Assessing the Value of On-line Information Using a Two-sided Equilibrium Search Model in the Real Estate Market

Paul E. Carrillo[*]

May 2005

Work in progress. Comments are welcome.

Abstract

Improvements in information contained in on-line advertisements should have important effects in buyers' and sellers' search behavior and in equilibrium market outcomes. We evaluate such effects in the Real Estate Market (REM) using a two-sided equilibrium search model. The model incorporates, in an equilibrium context, five very important characteristics of the REM: a) buyers' and sellers' search behavior, b) heterogeneity in agents' motivation to trade, c) transaction costs, d) a trading mechanism with posting prices and bargaining, and e) the availability of an online advertising technology. We estimate the model using Maximum Likelihood methods and Multiple Listing Services data. Our estimates suggest that, on average, only 3% of the relevant information that home-buyers collect before making a purchase decision is obtained through on-line ads. Furthermore, we use the estimated model to conduct counterfactual experiments and find that, improvements in online information displayed by Real Estate ads decrease equilibrium prices but increase the time that a property stays on the market.

Keywords: Advertising, Equilibrium Search Models, Internet, Real Estate Market.
JEL Codes: C51, D58, D80, D83.

1 Introduction

The last decade has witnessed an explosive growth in the use of the Internet. Not only the number of users has increased, but also the amount of information received by a web-user in an average session has notably improved. For example, Web Site Optimization, LLC (2005) report that the median amount of information per unit of time transmitted through the web has raised more than ten times during the last five years, providing users with more detailed and accurate information.[1] The information transmitted to web users is limited mostly by technological constraints (speed connections and specialized software) which have been rapidly changing over the last five years, and, according to Vanston et. al (2005), will be significantly less restrictive in the next decade.[2] These technological changes foreseen for the future should affect buyer's and seller's behavior in many electronic markets.

Different types of electronic markets have emerged due to the growing presence of on-line resources in everyday life. For example, in the US, almost every retail store provides their customers with the option of shopping online; there are many other electronic marketplaces where individual buyers and sellers may bid/bargain and trade products online; on the other hand, in markets of highly differentiated products (such as the Real Estate Market), sellers choose to use the web as an advertising platform to show potential buyers the characteristics and prices of their products. Changes in the amount of information displayed online should affect each of these markets differently. That is, markets of relatively

---

[1]According to the Web Site Optimization February 2005 Bandwith Report (which uses Nielsen//Netratings data), in October 1999, the median web-user used a modem connection to receive up to 28.8 kilobytes of information per second (Kbps); in January 2005, the same median user had a connection of at least 300 Kbps.

[2]According to Vanston et. al (2005), by 2006, U.S. broadband penetration will likely be above 50% and a shift to data rates of 24 Mbps to 100 Mbps will have begun. By 2010, 75% broadband penetration is likely, with 10% to 20% of households subscribing to very high-speed-broadband.

non-differentiated goods, whose attributes are well known by the public, should not change significantly if additional information is displayed. However, markets of highly differentiated products with expensive offline search costs should benefit greatly from this new technology, as buyers receive more information about the specific attributes of the products and make more informed decisions.

The widespread use of the Internet has drawn the attention of many researchers who, using micro-data, have attempted to explain how the Internet has affected prices, search behavior, and offline markets.[3] One can think, however, of two elements that determine the importance of electronic markets relative to offline markets: a) the size of the online markets, which is proportional to number of consumers who have Internet access, and b) the amount and quality of the information that Internet marketplaces provide to consumers. Clearly, changes in each of these elements should affect the electronic markets differently, and we are not aware of any study that has attempted to separate these effects. In this research paper, we attempt to explain how improvements in the information technology affect buyer's and seller's behavior in online markets of highly differentiated products, and to assess the economic value of such technology.

We shall focus our attention on the Real Estate Market (REM) for several reasons. First, a housing unit is an example of a perfectly differentiated product because it can be uniquely described by a large set of characteristics.[4] Second, housing units, in the vast majority of cases, are advertised through the Internet. Third, buyers incur remarkable high search costs in the REM which are significantly reduced with the use of the internet.[5] Fourth, online

---

[3]See, for example, Brown and Goolsbee (2002), Goolsbee (2000), and D'urso (2002).

