Seller's (Mis)Fortune in the Housing Market: Directed Search in Online Real Estate Platforms

Jean-Louis Barnwell* & Jean-François Fournel†

April 21, 2022

WORKING PAPER

Latest version available here.

Abstract

Algorithm-based market valuations for houses, such as the Zillow's Zestimate, impact trading outcomes in the housing market. Sellers who advertise an asking price below their Zestimate increase buyers' search intensity, shorten their time on market, and reduce their sales price, irrespectively of sellers' preferences and home characteristics. Using data about the home selling process on Zillow in the Seattle metropolitan area, we estimate the cost associated with this tradeoff is $3,600 (0.75%) of the house sales price for one fewer day on market. Despite this high cost, we show that it is still rational for a seller to advertise an asking price below the Zestimate if there exists a distance between his reservation value and the Zestimate. Our model implies that heterogeneity in house trading outcomes can arise from homogenous sellers' (mis)fortune in receiving a low or a high Zestimate.

*McGill University. Email: jean-louis.barnwellmenard@mail.mcgill.ca

†McGill University. Email: jean-francois.fournel@mcgill.ca

1 Introduction

Digital innovations are changing how buyers, sellers, and real estate agents interact in the housing market. The introduction of online real estate platforms such as Zillow (www.zillow.com) have transformed the search and matching process that precedes housing transactions. The rising popularity of these platforms is attributed to several factors. First, they contribute to lowering search costs by sharing detailed information about houses for sale in a centralized environment. Some platforms provide estimates for the market valuation of houses, whether they are on the market or not. For example, Zillow's Zestimate predicts market values through a proprietary machine learning algorithm which combines house and neighborhood characteristics, market trends, and past sales of comparable homes.

Publicly available house valuations have become central to the transaction process. However, their impact (if any) on key market outcomes is still not well understood. In this paper, we study if and how algorithm-based market valuations affect sellers' strategies (the decision about the initial asking price), buyers' search intensity, and the outcomes of the bargaining process (the expected time on market and sales price). We document that advertising an asking price below the Zestimate is associated with more virtual visits, a shorter time on market, and a lower sales price, irrespectively of sellers' preferences and home characteristics. Our estimates suggest that the cost associated with this tradeoff is 0.75% of the house's sales price for one fewer day on market. To rationalize why some sellers would chose to advertise an asking price below the Zestimate despite this high cost, we propose a theoretical model of directed search in the spirit of Albrecht, Gautier, and Vroman (2014, 2016). We show that it is sufficient to introduce a distance between a seller's reservation value and the Zestimate for the model to replicate the observed sellers' strategies, and the resulting market outcomes. More importantly, our model implies that heterogeneity in trading outcomes does not only arise from differences in sellers' preferences, but can also be the result of homogenous sellers (mis)fortune in receiving a low or a high Zestimate.

Our empirical analysis exploits data on daily listing views for each property advertised on Zillow in the Seattle metropolitan area. We find evidence of directed search in online real estate platforms. Houses advertised with an asking price below the Zestimate receives more listing views from potential buyers. These additional listing views are associated with a shorter number of days on the market and a sales prices above the initial asking price. We want to identify the causal links between sellers' initial decision about the asking price, buyers' search intensity, and the resulting terms of trade. Naturally, the asking price can reflects unobservable home characteristics (such as a leaky roof or bad foundations) or sellers' impatience.

To account for the presence of these potential confounders, we propose an instrumental variable (IV) strategy. We use the fact that some sellers arbitrarily choose to round the asking price in multiples of \$100 to recover exogenous variation in the asking price. Our IV estimates suggest that reducing the asking price by 1% relative to the Zestimate leads to 11.48 (4.22%) more listing views on day 2, 1.125 (1,73%) fewer days on the market, and a decrease in the sales price of \$4,046 (0.84%) for the average property. This illustrates the tradeoff faced by sellers when deciding on an asking price: either advertise on-the-low to generate more virtual visits and a faster sale, or advertise on-the-high to get a higher sales price. Our results imply that reducing time on market by one day costs \$3,598 (0.75%) in sales price.

To rationalize why some sellers' advertise an asking price below the Zestimate in the presence of such high costs, we propose a theoretical model that expands on the work of Albrecht, Gautier, and Vroman (2014, 2016). We show that it is sufficient to introduce a distance between a seller's reservation values and the Zestimate to generate the observed sellers' strategies. In our setup, buyers and sellers receive a public signal about the market value of the house, but sellers' private value is unknown to potential buyers. We assume that sellers use the asking price as a signalling device to which buyers' respond using an optimal directed search strategy. Intuitively, sellers have the choice to either advertise on-the-low (below the market valuation) to attract more buyers, or advertise on-the-high (above the market valuation) to increase the expected sales price. Prior to meeting potential buyers, sellers would like to advertise on-the-low to increase the probability of meeting one or several buyers, but would have preferred to have advertised on-the-high once matches are realized.

We consider the case of two identical sellers with identical properties (i.e. identical reservation values), where one received received a good market valuation (above the common reservation value) and the other a bad one (below the common reservation value). In equilibrium, it is optimal for both sellers to advertise their house at their reservation value. Under weak conditions, the fortunate sellers attracts strictly more buyers than the unfortunate seller. This happens because the fortunate seller is advertising an asking price below his Zestimate while the unfortunate seller advertises an asking price above his Zestimate. Since both are offering the same increasing bid auction, the fortunate seller wins on all fronts, enjoying both a higher probability of selling and a higher expected price than the unfortunate seller.

We contribute to several trends of literature. Our findings relate to Lu (2018), Lu (2019), Yu (2020), and Fu, Jin, and Liu (2022), who study the impact of publicly available market valuations in the housing market. Their results suggest that Zestimates impacts asking price decisions, time on market and the sales price. We contribute to this literature by providing an explicit mechanism that explains how the Zestimate impacts those trading outcomes. Our research is also related to the growing literature studying search frictions and directed search in the housing market.[1] To our knowledge, we are the first to provide formal empirical evidence of directed search in online real estate platforms. Our analysis complements the work of Han and Strange (2016) who shows that low asking prices generate more bids in the housing market. Finally, we contribute to the literature that study the impact of artificial intelligence on human interactions and decision making.[2]

The paper is organized as follows. Section 2 describes our data and provides empirical evidence of directed search in online real estate platforms. In Section 3, we propose a theoretical model that rationalizes our empirical findings and extends our understanding of real estate markets in the digital era. We conclude in Section 4.

[1]See Wheaton, 1990; Carrillo, 2012; Genesove and Han, 2012; Diaz and Jerez, 2013; Albrecht, Gautier, and Vroman, 2016; Anenberg and Bayer, 2020; Piazzesi, Schneider, and Stroebel, 2020.

[2]See Cowgill (2018) and Hoffman, Kahn, and Li (2018) (labor market hires), Kleinberg et al. (2018), Ludwig and Mullainathan (2021), and Stevenson and Doleac (2021) (justice system), or Assad et al. (2020) and Calvano et al. (2020) (competition).

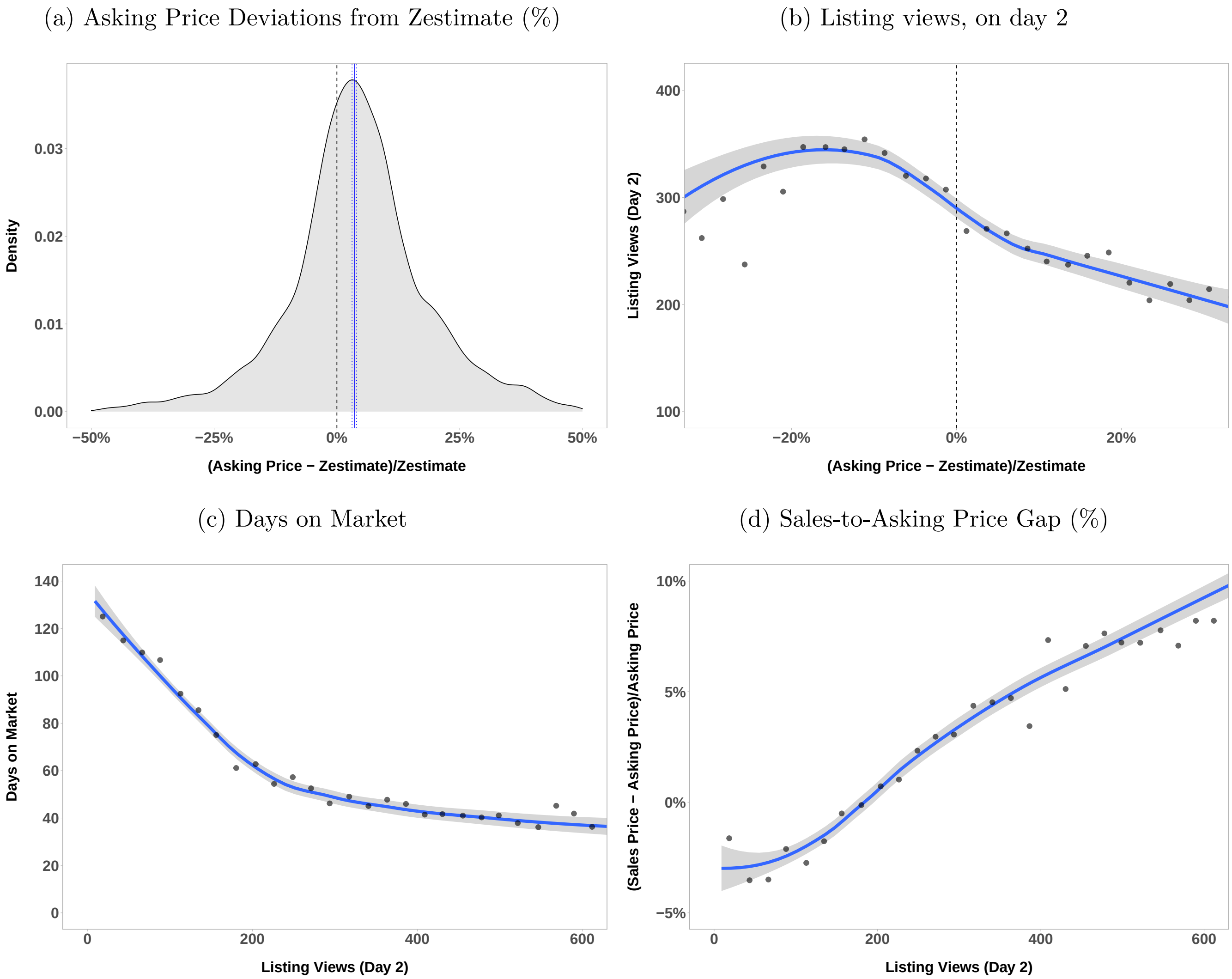
2 Evidence of Directed Search in Online Real Estate Platforms

We provide evidence of directed search in online real estate platforms for house transactions in the Seattle metropolitan area in 2015-2016. The timing of a particular transaction is as follows. First, sellers observe their property's Zestimate on Zillow's website, prior to listing the house. Throughout, we will consider this off-market valuation to be predetermined and exogenous from the point of view of all agents involved in the transaction process. Next, sellers advertise their property and post an asking price. We consider the decision of the asking price, conditional on the Zestimate, to be a signalling device used by sellers. Finally, buyers observe both the asking prices and the Zestimates, visit houses, and make offers. If buyers are using an optimal directed search strategy then sellers face the following tradeoff when advertising their home: either advertise on-the-high and increase the likelihood of getting a high price, or advertise on-the-low to increase the number of visits from buyers, leading to a higher probability of receiving simultaneous offers and a faster sale.

We seek to identify these causal links tying the initial seller's strategy to the observed market outcomes, through shifts in buyers' search intensity. We propose a model in two steps. First, we identify the causal effect of the asking price on buyers' search intensity, measured by virtual visits (i.e. listing views). Then we take predicted listing views from the first stage to identify the causal effect of directed search on key market outcomes, such as the sales price or the time on market. This identification strategy allows us to link sellers' strategy to market outcomes through changes in buyers' search intensity, which illustrate the tradeoff faced by sellers when choosing an asking price.

Figure 1 presents some initial evidence that these forces are indeed present in the Seattle housing market. Panel (a) depicts the distribution of the difference in percentage between the initial asking price and the Zestimate taken in the month prior to entry. While on average, sellers tend to advertise above the Zestimate, a large share of sellers choose to advertise below. To motivate our story that sellers are indeed using the Zestimate as an anchor point when signalling with the asking price, and highlight the tradeoff mentioned above, we turn to Panel (b), (c), and (d). Panel

Figure 1: Evidence of Directed Search

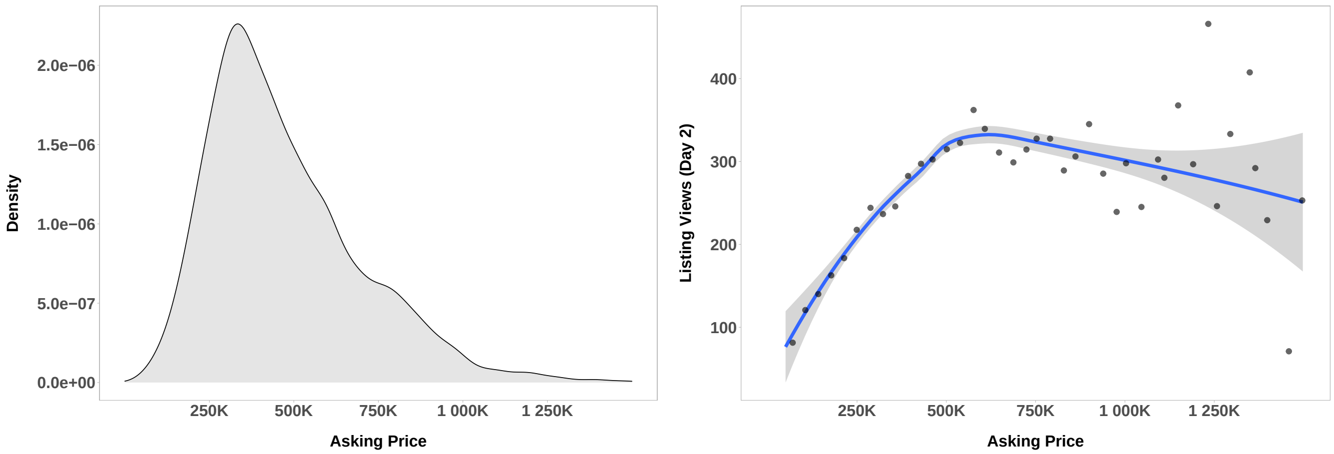


Notes: Sample based on 4,322 houses sold in the Seattle metropolitain area in 2015 and 2016. Panel (a) depicts the kernel density estimate of the percentage difference between the asking price and the Zestimate, the latter being evaluated on the month prior to entry. The solid and dotted blue lines represent the distribution mean and confidence interval at 95%. Panel (b) depicts the fit of a local polynomial regression between virtual visits on day 2 and the percentage difference between the asking price and the Zestimate. Panels (c) and (d) illustrate the fit of local polynomial regressions of the number listing views on day 2 (y-axis) on two key market outcomes: time on market in days, and the percentage difference between the sales price and asking prices (x-axis).

(b) depicts buyers' response to the signal in terms of search intensity, captured by the number of virtual visits to the house. We see a clearly decreasing pattern over most of the range of sellers strategies. For houses listed more than 20% below their Zestimate, we do observe a reversal of the effect: buyers visits increase with price. One possible explanation for this is that for such levels of

Figure 2: Buyers' Search Activity Unconditional on the Zestimate

(a) Asking Prices (b) Listing views, on day 2



Notes: Sample based on 4,322 houses sold in the Seattle metropolitain area in 2015 and 2016. Panel (a) depicts the kernel density estimate of the asking prices. Panel (b) depicts the fit of a local polynomial regression between virtual visits on day 2 and the the asking price.

discount, some buyers interpret the signal as revealing some unobserved problem with the house, for example a leaky roof or bad foundations. Next, Panel (c) and (d) show that these virtual visits translate into better outcomes for the seller. We see that more listing views correlates negatively with time on market (i.e. lead to a fast sale). In addition, more listing views correlates positively with the sales-to-asking price gap. Taken together, those facts suggest that sellers use the asking price as a signalling device, that there is a response from buyers in terms of virtual visits, and that those extra visits impact market outcomes.

To emphasize the importance of the Zestimate in this signalling game, we consider the alternative story that the asking price itself, not anchored to the Zestimate, is a sufficient signalling device. It is interesting to see what happens to the observed directed search patterns in this case. Results are presented in Figure 2. It is easy to see from Panel (b) that the patterns observed above vanish: we do not find any evidence that a low asking price generates more visits from buyers. In fact, buyers' search intensifies with the asking price over a large portion of the support, and declines only for luxurious houses. We take these findings as evidence towards our story of directed search.

