Match Quality and House Price Dispersion: Evidence from Norwegian Housing Auctions*

André K. Anundsen
Housing Lab, Oslo Metropolitan University

Arne Lyshol
BI Norwegian Business School

Plamen T. Nenov
Norges Bank

Erling Røed Larsen
Housing Lab, Oslo Metropolitan University
and BI Norwegian Business School

Abstract

Assessing the quantitative relevance of match quality and search frictions for house price dispersion is important for understanding house price formation and the importance of uninsurable housing wealth shocks. In this paper, we use a unique auction-level data set combined with a structural model of the housing transaction process that includes frictional arrival and endogenous entry of buyers into bidding, as well as information frictions between buyers and sellers to determine the importance of buyer taste heterogeneity for house prices and price dispersion. We find that quality differences matter greatly for house price dispersion, well beyond what "standard" hedonic pricing models may suggest, while buyer taste heterogeneity accounts for essentially all remaining price dispersion. Our structural model implies that list prices often deviate substantially from seller valuations, lending support to theories of list price determination that feature strategic interactions between sellers/agents and buyers.

Keywords: structural estimation, search frictions, buyer-specific valuations, bidding wars, transaction costs, list price, selection effects, informational mismatch

JEL classification: R21, R31

*We want to thank our discussant Anthony Lee Zhang, as well as Alina Arefeva, Mike Elsby, Kasper Roszbach and seminar and conference participants at BI Norwegian Business School, Norges Bank, and the 2021 ASSA meetings for valuable comments and suggestions.

1 Introduction

Housing markets are characterized by substantial price dispersion even among properties that share common features, such as location, type, size, and number of rooms. One possible driver of this residual price dispersion is quality differences, which remain unobserved to researchers, but which are observable to owners and potential buyers and are reflected in sale prices. Nevertheless, even when one is able to better account for such quality differences, using unit fixed effect approaches, there is still substantial price variation that remains unexplained. This unexplained price variation suggests other possible drivers. One such alternative is the frictional process of searching and matching between properties and buyers with heterogeneous tastes. This is a very natural candidate since housing transactions are characterized by substantial trading delays and informational frictions.[1]

Understanding the importance of frictional matching for house prices and price dispersion goes beyond understanding the workings of housing markets. Since housing wealth tends to comprise a substantial share of household wealth, the nature of house price dispersion may determine whether home owners are exposed to potentially large uninsurable (housing) wealth shocks. Indeed, if frictional matching plays an important role in driving house price dispersion, then a buyer who has been lucky to match well with her current home (and has paid a relatively high price for it) might later end up as the unlucky seller of that same property when she has to either accept a lower price from a buyer that does not value the property as much or continue searching and possibly facing costly delays.

In this paper we use a unique auction-level data set for Norway combined with a structural model of the housing transaction process to determine the contribution of match quality due to buyer taste heterogeneity and frictional matching to house prices and price dispersion. We find that buyer taste heterogeneity can account for essentially all of the price dispersion that is not accounted for by quality difference. Nevertheless quality differences play the main role in driving house price dispersion, even well beyond what hedonic pricing models suggest. Moreover, our structural model implies that there is an important *causal* effect of the number of bidders on house prices and that list (or asking) prices often deviate substantially from seller valuations.

Our data comprise detailed bidding-logs and information on viewers from all sales handled by two of the largest realtors in Norway for the period 2010–2015. The data set includes a unique bidder identifier, which allows us to compute the number of bidders associated with each transaction but also to follow the bidding behavior of each unique bidder in the auction. In addition, we have information on the list and sale dates, the list price and sale price, as well as standard hedonics. In Norway, all bids and acceptances of

[1] There is a large literature in labor economics which argues that search and matching frictions are important drivers of residual wage dispersion in labor markets (e.g. Postel-Vinay and Robin (2002)).

bids are legally binding. This allows us to identify one specific date upon which transfer of ownership is essentially determined.

Our structural model of the housing transaction process incorporates some features of the Norwegian institutional context, but is also applicable to housing transactions in other countries with institutional arrangements that may facilitate an auction-like format of transactions in some cases. It features an entry game, whereby a number of potential buyers (drawn randomly according to a distribution of viewers) observe the quality of a property and determine how well the attributes are aligned with their idiosyncratic tastes. In addition, they observe a signal of the seller's reservation value in the form of the list price. Importantly, we allow for some decoupling between the seller reservation value and the list price in the form of a systematic mean bias of the list price from the seller reservation value, as well as some dispersion between the two. Therefore, list prices in our model are a biased noisy signal of the seller reservation value. This property of the list price incorporates a number of theories of list price formation proposed in the literature.

Upon observing this information about the property and the seller, buyers endogenously choose whether to enter a bidding stage by paying a specific transaction cost or to walk away. Importantly, the transaction cost does not scale with the quality of the property, so that a potential buyer is more likely to enter bidding if the property has higher quality. In addition, buyers are more likely to bid if they have sufficiently higher idiosyncratic taste, or because the list price is lower. Since buyer entry decisions are strategic substitutes, a larger number of potential buyers (i.e. viewers) leads to a lower probability of buyer entry, other things equal.

At the bidding stage buyers observe the true seller reservation value. Due to imperfect information at the entry stage, some buyers may actually have a lower valuation than the seller, a situation we call *informational mismatch*. Indeed, it can be the case that there are multiple buyers entering the bidding stage but the property does not end up transacting in the end. Depending on the number of buyers who have chosen to enter and who have a higher valuation than the seller reservation value, the transaction price is either determined in a reduced-form negotiation process in the case of a single buyer or through a second-price auction.[2] These features of the model lead to rich selection effects due to the interaction of quality with buyer and seller heterogeneity and the underlying frictions, justifying a fully structural estimation.

We estimate our model using Simulated Method of Moments (SMM). Since our data span a large number of different locations, types of housing, and time periods, we define a number of segments based on property type (Small vs. Large apartments vs. Houses) and also time period (each year in our sample). Given the need for segmentation and given data coverage, we focus attention on segments in Oslo, the capital city of Norway. For

[2]If no buyers choose to enter the bidding stage, then the property ends up not selling.

each segment, we construct moments based on an artificial data set of sales, in which we pre-calibrate the viewer distribution to the empirical viewer distribution for that segment in our data set. For our estimation, we let the mean and standard deviation of quality vary by segment, while the dispersion in buyer idiosyncratic tastes, the mean and dispersion of seller reservation values, the mean and dispersion of the list price bias or "wedge" from the seller reservation value, and the transaction cost are time-invariant. All parameters vary by housing type. Our identification of the dispersion in buyer-specific tastes relies mainly on information contained in several key moments, including the price premium, defined as the log difference between the sale price and list price, the probability of sale, and the probability of a contested auction with 5 or more bidders bidding for a property.

The model matches the targeted moments very well, despite the substantial over-identification in our baseline estimation (there are in total 54 moments against 18 parameters for each segment). It also performs well against a number of non-targeted moments, including direct measures of buyer and seller valuations. Specifically, we use the detailed bidding information from our data set to construct a lower and upper bound on the valuations of bidders who participate in auctions (i.e. there are at least two bidders) and who do not end up submitting the highest final bid, i.e. they "lose" the auction. Similarly, we compute a proxy for the seller reservation value by considering bidding logs where the seller makes a counter-bid.

We estimate a dispersion parameter for buyer-specific tastes of between 0.026 and 0.031, depending on the type of housing, which corresponds to an inter-quartile range of between 0.04 and 0.048.[3] In contrast, we find that the dispersion in seller reservation values is substantially larger for Houses compared to Apartments. On the other hand, there is substantial dispersion between list prices and seller reservation values in the case of Apartments with an additional small negative mean bias in list prices, which lends support to theories of list price determination that feature strategic elements.

Our estimated model parameters imply a much more important role of quality differences for price dispersion than suggested by "standard" hedonic pricing models. Indeed, fully accounting for quality differences in our model implies an R^2 statistic 98 to 99%. In contrast, the typical hedonic regression tends to have an R^2 of around 80 to 90%. Therefore, there much of the "residual" house price variation from hedonic pricing models is likely due to unobservable quality differences. However, we find that the remaining variation in prices can be almost fully accounted for by match quality (determined by the dispersion in idiosyncratic buyer tastes). Specifically, our estimated model implies that match quality accounts for around 2.5 percentage points of the observed house price dispersion in Oslo during 2010-2015. Moreover, match quality also impacts the observed average price level, contributing around 4 to 5 percentage points to the observed average prices in Oslo during 2010-2015. Beyond match quality, the dispersion in seller reser-

[3]We assume that buyer idiosyncratic tastes follow a type I extreme value distribution.

vation values affects the observed house price dispersion for Houses with a much more limited effect for Apartments.

We also show by estimating reduced-form regressions of the effect of number of bidders on prices using simulated transaction data from our estimated model that there is an important *causal* effect of number of bidders on prices even after fully controlling for quality. Since bidder entry correlates positively with quality in our model, other things equal, not accounting for quality leads to a substantial bias in the estimated effects of bidders on prices. Whenever we fully account for quality, we recover a causal effect of one more bidder of around 0.01 log points. This is surprisingly close to reduced-form estimates in the literature, which provides one more independent validation for our structural model.

Related Literature

Our paper bridges two large strands of literature. First, our focus on understanding the contribution of match quality to house price dispersion and the assumption of stochastic arrival of (potential) buyers relate our paper to a growing literature on search frictions in housing markets (Wheaton (1990), Krainer (2001), Novy-Marx (2009), Caplin and Leahy (2011), Genesove and Han (2012), Carrillo (2012), Anenberg and Bayer (2013), Diaz and Jerez (2013), Head et al. (2014), Ngai and Tenreyro (2014), Nenov et al. (2016), Guren (2018), Guren and McQuade (2019), Ngai and Sheedy (2019), Moen et al. (2019), Piazzesi et al. (2020), Grindaker et al. (2021), Kotova and Zhang (2021), Rekkas et al. (2021) among others).[4] Second, our explicit modeling of the transaction process, whether through bargaining or as a second-price auction combined with a costly entry decision by potential buyers brings our paper close to the literature on auctions versus negotiations in housing markets (Ashenfelter and Genesove (1992), Mayer (1995), Lusht (1996), Merlo and Ortalo-Magne (2004), Genesove and Hansen (2014), Han and Strange (2014), Chow et al. (2015), Genesove and Han (2016), Arefeva (2020)).[5]

Our paper is most closely related to recent work by Genesove and Han (2016) and Rekkas et al. (2021). Genesove and Han (2016) use survey data on number of bidders combined with data on sale and list prices and a set of standard hedonics to estimate the dispersion in buyer-specific valuations using a semi-structural approach. The key moment the authors use for identification is the (reduced-form) coefficient of number of bidders on (log) prices, which in a static auction setting with two or more bidders and independent private values maps into the dispersion in buyer-specific tastes.[6] We differ from and complement this important paper in a number of ways. First, we rely on a fully-specified

[4]See Han and Strange (2015) for a review of this literature.

[5]See also McAfee and McMillan (1987), Levin and Smith (1994), and Bulow and Klemperer (1996, 2009) for theoretical treatments of auctions versus negotiations in the presence of entry costs.

[6]The authors report dispersion estimates for a number of underlying distributions for buyer-specific tastes. For a type 1 extreme value distribution they report an inter-quartile range of around 0.09. Our estimates suggest a lower interquartile range of between 0.04 and 0.048 depending on the type of housing.

structural model of the transaction process, which can account for selection effects on the set of transacted properties, due to the interaction of quality differences as well as seller heterogeneity and information frictions with the endogenous buyer entry into bidding. Second, our identification of buyer-specific taste dispersion relies on a different set of moments, while we use the reduced-form coefficient of number of bidders on prices to validate our estimated model. This is important for consistently identifying this key parameter, since as we point out in Section 5.3.3, the reduced-form coefficient of number of bidders on prices appears to be very sensitive to fully accounting for quality differences across objects.

In independent and contemporaneous work Rekkas et al. (2021) study a rich structural housing search model, which they estimate using housing transaction data from Vancouver. Some of the features of their model are qualitatively similar to ours. For example, their model features price posting and directed search while in our model list prices carry information about seller reservation values and thus affect the entry decisions of potential buyers. There are also important differences, however. For example, their model is dynamic, which allows them to explain the dynamics of list prices and why list prices appear sticky over time. On the other hand, their model abstracts away from the possibility of auctioned sales or informational mismatch whereby there is no trade despite entry of bidders. These differences imply that the two models are complementary and can be applied to housing markets with different institutional settings. Consequently, the data used to discipline the two models are also different, with Rekkas et al. (2021) relying on transaction-level data, while we exploit a unique auction-level data set with full bidding logs. Still, despite these differences, the two papers reach a similar conclusion, namely that quality differences (or property heterogeneity) play the main role in explaining house price dispersion with search frictions/match quality having secondary effects.

A number of other papers include match quality in quantitative models of the housing market with frictional search and matching. For example, Ngai and Tenreyro (2014), Guren and McQuade (2019), and Ngai and Sheedy (2019) calibrate distribution parameters for aggregate match-specific quality. Carrillo (2012) and Guren (2018) separate buyer idiosyncratic match quality from aggregate match quality, and estimate and calibrate parameters, respectively. However, unlike us they focus on other details of the transaction process, abstracting away from the possibility of an auctioned sale as well as the link between the list price and the seller reservation value. In addition, we use a much more detailed bidding level data to estimate the parameters in our model. Kotova and Zhang (2021) study empirically and theoretically the link between time-on-market and residual price dispersion. Their main insight is that shocks to liquidity supply, such as the flow of buyers relative to sellers, move time-on-market and residual price dispersion in the same direction, while shocks to liquidity demand, such as the average holding cost of sellers, move those two quantities in opposite directions. Our findings that both buyer and seller preference heterogeneity matters for (residual) price dispersion is consistent with their findings. Where we differ is in our explicit focus on modeling the details of the transaction mechanism with auctions with multiple bidders being important in the price determination process, while in their framework prices are solely determined through bargaining. Moreover, we are not explicitly interested in the determinants of market liquidity and instead focus attention on understanding the drivers of price dispersion through a structural estimation of a fully-specified structural model.[7]

Finally, our study is related to a growing literature on how list prices are set. This literature shows that the list price does not necessarily equal the seller reservation value (Horowitz, 1992; Taylor, 1999), although list prices often are equal to sale prices, a finding that Han and Strange (2016) argue demonstrates the importance of the list price for providing information to buyers about seller valuations.[8] Conceptually, deviations of list prices from (revealed) seller reservation values may be consistent either with some strategic considerations involved in setting the list price or with seller dynamic learning effects. Regarding the former theory, a growing literature suggests that strategic elements are involved in setting the list price. According to that literature, sellers seek to achieve the highest price by balancing a herding effect (Ku et al., 2006, Simonsohn and Ariely, 2008)) and an anchoring effect (Beggs and Graddy, 2009, Northcraft and Neale, 1987, Bucchianeri and Minson, 2013). Anundsen et al. (2020) find that in Norway some sellers set a list price that is, in fact, lower than the object's appraisal value and they show that a certain type of realtors tends to be associated with setting such a list price. They find that a low list price reduces the sale price due to the anchoring effect. In contrast, Repetto and Solis (2019) find that a house that is listed with a list price just below a round million (Swedish kronor) achieves a sale price that is three to five percent higher than otherwise expected. Our findings of a negative mean (log) wedge between seller reservation values and list prices, as well as a positive dispersion in that wedge are consistent with strategic considerations playing a role in setting the list price.

Turning to the learning effects theory, in an already classic paper, Merlo and Ortalo-Magne (2004) show that list price reductions are rare, but when they happen, they tend to be large and triggered by few bids. Merlo et al. (2015) argue that list prices are sticky, but that a model with small menu costs can account for the phenomenon. Related to that, Anenberg (2016) argues that seller uncertainty about prices and learning is important for housing market dynamics. Herrin et al. (2004) find that in thin markets

[7]Nenov et al. (2016) conjecture that differences in the dispersion of buyer-specific tastes for different housing types may be an important driver of transaction seasonality differences across the housing types. Our parameter estimates by housing type are consistent with this conjecture.