[4]Notice that, because housing units cannot share the same exact geographical location, even two identical units (in all other dimensions) are differentiated.

[5]For instance, according to the 2004 National Association of Realtors Profile of Home Buyers and Sellers,

housing sites have notably improved the amount of information they display by incorporating pictures and virtual tours to the existing Multiple Listing Services (MLS), and important improvements are forecasted for the near future. Finally, the housing market is one of the largest and most important in the US economy.[6]

Classical economic theory cannot adequately evaluate the effects of improvements in online information technology in the housing market because some assumptions underlying the traditional supply and demand model are invalid in the presence of market imperfections such as transactions costs or asymmetric information. For example, when there are high search costs, house buyers have market power, because it is costly for sellers to keep their house on the market; and sellers have market power because buyers incur pecuniary and non-pecuniary expenses when visiting properties. Thus, it is not suitable to assume that buyers and sellers are price-takers, the fundamental assumption required to apply the traditional model of supply and demand. To appropriately assess these effects, it is important to specify a model that depicts individual behavior of buyers and sellers that is consistent with the nature of the imperfections of the REM.

In this paper, we specify and estimate an equilibrium two-sided search model that depicts many of its real-life features. The theoretical model modifies the framework of existing equilibrium search models in the labor literature to capture the unique nature of the REM.[7] To our knowledge, it is the first attempt in the literature to model in an equilibrium context five very important characteristics of the REM: a) buyers' and sellers' search behavior, b)

the use of the internet in searching for a home has risen along with the level of internet penetration, rising from 2 percent of buyers in 1997 to 74 percent in 2004.

[6]According to the US Census and the National Association of Realtors, new and existing home sales in the US were approximately $1.5 trillion in 2003.

[7]A revision of these papers is presented in the next section.

heterogeneity in agents' motivation to trade, c) transaction costs, d) a trading mechanism with posting prices and bargaining, and e) the availability of an online advertising technology.

In our model, both buyers and sellers are infinitely lived agents who search simultaneously for a potential trading partner. Buyers and sellers are heterogeneous, as some individuals are more motivated to trade than others. Every seller uses the internet to advertise her product (housing unit) displaying online its posting price and some of its features. Each period, buyers use the internet to sample price postings (ads), learn the characteristics of the unit, and decide whether or not to visit a seller. There is an idiosyncratic random buyer-home value that can be learned fully only when a potential buyer visits a home. Part of this value, however, is revealed to the buyer when she looks at the ad on the internet. Sellers wait for a potential buyer to visit their property. When two potential trading partners meet, they play a well defined bargaining game, and trade may or may not occur. If a buyer and a seller engage in trade, they leave the market forever; otherwise, they return to the market and search for a trading opportunity next period. Sellers' choose a posting price and a reservation value (price floor) given the characteristics of their home and their motivation to trade. Buyers, on the other hand, decide when to visit a property and when to purchase a home given their motivation to trade. An endogenous equilibrium in these optimal buyers' and sellers' strategies defines the solution to the theoretical model.

To estimate our theoretical model, we follow the growing literature on estimation of equilibrium search models and use maximum likelihood methods.[8] The data used for estimation consists of Multiple Listing Services data for real estate transactions in Charlottesville City and Albemarle County (VA) during the years 2000 through 2002.

---

[8]See, for example, Eckstein and Wolpin (1990), Kiefer and Neumann (1994) and Bunzel et. al. (2001).

Our estimates suggest that, on average, only 3% of the relevant information that home-buyers collect before making a purchase decision is obtained through on-line ads. Furthermore, we use the estimated model to conduct counterfactual experiments and predict how improvements in the information technology affect the average posting price, transaction price, and time until a property is sold. We find that additional online information decreases market prices. For instance, if buyers could obtain from an online ad all the relevant information they need to make a purchase decision, posting and transaction prices would decrease by 2.9% and 3.1%, respectively; on the other hand, the average time that a home stays on the market would increase in about four days.

In the next section, we compare our research to the existing literature. The third section presents the theoretical model. In the fourth section, we introduce the data and a preliminary reduced form analysis of the data. Section 5 discusses the estimation methods. The sixth section presents our results, and the last one concludes.