The Zestimate provides some useful information to buyers and sellers about the value of any given house, and deviation from this common knowledge is what drives directed search.

2.1 Data

We assemble a rich dataset of houses advertised on Zillow in the Seattle metropolitan area in 2015-2016. The dataset is constructed by linking the street address of houses that entered the Multiple Listing Service (MLS) records as of January 2016 to their property specific listing on Zillow's website. House specific, and transaction specific data were retrieved through web scrapping realized in late 2016 and early 2017. The depth and the quality of the information retrieved provides new and instructive facts about the selling process in online real estate platforms.

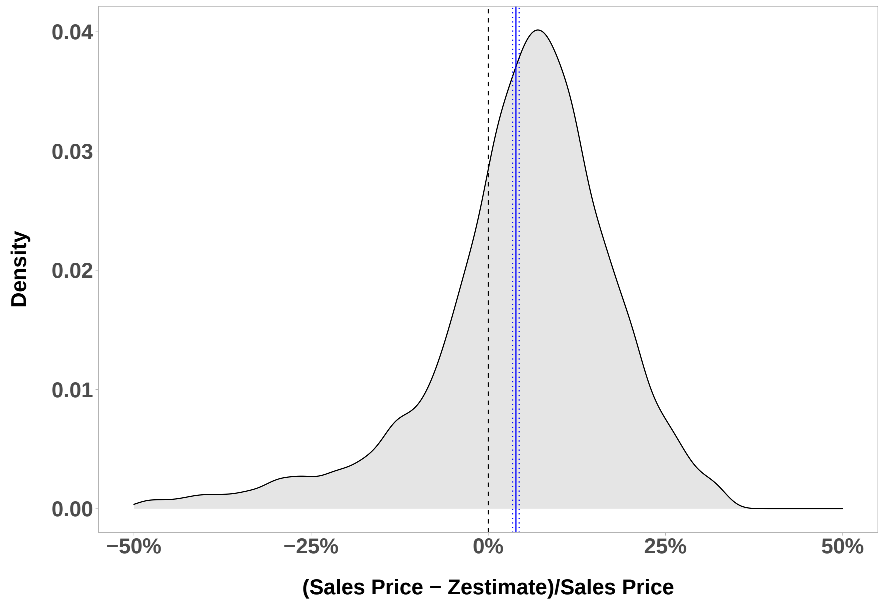
Our dataset includes two unique sources of information which we leverage in our estimation of the directed search mechanism. First, we observe monthly historical Zestimates series[3] which proxy each property's valuation through time. These valuations are estimated using Zillow's proprietary algorithm and aim at predicting the transaction price of each property.[4] We observe these valuations for houses, whether on the market or not, in a period prior to the moment they are put up for sale. For the purpose of this study, we focus on the Zestimate estimated one month prior to entry, that is when the house is off the market. We believe that this summarizes best the information available to sellers at the time when the entry decision is made and the asking price is determined (daily Zestimates are not available). This also allow us to treat the Zestimate as predetermined to the transaction process.[5] Zillow constantly improve their algorithm and update all historical series accordingly. The precision of the Zestimate and its ability to predict prices is one of Zillow's trademark. Figure 3 depicts the kernel density of the error rate, Zillow's preferred measure of precision. We notice that their algorithm on average underestimate sales prices and that there is significant dispersion: some Zestimates can be as far as 50% off the sales price. We estimate the median (absolute) error rate to be around 9.2% using the last Zestimate that is available before a

[3]See Panel (c) of Figure 5 in Appendix A.

[4]The algorithm factors in house characteristics, past sales, market conditions, market trends, and also a set of comparable houses valuations. *Source*: https://www.zillow.com/z/zestimate.

[5]Recent literature highlights the fact that the Zestimate updates once the asking price is advertised, leading to an endogeneity issue. See Fu, Jin, and Liu (2022).

Figure 3: Zestimate Algorithm Precision



property enters the market. In 2022, Zillow advertised instead a median error rate of 2% for on-market and 6.9% for off-market properties. This provides evidence that significant improvements have been made to the algorithm over time. Since our data was collected up to one year following the transactions, we do not observe the exact value that was available at the time when the house was initially listed. We are aware of one substantial change to their algorithm between the period that covers our study and data collection, and all historical Zestimates were updated following this improvement (before data collection).

We also observe the historical daily series for the number of listing views, starting on the day each house entered the market up to the day the sale is finalized. To fix ideas, a listing view occurs when a potential buyer clicks on a particular listing on Zillow's website to access the house specific information page.[6] Some information is available to potential buyers prior to clicking on a particular listing, for example the asking price, the number of bedrooms and bathrooms, the squared footage, and a front picture of the house. The Zestimate however is only accessible after clicking on the listing, along with additional pictures and other details about the property. Throughout, we focus

[6]See Panel (a) and (b) of Figure 5 in Appendix A.

on listing views on day 2 to measure buyers' interest in a given house. This deals with two selection issues. As we do not observe the exact time of the day each house is posted, using listing views on day 1 leads to biased parameter estimates as houses posted in the morning would have mechanically more views than those posted in the evening. Another issue arise if we use listing views beyond day 2. Houses frequently sell within the first week and potential buyers may decide not to click on a house with a pending offer. To our knowledge, this is the first work that features such a precise and direct measure of buyers' search intensity in the housing market.[7]

Our data also includes home characteristics, such as the number of bedrooms or squared footage, and transaction specific variables. In particular, we observe all sales prices and the full history of asking prices if for example the owner changed it over time. We keep the initial asking price as the relevant value for this study, and consider updates to be part of the sales process. We also observe the timing of the selling process and derive our measure of time on market as the number of days between the moment the house is put up for sale and when the transaction is completed. Our sample is restricted to single-family residences and condos that were transacted in the Seattle metropolitan area between 2015 and 2016. We further restrict the sample in the following ways. Properties with a Zestimate below the 1st and above the 95th percentile of the distribution were removed to limit the impact of outliers and very luxurious houses on our results. We also remove houses for which the sales- or the asking price-to-Zestimate ratio is below 0.5 or above 1.5. These houses were poorly estimated by the algorithm and the Zestimate doesn't carry any useful information with respect to their value. We remove outliers in our other key variables (listing views, time on market, asking price, and sales price) by censoring the first and last percentiles of the respective distributions. Finally, we remove houses that exited and reentered the market before the final sale occurred (about 7% of the remaining sample). Our final dataset contains information on 4,322 transactions.

[7]This information is no longer available on Zillow, and was replaced by cumulative listing views over the full period each house was on the market. This is problematic as house that stay for a long time on the market will mechanically have more listing views in total. The change to the this new measure of listing views occurred as we were still scrapping data, prompting us to stop data collection. As we were scrapping data more or less chronologically and that the change was unpredictable, we believe that this doesn't affect our results in a meaningful way.

Table 1: Summary Statistics

	Asking price $\leq$ Zestimate		Asking price > Zestimate	
	Mean	Std. Dev.	Mean	Std. Dev.
Asking price	437,094	210,540	492,325	230,424
Sales price	456,730	221,552	497,128	233,911
Zestimate	480,241	222,579	443,275	207,960
Days on market	61.0	51.6	67.2	54.3
Listing views, on day 2	320.6	196.2	247.9	163.5
Living space, in sq. ft	1949.9	985.5	1962.1	1066.2
Nb. of bedrooms	3.2	1.2	3.1	1.1
Nb. of bathrooms	2.1	1.0	2.1	1.0
N	1,507		2,815	

Notes: Sample based on 4,322 properties sold in the Seattle Metropolitan Area in 2015 and 2016.

Table 1 reports summary statistics for all transactions included in our final dataset. We split the summary statistics by sellers strategies, broadly defined as "advertised above" and "advertised below" their Zestimate. We observe that roughly 35% of sellers choose an initial asking price below or equal to their Zestimate while 65% chose to advertise above. The first group of sellers exhibit on average an asking price 11% below their Zestimate and a sales price 8% below their Zestimate. This could suggest that multiple offers leading to overbidding is frequent for these properties. Houses advertised below their Zestimate sell on average 9% faster, and receive 29% more virtual visits (on day 2), while their asking price and the sales price are on average very similar, suggesting that overbidding is not as frequent.

2.2 Empirical Strategy

We now estimate the directed search mechanism described above. The estimation is in two stages. In the first stage, we estimate the following directed search equation:

$$\ln(ListingViews_j) = \beta_0^I + \beta_1^I \ln(AskingPrice_j) + \beta_2^I \ln(Zestimate_j) + X_j\beta_3^I + \epsilon_j^I, \qquad \text{(Model I)}$$

where $ListingViews_j$ is the number of virtual visits to house j on day 2. The coefficient of interest is β_1^I, the elasticity to the asking price of buyer's search. We include Zestimates in the model, since previous evidence suggests that sellers' strategies are chosen conditionally on the Zestimates. Another reason for including them is that they account for several unobserved factors to the econometrician that could affect buyers and sellers together while being picked up by Zillow's algorithm (in some sense correcting an omitted variable bias). X_j contains other observable characteristics, such as living space and the number of bedrooms. We use the predicted values from Model I (i.e. our estimate of the intensity of buyers' search generated by sellers' strategies) and estimate its impact on key market outcomes. Let y_j be one such market outcome, for example time on market. We estimate the following second stage equation:

$$\ln(y_j) = \beta_0^{II} + \beta_1^{II} \ln(\widehat{ListingViews_j}) + \beta_2^{II} \ln(Zestimate_j) + X_j\beta_3^{II} + \epsilon_j^{II}. \qquad \text{(Model II)}$$

The parameter of interest is β_1^{II}, the buyer's search elasticity of outcome y. Both estimates together have an important economic interpretation. First, they allow us to understand the causal links between sellers' strategies, buyers' responses, and the resulting market outcomes. They also measure the tradeoff faced by sellers, allowing us to quantify how costly it is to reduce the asking price relative to the Zestimate in terms of expected sales price versus expected time on market.

Our identification of the causal effect relies on the fact that the decision about the asking price is orthogonal to unobservables in both Model I and II. This is of course unrealistic. The decision about the asking price depends on factors that are observed by both agents, but unobservable to Zillow or to the econometrician. These unobservables can include for example house conditions that are revealed at the inspection stage, but not disclosed on the house's listing. This leads to an endogeneity issue as these confounders could affect both the asking price, buyers' response, and the resulting terms of trade. We propose an instrumental variable approach to identify the causal effects in Model I and II. The identifying assumption is that the instrument is relevant and satisfy the exclusion restriction in both models, since predicted values from Model I are used in Model II. Our instrumental variable strategy leverages the fact that some sellers arbitrarily round prices in multiples of \$100 when posting an asking price, while others do not, providing an exogenous variation in the asking price that we exploit for identifying the causal effect in Model I. We discuss the construction of the instrument, it's relevance, and address potential threats to its validity in what follows.

2.3 Identification

2.3.1 Two-Digit Rounding as an Instrument for the Asking Price

We propose a new instrumental variable to correct for the endogeneity of the asking price in our directed search model. We leverage the fact that some sellers arbitrarily choose to round the asking price to the two-digit level while others do not. The idea is that the choice about rounding at the two-digit level is not correlated with unobservables in both models, but does trigger the response in buyers' search intensity. Specifically, the instrument is:

$$\mathrm{IV}_j = \mathbb{1}\{\textit{AskingPrice}_j\text{'s Right Two-Digit are ``00''}\},$$

where $\mathbb{1}\{\cdot\}$ is the indicator function. For example, the IV takes a value of 1 if the initial asking price of a house is \$499,900, and a value of 0 if it is instead \$499,850 or \$499,899.

The exclusion restriction of our instrumental variable strategy relies on the idea that variations under \$100 of the initial asking price for a durable good like a house (which sells for \$480,000 on average in our sample) is too small to result from unobservable home characteristics or heterogeneity in sellers' preferences. Our identifying assumption is that two identical houses could be sold by two identical sellers, where one arbitrarily chooses to round the asking price at the last two digits, say at \$499,900 while the other chooses to advertise \$499,899. If this rounding choice is arbitrary, then our instrumental variable provides exogenous variation in the initial asking price that can be used to identify the causal effects in our directed search model.

Our instrumental variable strategy relies nevertheless on a decision taken by the seller. Our exclusion restriction would be violated if for example the decision about rounding lead to discontinuities in realized market outcomes. There are to our knowledge two major threats to identification that need to be addressed at this point. First, recent literature shows that sellers' use rounding to signal their bargaining position or their degree of loss-aversion.[8] We refer to this phenomenon as "cheap talk signalling". It would be problematic in our context if this phenomenon occurred at the two-digit level of rounding. However, we believe that for a durable good like a house, signalling a weak bargaining position would require more aggressive rounding, say at the $10,000 or the $25,000 levels. The second threat to our instrumental variable strategy is what we call "charm pricing" or "psychological pricing". The idea is that sellers take advantage of buyers' left-digit bias when setting the asking price to generate discontinuities in realized market outcomes.

We discuss in more details the implications of "cheap talk signalling" and "charm pricing" in what follows. We then test our exclusion restriction formally using the methodology in Backus, Blake, and Tadelis (2019). We want to verify empirically if our instrumental variable generates discontinuities in observed market outcomes which could be used by sellers to discretely increase their revenue by changing the asking price marginally. Finally, we address the relevance of our instrument and test for weak identification.

2.3.2 Threats to Identification: Cheap Talk and Charm Pricing

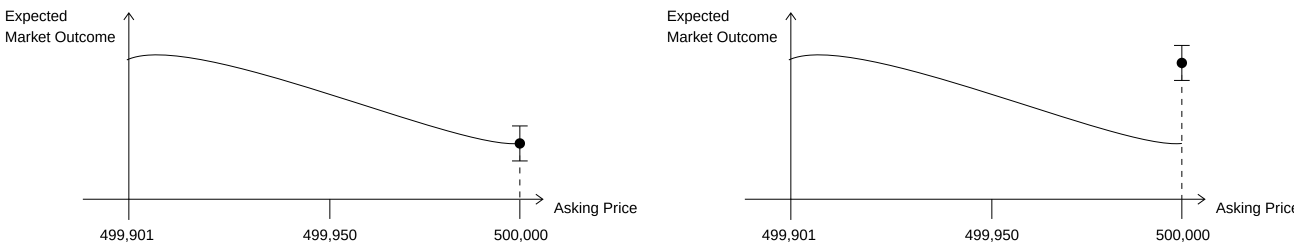
Cheap talk – The first threat to our instrumental variable strategy relates to "cheap talk signalling". Empirically, we are concerned that rounding leads to point discontinuities in the observed market outcomes of two houses that are identical in all respect but the rounding of the initial asking price. If such point discontinuities exists, then sellers can take advantage of them to increase revenues and the decision to round or not becomes strategic (hence endogenous to the pricing game). To fix idea, Figure 4 presents an example of what "cheap talk signalling" could look like empirically.

"Cheap talk signalling" can occur if buyers and sellers agree that rounding the asking price constitutes a credible signal about the seller's willingness to sell or about the unobserved quality of its

[8]See for example Backus, Blake, and Tadelis (2019) for evidence from Ebay auctions.

Figure 4: Asking Price Rounding and Point Discontinuity in Expected Market Outcome

(a) Absence of Point Discontinuity (b) Presence of Point Discontinuity



Notes: Panel (a) shows a case where there is no point discontinuity at $r = \$500,000$, while Panel (b) presents a case where a point discontinuity exists at $r = \$500,000$.

property. In our setup, this means that sellers would use the distance to the Zestimate together with rounding as a signalling device. Buyers then respond to rounding by discretely increasing demand for these houses, which leads to a discrete jump in market outcomes at the point of discontinuity.

Backus, Blake, and Tadelis (2019) find evidence of point discontinuities in their study of online auctions. Using millions of bargaining interactions on eBay's Best Offer platform, they find that goods initially advertised at prices in multiples of $100 received lower offers and were more likely to sell. In this instance, rounding was used as a device to signal a weak bargaining position to which buyers responded. They also study the effect of rounding in the housing market and find similar evidence. However, the level of rounding that is considered is at the $50,000 level, several orders of magnitude above our instrumental variable.

Another source of "cheap talk signalling" concerns aversion to losses.[9] If sellers round the asking price to signal loss aversion, then rounding would be correlated to sellers' characteristics which are captured by unobservables in our model. This would also lead to discontinuities in observed market outcomes and invalidate our identification strategy.