[8]There are many determinants of the list price, e.g. the market situation (Haurin et al., 2013) and loss aversion (Genesove and Mayer, 2001, Liu and van der Vlist, 2019). Han and Strange (2015) review the literature on the microstructure of the housing market and the role played by the list price.

well-informed sellers are less likely to change the list price.[9] Even though we do not explicitly include seller learning effects (for example, through the observed flow of bidders during the duration of the auction), our findings of a positive dispersion in the wedge between the list price and seller reservation value is consistent with a learning theory. The negative bias in the wedge is a bit harder to rationalize via a learning theory, however, since learning should lead to both upward and downward revisions in seller valuations.

The rest of the paper is organized as follows. We start by describing the institutional details of the Norwegian housing market, before proceeding to lay out the structural model in Section 3. Section 4 presents the data and details the estimation approach. Section 5 presents and discusses the results and counterfactuals. Finally, Section 6 provides brief concluding comments.

2 Institutional setting

In this section we provide a brief overview of the institutional background of the Norwegian housing market. A more detailed description can be found in Anundsen et al. (2020).

When a seller decides to sell her house in Norway, she typically contacts a realtor or several realtors. When the seller and the realtor reach an agreement in which the seller becomes the realtor's client, they initiate talks on when to post the advertisement, what list price to announce, and dates for public showing(s) (open houses). The realtor charges a commission that amounts to around 2% of the sell price. It is not common to hire a buyer-agent. In fact, even though this practice can be found in other countries, buyer-agents are almost non-existent in Norway. The realtor is required by law to take into account the interest of both the seller and the buyer, and is obliged to give advice to both seller and buyer on issues that may impact the selling process. There is detailed regulation of what is required of the realtors and the information they need to furnish the buyer with. However, the realtor is hired by and paid for by the seller, so it is in the best interest of the realtor to make sure the client obtains a satisfactory sale price within any time constraints imposed by the seller.

Together with the realtor, the seller decides on a list price. Having decided on the list price, the realtor lists the house for sale, typically using the nationwide online service Finn.no and national and local newspapers. The advertisement states the time of potential public showings. It is common to have one or two showings. The realtor, or a realtor assistant, is always present at showings, and the realtor answers questions to the

[9] Anundsen et al. (2020) also present evidence consistent with seller learning, albeit sellers seem to do so only modestly. Beyond housing, studying eBay data, Einav et al. (2015) find patterns that indicate that sellers experiment and learn what strategies to use, but also that sellers of similar items use different strategies.

potential buyers who inspect the unit. The public showing typically lasts for one hour. Most often, the seller is not present.

The auction is arranged as an ascending bid English auction. Bids are placed by telephone, fax, or electronic submission using digital platforms. The realtor informs the participants (both active and inactive) of developments in the auction. Each and every bid is legally binding and each and every acceptance of a bid is legally binding. When bidders make their first bid, they typically submit proof of financing.[10] The seller may decline any bid. If the seller declines a bid above the posted list price, and announces a new showing without adjusting the list price, the realtor risks being reported to the authorities for knowingly mispricing the unit.[11] The implication is that even though the list price is not a reservation price, and thus allows the use of a strategic element, most realtors and sellers shun illegitimately low list prices. Therefore, the list price contains useful information about the seller's reservation price.

When the auction is completed, every participant in the auction is entitled to see the bidding log, which provides an overview of all the bids that were placed during the auction. If no sale takes place, the seller advertises new showings. The time-on-market (TOM) in Norway is typically low, and in the capital, Oslo, it is often not more than three or four weeks. However, when the unit stays on the market for a longer time, the sales process tends to transform from an auction type to a negotiation between the seller and prospective buyers. The low TOM in Oslo is also associated with relatively quick bidder entry. Table 1 shows the distribution of the difference in entry times between the first and last bidder (based on the timing of their first bid) for Oslo rounded to the nearest hour.[12]. The median difference in entry times is 2 to 3 hours depending on the type of housing. Even at the 70th percentile of transactions, the difference in entry times is less than 24 hours. Therefore, most housing auctions in Oslo tend to feature near-simultaneous entry by bidders.

3 Model

We set-up a stylized (partial) equilibrium model of buyer entry and simultaneous bidding. The model is motivated by the Norwegian institutional context, but is also applicable to housing transactions in other countries with institutional arrangements that may facilitate an auction-like format of transactions.

[10] This practice is cloaked in some technicalities since bidders do not want to inform the realtor of the maximum financing available to them.

[11] In practice, such reports are rare. There have been some cases with claims of specific realtors being associated with artificially low list prices.

[12] This distribution is computed for transactions with at least two bidders

Table 1: Distribution of difference in entry times between first and last bidder in Oslo, Norway (in hours).

Type of unit	10th pctl	20th pctl	30th pctl	40th pctl	50th pctl	60th pctl	70th pctl	80th pctl	90th pctl
All	0	1	1	2	3	12	18	42	119
Small apt.	0	1	1	2	3	13	20	45	124
Large apt.	0	0	1	2	2	7	17	40	139
House	0	1	1	2	2	4	15	24	93

Note: The table shows distribution of the difference between the entry times of the first and last bidder in rounded hours for transactions with at least two bidders in Oslo during 2010-2015. The entry time of a bidder is determined by the time of their first bid. Small apartments have up to 2 rooms, while large apartments have 3 or more rooms. Houses are defined as all remaining housing units, which consist of row and semi-detached houses, and single-family homes.

3.1 Basic set-up

There is a single housing unit ("the house") for sale owned by an agent ("the seller"), and a large pool of potential buyers ("the buyers"). All agents in the economy are risk-neutral and utility is transferable. Motivated by the Norwegian institutional context, we abstract away from dynamics, such as making several sale attempts, future re-sale possibilities, learning, and so on by modeling the housing sale as taking place in one period.[13]

The seller. We assume that the seller cares only about selling the house for a price at or above her reservation value and abstract away from the costs associated with the sale process and any actions by the seller, such as the setting of a list price or signaling of private information, which we instead model in reduced-form.[14]

The seller's reservation value is given by

$$\tilde{v}(\theta, e) = \exp\{\theta + e\}, \tag{1}$$

where θ denotes the quality of the house, and e is an idiosyncratic seller preference. For example, sellers may find their own house to be more or less valuable than the average buyer (respectively, $\bar{e}$ is positive or negative), because of a particular selection into who chooses to sell their house (for example, to move as in Ngai and Sheedy (2019)). In addition, if some sellers are relatively impatient to sell, because of a moving shock, for example, that will be reflected in a lower average reservation value for the sellers. On the other hand, e may reflect a dynamic option value of selling later on. For example, heterogeneity in buyer tastes, which we describe below, implies that seller reservation values should be positive on average ($\bar{e}$ is positive), much like in any standard search model (McCall, 1970). Finally, the dispersion in seller preferences may also be driven by seller uncertainty about the value of the house (as in Anenberg (2016)).

[13]Such dynamic considerations would be reflected in some of our estimated parameters, as we explain in Section 5.

[14]Also we abstract from any agency problems between the seller and realtor.

The list price is assumed to be an imperfect signal of the seller's reservation value that buyers observe prior to their decision to enter bidding (see below). Specifically, we assume that the list price is given by

$$a = \tilde{v}(\theta, e) \times \eta, \tag{2}$$

where $\ln \eta \sim N\left(\bar{\eta}, \sigma_{\eta}^{2}\right)$ reflects any differences between the list price and seller's reservation value. Therefore, we assume that there is asymmetric information between buyers and sellers regarding seller valuations. As discussed in the related literature, this opens up the possibility for signaling of reservation values and strategic behavior by the seller through list prices. We abstract from explicitly modeling these signaling incentives by directly modeling the "wedge" between the seller reservation value and the list price that would arise with an endogenous choice by the seller. Notice that in the context of a signaling game between the buyer and seller, $\overline{\eta}$ will reflect the mean signal distortion that sellers engage in as in signal-jamming models à la Holmström (1999). As in these models, here buyers would rationally expect such a distortion and would adjust the observed signal when making inferences about the seller reservation value.

Next, the variance σ_{η}^{2} will reflect the signal informativeness, which in a signaling model would be endogenously determined by the nature of equilibrium played, i.e. separating, semi-separating, or pooling. For example, if σ_{η}^{2} is relatively small, then list prices would be very informative about seller valuations, as would be the case in a separating equilibrium. Conversely, if σ_{η}^{2} is large, then list prices would be uninformative about seller valuations as would be the case in a pooling equilibrium.

The buyers. The number of potential buyers is given by, B_p.[15] Buyers have preferences over the house that are comprised of a common component and an idiosyncratic or match-specific component. The common component reflects quality differences between houses, such as location, type, size, and age, but also whether the house has been recently renovated, distance to local (dis)amenities, a good view, etc. The idiosyncratic component reflects horizontal differences that are buyer-house specific, such as relative proximity to a buyer's workplace, relative appeal of the house, the housing layout, lighting, and interior decorations, relative preferences for certain types of local (dis)amenities, etc. Therefore, we assume that a buyer's valuation is given by,

$$w\left(\theta, u_i\right) = \exp\left\{\theta + u_i\right\} = \tilde{v}(\theta, u_i), \tag{3}$$

where u_i denotes the idiosyncratic taste of buyer i. We assume that buyers observe θ perfectly, so there is no information asymmetries about the quality of the unit. Therefore,

[15]In our structural estimation, we will equate this number to the number of viewers that come to showings.

we abstract away from any inferences buyers may draw from the number of other bidders present or from winning the bidding (see below). While interesting, we view these issues as secondary to our main goal of understanding the role of match quality in house prices. Also, we assume that u_i is private information for each buyer. This assumption is realistic, since each buyer evaluates the available information about their idiosyncratic preferences privately.

We assume that in the population of houses, sellers, and buyers, θ, e, and the u_i's are distributed independently of each other. Moreover, θ, and e, follow normal distributions[16]

$$\theta \sim N\left(\bar{\theta}, \sigma_\theta^2\right), \tag{4}$$

and

$$e \sim N\left(\bar{e}, \sigma_e^2\right), \tag{5}$$

respectively. Finally, u_i is assumed to be distributed according to a Type I Extreme value distribution with tail probability $\Pr\{u_i > x\} = 1 - \exp\{-\exp\{-x/\sigma_u\}\}$. Therefore σ_u parameterizes the dispersion in buyer idiosyncratic tastes. For that reason we refer to σ_u as the "dispersion" in buyer-specific tastes below.

We define

$$\tilde{a} \equiv \ln a - \theta - \bar{\eta} = e + \ln \eta - \bar{\eta}, \tag{6}$$

so that

$$\tilde{a} \sim N\left(e, \sigma_\eta^2\right) \tag{7}$$

Therefore, given these distributional assumptions, the informational assumptions for the buyers, and the assumption that buyers are fully rational, it follows that the posterior belief about e for each buyer, given the list price a is

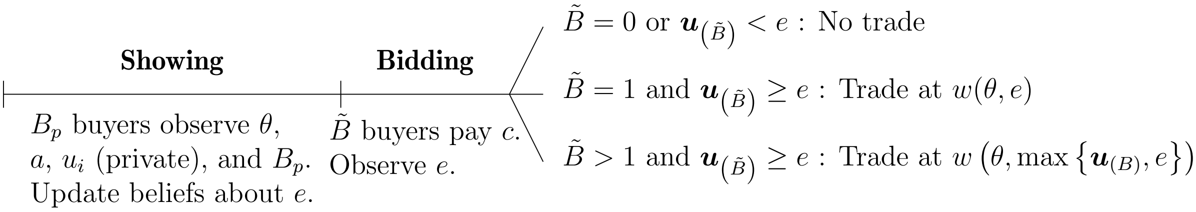
$$e|a \sim N\left(\frac{1/\sigma_e^2}{1/\sigma_e^2 + 1/\sigma_\eta^2}\bar{e} + \frac{1/\sigma_\eta^2}{1/\sigma_e^2 + 1/\sigma_\eta^2}\tilde{a}, \frac{1}{1/\sigma_e^2 + 1/\sigma_\eta^2}\right). \tag{8}$$

Price determination

After observing θ, a, and their u_i's, potential buyers choose simultaneously whether to enter a bidding stage. To enter that stage, a buyer must pay a cost of $c > 0$. If the buyer does not pay the entry cost she collects an outside option, which is normalized to 0. Therefore, the cost c will reflect both the true cost associated with the process

[16] Assuming a log-normal value for quality matches well the observed distributions of appraisals and sale prices in the data. The assumption of normally distributed seller valuation is made for tractability, since it simplifies the posterior distribution of e given the observed list price a.

Figure 1: Model time-line.



of buying a house, as well as an opportunity cost arising from the buyer's true outside option. After paying the cost c the buyer also learns the true value of e.

We let $\tilde{B}$ denote the number of buyers that enter the bidding stage. Below, it will also be useful to define $B = \tilde{B} - 1$ as the number of other buyers entering the bidding stage from the perspective of a buyer that has chosen to enter. We assume that a buyer who chooses to enter observes the realization of B perfectly, so that she knows how many opponents she is bidding against.[17]

If $\tilde{B} = 0$, then no transaction takes place. If $\tilde{B} = 1$ (so $B = 0$), there is no auction, and instead we assume that if the surplus from trading is positive, then the price is determined via a negotiation process. In particular, we assume that the buyer makes a take-it-or-leave-it offer to the seller, so the price equals the seller's reservation value.[18] If $\tilde{B} > 1$ (or $B > 0$), we assume that the price is determined in a second-price auction with a reserve price of $\tilde{v}(\theta, e)$.

Notice that given the *ex ante* information asymmetry, it is possible that even if $\tilde{B} > 0$, no trade takes place, since *ex post*, all buyers who enter the bidding stage may have lower valuation than the seller reservation value. This will be the case if $\boldsymbol{u}_{(\tilde{B})} < e$, where $\boldsymbol{u}_{(\tilde{B})}$ denotes the largest order statistic of $\boldsymbol{u} = (u_1, u_2, \ldots, u_{\tilde{B}})$. We can call this a situation of *informational mismatch*. Whenever there is no informational mismatch, and given the specific assumptions on the transaction process, we can write the transaction price as

$$p(\theta, \boldsymbol{u}, e) = \begin{cases} w(\theta, e) & , \tilde{B} = 1 \\ w\left(\theta, \max\left\{\boldsymbol{u}_{(B)}, e\right\}\right) & , \tilde{B} > 1 \end{cases}, \tag{9}$$

where $\boldsymbol{u}_{(B)}$ is the second (largest) order statistic of $\boldsymbol{u}$. Figure 1 presents a time-line of events in the model.

Buyer payoffs

Let $W(B, \theta, e, u_i)$ denote the expected payoff for buyer i who has entered the bidding stage, given that a total of B other buyers are present at that stage and given that

[17] This is without loss of generality given the second-price auction assumption and given that buyers are risk neutral in the price (see McAfee and McMillan (1987)).

[18] Allowing for price determination via Nash bargaining leads to structural estimates of the bargaining strength of the buyer close to one.

$u_i > e$. Let $\boldsymbol{u}_{-i} = (u_1, u_2, \ldots, u_{i-1}, u_{i+1}, \ldots, u_{\tilde{B}})$ denote the vector of buyer-specific valuations that exclude buyer i's valuation. Then, in the case when $B = 0$, we have,

$$W(0, \theta, e, u_i) = (w(\theta, u_i) - w(\theta, e)). \tag{10}$$

In the case when $B \geq 1$, we have

$$W(B, \theta, e, u_i) = \Pr\left\{(\boldsymbol{u}_{-i})_{(B)} < u_i\right\} E\left[w(\theta, u_i) - p(\theta, \boldsymbol{u}, e) \middle| (\boldsymbol{u}_{-i})_{(B)} < u_i\right], \tag{11}$$

or using equation (9),

$$\begin{aligned} W(B, \theta, e, u_i) &= \Pr\left\{(\boldsymbol{u}_{-i})_{(B)} < u_i\right\} \times \\ &\left\{w(\theta, u_i) - E\left[w\left(\theta, \max\left\{(\boldsymbol{u}_{-i})_{(B)}, e\right\}\right) \middle| (\boldsymbol{u}_{-i})_{(B)} < u_i\right]\right\}. \end{aligned} \tag{12}$$

It is straightforward to show that W is increasing in θ and u_i and decreasing in e, and B.