2   Literature review

There is a growing literature that seeks to explain how the use of the internet has affected online and offline markets. For example, Brown and Goolsbee (2002) examine a relative homogeneous product (term life insurance) and conclude that the use of the internet has reduced term life prices by 8 to 15 percent. In a different study, Goolsbee (2000) suggests that applying existing sales taxes to Internet commerce might reduce the number of online buyers by up to 24 percent. D'urso (2002) studies the impact of internet use on the duration of search in the housing market, and, using Instrumental Variable Quantile Regression methods she finds that it's use increases the duration of home search relative to employing other

conventional search methods. While these previous studies identify the global effect of internet use in some specific markets, they do not assess how improved information technology may influence buyers' and sellers' behavior in these online markets.

To a certain extent, our work is related to Anderson and Renault (2004)'s research, since both studies explore the role of informative advertising. However, while Anderson and Renault (2004) characterize the monopolist's decision of advertising prices and/or certain information about the product characteristics, we treat the seller's decision of what to advertise as exogenous.[9]

The theoretical model was motivated by the economic theory of search. Standard search models assume identical agents search sequentially, drawing price offers from a known stationary distribution. They show how the characteristics of the market -discount rates, distribution of prices, and transaction costs- affect search behavior, and have been widely used to describe posting and transaction prices in the REM.[10] Many variants of this standard model have been created, but these typically do not explain the origin of the price distribution and are inconsistent with rational conduct.[11]

Using assumptions about the sellers' optimal behavior and agents' heterogeneity, Burdett and Judd (1983), Albrecht and Axell (1984), and Mortensen (1990) have introduced

---

[9]This may not be a bad assumption for the Real Estate Market, because, in order to be listed in the Multiple Listing Services, a seller must provide certain required information (such as the list price and the square footage, for example).

[10]Yinger (1981) was one of the first to use a search model in a formal analysis of real estate markets. Yavas (1992) presents a single period matching model examining the impact of buyers', sellers', and brokers' search costs on transaction prices. Yavas and Yang (1995) extend the Yavas (1992) paper introducing asking prices in the same single-period matching model. Horowitz (1992) models sellers' behavior in the framework of a one-sided search model where list and reservation prices are endogenous. Arnold (1999) presents a one sided search-bargaining model where asking prices affect the rate at which potential buyers visit the property.

[11]That is, if individuals are homogeneous and face the same price distribution, they will have the same reservation price. Rational price-posting sellers would then post only that price, leading to a degenerate price distribution. If there is a degenerate price distribution, agents will not search, but search is precisely what these models are trying to explain (Diamond 1971).

theoretical models that determine non-degenerate price distributions and search behavior as an equilibrium outcome.[12] However, these models assume that only one side of the market searches, and cannot adequately describe a market where both buyers and sellers search simultaneously (for example, the REM).

Mortensen and Wright (2002) introduce a two-sided search model where buyers and sellers with heterogeneous preferences simultaneously search and bargain over the price of an indivisible object. Agents have perfect information about each other's preferences at the time they bargain, and the reservation prices, the size of the market, and the price distribution arise as an outcome of the equilibrium. While their model successfully describes search behavior in a two-sided market, it does not explain two very important features of the REM: heterogeneity in the agents' motivation to trade, and posting prices and bargaining.

We introduce heterogeneity in the agents' motivation to trade by assuming that the buyer's and seller's intertemporal discount factors are random variables.[13] To introduce posting prices in our equilibrium search model we assume (as in Chen and Rosenthal 1996)

---

[12]Burdett and Judd (1983) assume that identical agents sample a subset of price offers. Agents with more than one offer are able to bargain for a price below their reservation price. This induces the sellers to set their prices using a mixed strategy that generates an endogenous distribution of prices in the market. Albrecht and Axell (1984) introduce heterogeneity in workers and firms (i.e. buyers and sellers). They assume that firms have different levels of productivity and workers differ in their valuation of leisure. Workers look sequentially for a job, and as firms maximize their profits, it may be optimal for heterogeneous firms to have different wage policies. Mortensen (1990) models the labor market by assuming that workers search while employed and change jobs if they sample a wage realization greater that the current wage. Firms offering high wages are able to keep workers for a longer period of time than firms offering low wages, but their per-period surplus is smaller. Therefore, firms face a trade-off between offering short and long run profits. In equilibrium, different wage policies are optimal for each firm, giving rise to an endogenous distribution of wages.