Charm pricing – The second threat to our instrument validity is "charm pricing" or "psychological

[9]See Genesove and Mayer (2001) or Ross and Zhou (2021) for examples.

pricing". The idea is that sellers can take advantage of buyers' left-digit bias by advertising slightly below large round numbers, say $499,999 instead of $500,000. Evidence of "charm pricing" is well documented in the housing market[10] for multiples of large numbers, for example $50,000 or $100,000. We are not aware of any study that reports similar effects at the level of rounding that we use to create our instrumental variable. Nevertheless, it may be an issue if for example a large share of the rounded asking prices coincide with or are concentrated around these multiples of $50,000 or $100,000 at which we suspect "charm pricing" operates. Empirically, "charm pricing" would result the same types of discontinuities in observed market outcomes as "cheap talk signalling".

2.3.3 Testing for Point Discontinuity

We propose a formal test to address the validity of our instrumental variable strategy. Specifically, we want to know if rounding at the two-digit level leads to point discontinuities in observed market outcomes. We adapt the methodology in Backus, Blake, and Tadelis (2019) to our setup. The idea is to compare the expected market outcomes of two nearly identical houses, one priced at the rounding discontinuity and the other priced slightly off the discontinuity. If at the limit both houses have (statistically) similar outcomes, for example in terms of listing views, time on market, and sales price, then we take it as evidence against the presence of "charm pricing" or "cheap talk signalling". On the other hand, observing significant differences would invalidate our instrumental variable strategy.

Due to the limited size of our data sample, we depart from Backus, Blake, and Tadelis (2019) in the following way. Since each $100 bin is not populated enough to estimate a different functional form around each discontinuity, we pool observation across bins to estimate a single non-parametric functional form and allow for bin-specific dummies. We also exclude bins that are not well populated from the analysis.[11] We then test for discontinuity in the same way as in Backus, Blake, and Tadelis (2019), that is, we test for significance of our instrumental variable in explaining differences in market outcomes at the limit.

[10] See Allen and Dare (2004), Beracha and Seiler (2015), or Repetto and Solís (2020) for examples.

[11] We require bins to have at least 10 observations at the discontinuity and 10 observations off the discontinuity. In a robustness check, we further restrict to bins with 20 observations on and off the discontinuity.

Let $r \in R$ be finitely many discontinuities. In our setup, one such discontinuity occurs when the lasts two digits of the asking price are zeros. For our instrumental variable strategy to be valid, the expected market outcomes evaluated at discontinuity r must be the same as the expected market outcome evaluated slightly off discontinuity r. Let a be the asking price and define ψ_r as

$$\psi_r(\Delta) = E(y \mid a = r) - E(y \mid a = r - \Delta).$$

Then our identifying assumption requires that

$$\lim_{\Delta \to 0} \psi_r(\Delta) = 0.$$

Define the neighborhood around discontinuity r to be the interval $B_r = (r - 100, r]$. Let the expected market outcome of house j in bin B_r be

$$E(y_{jr} \mid a_j, r) = f(\text{dist}_{jr}) + \beta_r^{bin} \cdot \mathbb{1}\{a_j \in B_r\} + \beta_r^{round} \cdot \mathbb{1}\{a_j = r\},$$

where dist_{jr} is the distance between the asking price a_j and the discontinuity r and $f(\cdot)$ is a continuous function. The expected outcome of a seller which rounded the asking price is then

$$E(y_{jr} \mid a_j, r) = f(0) + \beta_r^{\text{bin}} + \beta_r^{\text{round}},$$

while the outcome of any other seller is

$$E(y_{jr} \mid a_j, r) = f(-\Delta) + \beta_r^{\text{bin}}.$$

Differencing we get

$$\psi_r(\Delta) = f(0) - f(-\Delta) + \beta_r^{\text{round}}.$$

From the continuity of $f(\cdot)$, it is easy to see that

$$\lim_{\Delta \to 0} \psi_r(\Delta) = \beta_r^{\text{round}},$$

that is, the effect of rounding on the market outcomes of two "identical" houses is equal to β_r^{round}. Provided that consistent estimators for $\widehat{\beta}_r^{\text{round}}$ can be obtained, we can test for the joint null hypothesis $H_0 : \beta_r^{\text{round}} = 0, \forall r \in R$ to test for the presence of "cheap talk signalling" or "charm pricing".

Following Backus, Blake, and Tadelis (2019), we estimate the following partially linear regression model

$$y_{jr} = f(\text{dist}_{jr}) + \sum_{r \in \mathcal{R}(\tau)} \Big[\beta_r^{\text{bin}} \cdot \mathbb{1}\{a_j \in B_r\} + \beta_r^{\text{round}} \cdot \mathbb{1}\{a_j = r\}\Big] + X_{jr}\delta + \epsilon_{jr}. \tag{1}$$

Our estimator is restricted to bins that are populated with enough observations, for example, only bins with at least τ observations at the discontinuity and τ observations off the discontinuity enter the estimation.[12] The outcomes of interest are (the log of) the number of virtual visits, measured by listing views on day 2, the sales price, and the number of days the house stayed on the market. Since we are interested in the effect of rounding on "identical houses", we include a full set of observed house characteristics (living space and dummies for the number of bathrooms and bedrooms), the log of the Zestimates one month prior to entry, location fixed effects (zip codes) and time fixed effects (year-month-day of the week). Local identification of our point discontinuities requires $f(\cdot)$ to be flexibly estimated. We estimate the model in (1) using a penalized splines estimator. We keeps bins B_r which contain at least $\tau = \{10, 20\}$ observations at the discontinuity and off the discontinuity to ensure we have enough observations to separately identify the β_r^{round} and the β_r^{bin}.

Table 2 reports the test statistics and p-values for the joint null hypothesis $H_0 : \beta^{\text{round}} = 0,\ \forall r \in \mathcal{R}(\tau)$. The chosen thresholds are $\tau = \{10, 20\}$. We cannot reject the hypotheses at conventional levels. These findings are reassuring for two reasons. First, we cannot find evidence that buyers respond to rounding by discretely increasing or decreasing the rate at which they visit houses.

[12] Mathematically, the set of discontinuities that are considered is

$$\mathcal{R}(\tau) = \Big\{r \in R : \sum_j \mathbb{1}\{a_j = r\} \geq \tau \ \& \ \sum_j \mathbb{1}\{a_j \in B_r, a_j \neq r\} \geq \tau\Big\}.$$

Table 2: Joint Test for Point Discontinuities

	Market outcome y_{jr} (in log)		
	Listing views	Days on market	Sales price
a) $H_0 : \beta_r^{\text{round}} = 0,\ \forall r \in \mathcal{R}(10)$			
Wald test	8.93	17.32	16.15
P-value	0.99	0.75	0.81
b) $H_0 : \beta_r^{\text{round}} = 0,\ \forall r \in \mathcal{R}(20)$			
Wald test	4.20	6.18	16.15
P-value	0.84	0.63	0.72

Notes: Both sets of results test for the joint null hypothesis that the β_r^{round} parameters are 0. The test in Panel a) is based on the estimation of 22 discontinuity parameters. The test in Panel b) is based on the estimation of 8 discontinuity parameters. The estimation of both models are available in Table 7 and 6 of Appendix B.

Second, we cannot find evidence that rounding generates changes in market outcomes. In other words, if sellers are indeed using rounding as a signalling device, then it is failing at generating more demand or changing market outcomes. This would happen if buyers do not agree that rounding the last two digits carries any relevant information about the seller or the house.

Our methodology does not allow us to rule out "charm pricing" or "cheap talk signalling" formally. Still, we believe that our instrument satisfies the exogeneity assumption. The idea is that a much more effective signalling tool is available to sellers: setting the asking price below or above the Zestimate is enough to signal their willingness to sell, their house's conditions, loss aversion, or a strong bargaining position. In this context, rounding at the two-digit level does not generate a large enough price variation to be used strategically.

2.3.4 Testing for Weak Instrument

We next address the relevance of our instrument. In particular, we are concerned that variations in the asking price of under \$100 are not large enough to generate a response from buyers. Another concern is that rounding is uncorrelated with the asking price. This would occur if for example rounding up and rounding down was equally likely. Both of these phenomena would lead to a prob-

Table 3: First-stage Regression

	Asking price (in log)
IV	-0.03***
	(0.005)
Zestimate (in log)	✓
Controls	✓
F-Statistic (IV)	29.44
R-squared	0.92
Observations	4322

Notes: Controls include house characteristics (living space and dummies for the number of bathrooms and bedrooms), location fixed effects (ZIP codes) and time fixed effects (year-month-day of the week). Robust standard errors are in parenthesis. Significance level: *** p<0.01, ** p< 0.05, * p< 0.1.

lem of weak identification. To address this issue, consider the following 1^{st} stage equation,

$$\ln(AskingPrice_j) = \gamma_0 + \gamma_1 IV_j + \gamma_2 \ln(Zestimate_j) + X_j\gamma_3 + u_j. \qquad (1^{\text{st}}\text{ Stage})$$

We want to test for the significance of γ_1. OLS estimates are presented in Table 3. The parameter estimate for γ_1 is -0.03 and is statistically significant at conventional level. Our findings suggest that rounding is associated with a decrease in the asking price relative to the Zestimate, that is, sellers typically round down rather than round up. This reinforces the idea that rounding generates sufficient variation in the asking price to identify the causal link between sellers' strategies and buyers' search intensity.

2.4 Results

We estimate two specifications for our directed search mechanism. First, we provide OLS estimates of Model I and II. Second, we estimate Model I using our instrumental variable and use predicted values to estimate the effect of directed search on market outcomes (i.e. this is in practice a three-stage least square estimation). Results from the estimation of the directed search equation are presented in Table 4. There are a few things to note about these results. First, the effect of a

Table 4: Directed Search Estimation

	Listing views (in log)	
	OLS	IV
Asking price (in log)	-0.84***	-4.22***
	(0.085)	(0.971)
Zestimate (in log)	✓	✓
Controls	✓	✓
R-squared	0.39	0.37
Observations	4,322	4,322

Notes: Controls include house characteristics (living space and dummies for the number of bathrooms and bedrooms), location fixed effects (ZIP codes) and time fixed effects (year-month-day of the week). Robust standard errors are in parenthesis. Significance level: *** p<0.01, ** p< 0.05, * p< 0.1.

raise in the asking price on listing views (conditional on Zestimates) has the expected sign. OLS estimates suggests that increasing the asking price by 10% is associated with a decrease of 8.4% in the number of virtual visits (an inelastic response). Our causal estimate of directed search suggests a much stronger effect. Using instrumental variable techniques, we estimate the elasticity to the asking price to be -4.22 (an elastic response), around five times the estimate obtained through OLS. The attenuation bias is also working in the right direction. To see why, consider a seller who sets a low asking price because his house is in serious need of repairs. After careful observation of the front picture of the house, some of the unappealing features of the house are revealed to buyers. Some of them then choose not to click on the listing to access the house's specific file, mitigating the increase in demand that is generated from the low price. In this case, the causal effect of lowering the price is larger in magnitude than the observed effect, which is exactly what we find in Table 4. We believe that these findings are consistent with optimal directed search: buyers' respond to the asking price by disproportionately allocating towards sellers which offer a lower initial asking price for a given Zestimate and keeping all other characteristics constant.

We now turn to the impact of sellers' strategies on market outcomes that occur through changes in buyers' search intensity. We use the empirical strategy presented above to highlight the tradeoff

Table 5: Effect of Directed Search on Market Outcomes

	Days on market (in log)		Sales price (in log)	
	OLS	IV	OLS	IV
Listing views (in log)	-0.18***	–	-0.01	–
	(0.013)		(0.004)	
Directed search	–	-0.41***	–	-0.20***
		(0.152)		(0.044)
Zestimate (in log)	✓	✓	✓	✓
Controls	✓	✓	✓	✓
R-squared	0.54	0.51	0.92	0.92
Observations	4,322	4,322	4,322	4,322

Notes: Controls include house characteristics (living space and dummies for the number of bathrooms and bedrooms), location fixed effects (ZIP codes) and time fixed effects (year-month-day of the week). Robust standard errors are in parenthesis. Significance level: *** $p<0.01$, ** $p< 0.05$, * $p< 0.1$.

faced by sellers when posting the initial asking price. The mechanism we have in mind is as follows. Sellers choose an initial asking price that generates a response from buyers. For example, we have shown that lowering the asking price leads to an increase in buyers' search intensity. Then, buyer search intensity together with the asking price determine market outcomes. For example, advertising on the low increases buyers' arrival rate which leads to a fast sale. On the other hand, we expect that lowering the asking price leads to a lower expected sales price. Our results are presented in Table 5. We present two sets of findings. First we regress listing views on time on market and sales price to get a sense of the patterns that are present in the data. The OLS coefficient for time on market is estimated at -0.18, while for sales price, we do not recover a statistically significant estimate. Second, we regress predicted listing views from Model I on the same market outcomes, to track the effect of initial strategies and illustrate the tradeoff faced by sellers.

We first focus on our findings for time on market. We estimate an elasticity of -0.41 which is statistically significant at conventional levels. We do not interpret this coefficient directly. Instead, we use this estimate to track the effect of seller's initial decision about the asking price on time on market. Consider an increase of the asking price of 1%. Estimates from Model I suggest that buyers reduce their search intensity by 4.22% in response to the higher asking price. Time on market is then affected by this change in search intensity. Our estimates from Model II suggests that a decrease in predicted listing views of 4.22% is associated with an increase in time on market of 1.73%. Putting both together suggests that increasing the asking price by 1% leads to an increase in days on market by 1.73% through the change in buyers' search intensity.

With this in mind, we now turn to the interpretation of the elasticity to directed search of the sales price. We estimate this elasticity to be -0.20. This result seems counterintuitive since we would expect that more buyer visits lead to a higher sales price. We observe the inverse effect since the increase in buyers' search intensity is initiated from a change in the asking price. The direct effect from the increase in the asking price dominates the buyers' reallocation effect that leads to more visits. We repeat the same exercise as above. Consider an increase in the asking price of 1%. This increase is associated with a decrease in search intensity of 4.22%. We now consider how this change in predicted listing views impacts the sales price. A decrease in predicted listing views by 4.22% is associated with an increase in the sales price of 0.84%. Putting both together, we find that increasing the asking price by 1% leads (through directed search) to an increase in the sales price of 0.84%.

We can use the estimates above to get a sense of the tradeoff faced by sellers when setting the asking price. In average in our sample, houses have an initial asking price of $471,700, a sales price of $481,650, receive 273 listing views on day 2, and sell in 65 days. We are interested in the effect of an increase in the asking price of 1%, that is an increase of $4,717, on market outcomes. Listing views would decrease by 4.22% following an increase in the asking price of 1%, that is, sellers would received 11.48 fewer virtual visits on day 2.[13] The change in the initial asking price can also be tracked to changes in expected time on market and expected sales price. Time on market would increase by 1.125 days and the sales price would increase by $4,046. This implies that the sales-to-asking price gap would decrease by $671, suggesting a lower likeliness of overbidding as the asking price increases. We use these values to price the tradeoff faced by sellers. Our results imply that the cost of reducing time on market by one day is $3,598, representing 0.75% of the average sales

[13] Using back-of-the-enveloppe calculation, this implies that sellers value receiving one additional listing view at $413.50 in terms of initial asking price.

price.

Our estimates imply a very large cost of reducing market duration through changes in the asking price. How can we rationalize that some sellers choose an asking price below the Zestimate? One possible explanation is that some sellers are more motivated or impatient than others (Albrecht, Gautier, and Vroman, 2016). A competing explanation concerns agents facing a joint buyer-seller problem: liquidity constrained agents must sell their current home in order to purchase a new one (Moen, Nenov, and Sniekers, 2021; Anenberg and Bayer, 2020). We find it implausible that these two explanations alone can fully rationalize why some sellers choose an asking price below the Zestimate despite the large estimated cost associated with reducing time on market. In the next section, we propose a theoretical model that contributes to explaining sellers' observed behavior.