Lemma 1. *A buyer's expected payoff W in the bidding stage is increasing in θ and u_i and is decreasing in e, and B.*

Proof. See Appendix B. □

Buyer entry decisions

Since the idiosyncratic preferences u_i are private information, buyers do not know the exact number of other buyers that will enter the bidding stage at the time of their entry decision but instead (rationally) anticipate that this number will follow a certain (endogenously determined) distribution. Moreover, by Lemma 1, buyers will choose whether to enter the bidding stage according to a cutoff rule. To see this, note that a buyer with idiosyncratic taste u will choose to enter iff

$$E_B\left[\mathbb{1}_{\{u>e\}} W(B, \theta, e, u) \,|a\right] \geq c, \tag{13}$$

where, $E_B[.]$ denotes expectation with respect to the (endogenously determined) distribution of B.[19] Since W is increasing in u, it follows that the left-hand side of Eq. (13) is also increasing in u, so that there exists a unique idiosyncratic valuation threshold, denoted by $\hat{u}(\theta)$, which satisfies

$$E_B\left[\mathbb{1}_{\{\hat{u}(\theta,a)>e\}} W(B, \theta, e, \hat{u}(\theta, a)) \,|a\right] = c, \tag{14}$$

[19] We assume that a buyer who is indifferent between entering and not chooses to enter the bidding stage.

such that a buyer enters iff $u_i \geq \hat{u}(\theta, a)$. We can further simplify Eq. (14) using the observation that a buyer with the threshold valuation $u_i = \hat{u}(\theta, a)$ has the highest valuation among $B+1$ buyers (for $B \geq 1$) with zero probability. Put differently, a buyer with the marginal idiosyncratic valuation $\hat{u}(\theta, a)$ never expects to win an auction, and hence, always receives a payoff of zero in that contingency. The only contingency when the buyer receives a positive payoff (that would compensate for his entry cost c in expectation) is when there are no other buyers, i.e. $B = 0$ and $e < \hat{u}(\theta, a)$, so that there is a positive surplus from trading. Therefore, Eq. (14) becomes

$$E_B\left[\mathbb{1}_{\{\hat{u}(\theta,a)>e\}} W(B, \theta, \hat{u}(\theta))\,|a\right] = \Pr\{e < \hat{u}(\theta, a)\,|a\} \Pr\{B = 0|a\} \times \{w(\theta, \hat{u}(\theta, a)) - E[\tilde{v}(\theta, e)\,|a, e < \hat{u}(\theta, a)]\} = c. \tag{15}$$

Hence, the expected payoff for the marginal buyer who enters the bidding stage equals the probability that trade will take place multiplied by the probability that there will be a negotiated sale and the difference between the buyer's valuation and his expectation of the seller's valuation given the list price.

3.2 Buyer entry game

Given this set-up, we can define a symmetric (Bayesian) Nash equilibrium of the buyer entry game as follows.

Definition. *Given values of θ, a, and B_p, a symmetric (pure strategy) Bayesian Nash Equilibrium of the buyer entry game consists of a buyer entry decision $\chi(u) \in \{0, 1\}$ and a distribution of entering buyers, $\tilde{B}$, such that $\chi(u) = 1$ iff condition (13) is satisfied, and the distribution $\tilde{B}$ reflects the entry decision $\chi(u)$.*

Given that entry follows the cutoff rule from Eq. (14), it follows that prior to drawing the idiosyncratic preferences, the probability, q, of any given buyer entering the bidding stage is given by

$$q(\theta, a) = \Pr\left\{E_B\left[\mathbb{1}_{\{u_i>e\}} W(B, \theta, e, u_i)\,|a\right] \geq c\right\} = \Pr\{u_i \geq \hat{u}(\theta, a)\}, \tag{16}$$

with $\hat{u}(\theta, a)$ satisfying Eq. (14). Since idiosyncratic draws are i.i.d., the (ex ante) distribution of entering buyers, $\tilde{B}$, is therefore a Binomial distribution with parameters B_p and $q(\theta, a)$. Similarly, the distribution of *other* buyers entering the bidding stage, B, is Binomial with parameters $B_p - 1$ and $q(\theta, a)$. We summarize these observations in the following equilibrium characterization result.

Proposition 1. *Given values of θ, a, and B_p, there is a unique symmetric pure strategy Bayesian Nash Equilibrium of the buyer entry game characterized by:*

- *A cutoff valuation $\hat{u}(\theta, a)$ that satisfies Eq. (14), such that a buyer enters and bids iff $u \geq \hat{u}(\theta, a)$.*
- *An (ex ante) distribution of entering buyers which is Binomial with parameters B_p and $q(\theta, a)$, where q satisfies Eq. (16).*

Figure 2 illustrates the shape of the cutoff rule $\hat{u}$ in the top row (that is its dependence on θ and a) and its dependence on model parameters in the remaining panels. Not surprisingly, given Lemma 1, $\hat{u}$ is decreasing in θ. Intuitively, given a fixed entry cost for bidding, a house that is more valuable (irrespective of the buyer's own idiosyncratic valuation) raises the buyer's expected payoff from bidding for the house and thus induces entry by buyers with lower idiosyncratic valuations. Put differently, higher quality houses attract more bidders. The cutoff $\hat{u}$ is also increasing in a, so that a higher list price acts on a buyer's entry decision the same way as lowering the quality of the house. This is because a signals a higher seller valuation e to buyers. Raising the average seller valuation $\bar{e}$ has a similar effect on $\hat{u}$ (bottom, right panel). Also, unsurprisingly, a higher bidding cost, c, raises the cutoff value $\hat{u}$, since buyers must expect to have a higher payoff from bidding to counter the higher bidding cost. $\hat{u}$ is also increasing in the number of potential buyers, B_p. Intuitively, a higher value of B_p raises the expected number of bidders, $\tilde{B}$, which lowers the payoff to any single buyer from entering the bidding stage and raises the cutoff $\hat{u}$. This congestion effect is generally present in any entry game.

More interestingly, $\hat{u}$ is increasing in the dispersion of buyer idiosyncratic tastes. Intuitively, a higher dispersion of buyer tastes raises the probability of having an extremely high idiosyncratic valuation, which, in turn, raises the probability of any single buyer entering the bidding stage. This results in a higher number of expected buyers, similarly to the effect of B_p, thus, also raising the cutoff.

Increasing the dispersion of the list price has an ambiguous effect on $\hat{u}$, since it depends on the position of the list price. For a low list price that lets the buyer update towards a lower value of e than the mean, higher dispersion σ_η, which makes the list price less informative about the seller reservation value, implies that the buyer relies more on his prior mean (i.e. $\bar{e}$). Consequently, the buyer expects a higher reservation value, and the value of $\hat{u}$ increases. The opposite happens when the list price is high and the informativeness of the list price goes down, so that the buyer relies more on the prior mean to form expectations about e.

A higher dispersion of the seller reservation value has several effects. First, similar to the effect of changing σ_η it affects the inference of the buyer about e, making the buyer weight more the list price when inferring e. As with the effect of changing σ_η, the position of the list price would matter for the direction of this effect. Second, it makes more extreme values of e more likely by moving mass to the tails of the seller reservation value distribution. Depending on the position of $\hat{u}$ relative to a, this may increase or lower the expected payoff from entering the bidding stage. In practice, however, this effect is quantitatively small, as shown in the figure (bottom, left panel).

4 Data and Estimation

4.1 Data

We use bidding-logs and information on viewers and showings from all sales handled by two of the largest real estate agencies in Norway, DNB Eiendom and Krogsveen over the period 2010–2015. We consider the Oslo housing market, excluding units that belong to a housing co-operative ("co-op units").[20]

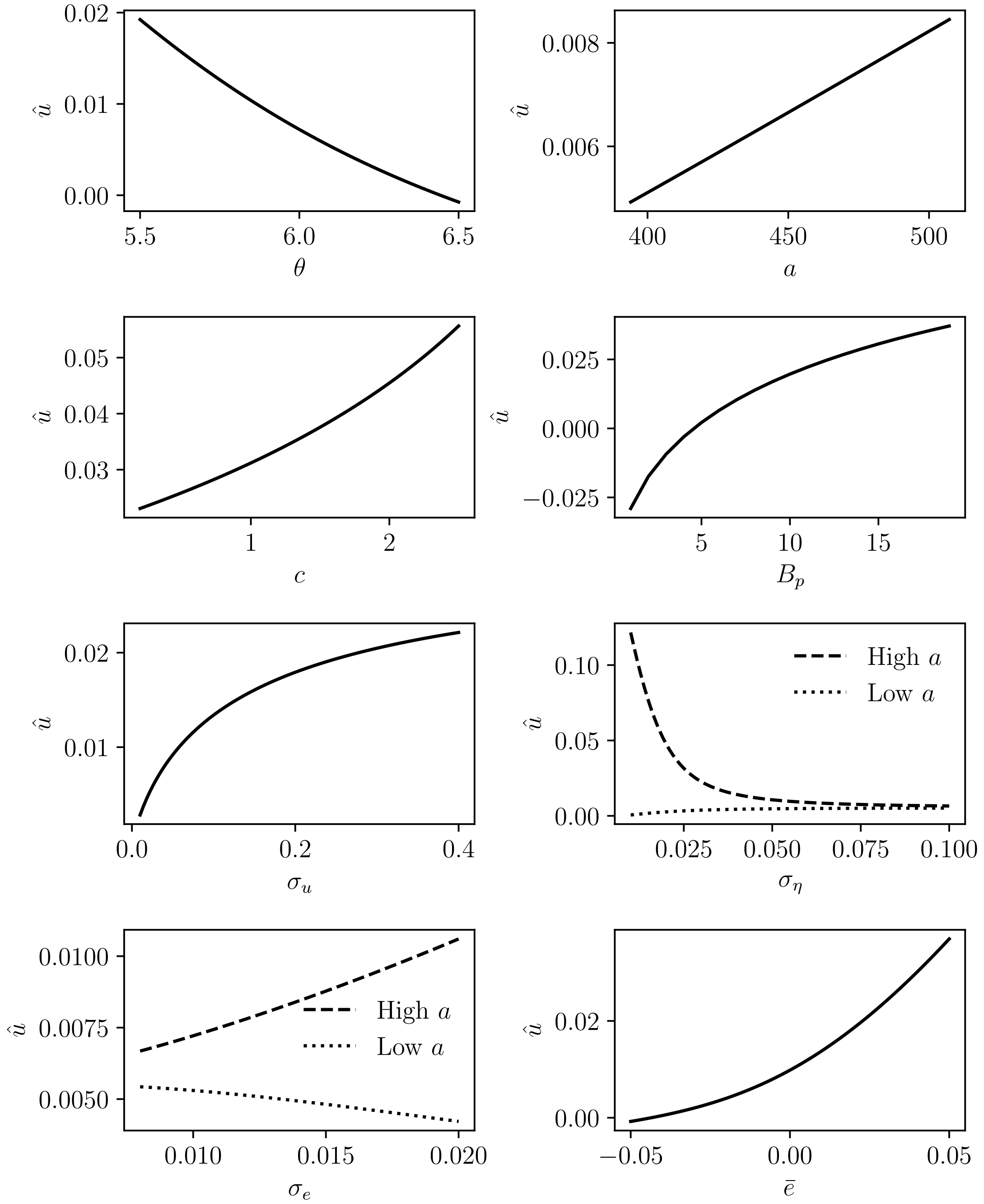
The data sets contain information on each bid placed in every housing sale handled by these real estate agencies. The data sets also include a unique bidder id, allowing us to compute the number of bidders in each auction. Additionally, both data sets contain information on the list price, the exact sales date, the exact date when the unit was listed for sale, attributes of the unit (size, address, unit type).

Before trimming, the data set contains 14,034 transactions and 133,224 bids. We start by dropping units that have missing information on sell price, list price, or size. We then truncate on the 1^{st} and 99^{th} percentiles of the sell, list, and size distributions. We also drop all sales of units that are transacted more than three times during our sample period, and units that have a negative time-on-market. If the bid-amount is missing, we drop this bid. We also drop bids that are lower than 80% of the list price, unless that is the winning bid. After this, we exclude some auctions with missing expiration dates and bids expiring before they are received. Finally, we remove auctions for which the distance (in days) between expiration of the previous bid and receiving a new bid is very long (99.5th pct) or short (0.5th pct), and we truncate on the 1^{st} and 99^{th} percentiles of the number of days elapsed between a realtor is hired and the unit is listed for sale. After trimming, we are left with 12,028 transactions and 116,111 bids. The number of observations exluded in each step and for each firm are described in Table A.1 in Appendix A.

For most of the analysis, we use data from both firms. However, the Krogsveen data set has information on the number of people showing up at the open house, as well as how many showings have been arranged before a sale takes place. That information is not

[20] Housing co-operatives were established in the post WW2 period to stimulate homeownership. The different properties in the co-op are owned by the housing co-operative, whereas members of the co-op buy shares which entitle them to occupy a specific unit in the co-op. The reason why we exclude co-ops from our analysis is that the selling process for co-op units differs from that of self-owned units, and thus from our model in Section 3. Specifically, co-ops always carry a "right of first refusal", which gives the option for members of the co-op (based on a seniority ranking) to purchase a unit at the highest accepted bid without taking part in the bidding. Bidders participating in an auction for a co-op units, therefore, also have to take into account the possibility that they may not buy the unit even if they placed the highest bid.

Figure 2: Comparative statics for $\hat{u}$.



Note: Parameterized with $\sigma_u = 0.0297$, $\sigma_e = 0.0140$, $\sigma_\eta = 0.0638$, $c = 0.1040$, $\bar{e} = -0.0101$, $\theta = 6.0332$, and $B_p = 6$. High (low) list prices, a, specified by increasing (decreasing) e by $9\sigma_e$.

contained in the DNB data, so we use the Krogsveen data for all calculations involving number of viewers and number of showings.

We use the bidding-level data to extract information on time-on-market, sell prices, list prices, and the spread between sell prices and list prices. In addition, we construct measures of the number of viewers at the showing, and number of bidders.

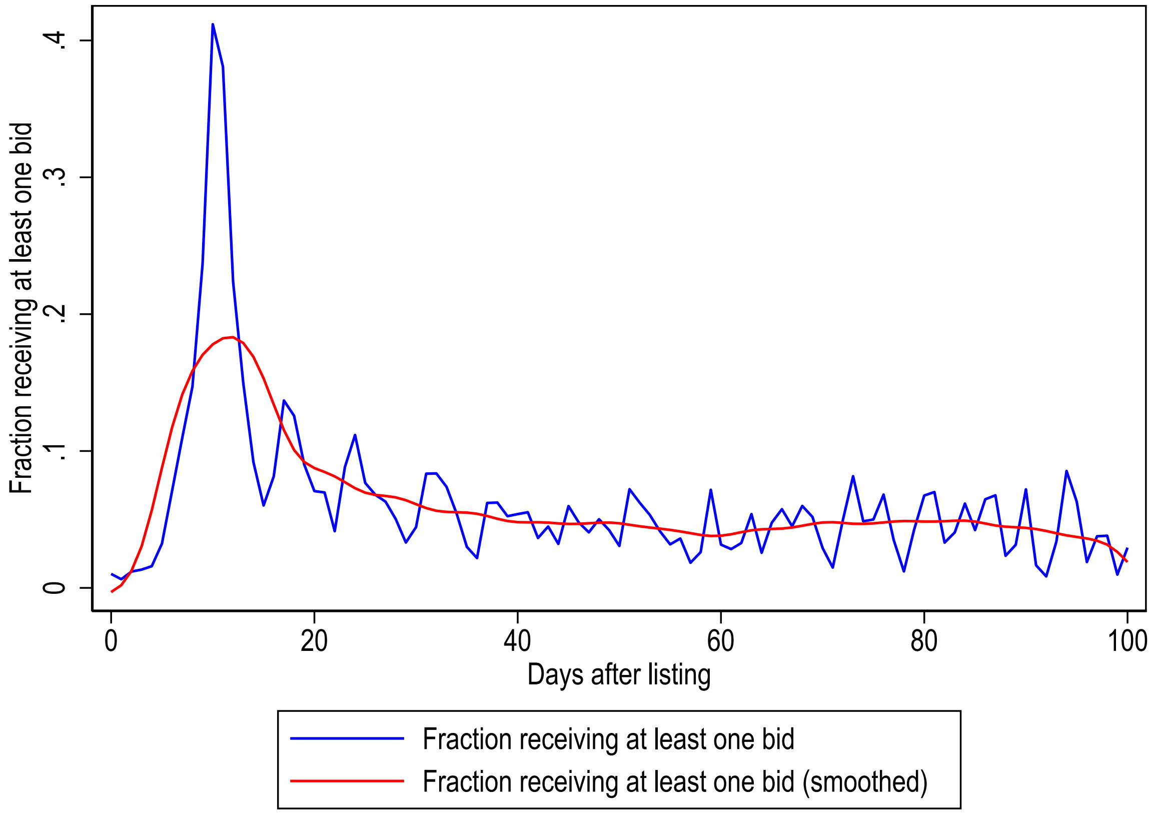
Table 2 summarize the data. The first two columns show the summary statistics for the full sample, the next two columns show summary statistics for sales involving only one bidder, whereas the final two columns show summary statistics for sales with two or more bidders.