[13]The level of the agents' motivation to trade is clearly one of the most important elements in the housing market. However, there is little literature that addresses this topic. Glower et al. (1998) explain how sellers' motivation fits into the standard (one-sided) search model. Then they use a small sample of sellers to explain the role of their "motivation" in determining selling time, list price and sale price and use their empirical results to test the theoretical hypothesis. They conclude that motivation affects the expected time in the market and the sale price, but not the posting price. Our research addresses the same issue using a general equilibrium approach.

that the list price constitutes a price ceiling and a commitment device. Unlike Chen and Rosenthal (1996), however, we use a simpler bargaining game that allows us to solve for the seller's list price and reservation value analytically.

Because equilibrium search models provide a natural interpretation of interesting market phenomena, the estimation of such models has received considerable attention. Eckstein and Wolpin (1990) estimate a generalization of the Albrecht and Axell model.[14] They use assumptions about the distribution of preferences and technology to identify the parameters of their model using workers' data only. Their estimated model fails to conform to the data, and measurement error accounts for almost all of the dispersion in wages. Kiefer and Neumann (1994) estimate a version of the model of Mortensen (1990). Bunzel et. al. (2001) estimate a version of Mortensen (1990) with several variations that fit the data better (measurement error and heterogeneity in firms productivity). The empirical implementation of these models is still developing. Future research needs to address identification issues as well as generate models that fit the data more closely. The estimation of our structural model adds to the growing literature related to estimation of equilibrium search models.

3 Theoretical Model

3.1 The market

There exists a market with risk-neutral, infinitely lived agents. The agents are households who either are actively searching for a home (buyers), or who have a home for sale (sellers). Agents are alike except for how motivated they are to trade.

To model this heterogeneity, we define β_b and β_s as the buyers' and sellers' value of

---

[14]Canals and Stern (2001) provide a comprehensive survey of empirical search models.

an opportunity to trade in the next period relative to the same opportunity in the current period and assume that they are random variables with constant distributions K_b and K_s, respectively. Formally, each of these discount factors are formed by two components; that is, $\beta_b = \beta^o \beta_b^*$ and $\beta_s = \beta^o \beta_s^*$. The first component $\beta^o = \frac{1}{1+r}$ discounts the future using a discount rate r that is common to all buyers and sellers. On the other hand, the second components, β_b^* and β_s^*, are random variables that capture idiosyncratic differences in buyers' and sellers' motivation when buying or selling their home. Note that the lower a household's β the more motivated and eager it is to engage in trade.

A home is considered to be an indivisible good that can be described fully by a vector of characteristics X from which both buyers and sellers derive utility. Define s_f as the per-period utility flow that sellers obtain by owning this good, and let $s = \frac{s_f}{1 - \beta^o}$ be the level of lifetime utility that sellers obtain by owning this good; furthermore, let $s = X\gamma$, where γ is a vector of parameters. Notice that s represents the quality of the home and that properties with different features X may provide sellers with the same level of utility s. To model heterogeneity in homes' characteristics, we assume that s is distributed according to an exogenous distribution Ψ, which is common knowledge to every agent in the market.

The lifetime utility that properties provide to buyers varies for each buyer-home combination. To model this assumption, we let $\tilde{b}$ be a random buyer-home match value that captures the lifetime utility a specific buyer derives from owning one particular property. Furthermore, let $\tilde{b}$ depend on the home quality and two independent (from each other and for any home-and-buyer combination) mean zero random errors b^o and b^u. That is, $\tilde{b} = \delta s + b^o + b^u$, where δ is a scalar parameter and b^o and b^u are random variables with exogenous cumulative distributions G_o and G_u, respectively. Notice that the parameter δ captures average percent

9

differences in the properties' valuations between buyers and sellers, and realizations of $\tilde{b}$ correspond to the specific value that one buyer assigns to one particular property.[15]

A seller joins the housing market by placing an ad that informs all the other agents: a) that her home is for sale, b) the posting price and, c) the home characteristics. As in Horowitz (1992), and Chen and Rosenthal (1996a and 1996b), we assume that the posting price constitutes a price ceiling and a commitment device; that is, if a potential buyer wants to buy the product at the posting price, the seller is obligated to engage in trade.[16]

Buyers search "home for sale" ads sequentially. Every period, a buyer samples an ad that provides her information about the home's characteristics X, its posting price p_s, and b^o (a fraction of her total random buyer-home match value). When a buyer observes an ad, she decides whether or not to visit the home. If she decides to visit the home, the buyer tours the house and b^u is revealed to her.

