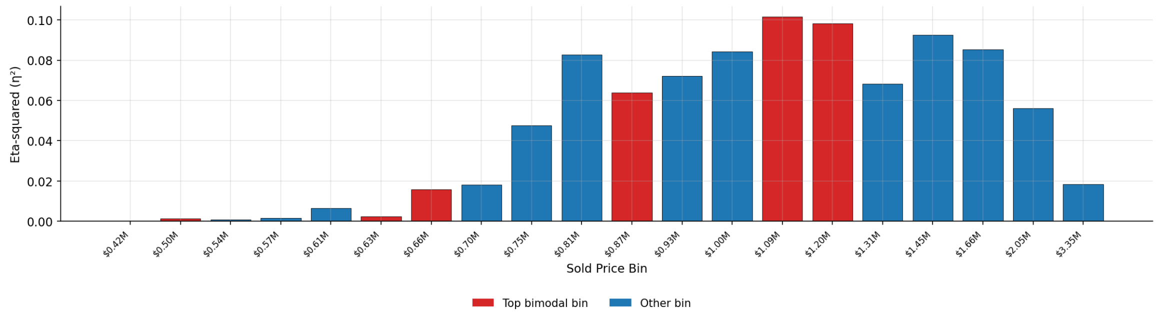
(b) Dip statistic (left axis) against median sold-to-ask ratio (right axis) across bin midpoints.

Note: Bimodality peaks in the \$0.9M–\$1.3M range, coinciding with the peak of the aggregate hump in Figure **??**.

D.9 Additional Regression Tables

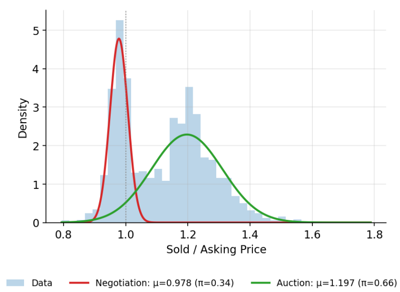
Table **??** in the main text reports how the return to underpricing varies by seller type through the interaction coefficients $\hat{\delta}_k$ in Equation (**??**). That specification shows that strategic con-

Figure D16: Property-Type Variance Explained (η^2) per Sold-Price Bin

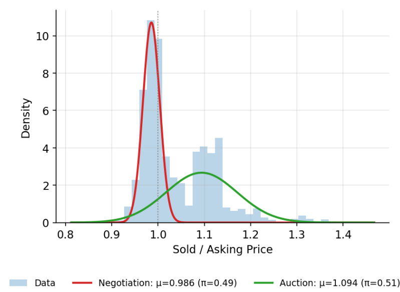


Note: Each bar is the one-way ANOVA η^2 of sold-to-ask ratio on property type (Detached vs. Townhouse vs. Condo/Apt) within a sold-price bin. Red bars mark the top bimodal bins (from Figure **??**). Low η^2 in red bins confirms that the bimodality is not driven by property-type composition.

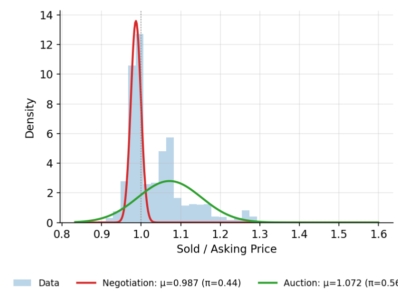
Figure D17: GMM Two-Component Fits — Peak Bimodal Bins