The average sell-list spread is 6.1 percent and it is considerably higher for sales with multiple bidders than for sales involving only one bidder. There are about 14 viewers on average for sales involving multiple bidders, and approximately a third of the viewers end up placing a bid. The ratio of bidders to viewers is similar in the one-bidder auctions and these auctions have considerably fewer viewers. It is common to arrange two auctions and the average time on market is only 14 days. Only 12 percent of the multiple-bidder auctions end up with a sell price below the list price. For one-bidder auctions, about 39 percent of the units are sold with a sell price lower than the list price. The one-bidder and multiple-bidder auctions are relatively similar in terms of sell price, list price, size, time-on-market, and fraction of units that are apartments.

Since the model in Section 3 is static, it is best suited to describe a single attempt at selling a house. Moreover, we are interested in the first such attempt, since the first attempt reflects the bulk of the transactions that take place. We identify the first sale attempt as the 4-week period from the listing date. We choose this cutoff, since the estimated hazards of receiving a bid flatten after these periods, as Figure 3 shows. Therefore, any auction-like transaction process with multiple bidders and bids likely take place prior to these cutoffs, and if a housing unit does not sell by this period, it either goes through a subsequent re-listing resulting in a new sale attempt or continues receiving sequential offers from single buyers, which is a transaction process that our model does not represent well. The probability of receiving a bid at the first attempt is 0.87 and the probability that a sale occurs at the first attempt is 0.81. Conditional on receiving at least one bid, the probability of selling at the first attempt is 0.94.

In Figure 4, we display time developments in four key indicators; mean sell-list spread, mean time-on-market, mean number of bidders, and the percentage of units that are sold at the first attempt. It is evident that there are cyclical movements in the Oslo housing market over the period we consider. In particular, there is notable drop in the sell-list spread in 2013, which is associated with a longer time-on-market, fewer bidders, and a substantial drop in the percentage number of units sold at the first sales attempt.

Figure 3: Bidding hazard.



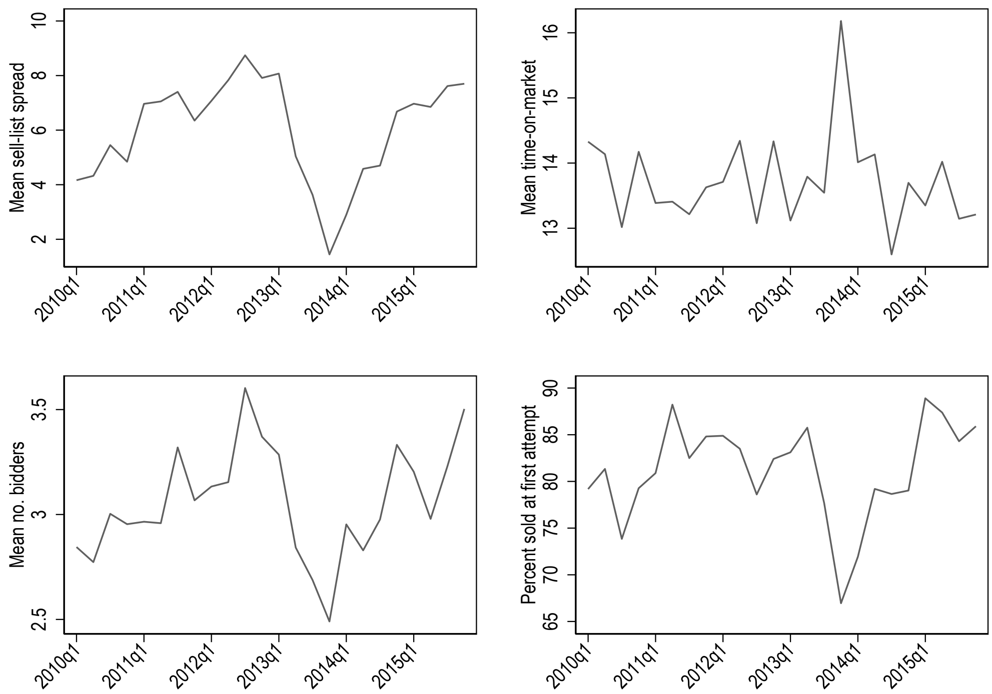
Note: The figure plots the estimated fraction of unsold properties that receive a bid in a given day against the number of days since the listing date for properties in Oslo sold by the realtor firm Krogsveen. The red line is LOESS-smoothed with a smoothing parameter of 0.2.

4.2 Estimation method

Next, we describe our estimation procedure, a number of parametric assumptions we make, and also discuss the features of the data that we use to identify our structural parameters of interest.

Since our data span a large number of different locations, types of housing, and time periods, and each of these characteristics would have a direct effect on the underlying distribution of quality, the distribution of potential buyers, and potentially all other model parameters, we define a number of sub-markets or segments and let a number of parameters of the estimated model vary by segment. Specifically, we define segments based on housing type and size, and time period. For time periods, we consider separately every year in our sample between 2010-2015, while for housing type and size we define three categories, namely Small apartments, Large apartments and Houses. Small apartments have up to 2 rooms, while large apartments have 3 or more rooms. Houses are defined as all remaining housing units, which consist of row houses, semi-detached houses, and single-family houses.

Figure 4: Housing market developments in Oslo. 2010–2015.



Note: The upper left panel plots the mean sell-list spread, the upper right panel shows the mean time-on-market, the lower left panel displays mean number of bidders, whereas the lower right panel show the percentage number of units sold at the first sales attempt (within four weeks of the listing date). The time period is 2010–2015.

Table 2: Summary statistics for auction-level data. Segmentation on one versus multiple bidders, 2010–2015

Variable	Full sample Mean	Full sample Std.	One bidder Mean	One bidder Std.	Multiple bidders Mean	Multiple bidders Std.
Sell (in 1,000 USD)	627.07	288.98	640.31	293.65	622.56	287.26
List (in 1,000 USD)	594.39	278.09	636.1	292.62	580.19	271.53
Square footage	915.55	488.34	941.45	492.03	906.73	486.8
Sell-List spr. (in %)	6.11	8	.85	5.07	7.9	8.03
No. bidders	2.98	2.09	1	0	3.66	2.02
No. bids	9.83	7.82	2.8	1.84	12.23	7.64
No. viewers	11.61	8.91	5.53	4.09	13.57	9.16
Bidders per viewer	.33	.2	.3	.24	.34	.19
No. public showings	1.97	.98	1.9	1.19	1.99	.9
Viewers per showing	6.81	6.08	3.23	2.74	7.86	6.38
Time-on-market (days)	13.69	7.3	13.79	8.1	13.66	7
Perc. with sell < list	18.53		39.09		11.54	
Perc. apartment	78.37		78.89		78.2	
Prob. sell at first att.	0.82					
Prob. bid at first att.	0.87					
Prob. sell if bid at first att.	0.94					
No. auctions	12,044		2,492		9,552	

Note: The table shows summary statistics for auction-level data from DNB Eiendom and Krogsveen over the period 2010–2015. Since the data from DNB Eiendom do not contain information on number of viewers, these measures are calculated using only data from Krogsveen. We distinguish between units sold in one-bidder auctions and units sold in multiple-bidder auctions. For each of the segments, the table shows the mean and standard deviation (Std.) of a selection of key variables. NOK values are converted to USD using the average exchange rate between USD and NOK over the period 2010–2015, in which the exchange rate was $USD/NOK = 0.1639$

We estimate the model parameters using Simulated Method of Moments.[21] Specifically, for each segment we construct a large number of artificial data-sets with size equal to the number of observed sales in each segment. Each observation in these artificial data-sets corresponds to a distinct instance of the buyer entry game described in Section 3. Therefore, for each distinct dwelling in the artificial data-sets we draw a unique combination of quality θ, seller reservation value e, list price a, as well as the number of potential buyers, B_p, and a vector of idiosyncratic preferences for these buyers, $\boldsymbol{u}$. Based on these factors, we compute the entry cutoff $\hat{u}$ according to Eq. (14), the set of potential buyers that enter the bidding stage given this cutoff, and the realized price, p, if a sale takes place, according to Eq. (9).

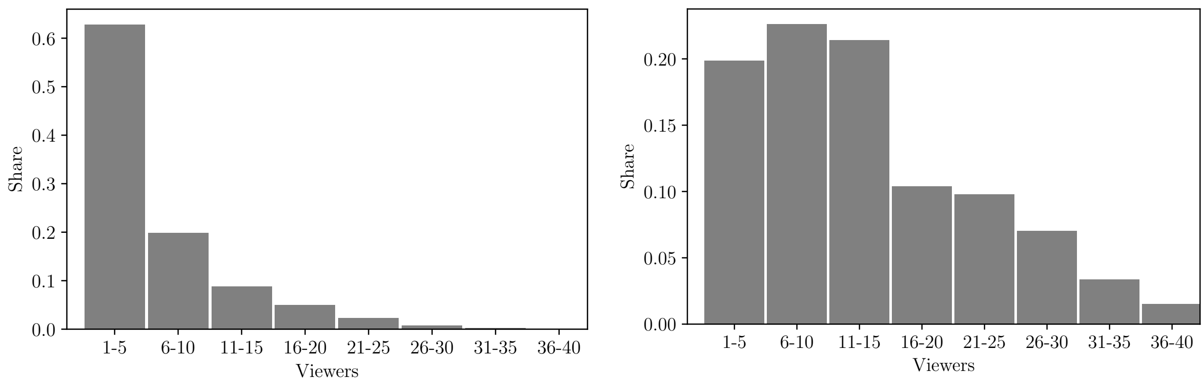
We draw the number of viewers, B_p, from each segment's empirical distribution of viewers. Figure 5 plots this distribution for small apartments in 2011 and houses in 2014. The probabilities are estimated by bins of number of viewers.[22]

[21] See the Appendix for details about the estimation algorithm.

[22] In the simulations, the probability of drawing any given number of viewers, within each bin, is assumed uniform.

Figure 5: Empirical viewer distribution.

(a) Small apartments, 2011 (b) Houses, 2014



In our estimation, we let the dispersion and mean of quality, θ, vary by segment, while for the remaining parameters – the dispersion in buyer idiosyncratic tastes, σ_u, the mean and dispersion of seller reservation values, ($\bar{e}$ and σ_e), the mean and dispersion of the log of the list price "wedge", η, ($\bar{\eta}$ and σ_η), and the (log of the) entry cost c – we impose time-invariance.[23]

4.3 Moments and identification

Next we discuss identification in our estimation framework. Formally, if we define a moment function from the space of parameters to the space of moments, (local) identification requires that the Jacobian matrix of the moment function evaluated at the true parameter values has rank equal to the number of parameters. More informally, each parameter should change the moments in a "unique" way. In practice, the moments have to also be "informative" about the parameters, so that the moment function is not close to flat around the true parameter values. These two considerations guide our choice of moments.

For each segment, we use 9 moments: (i) the mean (log) sale price ($E(\log p)$), (ii) the coefficient of variation of (log) price ($CV(\log p)$) and (iii) (log) list price ($CV(\log a)$), (iv) the mean price premium, defined as the difference between the mean (log) sale price and (mean) (log) list price ($\log p - \log a$), (v) and the mean price premium given only one bidder, (vi) the probability that the sale price is lower than the list price, as well as the sale probability – both (vii) unconditional, and (viii) conditional on there being at least one buyer who enters the bidding stage. Finally, (xi) we target the probability of 5 bidder or more entering the bidding stage. Overall, we have 54 moments per unit type and 18 parameters, so our model is substantially over-identified.

[23]We also perform the estimation by letting all parameters vary over time. The parameter estimates for this estimation are similar to our baseline parameters and are included in the Appendix.

Figure 6 shows how these moments respond to the model parameters for a specific time period for the time-invariant parameters σ_u, σ_e, σ_η, and $\log c$, while Figure 7 contains a similar set of moment comparative statics plots for the quality parameters $\bar{\theta}$ and σ_θ. Next we discuss briefly which moments are informative for each parameter and provide brief intuition for these effects based on the mechanics of our model.

For σ_u the informative moments are the mean and coefficient of variation of the sale price, the price premium, and also the sale probability and the probability that 5 or more bidders enter the bidding stage. Intuitively, a higher dispersion of idiosyncratic tastes increases both the mean price and price premium, as well as the probability that 5 or more bidders enter, since it becomes more likely for several very high-valuation buyers to enter bidding. It also increases the sale probability due to selection effects, since it is easier for a larger set of housing units to attract a buyer with a sufficiently high willingness to pay, despite the quality characteristics of the unit or the reservation value of the seller. Finally, price dispersion also increases, since buyers have more idiosyncratic tastes but also because of the aforementioned selection effects, whereby more lower quality units can also sell.

For σ_η the most informative moments are the sale probability, particularly the sale probability given at least one bidder and the probability that 5 or more potential buyers enter. Intuitively, by Eq. (8), a higher dispersion σ_η increases the posterior variance about the seller's reservation value, thus worsening the quality of information available to potential buyers when they make their bidding entry decisions. Consequently, there are more instances of informational mismatch, whereby buyers with idiosyncratic preferences lower than the seller's true reservation value enter the bidding stage, which reduces the share of transactions. A higher value of σ_η has an *ambiguous* effect on list price dispersion. Intuitively, a higher value of σ_η exerts two effects on list price dispersion. On the one hand, it mechanically increases list price dispersion. On the other hand, the informational mismatch due to higher σ_η interacts with the list price (see Figure 2 and the discussion preceding it), which leads to a selection effect for units with extreme values of the list price. This effect tends to lower the list price dispersion among *transacted* units. Finally, a higher dispersion in η tends to increase the probability that the sale price is below the list price. Intuitively, if there is no dispersion in η, the sale price can never be below the list price, since the list price equals the seller reservation value. With some dispersion in η the list price and the seller reservation value are decoupled, so that it is possible that the seller reservation value (and hence the sale price given at least one bidder with sufficiently high valuation) lies below the list price. This mechanism also explains the decrease in the price premium given at least one bidder as σ_η increases.

A higher value of σ_e has a similar effect to that of a higher value of σ_η on the sale probability and the probability of having 5 or more bidders for similar reasons. However, any selection effects due to informational mismatch are quantitatively weaker in that case.

Consequently, a higher value of σ_e unambiguously increases list price dispersion. It also increases the price premium, since it lowers the average (log) list price of units that end up transacting. Finally, it increases price dispersion, since having more units with more extreme values of the seller reservation value increase the probability that units sell with only one bidder. For these units the price equals the seller reservation value, so a higher dispersion of the seller reservation value mechanically increases price dispersion for those units.