After the buyer has visited the home, she meets the seller and both bargain over the transaction price. For simplicity, we adopt a reduced form representation of the bargaining process. With probability θ, the seller is not willing to accept counter-offers, and the posting price p_s constitutes a take-it-or-leave-it offer to the buyer. With probability $(1-\theta)$, the buyer

---

[15]The reader may notice that, for a given set of housing characteristics X, every seller receives the same lifetime utility $s = X\gamma$ from owning this particular house. On the other hand, we assume that the lifetime utility $\tilde{b}$ that buyers receive from owning the same good varies for each buyer-home combination. We are aware that, to the extent that today's buyers may become sellers in the future, this assumption may not be realistic. However, we choose to use this assumption for simplicity only, since it simplifies significantly the analytical solution to the model. A natural way to relax this assumption consists of including a random component in s, which does not change the nature nor the main results of our model.

[16]It is worth acknowledging, however, that houses do sometimes sell above the posting price. In certain cases, sellers are surprised by high demand and in others, sellers deliberately set low asking prices to foster competition among buyers. These cases, however, are a) infrequent (1.5% in our sample), b) almost always regarded as a seller who is "bargaining in bad faith", and, c) in some locations, ruled out by the existence of legal contracts that give the real estate broker the right to damages in the event that a seller does not agree to trade at the list price (Chen and Rosenthal 1996a). In any event, the assumptions of our paper rule out the possibility that the transaction price is above the posting price.

has the option to make a counter take-it-or-leave-it offer p_b to the seller. It is assumed that, once a buyer has visited a property, she has perfect information about the seller's preferences. That is, if she makes a counter take-it-or-leave-it offer, she will bid the seller's reservation value R_s (the minimum price at which she is willing to sell her property). During the meeting, the buyer decides whether she buys the home (paying either p_s or p_b), or searches again for a new ad (posting price) next period. When a buyer (seller) buys (sells) a property, she exits the market forever.[17]

3.2 The seller's problem

From a seller's point of view, trade occurs only if a buyer visits her property and is willing to trade, either at the posting price p_s or at her reservation value R_s. Let $q(p_s|s)$ (to be determined endogenously) be the rate at which buyers visit a particular seller who owns a type s property and has posted a price p_s; also define $\gamma_b(p_s|s)$ $(\gamma_s(p_s|s))$ as the probability that a buyer is willing to buy this property given that she has visited the home and did (did not) have the opportunity to make a counter offer (both to be determined endogenously).

Using these assumptions we define the seller's expected gain from trade and searching as

$$
\begin{aligned}
\Pi_s^e &= q\theta\left[\gamma_s(p_s - s) + (1 - \gamma_s)W_s\right] + \qquad\qquad\qquad\qquad\qquad (1)\\
&\quad q(1 - \theta)\left[\gamma_b(R_s - s) + (1 - \gamma_b)W_s\right] + [1 - q]W_s \\
&= q\left[\theta\gamma_s(p_s - s) + (1 - \theta)\gamma_b(R_s - s)\right] + [1 - q(\theta\gamma_s + (1 - \theta)\gamma_b)]W_s,
\end{aligned}
$$

---

[17]Instead, we could use a Rubenstein (1982)-type bargaining model where the surplus from trade is shared in fixed proportions between the buyer and the seller (such as in Chen and Rosenthal (1996a) or in Mortensen and Wright (2002), for example). Using this approach, however, we could not find analytical solutions for the optimal seller's posting and reservation prices.

and

$$W_s = \beta_s E[\max\{p - s, W_s\}], \tag{2}$$

where Π_s^e is the seller's expected profit from trade, W_s is her expected gain from searching (her value of search), and p is the random (from the seller's point of view) transaction price.

Equation (1) states that, in every period, there is $q\theta\gamma_s$ probability that a seller sells her home for the posting price and obtains $p_s - s$ profit when trading; with probability $q(1-\theta)\gamma_b$, trade occurs at the seller's reservation value, in which case her gains from trade are $R_s - s$; finally, if trade does not happen, she returns to the market and keeps her value of search W_s. As presented in equation (2), the seller's value of search is the discounted expected value of having an opportunity to trade next period, and it represents the utility that a seller obtains by staying in the market.