3 A Model of Directed Search

3.1 Setup

In this section, we propose a theoretical model to rationalize our empirical findings. Through this exercise, we aim at explaining how real estate online platforms can impact the housing market, in particular sellers. We expand on previous work done by Albrecht, Gautier, and Vroman (2016). In our setup, each seller has a reservation value for their home which is a function of both observed characteristics (for example, the number of bedrooms) and unobserved characteristics. We define this reservation value to be a function $s = g(Y) + \Delta$, where Y is a vector of observed characteristics and Δ represent unobserved characteristics. Real estate online platform on the other hand estimate the value of a house based on observed characteristics only using proprietary algorithms such as the Zestimate. They then release publicly this market value, denoted $z = g(Y)$,[14] which is observed by all sellers and buyers. Seller's offer a second price auction for their house with reservation price a (i.e. the asking price). Finally, buyers are following an optimal directed search strategy conditional

[14] Here we use the same function, $g(Y)$, to define both the reservation values of sellers and real estate online platforms' signal. Using a more flexible functional form will not affect results. To understand why, consider a function $z = h(Y)$. For any function $h(Y)$, we can always find a function $g(Y)$ such that $h(Y) = g(Y) + \epsilon$. Then rewrite sellers valuations as $s = h(Y) + \tilde{\Delta}$ where $\tilde{\Delta} = \epsilon + \Delta$ and we are done.

on observing asking prices a and market values z. This is a departure from Albrecht, Gautier, and Vroman (2016) which assume that all houses on the market are identical. In contrast, our model allows for houses not to be identical, and for z to summarize all of the houses' observed characteristics.

We model the market for all houses which received the same market value (or signal) $z \in Z$ as a one-shot game between $B(z)$ buyers and $S(z)$ sellers. We assume a large market for every possible signal $z \in Z$ such that $B(z)$ and $S(z)$ go to infinity keeping the market tightness $\theta(z) = B(z)/S(z)$ constant at each z. Throughout, we impose the assumption that buyers do not shop across house types such that a closed equilibrium exists at each z. In some sense, we allow for z to sort houses into categories, and that the choice of a category is exogenous to the buyer's decision process. This assumption is of course restrictive but not completely unrealistic: buyers often restrict their search using a variety of criteria that are captured by z, for example the number of bedrooms, the type of neighbourhood, or house valuations.[15] We impose this assumption nevertheless because it provides a useful simplification that helps us disentangling what happens within equilibrium (at a given z) from what happens across equilibria (as z changes). For example, our model provides insights about how optimal directed search operates within equilibrium, if for example some sellers advertise on-the-high while others advertise on-the-low. It also allows us to characterize how the housing market changes over house types. This improves on previous literatures that focused on similar houses. We show for example that the evolution of the market tightness and the distribution of buyer's valuations across house types are crucial in determining key market outcomes such as the time on market and the sale price.

3.2 Timing of the game

The timing of the game is as follows:

1. *State of the world* - All houses are assigned a market value $z \in Z$, observable by all potential

[15] The choice of a house type is usually driven by factors that are not easily adjusted by buyers in the short run, for example income or their level of personal debt which limits their ability to buy certain types of houses. We believe that this provides some justification for our assumption that a given buyer is not involved in several segments of the housing market at the same time.

buyers and sellers. This signal imperfectly measure sellers reservation values. We consider a simple functional form for s and z for expositional simplicity. We assume that the market value of any given house is given by:

$$z = \begin{cases} s - \Delta, & \text{with probability } q \\ s + \Delta, & \text{with probability } 1 - q \end{cases} \tag{2}$$

where s is the seller's reservation value. To put plainly, any given seller receives a market value that either undershoots or overshoots his reservation value by some deviation Δ. For a given z, this implies that sellers' reservation values are distributed according to the following probability mass function:

$$s = \begin{cases} z + \Delta, & \text{with probability } q \\ z - \Delta, & \text{with probability } 1 - q \end{cases} \tag{3}$$

We assume that there are $B(z)$ buyers and $S(z)$ sellers in the market for houses that received a market value of z.

2. *Sellers* - Upon observing z, each seller posts an asking price a.

3. *Buyers* - Each buyer observes all posted prices and chooses k houses to visit. There is no coordination among buyers and buyers do not shop across house types (i.e. they do not shop across z). Upon visiting a house, the buyer draws a match-specific value. Match-specific values are private information and are iid draws across buyer-seller pairs. Each buyer can bid on at most one house and chooses the house with the highest match-specific value, denoted by x. We assume that X is distributed as Uniform$(0, \mu)$, which implies that $F(x; \mu)$ has finite support and increasing hazard.[16] We allow the parameters of F to depend on z, i.e. $\mu = \mu(z)$ to allow for the market valuations to impact the distribution of buyers valuations. Buyers do not observe the number of other visitors to a house and do not interact with other buyers.

[16] This is a special case of Albrecht, Gautier, and Vroman (2016), which consider instead all distributions with increasing hazard.

4. At the chosen house, the buyer can either accept or reject the asking price a.

5. If no buyer visits, the seller retains its reservation value s.

6. If at least one buyer visits and accepts the asking price a, then there is a second-price auction with reserve price a among those buyers.[17] The seller is bound to transfer the house to the highest bidder.

If a transaction occurs, then the seller gets the transaction price p and the winning buyer with match-specific value x gets a payoff of $x - p$. If no sale occurs, then the seller retains the value of his house, s. All unsuccessful buyers get a payoff of zero. Under conditional directed search, sellers which receive the same market value z face a tradeoff when choosing an asking price: either advertise on-the-low (i.e. $a < z$) to attract more buyers, increasing the probability that a sale occur, or advertise on-the-high (i.e. $a > z$) to get a higher price.

Our model allows us to understand not only the tradeoff faced by sellers which received the same z, but also the impact of real estate online platforms on sellers across different z. For example, we consider what happens in the case where two sellers have the same reservation value s (i.e. their house is similar both in observed and unobserved characteristics), but different market valuations, say $z_1 < s$ and $z_2 > s$. We can show with a simple logical argument that receiving a favorable evaluation (i.e. $z > s$) can lead to a strictly better outcome than receiving an unfavorable one (i.e. $z < s$). Suppose for simplicity that both sellers choose to offer a second-price auction with reservation price $a_1 = a_2 = s$ and that market tightness is constant across house types (i.e. $\theta(z_1) = \theta(z_2), \forall z_1, z_2 \in Z$).[18] Then it has to be that, under directed search, seller 2 has a strictly higher expected revenue than seller 1. The intuition behind this result is as follows: both are offering a second-price auction with reservation price s, but seller 2 is perceived as advertising on-the-low, attracting more buyers within his segment than seller 1. Since the expected revenue in a second-

[17] Our results can be generalized to include more sophisticated mechanism, as in Albrecht, Gautier, and Vroman (2016). Since a revenue equivalence theorem holds for increasing bid auctions with reservation values, studying any of these auctions leads to the same outcome for sellers, although the equilibria can be slightly different. We focus on a second-price auction since it provides a unique equilibrium, which simplifies exposition.

[18] Note that we do not require $a_1 = a_2 = s$ to be an equilibrium of this game.

price auction is strictly increasing in the number of bidders, it has to be hat seller 2 is strictly better off. More generally, for any possible strategy $a_1 \in A$, seller 2 always has the option to choose $a_2 = a_1$ which earns a strictly higher expected payoff than seller 1. Let $\Pi_i(a)$ be the expected revenue for seller i offering a second-price auction with reservation value a. Then by the definition of a Bayesian Nash Equilibrium, we have that

$$\Pi_2(a^*) \geq \Pi_2(a) > \Pi_1(a), \forall a \in A,$$

that is, seller 2 is better off irrespective of seller 1's strategy.

We find that this simple result illustrate an important feature of real estate markets. Because market values made available by online platform cannot account for all unobserved features of houses, some sellers which receive a poor market evaluation are penalized compared to sellers of otherwise identical houses which received a good evaluation. We also show formally that this penalty is decreasing in the precision of the algorithm used to compute market values, which may partly explain why online real estate platform such as Zillow devote significant resources towards improving their proprietary algorithms. In what follows, we revisit the theory proposed by Albrecht, Gautier, and Vroman (2014) and Albrecht, Gautier, and Vroman (2016) and show how their equilibrium concept can be adapted to our setup. We then explore how conditional directed search affects key market outcomes such as the expected time on market and the expected price. Finally, we turn our attention to the case of two sellers with the same reservation value and two different market valuations to understand how real estate online platform affect market outcomes through their house evaluation algorithms.

3.3 Perfect market signal

3.3.1 Equilibrium

We lay down some fundamentals of the model by first exploring the case where all house characteristics are observed, and real estate online platform provide a perfect market value, that is $z = s$. Throughout, we consider the market equilibrium *at each* z to be the outcome of the game described above, played between $S(z)$ sellers and $B(z)$ sellers. We then use this benchmark to show how key market outcomes vary across equilibria, as z increases or decreases. We maintain the rather strong assumption that buyers do not shop across z, such that the market at each z and the resulting equilibrium are self contained.

Upon observing z, seller i offer a second-price auction with reserve price a_i. Directed search entails that the arrival rate of buyers depends not only on market tightness, but also on the distribution of asking prices. The rate at which buyers visit a given seller with asking price a_i, denoted θ_i, adjusts so that the value of visiting this seller is the same as that of visiting any other seller with a house of type z, which we denote $\overline{V}_z$. Our formulation of the maximization problem facing sellers follows that of Peters and Severinov (1997), Albrecht, Gautier, and Vroman (2014), and Albrecht, Gautier, and Vroman (2016), that is

$$\underset{a \in A}{max} \quad \Pi(a, \theta \mid z) \quad \text{subject to} \quad V(a, \theta \mid z) = \overline{V}_z, \tag{4}$$

where $\Pi(\cdot|z)$ and $V(\cdot|z)$ are the expected payoffs of sellers and buyers respectively, conditional on house type z. Following the notation in Albrecht et al. (2016), these are

$$\begin{aligned} \Pi(a, \theta|z) &= s + \theta \int_a^\mu (v(x; \mu) - s) e^{-\theta(1-F(x;\mu))} f(x; \mu) dx \\ &= s + (1 - e^{-\theta}) \int_a^\mu (v(x; \mu) - s) g(x; \mu, \theta) dx \end{aligned} \tag{5}$$

$$\begin{aligned} V(a, \theta|z) &= \int_a^\mu (x - v(x; \mu)) e^{-\theta(1-F(x;\mu))} f(x; \mu) dx \\ &= \int_a^\mu (1 - F(x; \mu)) e^{-\theta(1-F(x;\mu))} dx \end{aligned} \tag{6}$$

where

$$v(x; \mu) = x - \frac{1 - F(x; \mu)}{f(x; \mu)}$$

is the virtual valuation function and

$$g(x;\mu,\theta) = \frac{\theta\ e^{-\theta(1-F(x;\mu))} f(x;\mu)}{1-e^{-\theta}}$$

is the density of the highest drawn valuation by buyers visiting a particular house conditional on having at least one visitor. The solution to this problem is discussed in Lemma 1 of Albrecht, Gautier, and Vroman (2016). Sellers find it optimal to offer a second price auction with reservation price $a^* = z$. The corresponding Poisson arrival rate of buyers is the solution to $V(s,\theta \mid z) = \overline{V}_z$, that is $\theta^* = \theta(z)$. The intuition for that last equality is simple: because all sellers of a house of type z are playing the same strategy, it is optimal for buyers to spread equally across sellers such that the arrival rates are also equalized across sellers. It's straightforward to show that whenever the market tightness is fixed for a given z, then the only solution to the problem is $\theta^* = \theta(z)$.[19]

3.3.2 Time on market

Now that we have characterized the equilibrium of this simplified model, we examine how key outcomes such as expected time on market and expected price behave as we increase z. Note that the market value is tied to sellers reservation values in this setup since the signal is perfect. First, define the expected time on market as

$$E(T(z,\theta)) = \frac{1}{Pr(\text{House is sold} \mid z,\theta)} = \frac{1}{1-e^{-\theta(1-F(z;\mu))}}, \tag{7}$$

[19]Suppose for example that $\theta^* > \theta(z)$. Then we have that the implied total number of buyers is $B(\theta^*) = S(z)\cdot\theta^* > S(z)\cdot\theta(z) = B(z)$, a contradiction of the assumption that there are $B(z)$ buyers in the market for houses of type z. This implies that $\theta^* \leq \theta(z)$. Suppose on the contrary that $\theta^* < \theta(z)$. Then we have $B(\theta^*) = S(z)\cdot\theta^* < S(z)\cdot\theta(z) = B(z)$, another contradiction. We conclude that $\theta^* = \theta(z)$.

that is, time on market follows the Geometric distribution with parameter $p(z) = Pr(\text{House is sold} \mid z)$. Consider an increase in z. The the partial derivative with respect to z is

$$\frac{\partial E(T(z,\theta))}{\partial z} = \frac{\partial E(T)}{\partial \theta} \cdot \frac{\partial \theta}{\partial z} + \frac{\partial E(T)}{\partial z}, \tag{8}$$

$$= \frac{e^{-\theta(1-F(z;\mu))}}{(1-e^{-\theta(1-F(z;\mu))})^2} \cdot \left[\underbrace{-(1-F(z;\mu)) \cdot \frac{\partial \theta}{\partial z}}_{\text{Market tightness effect}} + \underbrace{\theta \cdot \frac{\partial F(z;\mu)}{\partial z}}_{\substack{\text{Buyer valuation}\\ \text{effect}}} \right]. \tag{9}$$

Whether this effect is positive or negative depends on two factors. First, how market tightness evolves with house types, i.e. the partial derivative of θ with respect to z. Second, it depends on how the conditional distribution of buyers' valuations evolve with z. To further our understanding, consider the following parametrization for $F(x;\mu)$,

$$X \sim \text{Uniform}(0, \mu(z))$$

with

$$\mu(z) = \omega \cdot z^{\alpha}, \quad \alpha \geq 0.$$

Provided $0 < z < \mu(z)$,[20] we can now write formally

$$F(z;\mu) = \left(\frac{1}{\omega}\right) \cdot z^{1-\alpha},$$

and

$$\frac{\partial F(z;\mu)}{\partial z} = \left(\frac{1-\alpha}{\omega}\right) \cdot z^{-\alpha}.$$

We now consider several examples and explore their implication for time on market.

Example 1: *The distribution of valuations does not depend on z, i.e. $\alpha = 0$.*

[20] One simple way to ensure this is satisfied at all possible $z \in Z$ is to pick $\{\omega, \alpha\}$ pairs such that that $\omega^{\frac{1}{1-\alpha}} > \sup(Z)$.

In this case, $F(x;\mu) = F(x), \forall x$. Then we have that

$$\frac{\partial F(z)}{\partial z} = \frac{1}{\omega} > 0, \quad \forall z \in Z.$$

This means that the second term in the parenthesis in equation (9) is positive. The overall effect of an increase in house quality on time on market depends on how market tightness evolves with z. If for example $\partial\theta/\partial z \leq 0$, then time on market increases with house quality. If on the contrary market tightness increases with z the effect on time on market becomes uncertain, depending on which effect dominates.

Example 2: *Buyers valuations is distributed as Uniform*$(0, \omega z)$*, i.e.* $\alpha = 1$.

In this case, we have $\partial F(z;\mu)/\partial z = 0$, hence the direction of the effect of an increase in market valuation on time on market depends on the evolution of market tightness only. One particularly interesting case occurs when market tightness is constant, that is $\partial\theta/\partial z = 0$. Then, time on market becomes constant across house types.

Example 3: *Buyers valuations is distributed as Uniform*$(0, \omega z^{\alpha})$*, with* $\alpha > 1$.

In this case, we have that $\partial F(z;\mu)/\partial z < 0$, hence the second term in equation (9) is negative. The overall effect of an increase in house quality on time on market depends on how market tightness evolves with z. If for example $\partial\theta/\partial z \geq 0$, then time on market decreases with house quality. If on the contrary market tightness decreases with z then the effect on time on market becomes uncertain, depending on which effect dominates.

3.3.3 Expected price

We consider next the effect of a change in market evaluation on the expected price. Following Albrecht et al. (2016), the expected price conditional on a sale is

$$E(P(z,\mu,\theta)) = \int_z^{\mu} v(x;\mu)h(x;\mu,\theta)dx \tag{10}$$

where

$$h(x; z, \mu, \theta) = \frac{g(x; \mu, \theta)}{\int_z^{\mu} g(t; \mu, \theta)dt}. \tag{11}$$

Define $H(x; z, \mu, \theta) = \int_z^x h(t; z, \mu, \theta)dt$ and recall from standard auction theory that

$$v(x; \mu) = v(z; \mu) + \int_z^x v'(t; \mu)dt. \tag{12}$$

Replacing in (10), the expected price can be rewritten as

$$\begin{aligned} E(P(z, \mu, \theta)) &= v(z; \mu) + \int_z^{\mu} h(x; z, \mu, \theta) \int_z^x v'(t; \mu)dtdx \\ &= v(z; \mu) + \int_z^{\mu} v(t; \mu) \int_t^{\mu} h(x; z, \mu, \theta)dxdt \\ &= v(z; \mu) + \int_z^{\mu} v'(t; \mu)(1 - H(t; z, \mu, \theta))dt. \end{aligned} \tag{13}$$

We want to see how the expected price varies with z. Using chain rule, the derivative of the expected price conditional on a sale is

$$\frac{\partial E(P(z, \mu, \theta))}{\partial z} = \underbrace{\frac{\partial E(P)}{\partial z}}_{\text{Direct effect}} + \underbrace{\frac{\partial E(P)}{\partial \mu} \cdot \frac{\partial \mu}{\partial z}}_{\text{Buyer valuation effect}} + \underbrace{\frac{\partial E(P)}{\partial \theta} \cdot \frac{\partial \theta}{\partial z}}_{\text{Market tightness effect}} . \tag{14}$$

We can show that the three derivatives in (14) are positive, using the following result.