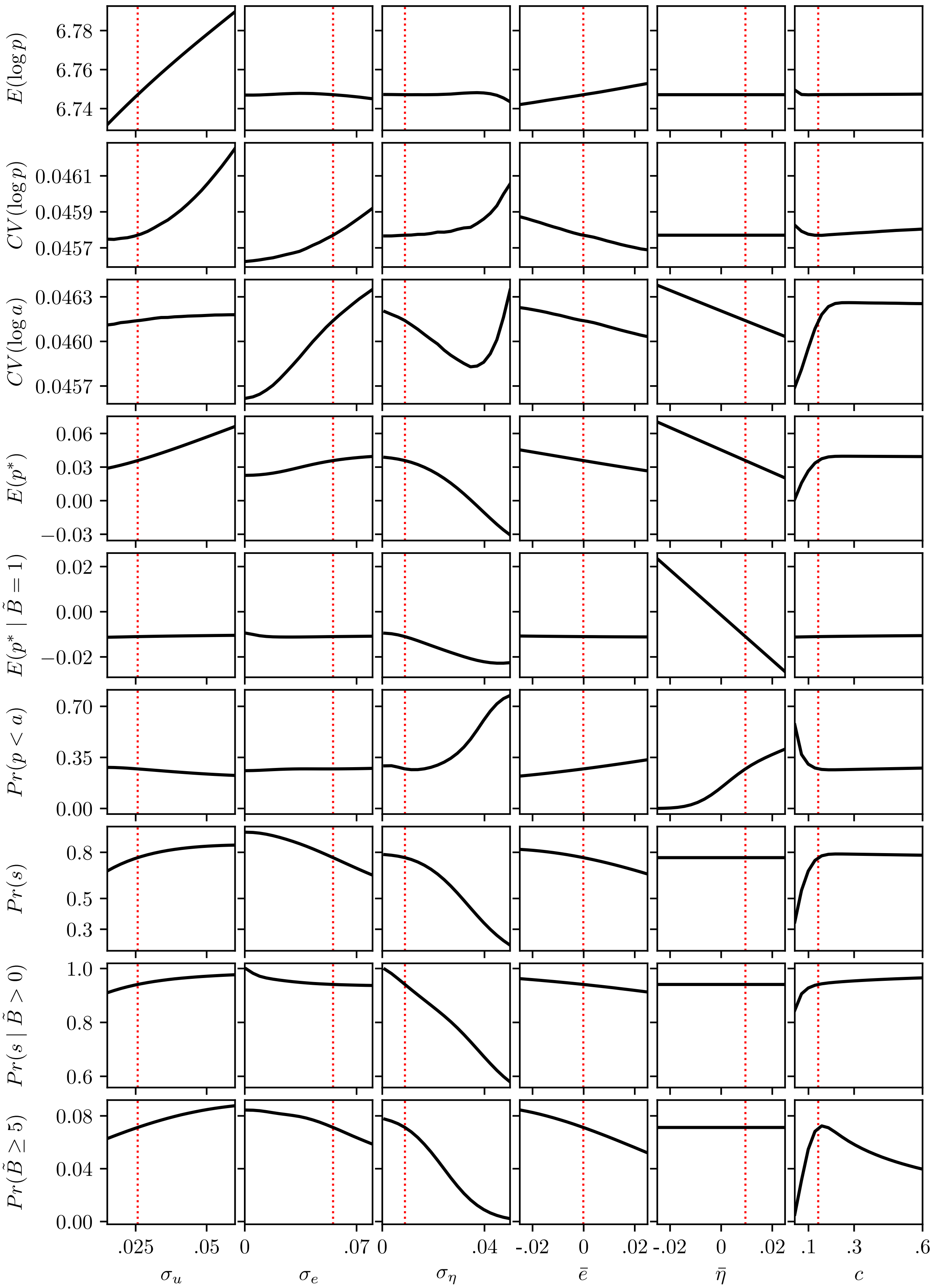
A higher entry cost c affects many of the moments. First, surprisingly, it exerts a non-monotone effect on the sale probability. The reason for the non-monotonicity is the following. On the one hand, since a higher value of c increases the cutoff $\hat{u}$ (see Figure 2), which, other things equal, lowers bidder entry and hence the sale probability. On the other hand, the increase in the cutoff $\hat{u}$ leads to entering buyers having valuations, which tends to reduce the impact of informational mismatch, leading to a larger probability of sale *given* bidder entry. For very low values of c the second effect dominates, while the first effect dominates for higher values of c.

An increase in the average seller reservation price, $\bar{e}$, reduces the sale probability, since it reduces the entry of buyers. It also increases the average sale price, since for all objects with a negotiated sale (i.e. only one bidder with valuation above the seller's reservation value) the price directly depends on the seller reservation value. However, a higher $\bar{e}$ decreases the price premium, since the average list price is more responsive to $\bar{e}$ than the average sale price. The average seller reservation price also affects price dispersion negatively. This is due to a selection effect the average seller reservation price exerts on the quality distribution of objects sold. Turning to the list price wedge, a change in $\bar{\eta}$ leads to a mechanical change in moments involving the list price. For example, a higher value of $\bar{\eta}$ decreases the price premium, since it increases the average (log) list price. It also increases the share of objects that sell below the list price.

Finally, regarding the quality distributions, the average quality parameter $\bar{\theta}$ has an effect on all moments, but particularly so on the average sell price, the probability of sale, as well as the probability that the sell price is below the list price. For the dispersion of quality, there are only two informative moments: the dispersions in the sell price and the list price.

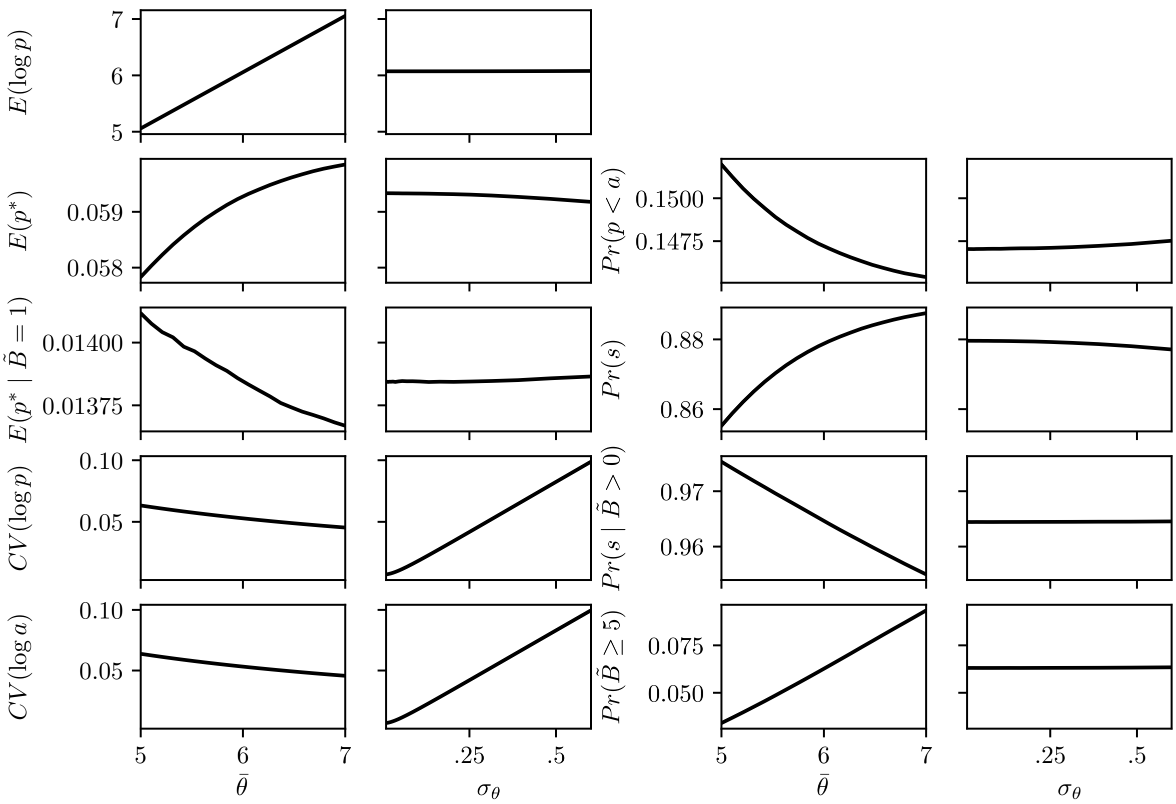
Figures 6 and 7 illustrate how identification is achieved by the moments we have chosen. Specifically, parameters influence the set of moments differently, and at least one moment responds to a specific parameter. For example, we can identify separately σ_θ from σ_u since the dispersion in the list price, $CV(\log a)$, responds to σ_θ but not to σ_u. Similarly, σ_u influences a number of other moments like the average sale price $E[\log p]$, and the probability of sale, $\Pr\{sale\}$, while σ_θ does not influence these moments.

Figure 6: Moment comparative statics.



Note: s denotes the sell probability. p^* denotes the price premium, $\log p - \log a$. The moments are computed for the following parameter values: $\sigma_\theta = 0.3068$, $\sigma_u = 0.0257$, $\sigma_e = 0.0553$, $\sigma_\eta = 0.0089$, $\bar{\theta} = 6.7146$, $\bar{e} = -0.0001$, $\bar{\eta} = 0.0095$, and $c = 0.1426$, also given by the vertical dotted lines.

Figure 7: Moment comparative statics for $\bar{\theta}$ and σ_θ.



Note: s denotes the sell probability. p^* denotes the price premium, $\log p - \log a$. The moments are computed for the following parameter values: $\sigma_\theta = 0.3171$, $\sigma_u = 0.0448$, $\sigma_e = 0.0095$, $\sigma_\eta = 0.0378$, $\bar{\theta} = 6.0170$, $\bar{e} = 0.0115$, $\bar{\eta} = -0.0154$, and $c = 0.0999$, also given by the vertical dotted lines.

5 Results

5.1 Parameter estimates and model fit

Table 3 presents the parameter estimates from our baseline estimation. Table 4 shows the simulated moments at the estimated parameters together with corresponding data moments for segments with the best and worst fit.[24] The simulated moments are for the most part very close to their data counterparts, even for the segments with the worst fit. This is reassuring since there is substantial over-identification in our estimation.

Table 3: Parameter estimates.

	Small apt.	Large apt.	House
σ_u	0.0268 (0.0080)	0.0312 (0.0156)	0.0257 (0.0106)
σ_e	0.0076 (0.0079)	0.0097 (0.0155)	0.0553 (0.0120)
σ_η	0.0439 (0.0073)	0.0392 (0.0128)	0.0089 (0.0110)
$\bar{e}$	-0.0008 (0.0069)	0.0166 (0.0117)	-0.0001 (0.0093)
$\bar{\eta}$	-0.0136 (0.0072)	-0.0111 (0.0133)	0.0095 (0.0099)
c	0.0781 (0.0061)	0.0744 (0.0111)	0.1426 (0.0095)

Note: Standard errors in parenthesis. σ_u, σ_e, σ_η, $\bar{e}$, $\bar{\eta}$, and c are constrained to vary only by housing type, whilst remaining parameters may vary between years.

We estimate a sizable and fairly similar dispersion parameter for buyer idiosyncratic tastes across all housing types.[25] In contrast, the dispersion in seller reservation values differs substantially across housing types and is substantially higher for houses compared to apartments. The estimated cost of bidding entry is close to zero across all housing types though about twice as high for houses compared to apartments. This is to be expected since transaction costs (in the form of brokerage fees and fees paid to the state) scale up with the average value of the property type.[26] The average of the seller-specific reservation value, $\bar{e}$ is estimated to be very close to zero apart from the large apartments segment where it is slightly positive, though with a substantial standard error. Overall, this suggests that on average there are no systematic differences between seller log reservation values and the dwelling's underlying quality (though with a substantial dispersion in the case of houses, as already explained).

Next, we examine the list price "wedge", η. There is substantial dispersion in η, particularly for apartments, reflecting a partial disconnect between list prices and seller valuations, and suggesting that list prices are imperfect signals for seller reservation val-

[24]Table A.2 in the Appendix includes the estimated parameters for the distribution of θ, while Table A.3 in the Appendix shows the simulated and data moments for all segments.

[25]These dispersion parameters translate into a distribution with inter-quartile range of between 0.04 and 0.048.

[26]In Norway, there is a 2.5 percent stamp duty, to be paid upon each transaction. Furthermore, all placed bids are binding, so that a buyer who bids always assumes that they may acquire the property (and pay the associated monetary transaction costs) at that bid.

ues in the case of apartments. Moreover, the estimated mean "wedge" is negative for apartments. This points to to a systematic downward bias in list prices relative to seller valuations, reflecting potential attempts by real estate agents to attract more bidders via a list price discount. This bias appears largest for the Small apartments segment. Both of these results suggest that strategic pricing plays an important role in the Oslo housing market, as emphasized by Anundsen et al. (2020).[27]

Table 4: Target moments for segments with best and worst fit.

	Best		Worst	
Moment	Model	Data	Model	Data
$E[\log p]$	6.747	6.757	6.064	6.074
$E[p^*]$	0.036	0.057	0.045	0.070
$E[p^* \mid B = 1]$	-0.011	0.007	0.012	0.014
$CV(\log p)$	0.046	0.046	0.053	0.053
$CV(\log a)$	0.046	0.046	0.054	0.054
$Pr(sale)$	0.765	0.753	0.916	0.856
$Pr(sale \mid \tilde{B} > 0)$	0.941	0.915	0.964	0.949
$Pr(\tilde{B} \geq 5)$	0.071	0.177	0.068	0.190
$Pr(p \leq a)$	0.271	0.190	0.197	0.129
Segment	House, 2012		Small apt., 2011	
Loss	0.0533		0.2293	

Note: The table shows targeted moments in the baseline estimation for individual segments with the greatest and smallest losses. $p^* = \log p - \log a$.

5.2 Validation

In Table 5, we compare the performance of the model against a number of non-targeted moments. The estimated model gets fairly close to most non-targeted data moments. However, the average number of bidders tend to be a bit higher than in the data and so does the average ratio of bidders to viewers. Similarly, the probability of having only one bidder tends to be a bit lower in the model than in the data. One explanation for these discrepancies is that the observed bidding data is naturally truncated from below given that in reality bidding entry is sequential rather than simultaneous as in our model and the auction format is of a dynamic ascending-bid (English) auction rather a simultaneous second-price auction. Therefore, in the data only lower valuation bidders who enter early would be reflected in the bidding records, whereas lower valuation bidders who come late to the bidding may choose to not bid at all if the current highest bid is already above their valuation and hence not appear in the bidding records.

Next, we validate the estimated model against a direct measure of buyer valuations. This is a particularly important validation of our model, since one of the main aims of the

[27] A complementary explanation for the large values of σ_η is that sellers update their valuation between the time the list price is set and the time the house is sold, as in Anenberg (2016).

Table 5: Non-targeted model moments.

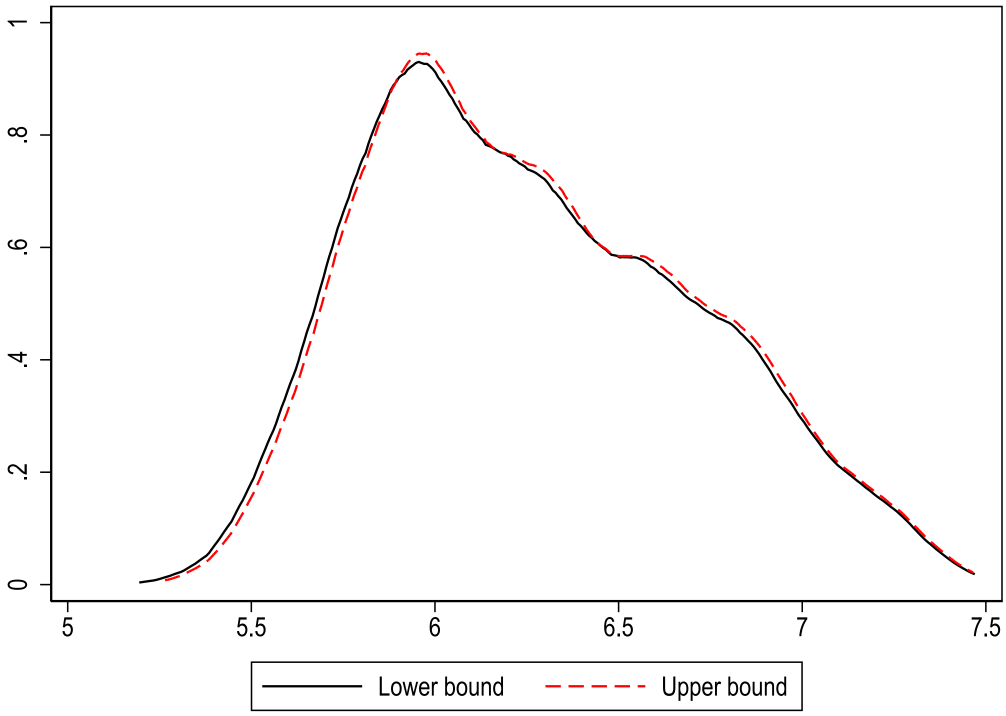
#	Type	Year	$E[\tilde{B}]$ Model	Data	$SD(\tilde{B})$ Model	Data	$Pr(\tilde{B}=1)$ Model	Data	$E[\tilde{B}/B_p]$ Model	Data
1	Small apt.	2010	2.689	2.684	1.467	1.976	0.232	0.296	0.629	0.429
2	Small apt.	2011	2.916	3.052	1.620	1.999	0.206	0.215	0.593	0.406
3	Small apt.	2012	3.037	3.222	1.665	2.308	0.189	0.222	0.581	0.371
4	Small apt.	2013	2.972	2.684	1.624	1.921	0.196	0.310	0.592	0.377
5	Small apt.	2014	3.896	2.999	1.883	2.109	0.094	0.250	0.429	0.308
6	Small apt.	2015	4.238	3.258	1.927	2.372	0.066	0.247	0.385	0.238
7	Large apt.	2010	2.660	2.681	1.511	1.728	0.253	0.251	0.523	0.397
8	Large apt.	2011	2.811	3.097	1.639	1.964	0.238	0.221	0.512	0.373
9	Large apt.	2012	2.761	3.154	1.588	2.200	0.240	0.233	0.523	0.343
10	Large apt.	2013	2.767	2.789	1.582	2.064	0.238	0.266	0.520	0.358
11	Large apt.	2014	3.679	2.717	1.835	1.717	0.112	0.254	0.385	0.274
12	Large apt.	2015	4.007	2.837	1.924	1.835	0.086	0.277	0.357	0.238
13	House	2010	3.085	3.211	1.740	2.108	0.205	0.218	0.556	0.367
14	House	2011	2.960	2.828	1.695	1.964	0.222	0.267	0.588	0.319
15	House	2012	3.049	3.019	1.719	2.017	0.209	0.231	0.576	0.345
16	House	2013	3.209	2.869	1.784	2.036	0.189	0.271	0.545	0.317
17	House	2014	3.811	2.915	2.005	1.966	0.129	0.245	0.393	0.274
18	House	2015	4.152	3.150	2.070	2.097	0.098	0.249	0.323	0.218