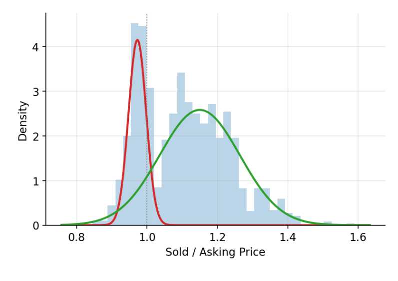
(a) Bin: $1.20M



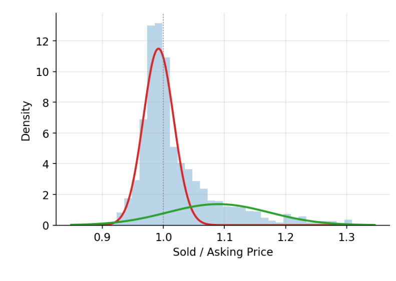
(b) Bin: $0.66M



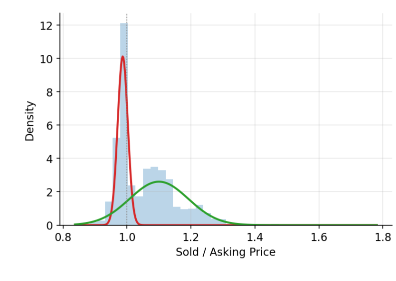
(c) Bin: $0.63M



(d) Bin: $1.09M



(e) Bin: $0.50M

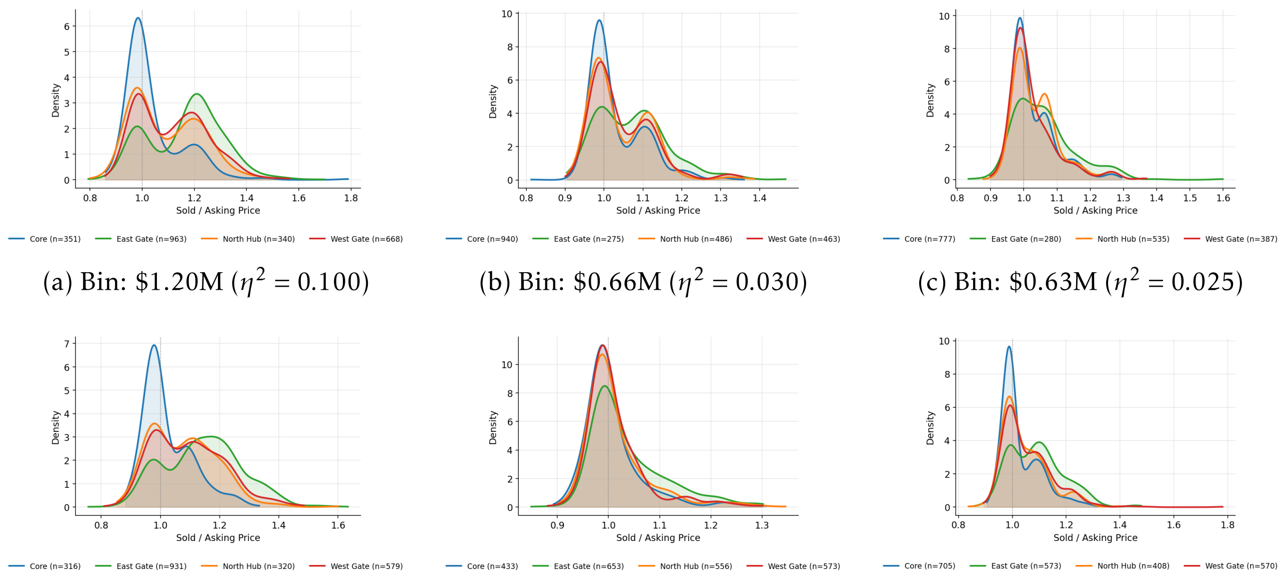


(f) Bin: $0.87M

Note: Each panel shows the histogram of sold-to-ask ratios for one of the six most bimodal bins, with the fitted GMM components: Negotiation Regime (red, lower mode) and Auction Regime (green, upper mode). Component means (μ) and mixing weights (π) are reported in each panel legend.

tent enters through underpricing depth rather than through seller-type levels alone. Table **??** supplements this result by reporting the corresponding direct seller-type effects $\hat{\beta}_k$—the predicted sold-to-ask ratio for each type when underpricing is set to zero, holding all controls fixed. Together, the two tables separate whether seller type shifts average pricing

Figure D18: Sold-to-Ask KDE by Zone — Peak Bimodal Bins



(a) Bin: \$1.20M ($\eta^2 = 0.100$)

(b) Bin: \$0.66M ($\eta^2 = 0.030$)

(c) Bin: \$0.63M ($\eta^2 = 0.025$)

(d) Bin: \$1.09M ($\eta^2 = 0.112$)

(e) Bin: \$0.50M ($\eta^2 = 0.030$)

(f) Bin: \$0.87M ($\eta^2 = 0.069$)

Note: Kernel density of sold-to-ask ratio by Toronto zone (Core, North Hub, East Gate, West Gate) within each of the six most bimodal sold-price bins. The η^2 in each subcaption is the variance explained by zone via one-way ANOVA. The two-mode structure persists within every zone, confirming that bimodality is not a geographic artifact.