PROPOSITION 1: Consider the equilibrium described above, where $(a^*, \theta^*) = (z, \theta(z)), \forall z \in Z$ and buyer's match specific valuations are distributed as Uniform$(0, \mu(z))$. Define the expected price as

$$E(P(z, \mu, \theta)) = v(z; \mu) + \int_z^{\mu} v'(t; \mu)(1 - H(t; z, \mu, \theta))dt,$$

where $v(z; \mu)$ is the virtual value function. Then for $0 < z < \mu$ and $\theta > 0$,

i. $\frac{\partial E(P)}{\partial z} > 0$,

ii. $\frac{\partial E(P)}{\partial \mu} > 0$,

iii. $\frac{\partial E(P)}{\partial \theta} > 0$.

The proof is available in the Appendix. Part (i) and (iii) use Lemma 1 from Albrecht, Gautier, and Vroman (2016), which we adapt to match our specific setup.

LEMMA 1: Let $X \sim \text{Uniform}(0, \mu)$. Then the cumulative distribution function $H(x; z, \mu, \theta) = \int_z^x h(t; z, \mu, \theta)dt$ satisfies first-order stochastic dominance with respect to z, μ and θ; that is

i. $z' > z \Longrightarrow H(x; z') < H(x; z), \quad \forall x \in (z, \mu)$,

ii. $\mu' > \mu \Longrightarrow H(x; \mu') < H(x; \mu), \quad \forall x \in (z, \mu)$,

iii. $\theta' > \theta \Longrightarrow H(x; \theta') < H(x; \theta), \quad \forall x \in (z, \mu)$.

The proof of Lemma 1 is also available in the Appendix. These results have implications for how the expected price behave as the market value increases. In particular, if the distribution of buyer's valuations doesn't deteriorate with z (i.e. $\partial \mu / \partial z \geq 0$), and market tightness doesn't decrease with z, then the expected price is increasing in z. It is interesting to consider also a special case where market tightness decreases with z which would happen if the ratio of buyers to seller is lower for more valuable houses. In this case, the effect on the expected price is ambiguous and depends on the magnitudes of the (1) direct effect, (2) the buyer valuation effect, and (3) the market tightness effects. Our model allows in extreme cases to generate a decrease in expected price following an increase in market value, provided that the drop in market tightness is sufficiently important. We believe that this is an important feature of the model, as these types of price fluctuations may occur in real data over some range of the market valuation distribution, and because market tightness is typically not fixed across house types.

One important feature of any model of directed search is that seller's must face a tradeoff when advertising asking prices: either advertising low to attract more buyers, sell faster and generate overbidding, or advertising high to get a higher expected price. In the next section, we show that our model generates these feature *within equilibrium* (at a given z). What our study of the perfect signal case highlights, is that this doesn't have to be true *across equilibria* (for different z). This has important implications for reconciling the theory with data: observing decreasing time on market or decreasing prices over some range of the the distribution of market valuations is not incompatible with the optimal directed search assumption. In what follow, we focus on within equilibrium dynamics, by considering the case where market valuations provide an imperfect signal about seller's reservation values.

3.4 Imperfect market signal

3.4.1 Equilibrium

We now focus on within equilibrium interactions by studying a very simple case where sellers can receive either a favorable market evaluation, $z = s + \Delta$ or an unfavorable one $z = s - \Delta$. Specifically, this simple model allows us to explain one important feature of the housing market: a large fraction of sellers choose to post an asking price below their house's market valuation while other choose to advertise above. We show that these observations can be rationalized by conditional directed search. We then show how the precision of the algorithm used to compute z affects market outcomes in this setup, and especially affect how buyers sort themselves across both types of sellers.

We consider the case of seller i which received a market valuation of z. Provided that he received a negative signal with probability q and a positive signal with probability $(1 - q)$, his reservation value is distributed as

$$s_i = \begin{cases} z + \Delta, & \text{with probability } q, \\ z - \Delta, & \text{with probability } 1 - q. \end{cases} \tag{15}$$

For simplicity, we will refer to both types of sellers as type-H and type-L respectively.[21] Albrecht, Gautier, and Vroman (2016) already show that under D1 refinement, the only Bayes Nash equilibrium of this game with two sellers type is a separating equilibrium[22] where

[21]Note that type-H sellers are those receiving the unfavorable market signal, i.e. $s_L < z < s_H$.
[22]We refer to their article for the complete proof.

1. All type-L sellers post $a_L = z - \Delta$;

2. All type-H sellers post $a_H = z + \Delta$;

3. Buyers believe that $a = z - \Delta$ signals type-L with probability 1;

4. Buyers believe that $a = z + \Delta$ signals type-H with probability 1.

We consider a continuum of these separating equilibria, one at each $z \in Z$. While by assumption buyers do not shop across z, directed search entails that they will allocate themselves among type-L and type-H at a given z. Suppose that a fraction r of buyers visit type-H sellers and a fraction $(1-r)$ visit type-L sellers. For a given q and θ,[23] the expected arrival rate of buyers for each type of sellers are $\theta_L = \frac{(1-r)\theta}{1-q}$ for type-L sellers and $\theta_H = \frac{r\theta}{q}$ for type-H. The expected payoff for a buyer who visits a type-L or type-H sellers are respectively

$$\begin{aligned} V_L(r|z) &= \int_{z-\Delta}^{\mu} (1 - F(x;\mu)) e^{-\theta_L(1-F(x;\mu))} dx \\ &= \frac{\mu}{\theta_L^2} \cdot \left(1 - e^{-\theta_L(1-\frac{z-\Delta}{\mu})} - \theta_L e^{-\theta_L(1-\frac{z-\Delta}{\mu})}\right), \end{aligned} \tag{16}$$

and

$$\begin{aligned} V_H(r|z) &= \int_{z+\Delta}^{\mu} (1 - F(x;\mu)) e^{-\theta_H(1-F(x;\mu))} dx \\ &= \frac{\mu}{\theta_H^2} \cdot \left(1 - e^{-\theta_H(1-\frac{z+\Delta}{\mu})} - \theta_H e^{-\theta_H(1-\frac{z+\Delta}{\mu})}\right), \end{aligned} \tag{17}$$

which depends on how buyers allocate between sellers through the respective arrival rates. In equilibrium, the following optimality condition must hold,

$$V_L(r^*|z) \geq V_H(r^*|z) \text{ with equality if } r^* > 0.$$

We try to understand how buyers allocate themselves across sellers as the market signal becomes

[23]For now, we assume that q and θ are fixed, i.e. $q(z) = q$ and $\theta(z) = \theta$ for all $z \in Z$.

less precise (i.e. as Δ increases). We can show that buyers allocate themselves disproportionately towards L-types when the market signal is imprecise. Moreover, as precision decreases, buyer's find it optimal to switch from type-H to type-L. Intuitively, as measurement error increases, the reservation value of type-L sellers decreases while that of type-H sellers increases at a given z which means that buyers can increase their expected payoff by switching from type-H to type-L. Proposition 2 summarizes these findings.

PROPOSITION 2: Consider the equilibrium described above. Let q be the share of type-H sellers and r be the share of buyers that visit type-H sellers. Then

i. $\Delta = 0 \Longrightarrow r^* = q$,

ii. $\Delta = 0 \Longrightarrow \theta_L^* = \theta_H^* = \theta$,

iii. $\Delta > 0 \Longrightarrow r^* < q$,

iv. $\Delta > 0 \Longrightarrow \theta_L^* < \theta < \theta_H^*$.

Proposition 2 suggests that the model indeed produces patterns of directed search as desired. Intuitively, buyers that allocate themselves towards type-H face a higher expected price for the house, but less competition since more buyers are allocated to type-L, increasing the odds of winning the auction. We derive a useful Corollary from Proposition 2.

COROLLARY 1: Consider the equilibrium described above. Let q be the share of type-H sellers and r^* be the share of buyers that visit type-H sellers in equilibrium. Then

$$\frac{\partial r^*}{\partial \Delta} < 0.$$

The proof of both statements are available in the Appendix. We now show how the allocation of buyers across seller's types affect the expected time on market and the expected price.

3.4.2 Time on market

We now explore how changes in the precision of the algorithm affects the expected time on market in our conditional directed search setting. Recall that conditional on type i, the expected time on market is

$$E(T(s_i, \theta_i) \mid type = i) = \frac{1}{1 - e^{-\theta_i(1-F(s_i;\mu))}},$$

for $i = \{H, L\}$. Differentiating with respect to Δ, we get

$$\frac{\partial E(T(s_i,\theta_i) \mid type = i)}{\partial \Delta} = \underbrace{\frac{\partial E(T \mid type = i)}{\partial \theta_i} \cdot \frac{\partial \theta_i}{\partial r^*} \cdot \frac{\partial r^*}{\partial \Delta}}_{\text{Reallocation effect}} + \underbrace{\frac{\partial E(T \mid type = i)}{\partial s_i} \cdot \frac{\partial s_i}{\partial \Delta}}_{\text{Direct effect}}. \tag{18}$$

The first term which we call the reallocation effect is the effect of buyers reallocating from type-H to type-L as Δ increases. Intuitively, as the algorithm looses in precision, the spread in sellers reservation values and the resulting asking prices increases. This drives some buyers to reallocate from type-H to type-L as, keeping θ constant, the value of visiting a type-L increases with Δ compared to that of visiting a type-H. The second term is the direct effect that comes from the change in asking price. For type-H, the increase in asking price mechanically increases time on market since offers that were previously accepted are now rejected, and conversely for type-L. We can write the above expression for type-H and type-L respectively,

$$\frac{\partial E(T(s_H,\theta_H) \mid H)}{\partial \Delta} = \frac{\partial E(T \mid H)}{\partial \theta_H} \cdot \frac{\theta}{q} \cdot \frac{\partial r^*}{\partial \Delta} + \frac{\partial E(T \mid H)}{\partial s_H}, \tag{19}$$

and

$$\frac{\partial E(T(s_L,\theta_L) \mid L)}{\partial \Delta} = -\frac{\partial E(T \mid L)}{\partial \theta_L} \cdot \frac{\theta}{1-q} \cdot \frac{\partial r^*}{\partial \Delta} - \frac{\partial E(T \mid L)}{\partial s_L}. \tag{20}$$

We can show that both the reallocation effect and the direct effect are positive for type-H and negative for type-L, leading to an increasing expected time on market for type-H and a decreasing time on market for type-L. We formalize this result in the following proposition.

PROPOSITION 4: Consider the equilibrium described above. Define the expected time on market of a type i seller to be

$$E(T(s_i,\theta_i) \mid type = i) = \frac{1}{1 - e^{-\theta_i(1-F(s_i;\mu))}},$$

Then,

i. $\frac{\partial E(T|H)}{\partial \Delta} > 0$ for type-H sellers, and

ii. $\frac{\partial E(T|L)}{\partial \Delta} < 0$ for type-L sellers.

The proof is available in the appendix.

3.4.3 Expected price

We turn our attention next on the expected price, conditional on a sale. Recall that the expected price for a type i sellers takes the following form

$$E(P(s_i,\theta_i) \mid type = i) = v(s_i;\mu) + \int_{s_i}^{\mu} v'(t,\mu)(1 - H(t; s_i, \mu, \theta_i))dt \tag{21}$$

We are interested in understanding how a change in Δ affects the expected price for each type of sellers. Differentiating using chain rule, we get that

$$\frac{\partial E(P(s_i,\theta_i) \mid type = i)}{\partial \Delta} = \underbrace{\frac{\partial E(P \mid type = i)}{\partial \theta_i} \cdot \frac{\partial \theta_i}{\partial r^*} \cdot \frac{\partial r^*}{\partial \Delta}}_{\text{Reallocation effect}} + \underbrace{\frac{\partial E(P \mid type = i)}{\partial s_i} \cdot \frac{\partial s_i}{\partial \Delta}}_{\text{Direct effect}}. \tag{22}$$

It is easy to see that the direct effect and the reallocation effects work in opposite directions in this case. For example, consider a type-H seller. An increase in Δ is associated with a higher reservation value, which means that this seller now rejects offers at the lower end of the set of offers that were previously accepted, leading to a mechanical increase in expected price. At the same time, buyers find it optimal to reallocate to type-L sellers, which induces a decrease in market tightness for type-H, and less overbidding which decreases the expected price. On the other hand, a type-L seller would experience a decrease in his reservation value which would mean accepting offers he was previously rejecting while at the same time receiving more offers than before which leads to more overbidding. Mathematically, we can write

$$\frac{\partial E(P \mid H)}{\partial \Delta} = \frac{\partial E(P \mid H)}{\partial \theta_H} \cdot \frac{\theta}{q} \cdot \frac{\partial r^*}{\partial \Delta} + \frac{\partial E(P \mid H)}{\partial s_H},$$

for type-H sellers, and

$$\frac{\partial E(P \mid L)}{\partial \Delta} = -\frac{\partial E(P \mid L)}{\partial \theta_L} \cdot \frac{\theta}{1-q} \cdot \frac{\partial r^*}{\partial \Delta} - \frac{\partial E(P \mid L)}{\partial s_L}.$$

for type-L sellers. Using Lemma 1 and Corollary 1, it is easy to see that the the first and second terms in each expression are of the opposite signs. We can show however that the direct effect always dominates the reallocation effect, such that the effect from buyers reallocating from type-H to type-L mitigates only partially the direct effect from a change in algorithm precision.

PROPOSITION 4: Consider the equilibrium described above. Define the expected price of a type i seller to be

$$E(P(s_i, \theta_i) \mid type = i) = v(s_i; \mu) + \int_{s_i}^{\mu} v'(t, \mu)(1 - H(t; s_i, \mu, \theta_i))dt.$$

Then,

i. $\frac{\partial E(P|H)}{\partial \Delta} > 0$ for type-H sellers, and

ii. $\frac{\partial E(P|L)}{\partial \Delta} < 0$ for type-L sellers.

The proof is available in the Appendix. Intuitively, this result is necessary for our equilibrium to hold. If for example, an increase in Δ were to cause both and increase in expected time on market and a decrease in expected price for type-H sellers, then some type-H may find it profitable to deviate and advertise as type-L instead, destroying our equilibrium. Instead, Proposition 3 and 4 highlights that our model reproduce the tradeoff faced by sellers in the real estate market. By advertising a lower asking price, sellers receive more visits from buyers which leads to a higher probability of selling and a faster sale. On the other hand, advertising a high asking price leads to a higher expected price, a lower probability of selling and an extended time on market driven by buyers reallocating towards other sellers. Our model also generates other patterns observed in the data, namely that some sellers choose to advertise above their *Zestimate* while other choose to advertise below.

3.5 Seller's (mis)fortune in real estate online platforms

We want to provide additional insight as to how real estate online platforms can affect the market outcomes of sellers. We try to answer the concerns some sellers have raised in the media, specifically that algorithms such as the *Zestimate* misprice their home in such a way that it becomes impossible to sell at or above their reservation value. We use the model developed above and show that being overevaluated by online real estate platforms leads to a strictly better outcome than being underevaluated. We also show that improving the precision of the algorithm generating these market valuations decreases these inequalities such that there is a potential social value attached to precision that goes beyond profits of platform operators. To do so, we consider the case of two sellers, say A and B, with identical reservation values s but different market valuations advertised on the online platform. Specifically, we have in mind that seller A received an unfavorable market value while B received a favorable one, such that $z_A < s < z_B$. For simplicity, we assume that $z_A = s - \Delta$ and $z_B = s + \Delta$, that is, the deviation from the true seller's reservation value is symmetric around s, although our results hold with a more general specification. Recall that seller i's expected payoff from entering the market is

$$\begin{aligned}\Pi(s \mid z_i) &= s + (1 - e^{-\theta_i}) \int_s^{\mu_i} (v(x;\mu_i) - s) g(x; s, \mu_i, \theta_i) dx \\ &= (1 - e^{-\theta_i(1-F(s;\mu_i))}) \cdot \int_s^{\mu_i} v(x;\mu_i) h(x;\mu_i,\theta_i) dx + e^{-\theta_i(1-F(s;\mu_i))} \cdot s \qquad (23) \\ &= Pr(Sale \mid z_i) \cdot E(P \mid Sale, z_i) + (1 - Pr(Sale \mid z_i)) \cdot s\end{aligned}$$

We now show that when both sellers have the same reservation value, then seller B is strictly better off than seller A, enjoying both a higher expected price and shorter time on market. This result draws on the previous sections in that the equilibrium at z_A and at z_B are self contained, so we are comparing these two sellers across equilibria. In some sense, seller A is stuck with his bad valuation and cannot attract buyers as effectively as seller B. We summarize our findings in the following Proposition.