Note: The table shows non-targeted moments in the baseline estimation for individual segments. $p^* = \log p - \log a$.

estimation is to understand the contribution of buyer preference heterogeneity to house prices. Specifically, we use the detailed bidding information from our data set to construct a lower and upper bound on the valuations of bidders who participate in auctions (i.e. there are at least two bidders) and who do not end up submitting the highest final bid, i.e. they "lose" the auction. The lower bound is based on the highest bid that a losing bidder submits in the bidding log, while the upper bound is based on the bid that is submitted by another bidder subsequent to the bidder's highest bid. The reason for constructing both a lower and upper bound to the bidder valuation rather than treating the bidder's highest bid as their valuation is that jump bidding and counter-bidding is prevalent in the auction bidding logs. Indeed, around 25% of bids following a losing bidder's highest bid have a bid increment of at least 50,000 Norwegian krone, which represent more than 1% of the average sell price of units in our sample. Figure 8 plots the distribution of the logged lower and upper bounds of losing bidder valuations constructed as explained above. As the Figure shows, the upper and lower bounds are fairly close. Moreover, unconditionally, the upper-bound distribution is essentially a rightward shifted version of the lower-bound distribution.

Next, Tables 6 compares the model and data-derived bidder valuation distributions based on the mean and standard deviation of the log valuation. We make this comparison

Figure 8: Losing bidder valuations.



Note: The graph plots the lower and upper bounds of losing bidder valuations in the data, all in logs. The lower bound is based on the highest bid that a losing bidder submits in the bidding log, while the upper bound is based on the bid that is submitted by another bidder subsequent to the bidder's highest bid.

by housing type and year. Overall the model-generated bidder valuations have means and standard deviations in line with the data generated moments, particularly for the upper bound measures (denoted by UB in the table). For example, the standard deviation of the log losing bidder valuation is between 0.289 and 0.438 in the model and between 0.276 and 0.471 for the upper bound measure.[28]

Table 6: Non-targeted model moments for log valuations of auction losers.

#	Type	Year	Mean			Std. dev.		
			Model	Data		Model	Data	
				LB	UB		LB	UB
1	Small apt.	2010	6.030	5.958	5.951	0.302	0.753	0.331
2	Small apt.	2011	6.085	5.989	6.004	0.320	0.324	0.321
3	Small apt.	2012	6.181	6.049	6.056	0.308	0.589	0.321
4	Small apt.	2013	6.167	6.078	6.090	0.298	0.312	0.301
5	Small apt.	2014	6.178	6.083	6.095	0.304	0.314	0.307
6	Small apt.	2015	6.305	6.191	6.206	0.294	0.299	0.295
7	Large apt.	2010	6.141	6.137	6.152	0.398	0.415	0.412
8	Large apt.	2011	6.289	6.251	6.265	0.403	0.441	0.440
9	Large apt.	2012	6.349	6.209	6.223	0.409	0.451	0.447
10	Large apt.	2013	6.315	6.273	6.293	0.389	0.415	0.412
11	Large apt.	2014	6.331	6.276	6.290	0.413	0.416	0.413
12	Large apt.	2015	6.508	6.453	6.463	0.438	0.476	0.471
13	House	2010	6.668	6.644	6.660	0.289	0.278	0.276
14	House	2011	6.669	6.723	6.731	0.304	0.334	0.300
15	House	2012	6.761	6.761	6.773	0.305	0.300	0.298
16	House	2013	6.827	6.798	6.808	0.296	0.293	0.290
17	House	2014	6.804	6.832	6.841	0.304	0.313	0.301
18	House	2015	6.883	6.925	6.938	0.306	0.310	0.307

Note: The model-generated mean of (standard deviation of) log valuation is given by the conditional expectation (standard deviation) of $\log \tilde{v}(\theta, u_i)$, conditional on sale, $\tilde{B} > 1$, $u_i < \boldsymbol{u}_{(\tilde{B})}$, and $u_i \geq \hat{u}$. The data-generated moments are based on the highest bid that a losing bidder submits in the bidding log (LB), and on the bid that is submitted by another bidder subsequent to the bidder's highest bid (UB) for objects sold in the first attempt as defined in Section 4.1.

Finally, similar to the data-derived distribution of bidder valuations, we use our detailed bidding logs data to compute a proxy for the seller reservation value by considering bidding logs where the seller makes a counter-bid. Specifically, we take the first counter-bid the seller makes in that case. In our model, we equate this to either (i) a situation with entry of only one bidder or (ii) a situation with multiple bidder entry where at most one bidder has a valuation above the seller reservation value. These are precisely the situations in the data where a negotiated sale occurs. Table 7 compares the mean and

[28]The lower bound measure (denoted by LB in the table) is consistently below the upper bound and model-generated mean log valuation, which is consistent with the way the two measures are constructed. In terms of standard deviations, the LB and UB measures are quite similar apart from two of the small apartment segments (in year 2010 and 2012), where the standard deviation of the LB measure is around 2 times larger, which likely reflect substantial outliers in terms of low bids.

standard deviation of the model generated log seller reservation value given these events, and the mean and standard deviation of the log of the seller counter-bid in the data. The model-generated log seller reservation value has consistently lower mean compared to the data-generated counterpart but very similar standard deviation. The discrepancy in the mean values likely reflects strategic aspects of the negotiation between buyer and seller whereby the seller's first counterbid (i.e. the seller's first counteroffer) would consistently exceed her reservation value.

Table 7: Non-targeted model moments for log seller reservation values.

			Mean		Std. dev.	
#	Type	Year	Model	Data	Model	Data
1	Small apt.	2010	5.977	5.983	0.302	0.288
2	Small apt.	2011	6.023	6.079	0.321	0.287
3	Small apt.	2012	6.121	6.172	0.309	0.279
4	Small apt.	2013	6.111	6.190	0.299	0.257
5	Small apt.	2014	6.094	6.231	0.304	0.324
6	Small apt.	2015	6.214	6.372	0.294	0.314
7	Large apt.	2010	6.077	6.223	0.400	0.406
8	Large apt.	2011	6.215	6.244	0.404	0.340
9	Large apt.	2012	6.283	6.306	0.411	0.363
10	Large apt.	2013	6.250	6.396	0.390	0.421
11	Large apt.	2014	6.230	6.383	0.413	0.421
12	Large apt.	2015	6.388	6.520	0.437	0.464
13	House	2010	6.627	6.716	0.292	0.301
14	House	2011	6.624	6.696	0.306	0.318
15	House	2012	6.720	6.799	0.308	0.301
16	House	2013	6.788	6.827	0.299	0.338
17	House	2014	6.762	6.772	0.305	0.308
18	House	2015	6.845	6.960	0.306	0.296

Note: The model-generated mean (standard deviation) of the seller's reservation value is given by the conditional expectation (standard deviation) of $\log \tilde{v}(\theta, e)$, conditional on sale and either (i) a situation with entry of only one bidder (i.e. $\tilde{B} = 1$) or (ii) a situation with multiple bidder entry ($\tilde{B} > 1$), where at most one bidder has a valuation above the seller reservation value. The data-generated moments are based on the first counterbid a seller makes for objects where the seller makes a counterbid and which are sold in the first attempt as defined in Section 4.1.

5.3 Counterfactual exercises

Next, we use the estimated model in a number of counterfactual exercises. First, we assess the contribution of quality towards prices and price dispersion. Next, we assess the contribution of match quality via idiosyncratic buyer tastes for average prices and price dispersion. We also examine the importance of seller reservation values, as well as the informational frictions (the list price "wedge"). Third, we assess the effect of number of bidders on house prices in our model, and compare them against estimates from other studies.

5.3.1 Effects of quality

Table 8 shows the average log price and the standard deviation of log price by type of housing based on data from the estimated model together with the average and standard deviation of log price net of the common quality component θ. Several things stand out from this table. First, quality is the main driver of the average price and price dispersion according to the estimated model. This is not surprising as attributes, such as size and location are the main drivers of house prices in essentially all pricing models used in research and by practitioners. More interestingly, the unexplained variation from fully accounting for quality differences is much lower than that implied by R^2 statistics from "standard" hedonic regressions. Using the last row from the table to compute an implied R^2, one obtains values in the range of 98 to 99%. In contrast, the typical hedonic regression tends to have an R^2 of around 80 to 90%. Table 9 reports R^2s from such a typical model (Model 1 in that table). In contrast models that regress the log sell price on log list price (Model 2 in the table) or use an appraisal value as an explanatory variable (Model 3 in the table) produce much higher values of R^2 and in line with the implied R^2 from Table 8.

Therefore, when *fully* accounted for quality explains most of the variation in sale prices and to a similar extent as regressions that include the list price or appraisal values (which also fully account for quality). Put differently, estimating "standard" hedonic models leaves substantial "unexplained" variation due to *unobservable* quality differences. Still, even after fully accounting for quality, θ, there remains a sizable "residual" price component that is driven by buyer/seller preference heterogeneity and frictions.[29] Next, we turn to a decomposition of that residual price dispersion.

Table 8: Contribution of quality θ to mean and standard deviation of log sell prices.

	Small apt.	Large apt.	House
$E[\log p]$	6.1414	6.2921	6.7571
$E[\log p - \theta]$	0.0366	0.0484	0.0400
$SD(\log p)$	0.3064	0.4125	0.3034
$SD(\log p - \theta)$	0.0282	0.0290	0.0344
$Var(\log p - \theta)/Var(\log p)$	0.850 %	0.494 %	1.287 %

Note: The table reports the mean and standard deviation of log price and log price minus *heta* in the estimated model for the three different types of housing, as well as the ratio of the variance of log price minus theta over the variance of log price. Small apartments have up to 2 rooms, while large apartments have 3 or more rooms. Houses are defined as all remaining housing units, which consist of row and semi-detached houses, and single-family homes.

[29]Note that quality may still impact that residual price component indirectly due to selection effects.

Table 9: Adjusted R^2 for hedonic regression models.

	Model 1		Model 2		Model 3	
Type of unit	R^2	Obs.	R^2	Obs.	R^2	Obs.
Small apt.	0.846	5705	0.961	5952	0.959	3162
Large apt.	0.878	1616	0.969	1738	0.968	944
House	0.795	2044	0.932	2123	0.930	991

Note: The table reports the adjusted R^2 for 3 regression models with log of sell price as dependent variable. Model 1 is a standard hedonic regressions with log size, log size squared, indicator for lot size greater than 1000 square meters (for houses), log of common debt, year-by-month fixed effects, postal code fixed effects, and indicators for four construction periods. Model 2 has log of list price and log of common debt as regressors. Model 3 has log of appraised value and log of common debt as regressors. Each model is estimated for a specific type of unit in Oslo for the period 2010-2015. Small apartments have up to 2 rooms, while large apartments have 3 or more rooms. Houses are defined as all remaining housing units, which consist of row and semi-detached houses, and single-family homes.

5.3.2 Effects of preference heterogeneity and frictions

Table 10: Counterfactual effects on the mean "residualized" price, $\log \tilde{p}$.

Counterfactual	Small apt.	Large apt.	Houses	Note
Baseline	0.0366	0.0484	0.0400	Baseline mean
$\sigma_u \to 0$	−103.11%	−104.21%	−111.75%	No dispersion in buyer tastes
$\sigma_e \to 0$	2.10%	2.64%	−1.35%	No dispersion in seller reservation values
$\sigma_\eta \to 0$	2.43%	2.80%	1.23%	Symmetric information
$\bar{e} = 0$	0.96%	−22.34%	0.08%	Mean zero seller reservation value

Note: The table shows the percent change of mean log sale price less θ given the parameter estimates in Table 3, when changing parameters as indicated by the first column. The first row gives the mean "residualized" price (in log points) at the parameter estimates.

Table 11: Counterfactual effects on the dispersion of the "residualized" price, $\log \tilde{p}$.

Counterfactual	Small apt.	Large apt.	Houses	Note
Baseline	0.0282	0.0290	0.0344	Baseline std. dev.
$\sigma_u \to 0$	−87.62%	−89.62%	−55.62%	No dispersion in buyer tastes
$\sigma_e \to 0$	0.64%	0.87%	−18.71%	No dispersion in seller reservation values
$\sigma_\eta \to 0$	1.25%	2.40%	−0.32%	Symmetric information
$\bar{e} = 0$	−0.41%	5.29%	−0.04%	Mean zero seller reservation value

Note: The table shows the percent change of the standard deviation of log sale price less θ given the parameter estimates in Table 3, when changing parameters as indicated by the first column. The first row gives the standard deviation of the "residualized" price (in log points) at the parameter estimates.

Next, we examine the importance of buyer and seller preference heterogeneity and the informational frictions for prices and price dispersion. Following the discussion from the previous section, we focus on the "residualized" (log) price after removing the direct effects of quality on price. Therefore, we examine the object $\log \tilde{p} = \log p - \theta$. Tables 10 and 11 show how the mean and standard deviation of the "residualized" price changes for the different housing type segments in a number of counterfactual exercises. To assess the contribution of match quality to the mean and standard deviation of the price premium, we first reduce the idiosyncratic dispersion in buyer tastes, σ_u, to (close to) zero (Row 2 in Tables 10 and 11). In this case, the average premium is reduced by around 100 percent for all housing types, while the standard deviation goes down by between 55 and 90 percent, which translates into a reduction of between 0.02 and 0.025 log points. This is substantially larger than the contribution of the dispersion in the seller reservation values (Row 3), apart from houses, where σ_e also has a non-negligible effect. The list price wedge dispersion, σ_η (Row 4), has a negligible effect on both the mean and standard deviation of prices, suggesting a fairly limited quantitative role of informational mismatch at the estiamted parameters. Finally, lowering the mean seller valuation (Row 5) has a substantial effect on the average "residualized" price only for large apartments where this mean value was estimated to be relatively large in magnitude.

Overall, these exercises point to match quality being the main driver of "residualized" house price dispersion with match quality accounting for around 2.5 percentage points of the observed house price dispersion in Oslo during 2010-2015. Moreover, match quality is also important for the observed average price *level* and contributed around 4 to 5 percentage points to the observed average prices in Oslo during 2010-2015.

5.3.3 Effects of number of bidders on prices

Using our estimated model for each sub-market we generate simulated sales data and estimate the following regression

$$\log P_h = \alpha + \beta_1 \tilde{B}_h \, (+\beta_2 \theta_h) + \varepsilon_h, \tag{17}$$

in which $\tilde{B}$ denotes the number of bidders for property h. Therefore, we examine the effect of number of bidders on the final sale price using a reduced-form regression that is often estimated in the literature. Table 12 shows the coefficient estimates for the different segments. After fully controlling for quality, the effect of one more bidder on log prices is around 0.01. If instead we do not control for quality, we obtain substantially larger coefficient estimates of between .026 and 0.045. Therefore, the coefficient estimates with and without controlling for quality provide bounds on the estimated effect of number of bidders on prices in reduced-form regressions, depending on how successfully one can control for housing quality. Interestingly, the coefficient estimates after controlling for quality are quite close to estimates reported in other empirical studies. For example, Genesove and Han (2016) estimate a reduced-form coefficient of bidders on log prices of 0.011 for a North American housing market, while Hungria-Gunnelin (2013) estimates a coefficient of around 0.04 using data for Stockholm, Sweden. Finally, Anundsen et al. (2020) estimate an effect of between 0.02 and 0.03 for Norway. This provides another validation for our estimated model.