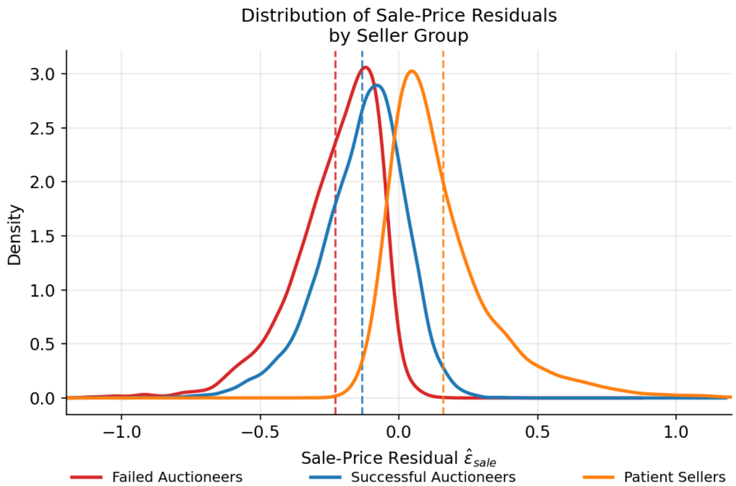
Table D4: **Signal Intensity and Seller-Type Composition by Outcome Group (Alternative Definition)**

	Successful Auctioneers	Failed Auctioneers	Patient Sellers
N	12,496	9,941	10,422
% Generic Listing	23.3%	30.0%	29.6%
Coordination signals	0.031	0.033	0.035
	(0.002)	(0.002)	(0.002)
Hard quality signals	0.963	0.959	0.984
	(0.009)	(0.010)	(0.010)

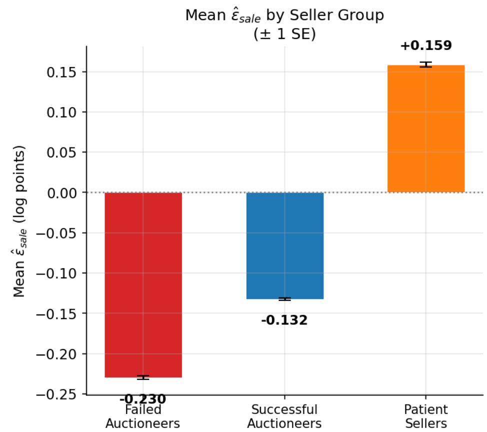
Note: This table repeats Table **??** under the alternative outcome-group definition in the right panel of Figure **??**: failed and patient sellers are classified using DOM in Q3 or Q4 rather than Q4 only; successful auctioneers remain DOM Q1/Q2 with $\widehat{\text{Resid}}_{\text{ask}} < 0$. The qualitative conclusion is unchanged: failed and successful auctioneers remain statistically indistinguishable on coordination and hard-quality signals. Coordination signals count mentions of offer-date language, hold-offers clauses, and virtual tours per listing. Hard quality signals count mentions of premium finishes (stainless steel, renovated, granite, hardwood, quartz, marble, pot lights, crown moulding). Failed vs. Successful Auctioneers remain statistically indistinguishable on both signal dimensions: Welch $t = 0.719$ ($p = 0.47$) for coordination signals; $t = -0.244$ ($p = 0.81$) for hard quality signals.

outcomes on its own from whether it changes the payoff to underpricing.