PROPOSITION 5: Consider two sellers, A and B, with identical reservations values s and market valuations $z_A = s - \Delta$ and $z_B = s + \Delta$. Let $\theta(z_A) = \theta(z_B) = \theta$ and $q(z_A) = q(z_B) = q$, for all $z_A, z_B \in Z$. Furthermore, assume that buyer's valuations are distributed as Uniform$(0, \mu(z_i))$. Then seller B's expected payoff is strictly higher than seller A, that is,

$$\Pi(s \mid z_B) > \Pi(s \mid z_A).$$

Proposition 5 holds more generally in the case where $q_A \geq q_B$ and $\theta_A \leq \theta_B$. The result takes into account the fact that the market for houses of type z_A and z_B are both in equilibrium, and that A and B are playing equilibrium strategies within their market, that is $a_i^* = s$. The full proof is provided in the Appendix.

We can also express Proposition 5 in terms of inequality between seller A and B. Define $\phi(\Delta) = \Pi_B(\Delta) - \Pi_A(\Delta)$ to be the level of inequality between seller A and B conditional on Δ. Then Proposition 5 leads to the following Corollary.

Corollary 2: Consider two sellers A and B with identical reservation values s, and market valuations $z_A = s - \Delta$ and $z_B = s + \Delta$. Let $\theta(z_A) = \theta(z_B) = \theta$ and $q(z_A) = q(z_B) = q$, for all $z_A, z_B \in Z$. Furthermore, assume that buyer's valuations are distributed as Uniform$(0, \mu(z_i))$. Then the function

$$\phi(\Delta) = \Pi(s \mid z_B(\Delta)) - \Pi(s \mid z_A(\Delta))$$

is strictly increasing in Δ, that is

$$\frac{\partial \phi}{\partial \Delta} > 0.$$

These results taken together have important implications for sellers. First, measurement errors in market valuations create winners and losers among sellers with otherwise identical reservation values. Second, these inequalities decrease with algorithm precision. This has far reaching implications on this market. One such implication is that measurement errors can affect the entry decision of sellers. For example, sellers with a good market signal could decide to enter the market, while sellers with a bad signal could decide to leave the market. The overall effect on entry, firms profits, or welfare (defined broadly as the sum of buyers and sellers payoffs) depends on several factors and is beyond the scope of this work. Nevertheless, we believe that we our model highlights the importance of improving market valuation algorithms.

4 Conclusion

Online real estate platforms like Zillow have become an integral part of the search and matching process in the housing market. The impact of algorithm-based market valuations on key trading outcomes remain unclear. In this paper, we try to understand how the Zestimate affect sellers' initial strategies, buyers' search intensity, and bargaining outcomes. This question is crucial as these platform are becoming commonplace internationally. Using data on daily listing views to proxy buyers' search intensity on Zillow, we find evidence of directed search in the Seattle real estate market. We show that sellers who advertise on-the-low (below the Zestimate) receive more virtual visits from potential buyers, sell faster, and at a lower price that those advertising on-the-high (above the Zestimate).

We propose an instrumental variable strategy to identify the causal links between sellers' strategies, buyers' search intensity, and market outcomes. We exploit the fact that some sellers arbitrarily choose to round the asking price to a multiple of \$100 to recover exogenous variation in the asking price. The idea is that such small variations in the asking price do not originate from a strategic choice made by sellers. We estimate that the cost of reducing market duration by one day is \$3,600 (0.75% of the sales price). We find it surprising that many sellers choose to advertise an asking price below the Zestimate under of such a high cost.

We hypothesize that heterogeneity in the attribution of market valuations rather than heterogeneous discount rates explains the observed dispersion in sellers' strategies. We propose a model in which each seller has a chance to get either a good or a bad market valuation for their house. Our model predicts that getting the favorable market valuation leads to a strictly shorter time on market, a higher expected sales price, and higher expected revenues. The misfortunate seller has no way to get rid of a bad signal, which can have lasting consequences on his wealth level. Interestingly, this mirrors concerns raised in traditional medias by sellers "stuck" with a bad Zestimate.[24] This raises questions about the role of online real estate platforms, suggesting that they are maybe closer to market makers than market observers.

While this paper provides an explicit mechanism to rationalize how algorithm-based market valuations impact house trades, some important questions remain to be answered in future research. One such question relates to the self-fulfilling (or looping) nature of market valuation tools.[25] Fu, Jin, and Liu (2022) show that the Zestimate updates in response to the initial asking price set by the seller. This could add another layer to sellers' strategies, if pricing on-the-high can increase buyers perception about the value of a home through the Zestimate updating process. Another interesting extension to our work involves linking market valuations to sellers' entry decision, which could potentially affect the aggregate supply of homes.

References

Albrecht, James, Pieter A Gautier, and Susan Vroman (2014). "Efficient entry in competing auctions". *American Economic Review* 104.10, pp. 3288–96.

– (2016). "Directed search in the housing market". *Review of Economic Dynamics* 19, pp. 218–231.

Allen, Marcus T and William H Dare (2004). "The effects of charm listing prices on house transaction prices". *Real Estate Economics* 32.4, pp. 695–713.

[24] *Source*: Kenneth R. Harney, "Zillow faces lawsuit over 'Zestimate' tool that calculates a house's worth", *The Washington Post*, May 10, 2017.

[25] See Malik (2020) for a theoretical framework.

Anenberg, Elliot and Patrick Bayer (2020). “Endogenous Sources Of Volatility In Housing Markets: The Joint Buyer–Seller Problem”. *International Economic Review* 61.3, pp. 1195–1228.

Assad, Stephanie et al. (2020). “Algorithmic pricing and competition: Empirical evidence from the German retail gasoline market”.

Backus, Matthew, Thomas Blake, and Steven Tadelis (2019). “On the empirical content of cheap-talk signaling: An application to bargaining”. *Journal of Political Economy* 127.4, pp. 1599–1628.

Beracha, Eli and Michael J Seiler (2015). “The effect of pricing strategy on home selection and transaction prices: An investigation of the left-most digit effect”. *Journal of Housing Research* 24.2, pp. 147–161.

Calvano, Emilio et al. (2020). “Artificial intelligence, algorithmic pricing, and collusion”. *American Economic Review* 110.10, pp. 3267–97.

Carrillo, Paul E (2012). “An empirical stationary equilibrium search model of the housing market”. *International Economic Review* 53.1, pp. 203–234.

Cowgill, Bo (2018). “Bias and productivity in humans and algorithms: Theory and evidence from resume screening”. *Columbia Business School, Columbia University* 29.

Diaz, Antonia and Belén Jerez (2013). “House prices, sales, and time on the market: A search-theoretic framework”. *International Economic Review* 54.3, pp. 837–872.

Fu, Runshan, Ginger Zhe Jin, and Meng Liu (2022). “Human-Algorithm Interactions: Evidence from Zillow.com”. Working Paper Series 29880.

Genesove, David and Lu Han (2012). “Search and matching in the housing market”. *Journal of Urban Economics* 72.1, pp. 31–45.

Genesove, David and Christopher Mayer (2001). “Loss aversion and seller behavior: Evidence from the housing market”. *The Quarterly Journal of Economics* 116.4, pp. 1233–1260.

Han, Lu and William C Strange (2016). “What is the role of the asking price for a house?” *Journal of Urban Economics* 93, pp. 115–130.

Hoffman, Mitchell, Lisa B Kahn, and Danielle Li (2018). “Discretion in hiring”. *The Quarterly Journal of Economics* 133.2, pp. 765–800.

Kleinberg, Jon et al. (2018). “Human decisions and machine predictions”. *The quarterly journal of economics* 133.1, pp. 237–293.

Lu, Guangli (2018). “Learning from Online Appraisal Information and Housing Prices”. *Available at SSRN 3489522*.

– (2019). “How Machine Learning Mitigates Racial Bias in the US Housing Market”. *Available at SSRN 3489519*.

Ludwig, Jens and Sendhil Mullainathan (2021). “Fragile Algorithms and Fallible Decision-Makers: Lessons from the Justice System”. *Journal of Economic Perspectives* 35.4, pp. 71–96.

Malik, Nikhil (2020). “Does Machine Learning Amplify Pricing Errors in Housing Market?: Economics of ML Feedback Loops”. *Economics of ML Feedback Loops (September 18, 2020)*.

Moen, Espen R, Plamen T Nenov, and Florian Sniekers (2021). “Buying first or selling first in housing markets”. *Journal of the European Economic Association* 19.1, pp. 38–81.

Peters, Michael and Sergei Severinov (1997). “Competition among sellers who offer auctions instead of prices”. *Journal of Economic Theory* 75.1, pp. 141–179.

Piazzesi, Monika, Martin Schneider, and Johannes Stroebel (2020). “Segmented housing search”. *American Economic Review* 110.3, pp. 720–59.

Repetto, Luca and Alex Solís (2020). “The price of inattention: Evidence from the swedish housing market”. *Journal of the European Economic Association* 18.6, pp. 3261–3304.

Ross, Stephen L and Tingyu Zhou (2021). *Loss Aversion in Housing Sales Prices: Evidence from Focal Point Bias*. Tech. rep. National Bureau of Economic Research.

Stevenson, Megan T and Jennifer L Doleac (2021). “Algorithmic risk assessment in the hands of humans”. *Available at SSRN 3489440*.

Wheaton, William C (1990). “Vacancy, search, and prices in a housing market matching model”. *Journal of political Economy* 98.6, pp. 1270–1292.

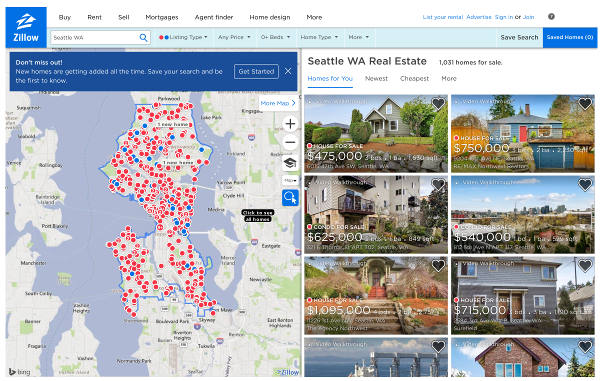
Yu, Shuyi (2020). “Algorithmic Outputs as Information Source: The Effects of Zestimates on Home Prices and Racial Bias in the Housing Market”. *Available at SSRN 3584896*.

Appendix

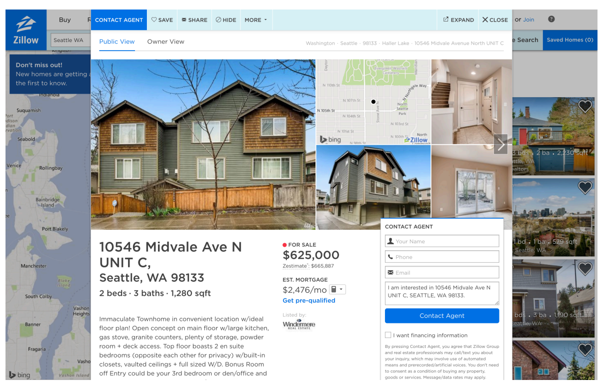
A. Additional Figures

Figure 5: Searching for a House on *Zillow*

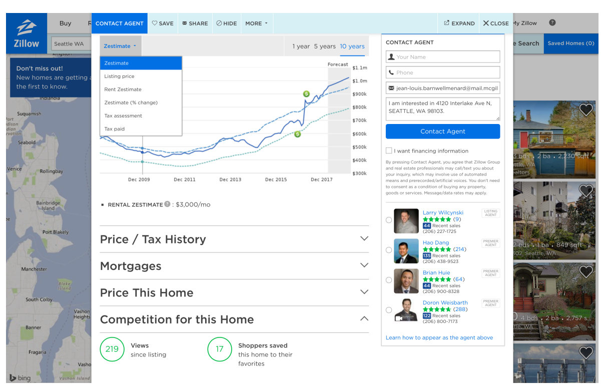
(a) Browsing through housing market segment



(b) Initial house listing view



(c) Historical Zestimate series



Notes: These figures provide a snapshot of the typical information available to buyers on *www.zillow.com* in 2016. The information available includes the number of houses that are up for sale at a given moment in time and their location (panel (a)), details about each listing which includes pricing and observable characteristics (panel (b)), and historical Zestimate time-series.

B. Additional Tables

Table 6: Point Discontinuity Estimates for $\tau = 20$

Discontinuity r	Listing views $\widehat{\beta}_r^{\text{round}}$	Days on market $\widehat{\beta}_r^{\text{round}}$	Sales price $\widehat{\beta}_r^{\text{round}}$	# Obs. $= r$	# Obs. $\in B_r$
250,000	0.34*	−0.19	0.06	21	47
	(0.189)	(0.134)	(0.041)		
300,000	−0.11	−0.06	0.00	28	76
	(0.154)	(0.110)	(0.034)		
350,000	0.04	−0.12	−0.00	37	69
	(0.157)	(0.112)	(0.034)		
400,000	0.01	0.03	0.00	31	64
	(0.162)	(0.115)	(0.035)		
450,000	−0.07	0.00	0.04	47	78
	(0.150)	(0.107)	(0.033)		
500,000	0.03	0.00	0.02	23	66
	(0.167)	(0.119)	(0.037)		
550,000	−0.01	0.20	−0.02	37	57
	(0.179)	(0.128)	(0.039)		
600,000	−0.05	−0.06	0.05	21	49
	(0.186)	(0.133)	(0.041)		

Notes: Based on a sample of 4,322 properties sold in the Seattle metropolitan area in 2015 and 2016. Columns (b), (c), and (d) report penalized splines estimates of the point discontinuities β_r^{round} in

$$y_{jr} = f(\text{dist}_{jr}) + \sum_{r \in \mathcal{R}(\tau)} \left[\beta_r^{\text{bin}} \cdot \mathbb{1}\{a_j \in B_r\} + \beta_r^{\text{round}} \cdot \mathbb{1}\{a_j = r\}\right] + X_j \delta + \epsilon_j$$

when the outcome of interest are listing views on day 2, days on market, and sales price respectively. Column (e) and (f) report the number of observations at the discontinuity and within each bin. Bins with at least 20 observations at the point discontinuity and 20 observations off the discontinuity are included in the regressions ($\tau = 20$). Column (a) reports the estimate at the discontinuity. Robust standard errors are reported in parenthesis. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Point Discontinuity Estimates for $\tau = 10$

Discontinuity r	Listing views $\widehat{\beta}_r^{\text{round}}$	Days on market $\widehat{\beta}_r^{\text{round}}$	Sales price $\widehat{\beta}_r^{\text{round}}$	# Obs. $= r$	# Obs. $\in B_r$
200,000	-0.09	-0.08	-0.13^{**}	13	24
	(0.265)	(0.189)	(0.058)		
240,000	0.01	0.02	0.01	12	24
	(0.262)	(0.187)	(0.057)		
250,000	0.33^{*}	-0.19	0.05	21	47
	(0.189)	(0.135)	(0.041)		
290,000	0.21	-0.07	-0.06	14	26
	(0.252)	(0.180)	(0.055)		
300,000	-0.12	-0.06	0.00	28	76
	(0.155)	(0.110)	(0.034)		
320,000	-0.04	-0.06	0.01	17	34
	(0.220)	(0.157)	(0.048)		
325,000	-0.04	0.05	-0.03	42	57
	(0.193)	(0.138)	(0.042)		
340,000	-0.11	0.10	0.01	17	45
	(0.202)	(0.144)	(0.044)		
350,000	0.04	-0.12	-0.00	37	69
	(0.157)	(0.112)	(0.034)		
360,000	0.29	0.41^{**}	-0.07	12	28
	(0.246)	(0.176)	(0.054)		
380,000	0.17	0.13	0.03	15	25
	(0.263)	(0.187)	(0.057)		
390,000	0.09	0.01	-0.03	13	28
	(0.244)	(0.174)	(0.053)		
400,000	0.01	0.03	-0.00	31	64
	(0.162)	(0.115)	(0.035)		
430,000	-0.19	-0.05	-0.02	12	24
	(0.262)	(0.187)	(0.057)		
440,000	0.06	0.07	0.04	12	23
	(0.268)	(0.191)	(0.059)		
450,000	-0.08	0.01	0.04	47	78
	(0.151)	(0.107)	(0.033)		
500,000	0.03	0.01	0.02	23	66
	(0.167)	(0.119)	(0.037)		
520,000	-0.02	-0.20	-0.04	11	21
	(0.281)	(0.201)	(0.061)		
550,000	-0.02	0.20	-0.02	37	57
	(0.180)	(0.128)	(0.039)		
600,000	-0.06	-0.05	0.05	21	49
	(0.187)	(0.133)	(0.041)		
650,000	0.18	-0.14	-0.00	26	40
	(0.213)	(0.152)	(0.047)		
700,000	0.14	-0.23	0.02	14	35
	(0.222)	(0.158)	(0.048)		

Notes: Based on a sample of 4,322 properties sold in the Seattle metropolitan area in 2015 and 2016. Columns (b), (c), and (d) report penalized splines estimates of the point discontinuities β_r^{round} in

$$y_{jr} = f(\text{dist}_{jr}) + \sum_{r \in \mathcal{R}(\tau)} \left[\beta_r^{\text{bin}} \cdot \mathbb{1}\{a_j \in B_r\} + \beta_r^{\text{round}} \cdot \mathbb{1}\{a_j = r\}\right] + X_j \delta + \epsilon_j$$

when the outcome of interest are listing views on day 2, days on market, and sales price respectively. Column (e) and (f) report the number of observations at the discontinuity and within each bin. Bins with at least 10 observations at the point discontinuity and 10 observations off the discontinuity are included in the regressions ($\tau = 10$). Column (a) reports the estimate at the discontinuity. Robust standard errors are reported in parenthesis. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

C. Proofs

a) LEMMA 1: We demonstrate that $H(x;z,\mu,\theta)$ satisfies first-order stochastic dominance with respect to z, μ, and θ. This is done by showing that the derivatives of $1-H(x;z,\mu,\theta)$ with respect to z,μ or θ are all positive for $0<z<x<\mu$. The proof for part i) and iii) is from Albrecht, Gautier, and Vroman (2016). Part ii) follows the same pattern.