Table 12: Counterfactual effect of number of bidders on sale prices.

Housing type	$\tilde{B}$	$\tilde{B} \mid \theta$
Small apt.	0.0349	0.0126
Large apt.	0.0454	0.0126
House	0.0257	0.0096

Note: The table shows the estimated coefficients of the regression $\log P_h = \alpha + \beta_1 \tilde{B}_h + \beta_2 \theta_h + \varepsilon_h$, without (first column) and with (second column) a quality control, θ, in the regression.

6 Conclusion

Using a unique auction-level data set for housing sales in the Norwegian capital Oslo, we estimate a structural model of the housing transaction process, which explicitly includes the (endogenously determined) possibility of negotiated versus auctioned sales. We find that quality matters substantially more for price dispersion than "traditional" hedonic pricing models would suggest, pointing to the importance of unobserved heterogeneity in the housing market. Beyond quality, buyer taste heterogeneity matters the most for any "residual" price dispersion, with the distribution of seller reservation values mattering only for some types of housing. Finally, there's a partial disconnect between list prices and seller reservation values, particularly for apartments, which implies that buyers face imperfect information about seller reservation values when making bidding decisions, leading to informational mismatch.

One important insight from our analysis is that match quality matters not just for price dipsersion but for average house prices as well. However, the relative importance of match quality for house prices versus price dispersion depends on the underlying bidding costs that potential buyers face. Larger bidding costs reduce bidder entry, thus leading to a lower average price, but also to higher price dispersion. Therefore, any policy that tries to "cool" down the housing market and house price appreciation by slowing down bidder entry via either macro-prudential tools or direct regulation of the housing market transaction process would have the unintended effect of increasing price dispersion and subjecting home owners to greater uninsurable housing wealth risk. Moreover, such policies may have spillover effects on the spending decisions of existing homeowners, and hence, aggregate economic activity. We find these issues to be important for future research on this topic.

References

Anenberg, E. (2016). “Information frictions and housing market dynamics.” *International Economic Review*, 57(4), 1449–1479.

Anenberg, E. and P. Bayer (2013). “Endogenous Sources of Volatility in Housing Markets: The Joint Buyer-Seller Problem.” Working Paper No. 18980, NBER.

Anundsen, A. K., E. Røed Larsen, and D.E. Sommervoll (2020). “Strategic price-setting and incentives in the housing market.” Tech. Rep. 1, Housing Lab, Oslo Metropolitan University.

Arefeva, A. (2020). “How auctions amplify house-price fluctuations.” *Available at SSRN 2980095*.

Ashenfelter, O. and D. Genesove (1992). “Testing for Price Anomalies in Real-Estate Auctions.” *The American Economic Review*, 82(2), 501–505.

Beggs, A. and K. Graddy (2009). “Anchoring Effects: Evidence from Art Auctions.” *American Economic Review*, 99(3), 1027–1039.

Bucchianeri, G. W. and J. A. Minson (2013). “A homeowner’s dilemma: Anchoring in residential real estate transactions.” *Journal of Economic Behavior & Organization*, 89, 76–92.

Bulow, J. and P. Klemperer (1996). “Auctions Versus Negotiations.” *The American Economic Review*, pp. 180–194.

Bulow, J. and P. Klemperer (2009). “Why do sellers (usually) prefer auctions?” *American Economic Review*, 99(4), 1544–75.

Caplin, A. and J. Leahy (2011). “Trading Frictions and House Price Dynamics.” *Journal of Money, Credit, and Banking*, 43, 283–303.

Carrillo, P. E. (2012). “An empirical stationary equilibrium search model of the housing market.” *International Economic Review*, 53(1), 203–234.

Chow, Y. L., I. E. Hafalir, and A. Yavas (2015). “Auction versus negotiated sale: evidence from real estate sales.” *Real Estate Economics*, 43(2), 432–470.

Diaz, Antonia and Belen Jerez (2013). “House Prices, Sales, and Time on the Market: A Search Theoretic Framework.” *International Economic Review*, 54, 837–872.

Einav, L., T. Kuchler, J. Levin, and N. Sundaresan (2015). “Assessing Sale Strategies in Online Markets Using Matched Listings.” *American Economic Journal: Microeconomics*, 7(2).

Genesove, D. and L. Han (2016). "The Thinness of Real Estate Markets."

Genesove, D. and J. Hansen (2014). "Predicting dwelling prices with consideration of the sales mechanism." Tech. rep.

Genesove, D. and C. Mayer (2001). "Loss aversion and seller behavior: evidence from the housing market." *Quarterly Journal of Economics*, 116(4), 1233–1260.

Genesove, David and Lu Han (2012). "Search and Matching in the Housing Market." *Journal of Urban Economics*, 72, 31–45.

Grindaker, M., A. Karapetyan, P. Nenov, and E. Moen (2021). "Transaction Sequencing and House Price Pressures." Mimeo.

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

Guren, A. M. and T. J. McQuade (2019). "How Do Foreclosures Exacerbate Housing Downturns?" Working Paper No. 26216, NBER.

Han, L. and W. C. Strange (2014). "Bidding wars for houses." *Real Estate Economics*, 42(1), 1–32.

Han, L. and W. C. Strange (2015). "The microstructure of housing markets: Search, bargaining, and brokerage." In *Handbook of Regional and Urban Economics, vol. 5*, edited by Henderson J. V. Duranton, G. and W. Strange, pp. 813–886. Elsevier.

Han, L. and W.C Strange (2016). "What is the role of the asking price for a house?" *Journal of Urban Economics*, 93, 115–130.

Haurin, D., S. McGreal, A. Adair, L. Brown, and J. R. Webb (2013). "List price and sales prices of residential properties during booms and busts." *Journal of Housing Economics*, 22(1), 1–10.

Head, A., H. Lloyd-Ellis, and H. Sun (2014). "Search, Liquidity and the Dynamics of House Prices and Construction." *American Economic Review*, 104, 1172–1210.

Hennessy, C. A. and T. M. Whited (2007). "How costly is external financing? Evidence from a structural estimation." *The Journal of Finance*, 62(4), 1705–1745.

Herrin, E., W., J. R. Knight, and C. F. Sirmans (2004). "Price cutting behavior in residential markets." *Journal of Housing Economics*, 13(3), 195–207.

Holmström, Bengt (1999). "Managerial incentive problems: A dynamic perspective." *The review of Economic studies*, 66(1), 169–182.

Horowitz, J. L. (1992). "The role of the list price in housing markets: Theory and an econometric model." *Journal of Applied Econometrics*, 7, 115–129.

Hungria-Gunnelin, R. (2013). "Impact of number of bidders on sale price of auctioned condominium apartments in Stockholm." *International Real Estate Review*, 16(3), 274–295.

Kotova, Nadia and Anthony L Zhang (2021). "Liquidity in residential real estate markets." Tech. rep., Working paper.

Krainer, J. (2001). "A Theory of Liquidity in Residential Real Estate Markets." *Journal of Urban Economics*, 49, 32–53.

Ku, G., A. D. Galinsky, and J.K. Murnighan (2006). "Starting low but ending high: A reversal of the anchoring effect in auctions." *Journal of Personality and Social Psychology*, 90(6), 975–986.

Levin, D. and J. L. Smith (1994). "Equilibrium in auctions with entry." *The American Economic Review*, pp. 585–599.

Liu, X. and A. J. van der Vlist (2019). "Listing strategies and housing busts: Cutting loss or cutting list price?" *Journal of Housing Economics*, 43, 102–117.

Lusht, K. M. (1996). "A comparison of prices brought by English auctions and private negotiations." *Real Estate Economics*, 24(4), 517–530.

Mayer, C. J. (1995). "A model of negotiated sales applied to real estate auctions." *Journal of urban Economics*, 38(1), 1–22.

McAfee, R. P. and J. McMillan (1987). "Auctions with a stochastic number of bidders." *Journal of economic theory*, 43(1), 1–19.

McCall, John Joseph (1970). "Economics of information and job search." *The Quarterly Journal of Economics*, pp. 113–126.

Merlo, A. and F. Ortalo-Magne (2004). "Bargaining over residential real estate: evidence from England." *Journal of urban economics*, 56(2), 192–216.

Merlo, A., F. Ortalo-Magne, and J. Rust (2015). "The home selling problem: theory and evidence." *International Economic Review*, 56(2), 457–484.

Moen, Espen R, Plamen T Nenov, and Florian Sniekers (2019). "Buying First or Selling First in Housing Markets." *Journal of the European Economic Association*. Jvz059.

Nenov, P. T., E. Røed Larsen, and D. E. Sommervoll (2016). “Thick-market Effects, Housing Heterogeneity, and the Determinants of Transaction Seasonality.” *The Economic Journal*, 126(598), 2402–2423.

Ngai, L. R. and K. D. Sheedy (2019). “The decision to move house and aggregate housing-market dynamics.” *Journal of the European Economic Association*.

Ngai, L. R. and S. Tenreyro (2014). “Hot and Cold Seasons in the Housing Market.” *American Economic Review*, 104, 3991–4026.

Northcraft, G. B. and M. A. Neale (1987). “Experts, Amateurs, and Real Estate: An Anchoring-and- Adjustment Perspective on Property Pricing Decisions.” *Organizational behavior and human processes*, 39(1), 213–237.

Novy-Marx, R. (2009). “Hot and Cold Markets.” *Real Estate Economics*, 37, 1–22.

Piazzesi, M., M. Schneider, and J. Stroebel (2020). “Segmented Housing Search.” *American Economic Review*, 110(3), 720–759.

Postel-Vinay, F. and J. Robin (2002). “Equilibrium wage dispersion with worker and employer heterogeneity.” *Econometrica*, 70(6), 2295–2350.

Rekkas, Marie, Randall Wright, and Yu Zhu (2021). “How Well Does Search Theory Explain Housing Prices.” Tech. rep., Working paper.

Repetto, L. and A. Solis (2019). “The Price of Inattention: Evidence from the Swedish Housing Market.” *Journal of the European Economic Association*. Forthcoming.

Simonsohn, U. and D. Ariely (2008). “When rational sellers face nonrational buyers: evidence from herding on eBay.” *Management Science*, 54(9), 1624–1637.

Storn, R. and K. Price (1997). “Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces.” *Journal of global optimization*, 11(4), 341–359.

Taylor, C. R. (1999). “Time-on-market as a signal of quality.” *Review of Economic Studies*, 66(3), 555–578.

Wheaton, W. (1990). “Vacancy, Search, and Prices in a Housing Market Matching Model.” *Journal of Political Economy*, 98, 1270–1292.

A Additional tables and results

Table A.1: Data trimming

Step:	Both firms		Krogsveen		DNB	
	Bids	Trans.	Bids	Trans.	Bids	Trans
Initial data set (excl. Coops)	60,217	6,003	73,007	8,031	133,224	14,034
Drop missing sell, ask, size	59,932	5,928	72,289	7,742	132,221	13,670
Truncate on 1^{st} and 99^{th} pct. of size, ask, sell	57,479	5,678	69,170	7,410	126,649	13,088
Drop all transactions of units with more than three sales	57,225	5,646	68,800	7,368	126,025	13,014
Drop if TOM < 0	57,150	5,630	68,800	7,368	125,950	12,998
Drop missing bids or if bid $<$ 80% of list price†	54,586	5,629	66,534	7,368	121,120	12,997
Additional constraints (see tablenotes)	50,621	5,122	63,188	6,922	113,809	12,044

Note: The table shows the different steps taken when trimming the data. We show the number of bids and transactions after each step, both for the full sample of both firms, as well as for the individual firms. †: We do not drop bids lower than 80% of the list price if it is the winning bid. The additional constraints are: We drop the entire auction (all bids) if one bid is missing expiration date and there are less than 5 bidders, or if more than one bid is missing the expiration date of the bid. In addition, drop only the bid with missing expiration date if there are more than 5 bidders and only one bid has missing expiration date. We drop the entire auction if the bid expires before it is received and there are less than 5 bidders, or if more than one bid expires before it is received. In addition, drop only the bid that expires before it is received if there are more than 5 bidders and only one bid expires before it is received. We drop the entire auction if the distance (in days) between expiration of the previous bid and receiving a new bid is very long (99.5th pct) or short (0.5th pct) and there are less than 5 bidders, or if more than one bid has a long/short distance (in days) between expiration of the previous bid and receiving a new bid. In addition, drop only the bid that has a long/short distance (in days) between expiration of the previous bid and receiving a new bid if there are more than 5 bidders and only one bid has a long/short distance (in days) between expiration of the previous bid and receiving a new bid. Finally, we truncate on 1st and 99th pct of number of days elapsed between ready for sale date and date of hiring the realtor.

Table A.2: Parameter estimates (all parameters).

	Small apt.	Large apt.	House
σ_u	0.0268 (0.0080)	0.0312 (0.0156)	0.0257 (0.0106)
σ_e	0.0076 (0.0079)	0.0097 (0.0155)	0.0553 (0.0120)
σ_η	0.0439 (0.0073)	0.0392 (0.0128)	0.0089 (0.0110)
$\bar{e}$	-0.0008 (0.0069)	0.0166 (0.0117)	-0.0001 (0.0093)
$\bar{\eta}$	-0.0136 (0.0072)	-0.0111 (0.0133)	0.0095 (0.0099)
c	0.0781 (0.0061)	0.0744 (0.0111)	0.1426 (0.0095)
$\sigma_{\theta,2010}$	0.3029 (0.0011)	0.4035 (0.0018)	0.2905 (0.0020)
$\sigma_{\theta,2011}$	0.3214 (0.0006)	0.4085 (0.0010)	0.3052 (0.0020)
$\sigma_{\theta,2012}$	0.3088 (0.0012)	0.4135 (0.0020)	0.3068 (0.0008)
$\sigma_{\theta,2013}$	0.2992 (0.0120)	0.3928 (0.0284)	0.2975 (0.0162)
$\sigma_{\theta,2014}$	0.3048 (0.0115)	0.4178 (0.0291)	0.3052 (0.0167)
$\sigma_{\theta,2015}$	0.2951 (0.0098)	0.4434 (0.0260)	0.3072 (0.0167)
$\bar{\theta}_{2010}$	5.9851 (0.0093)	6.0675 (0.0220)	6.6207 (0.0151)
$\bar{\theta}_{2011}$	6.0324 (0.0092)	6.2062 (0.0227)	6.6216 (0.0177)
$\bar{\theta}_{2012}$	6.1294 (0.0082)	6.2725 (0.0211)	6.7146 (0.0162)
$\bar{\theta}_{2013}$	6.1187 (0.0010)	6.2408 (0.0025)	6.7798 (0.0008)
$\bar{\theta}_{2014}$	6.1107 (0.0012)	6.2336 (0.0019)	6.7432 (0.0010)
$\bar{\theta}_{2015}$	6.2353 (0.0031)	6.3961 (0.0050)	6.8172 (0.0095)

Note: Standard errors in parenthesis. σ_u, σ_e, σ_η, $\bar{e}$, $\bar{\eta}$, and c are constrained to vary only by housing type, whilst remaining parameters may vary between years.

Table A.3: Targeted model moments and loss.