Figure D19: **Sale-Price Residuals by Seller Group**



(a) Distribution of $\hat{\epsilon}_{sale}$ for Each Seller Group



(b) Group means

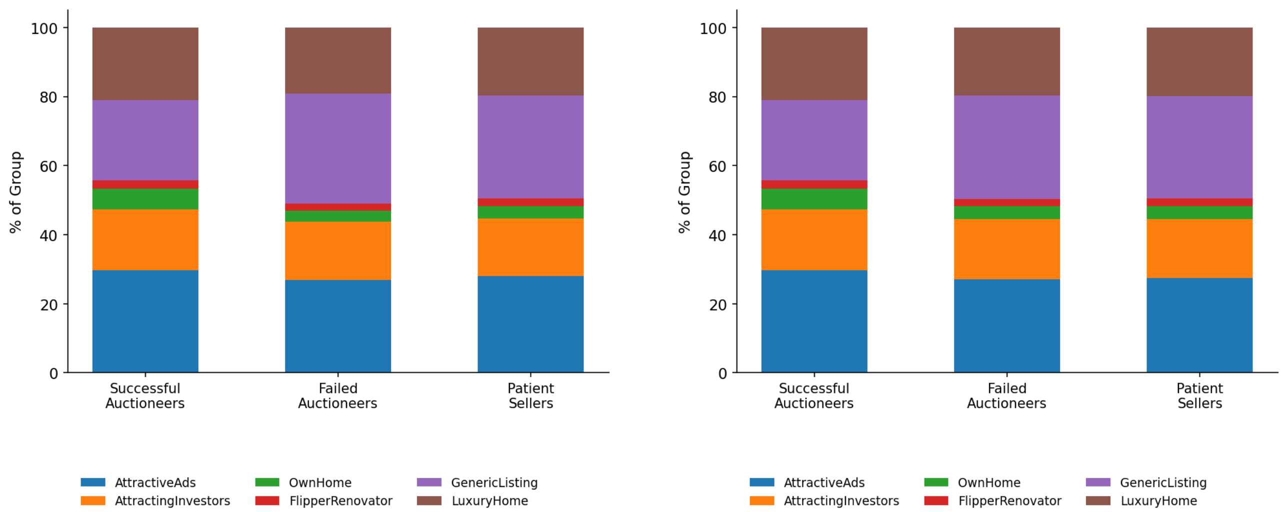
Note: Sale-price hedonic residual $\hat{\epsilon}_{sale} = \ln(\text{SoldPrice}) - \ln(\widehat{\text{SoldPrice}})$ from OLS with full structural, amenity, and fixed-effect controls ($R^2 = 0.768$). Failed Auctioneers: DOM Q4 and $\widehat{\text{Resid}}_{\text{ask}} < 0$; Successful Auctioneers: DOM Q1/Q2 and $\widehat{\text{Resid}}_{\text{ask}} < 0$; Patient Sellers: DOM Q4 and $\widehat{\text{Resid}}_{\text{ask}} > 0$. Panel (a) plots kernel densities; dashed vertical lines mark group means. Panel (b) reports means with ±1 SE error bars: Failed −0.230 (N = 4,677); Successful −0.132 (N = 12,496); Patient +0.159 (N = 5,640). All pairwise Welch t-tests in panel (a) are significant at $p < 10^{-200}$.

Table D5: **Direct Effects of Seller Type on Sold-to-Ask Ratio**

	Direct effect ($\hat{\beta}_k$)	
Seller type	Coef.	SE
GenericListing	(baseline)	
AttractiveAds	−0.0011	(0.0012)
LuxuryHome	−0.0012	(0.0013)
AttractingInvestors	0.0020	(0.0013)
OwnHome	−0.0046**	(0.0028)
FlipperRenovator	0.0017	(0.0033)

Note: OLS estimates of Equation (**??**) in the main text. Each $\hat{\beta}_k$ is the unconditional seller-type fixed effect on the sold-to-ask ratio relative to GenericListing, holding underpricing depth at zero. None of the direct effects are individually significant at conventional levels except OwnHome, consistent with the view that seller type alone does not predict outcomes; the strategic content enters through the interaction with underpricing (Table **??**). Standard errors (in parentheses) are HC3 heteroskedasticity-consistent (**MACKINNON1985305**). N = 43,273. $^{**}p < 0.01$.

Figure D20: **Seller-Type Composition by Outcome Group**



(a) Baseline definition

(b) Alternative definition

Note: Share of each dominant seller type within Successful Auctioneer, Failed Auctioneer, and Patient Seller groups. Both panels hold the asking-price-residual thresholds fixed: failed and successful auctioneers have $\widehat{\text{Resid}}_{\text{ask}} < 0$; patient sellers have $\widehat{\text{Resid}}_{\text{ask}} > 0$. Successful auctioneers are always defined by DOM in Q1 or Q2. The baseline panel (left) classifies failed and patient sellers using DOM in Q4 only; the alternative panel (right) classifies failed and patient sellers using DOM in Q3 or Q4, enlarging the slow-sale groups to test robustness. Chi-square test (Failed vs. Successful), baseline: $\chi^2 = 168.26$, $df = 5$, $p = 1.71 \times 10^{-34}$.