Proof: First, notice that

$$1-H(x;z,\mu,\theta)=1-\frac{\int_z^x g(t;\mu,\theta)dt}{\int_z^\mu g(t;\mu,\theta)dt} \tag{24}$$

$$=1-\frac{\int_0^x g(t;\mu,\theta)dt-\int_0^z g(t;\mu,\theta)dt}{1-\int_0^z g(t;\mu,\theta)dt} \tag{25}$$

$$=\frac{1-\int_0^x g(t;\mu,\theta)dt}{1-\int_0^z g(t;\mu,\theta)dt} \tag{26}$$

$$=\frac{1-e^{-\theta(1-\frac{x}{\mu})}}{1-e^{-\theta(1-\frac{z}{\mu})}} \tag{27}$$

i) The derivative with respect to z is

$$\frac{\partial(1-H(x;z,\mu,\theta))}{\partial z}=\frac{\theta}{\mu}\cdot\frac{e^{-\theta(1-\frac{x}{\mu})}e^{-\theta(1-\frac{z}{\mu})}}{\left(1-e^{-\theta(1-\frac{z}{\mu})}\right)^2}\cdot\left(e^{\theta(1-\frac{x}{\mu})}-1\right) \tag{28}$$

$$>0. \tag{29}$$

Recall that θ is the ratio of buyers to sellers, hence strictly positive. It is then trivial to show that $0<z<x<\mu$ implies

$$e^{\theta(1-\frac{x}{\mu})}>1$$

since $\theta(1-\frac{x}{\mu})>0$.

ii) The derivative with respect to μ is

$$\frac{\partial(1-H(x;z,\mu,\theta))}{\partial\mu} = \frac{\theta}{\mu^2}\cdot\frac{e^{-\theta(1-\frac{x}{\mu})}e^{-\theta(1-\frac{z}{\mu})}}{\left(1-e^{-\theta(1-\frac{z}{\mu})}\right)^2}\cdot\left(xe^{\theta(1-\frac{z}{\mu})}-x-ze^{\theta(1-\frac{x}{\mu})}+z\right) \tag{30}$$

Equation (30) is positive as long as

$$f(x) = xe^{\theta(1-\frac{z}{\mu})}-x-ze^{\theta(1-\frac{x}{\mu})}+z > 0, \quad 0<z<x<\mu. \tag{31}$$

First note that $f(z)=0$ and $f(\mu)>0$. Remains to show that the first derivative $f'(x)>0$ for $0<z<x<\mu$. Differentiating with respect to x yields

$$f'(x) = \frac{\theta z}{\mu}e^{\theta(1-\frac{x}{\mu})}+e^{\theta(1-\frac{z}{\mu})}-1 \tag{32}$$

$$> 0 \tag{33}$$

since $e^{\theta(1-\frac{z}{\mu})}>1$ and $\frac{z\theta}{\mu}e^{\theta(1-\frac{x}{\mu})}>0$ for $0<z<x<\mu$ and $\theta>0$. We conclude that the expression in equation (30) is positive for $0<z<x<\mu$ and $\theta>0$.

iii) The derivative with respect to θ is

$$\frac{\partial(1-H(x;z,\mu,\theta))}{\partial\theta} = \frac{1}{\mu}\cdot\frac{e^{-\theta(1-\frac{x}{\mu})}e^{-\theta(1-\frac{z}{\mu})}}{\left(1-e^{-\theta(1-\frac{z}{\mu})}\right)^2}\cdot\left(ze^{\theta(1-\frac{x}{\mu})}-z-xe^{\theta(1-\frac{z}{\mu})}+x+\mu e^{\theta(1-\frac{z}{\mu})}-\mu e^{\theta(1-\frac{x}{\mu})}\right) \tag{34}$$

The derivative in (34) is positive as long as

$$f(x) = ze^{\theta(1-\frac{x}{\mu})}-z-xe^{\theta(1-\frac{z}{\mu})}+x+\mu e^{\theta(1-\frac{z}{\mu})}-\mu e^{\theta(1-\frac{x}{\mu})} > 0, \quad 0<z<x<\mu. \tag{35}$$

First note that $f(z)=0$ and $f(\mu)=0$. Since $f(x)$ is continuous, remains to show that the second derivative $f''(x) < 0$ for $0 < z < x < \mu$. Differentiating twice with respect to x yields

$$f''(x) = -\frac{\theta^2}{\mu} \cdot \left(1 - \frac{z}{\mu}\right) \cdot e^{\theta(1-\frac{x}{\mu})} < 0 \tag{36}$$

which is negative as required. This concludes the proof.

b) PROPOSITION 1: We want to show that for $0 < z < \mu$ and $\theta > 0$, we have that

i. $\frac{\partial E(P)}{\partial z} > 0$,

ii. $\frac{\partial E(P)}{\partial \mu} > 0$,

iii. $\frac{\partial E(P)}{\partial \theta} > 0$.

Proof:

i) The partial derivative with respect to z is

$$\frac{\partial E(P)}{\partial z} = \int_z^\mu v'(t;\mu)\frac{\partial(1-H(t;z,\mu,\theta))}{\partial z}dt, \tag{37}$$

$$= 2\int_z^\mu \frac{\partial(1-H(t;z,\mu,\theta))}{\partial z}dt. \tag{38}$$

since $v(x;\mu) = 2x - \mu$ and $v'(x;\mu) = 2$. A sufficient condition for the expression in (38) to be positive is that $\frac{\partial(1-H(x;z,\mu,\theta))}{\partial z} > 0$, for $z < x < \mu$. By Lemma 1, $H(x;z,\mu,\theta)$ satisfies first-order stochastic dominance with respect to z, which implies that the terms inside the integral are all positive. Since $0 < z < t < \mu$ for all t, this implies that (38) is positive as required.

ii) The partial derivative with respect to μ is

$$\frac{\partial E(P)}{\partial \mu} = \frac{\partial v(z;\mu)}{\partial \mu} + \int_z^{\mu} \frac{\partial v'(t;\mu)}{\partial \mu}(1 - H(t;z,\mu,\theta))dt + \int_z^{\mu} v'(t;\mu)\frac{\partial(1 - H(t;z,\mu,\theta))}{\partial \mu}dt, \tag{39}$$

$$= 2\int_z^{\mu} \frac{\partial(1 - H(t;z,\mu,\theta)}{\partial \mu}dt - 1. \tag{40}$$

since $\frac{\partial v(x;\mu)}{\partial \mu} = -1$ and $\frac{\partial v'(x;\mu)}{\partial \mu} = 0$. A sufficient condition for (40) to be positive is

$$\int_z^{\mu} \frac{\partial(1 - H(x;z,\mu,\theta)}{\partial \mu} > \frac{1}{2}. \tag{41}$$

We first show that the statement is true for $z = 0$. Consider the function

$$f(\theta) = \int_0^{\mu} \frac{\partial(1 - H(x;0,\mu,\theta)}{\partial \mu}dx \tag{42}$$

$$= \int_0^{\mu} \frac{\theta}{\mu^2 \cdot (1 - e^{-\theta})} \cdot xe^{-\theta(1-\frac{x}{\mu})}dx \tag{43}$$

$$= \frac{\theta + e^{-\theta} - 1}{\theta\,(1 - e^{-\theta})}. \tag{44}$$

We want to show that $f(\theta) > \frac{1}{2}$ for all $\theta > 0$. The proof is done in two steps. First we show that

$$\lim_{\theta \to 0^+} f(\theta) = \frac{1}{2},$$

and then that for $\theta > 0$,

$$f'(\theta) > 0.$$

Taking the limit as θ approaches 0 yields

$$\lim_{\theta\to 0^+} f(\theta) = \lim_{\theta\to 0^+} \frac{\theta + e^{-\theta} - 1}{\theta\left(1 - e^{-\theta}\right)} = \frac{0}{0}, \tag{45}$$

$$\overset{L'H}{=} \lim_{\theta\to 0^+} \frac{1 - e^{-\theta}}{1 - e^{-\theta} + \theta e^{-\theta}} = \frac{0}{0}, \tag{46}$$

$$\overset{L'H}{=} \lim_{\theta\to 0^+} \frac{1}{2-\theta} = \frac{1}{2}. \tag{47}$$

The result is achieved by using l'Hospital rule twice. Remains to show that $f'(\theta) > 0$ for $\theta > 0$. Notice that

$$f'(\theta) = 1 - \frac{\theta^2}{\left(e^{\theta/2} - e^{-\theta/2}\right)^2}, \tag{48}$$

which is greater than 0 as long as

$$g(\theta) = e^{\theta/2} - e^{-\theta/2} - \theta > 0. \tag{49}$$

Note that $g(0) = 0$. To prove the inequality in (49) requires that $g'(\theta) > 0$ for all $\theta > 0$. We have that

$$g'(\theta) = \frac{1}{2}e^{\theta/2} + \frac{1}{2}e^{-\theta/2} - 1. \tag{50}$$

Since $g'(0) = 0$, we require that $g''(\theta) > 0$ for all $\theta > 0$. Differentiating one more time yields

$$g''(\theta) = \frac{1}{4}e^{\theta/2} - \frac{1}{4}e^{-\theta/2} \tag{51}$$

$$> 0 \tag{52}$$

for $\theta > 0$. We conclude that $f(\theta) > \frac{1}{2}$ for all $\theta > 0$.

We now extend this result for a generic choice of z. More precisely, we want to show that expression

(41) holds for $0 < z < \mu$. Consider the function

$$f(\theta) = \int_z^\mu \frac{\partial(1 - H(x; 0, \mu, \theta)}{\partial \mu} dx \tag{53}$$

$$= \int_z^\mu \frac{\theta}{\mu^2} \cdot \frac{e^{-\theta(1-\frac{x}{\mu})} e^{-\theta(1-\frac{z}{\mu})}}{\left(1 - e^{-\theta(1-\frac{z}{\mu})}\right)^2} \cdot \left(x e^{\theta(1-\frac{z}{\mu})} - x - z e^{\theta(1-\frac{x}{\mu})} + z\right) dx. \tag{54}$$

Let $X = \left(1 - \frac{x}{\mu}\right)$ and $Z = \left(1 - \frac{z}{\mu}\right)$. Rewriting the expression above in terms of X and Z gives

$$f(\theta) = \int_0^Z \theta \cdot \frac{e^{-\theta X} e^{-\theta Z}}{(1 - e^{-\theta Z})^2} \cdot \left(e^{\theta Z} - X e^{\theta Z} + X - e^{\theta X} + Z e^{\theta X} - Z\right) dX \tag{55}$$

$$= \frac{\theta}{(1 - e^{-\theta Z})^2} \cdot \int_0^Z \left(e^{-\theta X} - X e^{-\theta X} + X e^{-\theta X} e^{-\theta Z} - e^{-\theta Z} + Z e^{-\theta Z} - Z e^{-\theta X} e^{-\theta Z}\right) dX \tag{56}$$

$$= \frac{\theta(1 - e^{-\theta Z}) - (1 - e^{-\theta Z})^2 - \theta^2 Z(1 - Z) e^{-\theta Z}}{\theta(1 - e^{-\theta Z})^2}. \tag{57}$$

Define $\psi = \theta Z$. Then we have

$$f(\psi, Z) = \frac{\psi(1 - e^{-\psi}) - Z(1 - e^{-\psi})^2 - \psi^2(1 - Z)e^{-\psi}}{\psi(1 - e^{-\psi})^2}. \tag{58}$$

Notice that at $Z = 1$ (i.e. z = 0), we get

$$f(\psi, 1) = \frac{\psi + e^{-\psi} - 1}{\psi(1 - e^{-\psi})} > \frac{1}{2} \tag{59}$$

which was shown above. We now show that $f(\psi, Z) > f(\psi, 1)$, for $0 < Z < 1$ and $\psi > 0$. A sufficient condition is that the derivative of $f(\cdot)$ with respect to Z is negative for $\psi > 0$. Consider the function

$$g(\psi) = \frac{\partial f(\psi, Z)}{\partial Z} \tag{60}$$

$$= -\frac{e^{\psi} - (2+\psi) + e^{-\psi}}{\psi e^{\psi} - 2\psi + \psi e^{-\psi}} \tag{61}$$

We now show that $g(\psi) < 0$ for $\psi > 0$. The proof is in two steps. First we show that

$$\lim_{\psi \to 0^+} g(\psi) = 0,$$

and then that for $\psi > 0$,

$$g'(\psi) < 0.$$

Taking the limit as ψ approaches 0 yields

$$\lim_{\psi \to 0^+} g(\psi) = \lim_{\psi \to 0^+} \; -\frac{e^{\psi} - (2+\psi^2) + e^{-\psi}}{\psi e^{\psi} - 2\psi + \psi e^{-\psi}} = \frac{0}{0} \tag{62}$$

$$\stackrel{L'H}{=} \lim_{\psi \to 0^+} \; -\frac{e^{\psi} - 2\psi - e^{-\psi}}{(1+\psi)e^{\psi} - 2 + (1-\psi)e^{-\psi}} = \frac{0}{0} \tag{63}$$

$$\stackrel{L'H}{=} \lim_{\psi \to 0^+} \; -\frac{e^{\psi} - 2 + e^{-\psi}}{(2+\psi)e^{\psi} - (2-\psi)e^{-\psi}} = \frac{0}{0} \tag{64}$$

$$\stackrel{L'H}{=} \lim_{\psi \to 0^+} \; -\frac{e^{\psi} - e^{-\psi}}{(3+\psi)e^{\psi} + (3-\psi)e^{-\psi}} = 0. \tag{65}$$

Remains to show that $g'(\psi) < 0$ for $\psi > 0$. Notice that

$$g'(\psi) = -\frac{1}{\psi} + \frac{\psi}{(e^{\psi/2} - e^{-\psi/2})^2}, \tag{66}$$

which is smaller than 0 if

$$h(\psi) = e^{\psi/2} - e^{-\psi/2} - \psi > 0. \tag{67}$$

Note that $h(0) = 0$. To prove the inequality in (67) requires that $h'(\psi) > 0$ for all $\psi > 0$. We have that

$$h'(\psi) = \frac{1}{2}e^{\psi/2} + \frac{1}{2}e^{-\psi/2} - 1. \tag{68}$$

Since $h'(0) = 0$, we require that $h''(\psi) > 0$ for all $\psi > 0$. Differentiating one more time yields

$$h''(\psi) = \frac{1}{4}e^{\psi/2} - \frac{1}{4}e^{-\psi/2} \tag{69}$$

$$> 0 \tag{70}$$

for $\psi > 0$. We conclude that $f(\psi, Z) > \frac{1}{2}$ for all $0 < Z < 1$ and $\psi > 0$. This concludes the proof.

iii) The partial derivative with respect to θ is

$$\frac{\partial E(P)}{\partial \theta} = \int_z^{\mu} v'(t;\mu)\frac{\partial(1 - H(t; z, \mu, \theta))}{\partial \theta}dt, \tag{71}$$

$$= 2\int_z^{\mu} \frac{\partial(1 - H(t; z, \mu, \theta))}{\partial \theta}dt. \tag{72}$$

A sufficient condition for the expression in (72) to be positive is that $\frac{\partial(1-H(x;z,\mu,\theta))}{\partial\theta} > 0$, for $z < x < \mu$. By Lemma 1, $H(x; z, \mu, \theta)$ satisfies first-order stochastic dominance with respect to θ, which implies that the terms inside the integral are all positive. Since $0 < z < t < \mu$ for all t, this implies that (72) is positive as required.

c) PROPOSITION 2: We want to show that

i. $\Delta = 0 \Longrightarrow r^* = q$,

ii. $\Delta = 0 \Longrightarrow \theta^*_L = \theta^*_H = \theta$,

iii. $\Delta > 0 \Longrightarrow r^* < q$,

iv. $\Delta > 0 \Longrightarrow \theta_L^* < \theta < \theta_H^*$.