			$E[\log p]$		$E[p^*]$		$E[p^* \mid B=1]$		$CV(\log p)$		$CV(\log a)$		$Pr(sale)$		$Pr(sale \mid \tilde{B}>0)$		$Pr(\tilde{B} \geq 5)$		$Pr(p \leq a)$		
#	Type	Year	Model	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model	Data	Δ loss
1	Small apt.	2010	6.012	5.996	0.041	0.039	0.012	-0.004	0.051	0.051	0.051	0.051	0.909	0.789	0.959	0.930	0.043	0.122	0.214	0.242	0.1909
2	Small apt.	2011	6.064	6.074	0.045	0.070	0.012	0.014	0.053	0.053	0.054	0.054	0.916	0.856	0.964	0.949	0.068	0.190	0.197	0.129	0.1745
3	Small apt.	2012	6.163	6.151	0.046	0.080	0.012	0.017	0.050	0.049	0.051	0.053	0.918	0.860	0.966	0.955	0.079	0.217	0.190	0.109	0.2293
4	Small apt.	2013	6.150	6.146	0.045	0.046	0.012	0.000	0.049	0.048	0.049	0.050	0.922	0.797	0.964	0.921	0.072	0.140	0.194	0.236	0.1496
5	Small apt.	2014	6.164	6.159	0.061	0.044	0.012	0.002	0.050	0.051	0.050	0.053	0.895	0.796	0.987	0.926	0.173	0.180	0.123	0.202	0.1732
6	Small apt.	2015	6.295	6.295	0.066	0.072	0.012	0.019	0.047	0.048	0.048	0.052	0.886	0.895	0.992	0.966	0.216	0.217	0.103	0.131	0.1226
7	Large apt.	2010	6.111	6.100	0.036	0.039	0.009	-0.008	0.066	0.066	0.066	0.065	0.824	0.749	0.927	0.912	0.043	0.130	0.228	0.223	0.1626
8	Large apt.	2011	6.255	6.253	0.038	0.065	0.009	0.002	0.065	0.067	0.066	0.068	0.824	0.854	0.927	0.935	0.061	0.217	0.219	0.161	0.1763
9	Large apt.	2012	6.318	6.284	0.037	0.069	0.009	0.015	0.065	0.066	0.066	0.069	0.824	0.808	0.925	0.923	0.054	0.203	0.224	0.128	0.1884
10	Large apt.	2013	6.286	6.289	0.037	0.045	0.009	0.001	0.062	0.062	0.063	0.064	0.831	0.819	0.926	0.939	0.053	0.146	0.222	0.245	0.0830
11	Large apt.	2014	6.303	6.318	0.052	0.041	0.008	0.007	0.066	0.065	0.067	0.067	0.883	0.761	0.973	0.903	0.143	0.133	0.146	0.216	0.1581
12	Large apt.	2015	6.478	6.471	0.056	0.059	0.008	0.021	0.068	0.068	0.069	0.070	0.854	0.856	0.979	0.947	0.181	0.175	0.127	0.136	0.0647
13	House	2010	6.655	6.657	0.037	0.056	-0.011	-0.007	0.044	0.044	0.044	0.044	0.768	0.803	0.944	0.950	0.075	0.221	0.264	0.235	0.1202
14	House	2011	6.653	6.665	0.035	0.053	-0.011	0.005	0.046	0.047	0.046	0.046	0.759	0.811	0.940	0.957	0.064	0.144	0.281	0.200	0.0995
15	House	2012	6.747	6.757	0.036	0.057	-0.011	0.007	0.046	0.046	0.046	0.046	0.765	0.753	0.941	0.915	0.071	0.177	0.271	0.190	0.0982
16	House	2013	6.816	6.814	0.038	0.041	-0.011	-0.003	0.044	0.044	0.044	0.043	0.775	0.757	0.944	0.917	0.086	0.147	0.253	0.264	0.0533
17	House	2014	6.795	6.798	0.047	0.042	-0.012	0.009	0.045	0.045	0.046	0.045	0.807	0.710	0.960	0.899	0.160	0.185	0.187	0.239	0.1552
18	House	2015	6.877	6.888	0.051	0.057	-0.012	0.014	0.045	0.045	0.045	0.045	0.822	0.792	0.967	0.949	0.204	0.225	0.155	0.225	0.1429

Note: The table shows targeted moments in the baseline estimation for individual segments. $p^* = \log p - \log a$.

Table A.4: Parameter estimates (extended estimation).

#	Type	Year	σ_θ	θ	σ_u	σ_e	σ_η	$\bar{e}$	$\bar{\eta}$	c
1	Small apt.	2010	0.3032 (0.0092)	5.9630 (0.0183)	0.0321 (0.0048)	0.0115 (0.0026)	0.0508 (0.0050)	0.00 (0.01)	0.01 (0.00)	0.07 (0.01)
2	Small apt.	2011	0.3156 (0.0084)	6.0332 (0.0156)	0.0297 (0.0069)	0.0140 (0.0072)	0.0638 (0.0031)	-0.01 (0.01)	-0.01 (0.00)	0.10 (0.03)
3	Small apt.	2012	0.3070 (0.0099)	6.1722 (0.0158)	0.0024 (0.0076)	0.0328 (0.0045)	0.1036 (0.0104)	-0.06 (0.02)	-0.01 (0.00)	0.21 (0.04)
4	Small apt.	2013	0.2952 (0.0073)	6.1612 (0.0119)	0.0001 (0.0006)	0.0275 (0.0031)	0.0761 (0.0049)	-0.04 (0.00)	-0.00 (0.00)	0.17 (0.02)
5	Small apt.	2014	0.3129 (0.0071)	6.0797 (0.0139)	0.0381 (0.0041)	0.0019 (0.0012)	0.0758 (0.0031)	0.04 (0.01)	0.01 (0.00)	0.04 (0.01)
6	Small apt.	2015	0.3021 (0.0060)	6.2411 (0.0120)	0.0249 (0.0051)	0.0132 (0.0030)	0.0695 (0.0029)	0.01 (0.01)	-0.02 (0.00)	0.09 (0.02)
7	Large apt.	2010	0.4046 (0.0171)	6.0639 (0.0408)	0.0161 (0.0127)	0.0301 (0.0106)	0.0441 (0.0118)	-0.02 (0.02)	0.01 (0.00)	0.15 (0.06)
8	Large apt.	2011	0.4053 (0.0195)	6.2586 (0.0447)	0.0001 (0.0043)	0.0001 (0.0005)	0.0758 (0.0068)	-0.06 (0.12)	0.00 (0.10)	0.06 (0.21)
9	Large apt.	2012	0.4323 (0.0327)	6.2939 (0.0495)	0.0001 (0.0009)	0.0001 (0.0003)	0.0867 (0.0109)	-0.06 (0.03)	0.00 (0.01)	0.06 (0.03)
10	Large apt.	2013	0.3889 (0.0136)	6.2712 (0.0243)	0.0325 (0.0093)	0.0090 (0.0030)	0.0765 (0.0068)	0.01 (0.02)	0.00 (0.00)	0.07 (0.01)
11	Large apt.	2014	0.4305 (0.0147)	6.3243 (0.0236)	0.0001 (0.0008)	0.0132 (0.0032)	0.0643 (0.0099)	-0.02 (0.00)	-0.01 (0.01)	1.27 (0.37)
12	Large apt.	2015	0.4390 (0.0113)	6.4404 (0.0271)	0.0191 (0.0080)	0.0132 (0.0068)	0.0698 (0.0051)	0.01 (0.02)	-0.02 (0.01)	0.08 (0.04)
13	House	2010	0.2884 (0.0109)	6.5917 (0.0253)	0.0464 (0.0097)	0.0084 (0.0029)	0.0713 (0.0077)	0.01 (0.02)	0.01 (0.01)	0.08 (0.02)
14	House	2011	0.3112 (0.0128)	6.6224 (0.0233)	0.0278 (0.0118)	0.0082 (0.0041)	0.0543 (0.0057)	-0.01 (0.02)	-0.00 (0.01)	0.07 (0.02)
15	House	2012	0.3108 (0.0111)	6.7554 (0.0166)	0.0001 (0.0016)	0.0387 (0.0073)	0.0800 (0.0052)	-0.06 (0.02)	-0.01 (0.02)	0.23 (0.04)
16	House	2013	0.2984 (0.0096)	6.7696 (0.0290)	0.0290 (0.0145)	0.0141 (0.0068)	0.0591 (0.0080)	0.01 (0.02)	0.00 (0.01)	0.08 (0.04)
17	House	2014	0.3051 (0.0102)	6.7447 (0.0219)	0.0177 (0.0060)	0.0134 (0.0043)	0.0680 (0.0069)	0.01 (0.01)	-0.01 (0.01)	0.07 (0.02)
18	House	2015	0.3069 (0.0100)	6.8235 (0.0250)	0.0234 (0.0060)	0.0145 (0.0057)	0.0815 (0.0075)	0.02 (0.01)	-0.01 (0.01)	0.08 (0.03)

B Omitted proofs from model

Proof of Lemma 1

To show that W is increasing in θ, note that for $B=0$,

$$W(0,\theta,e,u_i)=(w(\theta,u_i)-w(\theta,e))=\exp\{\theta\}(\exp\{u_i\}-\exp\{e\}),$$

and similarly for $B>0$,

$$\begin{aligned}W(B,\theta,e,u_i)=\exp\{\theta\}\Pr\left\{(\boldsymbol{u}_{-i})_{(B)}<u_i\right\}\times\\ \left\{\exp\{u_i\}-E\left[\exp\left\{\max\left\{(\boldsymbol{u}_{-i})_{(B)},e\right\}\right\}\middle|(\boldsymbol{u}_{-i})_{(B)}<u_i\right]\right\}.\end{aligned}$$

Direct inspection of these expressions immediately implies that W is increasing in θ.

To show that W is increasing in u_i and decreasing in e, note that this is trivial for $B=0$. For $B>0$, we rewrite W as

$$\begin{aligned}W(B,\theta,e,u_i)=\exp\{\theta\}\Big[\Pr\left\{(\boldsymbol{u}_{-i})_{(B)}<u_i\right\}\exp\{u_i\}\\ -\Pr\left\{(\boldsymbol{u}_{-i})_{(B)}<e\right\}\exp\{e\}-\int_e^{u_i}\exp\{x\}\psi_B(x)\,dx\Big],\end{aligned}$$

where $\psi_B(x)$ denotes the probability density function of the largest order statistic of $(\boldsymbol{u}_{-i})$. Differentiating with respect to u_i, we get

$$\begin{aligned}\frac{\partial W}{\partial u_i}&=\exp\{\theta\}\left[\psi_B(u_i)\exp\{u_i\}+\Pr\left\{(\boldsymbol{u}_{-i})_{(B)}<u_i\right\}\exp\{u_i\}-\exp\{u_i\}\psi_B(u_i)\right]\\ &=\exp\{\theta\}\Pr\left\{(\boldsymbol{u}_{-i})_{(B)}<u_i\right\}\exp\{u_i\}>0.\end{aligned}$$

Similarly, differentiating with respect to e, we get

$$\begin{aligned}\frac{\partial W}{\partial e}&=\exp\{\theta\}\left[-\psi_B(e)\exp\{e\}-\Pr\left\{(\boldsymbol{u}_{-i})_{(B)}<e\right\}\exp\{e\}+\exp\{e\}\psi_B(e)\right]\\ &=-\exp\{\theta\}\Pr\left\{(\boldsymbol{u}_{-i})_{(B)}<e\right\}\exp\{e\}<0.\end{aligned}$$

Finally, to show that W is decreasing in B, first of all note that

$$W(0,\theta,e,u_i)\geq W(1,\theta,e,u_i),$$

since $\Pr\left\{(\boldsymbol{u}_{-i})_{(B)}<u_i\right\}\leq 1$ and $E\left[\exp\left\{\max\left\{(\boldsymbol{u}_{-i})_{(B)},e\right\}\right\}\middle|(\boldsymbol{u}_{-i})_{(B)}<u_i\right]\geq\exp\{e\}$.

For $B>1$, note first that

$$\Pr\left\{(\boldsymbol{u}_{-i})_{(B)}<u_i\right\}=\Psi(u_i)^B$$

is decreasing in B, where $\Psi(x)$ is the cumulative distribution function of a Type 1 extreme value distribution, so $\Pr\left\{u_{(B)} < u_i\right\}$ is decreasing in B. Second, note that

$$(\boldsymbol{u}_{-i})_{(B)} = \max\left\{u_1, u_2, u_3, ..., u_{i-1}, u_{i+1}, ..., u_{B-1}, u_B\right\} \geq \\ \max\left\{u_1, u_2, u_3, ..., u_{i-1}, u_{i+1}, ..., u_{B-1}\right\} = (\boldsymbol{u}_{-i})_{(B-1)},$$

and so

$$\exp\left\{\max\left\{(\boldsymbol{u}_{-i})_{(B)}, e\right\}\right\} \geq \exp\left\{\max\left\{(\boldsymbol{u}_{-i})_{(B-1)}, e\right\}\right\}.$$

These two observations directly imply that

$$W(B-1, \theta, u_i) \geq W(B, \theta, u_i),$$

for $B > 1$.

D Details on the estimation

Simulated method of moments

We describe in more detail the estimation procedure for the extension of the model, where all parameters are estimated year-by-year. We largely follow Hennessy and Whited (2007). For each segment, $j \in J$, we estimate the parameters, $\Phi_j = \{\sigma_u, \sigma_e, \sigma_\eta, \sigma_\theta, \bar{e}, \bar{\eta}, c, \bar{\theta}\}$, that satisfy

$$\bar{\Phi}_j = \underset{\Phi}{arg\,min}\, \bar{m}_K(X_j, \Phi) W_N^j \bar{m}_K(X_j, \Phi), \tag{D.1}$$

where W_N^j is a weighting matrix. We set it to the inverse of the covariance matrix of the data moments, V^j. The empirical variance-covariance matrix is computed by bootstrap separately for each segment. Also,

$$\bar{m}_K(X_j, \Phi) = \frac{1}{K} \sum_{k=1}^{K} m_k(X_j, \Phi), \tag{D.2}$$

where $m_k(X_j, \Phi) = m_n^d(X_j) - m_n^s(\Phi)$, a vector of empirical moments less their simulated counterparts. Each $m_k(X_j, \Phi)$ is computed from n simulated houses, equal to the number of houses sold in that segment. This number is typically too low for computational purposes. Therefore, the simulation is repeated K times for the same Φ. The moment used to compute the loss, $\bar{m}_K(X_j, \Phi)$, is the average moment of the K runs.

The model variance-covariance matrix is computed from

$$\left(1 + \frac{1}{K}\right) \left(J^T W_N^j J\right)^{-1} J^T W_N^j J \left(J^T W_N^j J\right)^{-1} \tag{D.3}$$

where J is the Jacobian matrix of $\bar{m}_K(X_j, \Phi)$ over Φ computed by central finite differences.[30]

For the baseline estimation, where some parameters are fixed across years within each housing type, we estimate the parameters as follows. For each group of segments, $g \in G$, we estimate the parameters, $\hat{\Phi}_g = \bigcup_{j \in g} \Phi_j$, that satisfy Equation (D.1) and specified time invariance constraints (e.g., $\sigma_u = \sigma_u^j \forall j \in G$). Now,

$$\bar{m}(X_g, \hat{\Phi}_g) = \text{vec}\left(\left\{\frac{1}{K_j}\sum_{k=1}^{K_j} m_{k_j}(X_j, \Phi_j)\right\}_{j \in g}\right) \tag{D.4}$$

with $m_{k_j}(\cdot)$ as before. The data variance-covariance matrix used for both the weighting matrix and for computing the model variance-covariance matrix is

$$V^g = \text{blockdiag}\left(\{V^j\}_{j \in g}\right) \tag{D.5}$$

where *blockdiag* generates a block diagonal matrix of its inputs. Note that since the off-block diagonal entries are zero, we are assuming that there is no correlation across years. The variance-covariance matrix is computed as in the baseline model.

Estimation algorithm

For each segment, with n sales, the parameters, Φ, are estimated as follows.

1. Make an initial guess for the parameters, $\Phi = \Phi_0$.
2. Initialise K datasets, with $K = \underset{k}{arg\,min}\{n \times K \geq N\}$. For each dataset,
   (a) Draw n of μ, θ, B_p, e, η.
   (b) Find the fixed-point of Eq. (15), $\hat{u}(\theta)$, with Brent's method.
   (c) For each house, draw B_p of u. Compute prices according to Eq. (9). Compute moments, $m_k(X_j, \Phi)$.
3. Average moments across the K datasets and compute $m_K(X_j, \Phi)$.
4. Compute loss function, $m_K(X_j, \Phi) W_N^j m_K(X_j, \Phi)$. According to the Nelder-Mead algorithm, evaluate the innovation in the loss function and, if required, updated the parameter guess, Φ, and repeat steps 2 through 4.[31]

[30] For parameters estimated close to or at range constraints, we compute the forward or backward finite difference.

[31] For some extensions we rely on the differential evolution algorithm due to Storn and Price (1997).