Proof: We start by considering the case where $\Delta = 0$. In this case, the solution to the buyers' optimality condition is

$$\frac{\mu}{\theta_L^2} \cdot \left(1 - e^{-\theta_L(1-\frac{z}{\mu})} - \theta_L e^{-\theta_L(1-\frac{z}{\mu})}\right) = \frac{\mu}{\theta_H^2} \cdot \left(1 - e^{-\theta_H(1-\frac{z}{\mu})} - \theta_H e^{-\theta_L(1-\frac{z}{\mu})}\right). \tag{73}$$

The unique solution is $\theta_L^* = \theta_H^* = \theta$, which occurs at $r^* = q$. In this case it is optimal for buyers to split between sellers in such a way as to keep market tightness equalized across seller types. This proves i) and ii). To show iii) and iv), it is sufficient to show that $\frac{\partial r^*}{\partial \Delta} < 0$ since $r^* = q$ at $\Delta = 0$ (this also proves Corollary 1).

Start from the buyers' optimality condition, that is

$$V_L(r^*, \Delta) - V_H(r^*, \Delta) = 0. \tag{74}$$

Differentiating with respect to r^* yields

$$\frac{\partial V_L}{\partial \theta_L} \cdot \frac{\partial \theta_L}{\partial r^*} + \frac{\partial V_L}{\partial \Delta} \cdot \frac{\partial \Delta}{\partial r^*} - \frac{\partial V_H}{\partial \theta_H} \cdot \frac{\partial \theta_H}{\partial r^*} - \frac{\partial V_H}{\partial \Delta} \cdot \frac{\partial \Delta}{\partial r^*} = 0. \tag{75}$$

Rearranging, we get

$$\frac{\partial r^*}{\partial \Delta} = \frac{\left[\frac{\partial V_H}{\partial \Delta} - \frac{\partial V_L}{\partial \Delta}\right]}{\left[\frac{\partial V_L}{\partial \theta_L} \cdot \frac{\partial \theta_L}{\partial r} - \frac{\partial V_H}{\partial \theta_H} \cdot \frac{\partial \theta_H}{\partial r}\right]}. \tag{76}$$

We now show that r^* is decreasing in Δ by considering each term in turn. The derivatives of V_H and V_L with respect to Δ are

$$\frac{\partial V_L}{\partial \Delta} = \frac{z(1+\theta_L)}{\theta_L} \cdot e^{-\theta_L(1-\frac{z-\Delta}{\mu})} > 0, \tag{77}$$

and

$$\frac{\partial V_H}{\partial \Delta} = -\frac{z(1+\theta_H)}{\theta_H} \cdot e^{-\theta_H(1-\frac{z+\Delta}{\mu})} < 0. \tag{78}$$

which imply that the numerator in (76) is negative. Then we have

$$\begin{aligned}\frac{\partial V_i}{\partial \theta_i} &= \frac{\partial}{\partial \theta_i}\left[\int_{s_i}^{\mu}(1-F(x;\mu))e^{-\theta_i(1-F(x;\mu))}dx\right] \\ &= -\int_{s_i}^{\mu}(1-F(x;\mu))^2 e^{-\theta_i(1-F(x;\mu))}dx \qquad (79) \\ &< 0\end{aligned}$$

for $i = \{H, L\}$ since the term inside the integral is positive at all x. Finally, we have that

$$\frac{\partial \theta_L}{\partial r^*} = -\frac{\theta}{1-q} < 0, \qquad \frac{\partial \theta_H}{\partial r^*} = \frac{\theta}{q} > 0 \tag{80}$$

which together with (79) implies that the denominator in (76) is positive. This concludes the proof.

d) PROPOSITION 3: We want to show that

i. $\frac{\partial E(T|H)}{\partial \Delta} > 0$ for type-H sellers, and

ii. $\frac{\partial E(T|L)}{\partial \Delta} < 0$ for type-L sellers.

Proof: First, we can write the expected time on market for a seller of type i as

$$E(T(s_i, \theta_i) \mid type = i) = \frac{1}{1 - e^{-\theta_i(1-F(s_i;\mu))}} \tag{81}$$

The derivative with respect to Δ is

$$\frac{\partial E(T(s_i,\theta_i) \mid type = i)}{\partial \Delta} = \frac{e^{-\theta_i(1-F(s_i;\mu))}}{\left(1-e^{-\theta_i(1-F(s_i;\mu))}\right)^2} \cdot \left[-(1-F(s_i;\mu)) \cdot \frac{\partial \theta_i}{\partial r^*} \cdot \frac{\partial r^*}{\partial \Delta} + \theta_i f(s_H;\mu) \cdot \frac{\partial s_i}{\partial \Delta}\right] \tag{82}$$

Solving for type-L and type-H separately yields

$$\frac{\partial E(T \mid H)}{\partial \Delta} = \frac{\theta_H e^{-\theta_H(1-F(s_H;\mu))}}{\left(1-e^{-\theta_H(1-F(s_H;\mu))}\right)^2} \cdot \left[-\frac{(1-F(s_H;\mu))}{r^*} \cdot \frac{\partial r^*}{\partial \Delta} + f(s_H;\mu)\right] > 0, \tag{83}$$

and

$$\frac{\partial E(T \mid L)}{\partial \Delta} = \frac{\theta_L e^{-\theta_L(1-F(s_L;\mu))}}{\left(1-e^{-\theta_L(1-F(s_L;\mu))}\right)^2} \cdot \left[\frac{(1-F(s_L;\mu))}{1-r^*} \cdot \frac{\partial r^*}{\partial \Delta} - f(s_L;\mu)\right] < 0. \tag{84}$$

It is easy to see that the expression in (83) is positive while the expression in (84) is negative since the term outside the parenthesis is positive and $\frac{\partial r^*}{\partial \Delta} < 0$. This concludes the proof.

e) PROPOSITION 4: We want to show that

i. $\frac{\partial E(P|H)}{\partial \Delta} = \frac{\partial E(P)}{\partial \theta_H} \cdot \frac{\theta}{q} \cdot \frac{\partial r^*}{\partial \Delta} + \frac{\partial E(P)}{\partial s_H} > 0$, and

ii. $\frac{\partial E(P|L)}{\partial \Delta} = -\frac{\partial E(P)}{\partial \theta_L} \cdot \frac{\theta}{1-q} \cdot \frac{\partial r^*}{\partial \Delta} - \frac{\partial E(P)}{\partial s_L} < 0$.

Proof: We focus on proving statement i). The proof of statement ii) can be formulated using a similar argument. First, recall that

$$\frac{\partial E(P)}{\partial \theta_i} = 2 \cdot \int_{s_i}^{\mu} \frac{\partial(1-H(t;s_i,\mu,\theta_i))}{\partial \theta_i} dt > 0, \tag{85}$$

$$\frac{\partial E(P)}{\partial s_i} = 2 \cdot \int_{s_i}^{\mu} \frac{\partial(1-H(t;s_i,\mu,\theta_i))}{\partial s_i} dt > 0, \tag{86}$$

for $i = \{H, L\}$, with

$$\frac{\partial(1 - H(x; s, \mu, \theta))}{\partial s} = \frac{\theta}{\mu} \cdot \frac{e^{-\theta(1-\frac{s}{\mu})}}{\left(1 - e^{-\theta(1-\frac{s}{\mu})}\right)^2} \cdot \left(1 - e^{-\theta(1-\frac{x}{\mu})}\right) \tag{87}$$

$$> 0. \tag{88}$$

and

$$\frac{\partial(1 - H(x; s, \mu, \theta))}{\partial \theta} = \frac{1}{\mu} \cdot \frac{e^{-\theta(1-\frac{z}{\mu})}}{\left(1 - e^{-\theta(1-\frac{s}{\mu})}\right)^2} \tag{89}$$

$$\cdot \left(s - se^{-\theta(1-\frac{x}{\mu})} - xe^{-\theta(1-\frac{x}{\mu})}e^{\theta(1-\frac{s}{\mu})} + xe^{-\theta(1-\frac{x}{\mu})} + \mu e^{-\theta(1-\frac{x}{\mu})}e^{\theta(1-\frac{s}{\mu})} - \mu\right) \quad > 0 \tag{90}$$

which imply that the terms in i. have opposite signs. Rewrite statement i. as

$$\frac{\partial E(P \mid H)}{\partial \Delta} = \frac{2A\theta}{\mu} \cdot \int_s^\mu \left[sB - sBe^{-\theta(1-\frac{t}{\mu})} - tBCe^{-\theta(1-\frac{t}{\mu})} + tBe^{-\theta(1-\frac{t}{\mu})} + \mu BCe^{-\theta(1-\frac{t}{\mu})} - \mu B + 1 - e^{-\theta(1-\frac{t}{\mu})} \right] dt \tag{91}$$

where $s = s_H$ and $\theta = \theta_H$ for simplicity, $A = \frac{e^{\theta(1-\frac{s}{\mu})}}{(1-e^{\theta(1-\frac{s}{\mu})})^2} > 0$, $B = \frac{1}{q} \cdot \frac{\partial r^*}{\partial \Delta} < 0$, and $C = e^{\theta(1-\frac{s}{\mu})} > 1$. We need to show that the integral in (91) is positive. Regrouping terms, we can rewrite the integral as

$$(sB - \mu B + 1) \cdot \int_s^\mu (1 - e^{-\theta(1-\frac{t}{\mu})}) dt + (-BC + B) \cdot \int_s^\mu te^{-\theta(1-\frac{t}{\mu})} dt \tag{92}$$

It is easy to see that (92) is positive for $0 < s < \mu$ and $\theta > 0$. First, notice that the terms inside the integrals are positive for $t > 0$, which implies that both integrals are positive. Next, we have

$$(sB - \mu B + 1) = -\frac{1}{q} \cdot \frac{\partial r^*}{\partial \Delta} \cdot (\mu - s) + 1 > 0 \tag{93}$$

since $\mu > s$ and $\frac{\partial r^*}{\partial \Delta} < 0$. Finally, we have

$$(-BC + B) = -\frac{1}{q} \cdot \frac{\partial r^*}{\partial \Delta} \cdot (e^{\theta(1-\frac{s}{\mu})} - 1) > 0 \tag{94}$$

since again $\frac{\partial r^*}{\partial \Delta} < 0$ and $e^{\theta(1-\frac{s}{\mu})} > 1$ for $\mu > s$ and $\theta > 0$. This concludes the proof of part i). A similar argument can be made to prove part ii).

f) PROPOSITION 5: We want to show that for two sellers A and B with market valuations $z_A = s - \Delta$ and $z_B = s + \Delta$, then $\Pi_B(s \mid \Delta) > \Pi_A(s \mid \Delta)$.

Proof: Consider the basic case where $\Delta = 0$. We have shown in the previous section that buyers allocate themselves according to $r_i^* = q$ in this simple case. This means that both sellers face the same arrival rate of buyers since

$$\theta_A = \frac{r^*}{q}\theta = \theta, \qquad \text{and} \qquad \theta_B = \frac{1-r^*}{1-q}\theta = \theta,$$

and the same buyers' valuation distribution, i.e. Uniform$(0, \mu(s))$. It follows in this case that

$$\Pi_A(s \mid \Delta = 0) = \Pi_B(s \mid \Delta = 0), \tag{95}$$

that is, both sellers are receiving the exact same revenue. To show that $\Pi_B > \Pi_A$ for $\Delta > 0$ it is sufficient to demonstrate that

$$\frac{\partial \Pi_B(s \mid \Delta)}{\partial \Delta} = \frac{\partial Pr(Sale \mid B)}{\partial \Delta} \cdot \frac{\partial E(P \mid B)}{\partial \Delta} > 0, \tag{96}$$

and

$$\frac{\partial \Pi_A(s \mid \Delta)}{\partial \Delta} = \frac{\partial Pr(Sale \mid A)}{\partial \Delta} \cdot \frac{\partial E(P \mid A)}{\partial \Delta} < 0. \tag{97}$$

More generally, we are able to show that

$$\frac{\partial Pr(Sale \mid B)}{\partial \Delta} > 0, \qquad \frac{\partial Pr(Sale \mid A)}{\partial \Delta} < 0,$$

and

$$\frac{\partial E(P \mid B)}{\partial \Delta} > 0, \qquad \frac{\partial E(P \mid A)}{\partial \Delta} < 0,$$

which together imply that seller B enjoys both a higher probability to sell his house and receives a higher price (in expectation) conditional on a sale than seller A.

First we consider the change in the probability of a sale occurring following an increase in measurement error,

$$\frac{\partial Pr(Sale|type = i)}{\partial \Delta} = \frac{\partial Pr(Sale)}{\partial \theta_i} \cdot \frac{\partial \theta_i}{\partial r^*} \cdot \frac{\partial r^*}{\partial \Delta} + \frac{\partial Pr(Sale)}{\partial \mu_i} \cdot \frac{\partial \mu_i}{\partial z_i} \cdot \frac{\partial z_i}{\partial \Delta}. \tag{98}$$

We consider each terms in turn. First, notice that

$$\frac{\partial Pr(Sale \mid type = i)}{\partial \theta_i} = (1 - F(s; \mu_i)) \cdot e^{-\theta_i(1-F(s;\mu_i))} > 0 \tag{99}$$

Since $\frac{\partial r^*}{\partial \Delta} < 0$ by Corollary 1, the sign of the first term depends on the sign of $\partial \theta_i / \partial r^*$. We have

$$\frac{\partial \theta_A}{\partial r^*} = \frac{\theta}{q} \qquad \text{and} \qquad \frac{\partial \theta_B}{\partial r^*} = -\frac{\theta}{1-q},$$

meaning that the first term in 98 is positive for seller B and negative for seller A. Next, we have that

$$\frac{\partial Pr(Sale)}{\partial \mu_i} = \theta_i f(s; \mu_i) \cdot e^{-\theta_i(1-F(s;\mu_i))} > 0. \tag{100}$$

Since $\frac{\partial z_B}{\partial \Delta} > 0$ and $\frac{\partial z_A}{\partial \Delta} < 0$, and assuming that buyers valuation do not deteriorate with z, that is $\frac{\partial \mu}{\partial z} \geq 0$, we have that the second term is positive for seller B and negative for seller A. We conclude that

$$\frac{\partial Pr(Sale \mid B)}{\partial \Delta} > 0, \qquad \frac{\partial Pr(Sale \mid A)}{\partial \Delta} < 0,$$

as desired.

We now turn to the effect on expected price conditional on a sale. Using chain rule, we have that

$$\frac{\partial E(P \mid Sale)}{\partial \Delta} = \frac{\partial E(P)}{\partial \theta_i} \cdot \frac{\partial \theta_i}{\partial r^*} \cdot \frac{\partial r^*}{\partial \Delta} + \frac{\partial E(P)}{\partial \mu_i} \cdot \frac{\partial \mu_i}{\partial z_i} \cdot \frac{\partial z_i}{\partial \Delta}. \tag{101}$$

We have shown already that the derivative of the expected price with respect to θ and μ are both positive (see Lemma 1). Using the same logic as above, it is easy to show that

$$\frac{\partial E(P \mid B)}{\partial \Delta} > 0 \quad \text{and} \quad \frac{\partial E(P \mid A)}{\partial \Delta} < 0.$$