The seller's problem consists of choosing the optimal reservation value R_s^* and posting price p_s^* that simultaneously maximize her expected profit and value of search.

First, let us solve the seller's search problem. For any given p_s^*, and using the assumptions of our bargaining game, we work out the expectation in equation (2) and obtain

$$W_s = \beta_s \left\{ q\theta\gamma_s(p_s^* - s) + q(1-\theta)\gamma_b(R_s - s) + [1 - q(\theta\gamma_s + (1-\theta)\gamma_b)]W_s \right\}.$$

Furthermore, following the solution techniques of standard search models, notice that any optimal seller's behavior necessarily implies that $R_s^* = W_s^* + s$.[18] We replace this optimality condition in the previous equation, rearrange, and solve for the optimal reservation value

$$R_s^* = \frac{\beta_s\theta(1 - \phi(p_s^*|s))p_s^* + (1 - \beta_s)s}{1 - \beta_s[1 - \theta(1 - \phi(p_s^*|s))]}, \tag{3}$$

---

[18]See, for example, Lipmann and McCall (1976).

where, for notational simplicity, we have defined $1 - \phi(p_s|s)$ as the probability that, given that the posting price is a take-it-or-leave-it offer to the buyer, a type s home sells for the posting price; that is: $1 - \phi(p_s|s) = q(p_s|s)\gamma_s(p_s|s)$.[19]

The previous result allows us to find the seller's optimal posting price p_s^*. To determine p_s^*, we substitute the optimality condition $(R_s^* = W_s^* + s)$ in equation (1) and obtain that, for any R_s^*,

$$\Pi_s^e = \theta q \gamma_s (p_s - s) + (1 - \theta q \gamma_s)(R_s^* - s).$$

Differentiating this equation with respect to p_s, we derive that the optimal seller's posting price p_s^* solves

$$p_s^* - R_s^* = \frac{1 - \phi(p_s^*|s)}{\phi'(p_s^*|s)}. \tag{4}$$

Theorem 1: As long as, (a) the hazard function $h(p_s^*|s) = \frac{\phi'(p_s^*|s)}{1-\phi(p_s^*|s)}$ is non-decreasing in p_s^*, and (b) ϕ is decreasing in s, the optimal seller's posting price and reservation value are defined by the unique pair $\{p_s^*, R_s^*\}$ that solves equations (3) and (4) simultaneously. For any seller, and for any β_s and $s : p_s^*(\beta_s, s) \geq R_s^*(\beta_s, s) \geq s$; in addition, these functions are increasing in both arguments. Proof in Appendix.

The two assumptions of Theorem 1 are not restrictive in any sense. Condition (a) is a commonly used standard assumption about the shape of the demand function that guarantees the existence of a unique solution in similar problems and is satisfied by many standard distributions, such as the normal, uniform, and exponential.[20] However, notice that $\phi(p_s^*|s)$ will be determined endogenously in our model. Thus, to guarantee that this assumption is

---

[19]Notice that $1 - \phi(p_s|s)$ resembles to a traditional demand function.

[20]This assumption is equivalent to $1 - \phi(p_s^*|s)$ being strictly log-concave, and is commonly used in the literature (for example, in Anderson and Renault 2004 and in Chen and Rosenthal 1996b). For a list of distributions satisfying this assumption as well as some of its other applications, see Bagnoli and Bergstrom (1989).

satisfied in equilibrium, we need to specify this function, and place certain restrictions in the exogenous distributions of our model (to be done in the next sections). Condition (b) makes the reasonable conjecture that, everything else constant, the higher the quality of a home, the higher the probability it is sold.[21]

The results of Theorem 1 are all intuititive. First, we expect that a rational seller would post a sale price that is at least as high as her reservation value; and that the minimum price at which she is willing to sell her product must be no less than her outside option (the utility that she gets by keeping the product). Second, the model predicts that both posting and reservation prices diminish as the seller's motivation to trade increases, and these predictions are consistent with other findings (Glower et al. 1998). Third, higher quality properties sell for higher prices. Finally, the monotonicity of $p_s^*(\beta_s, s)$ and $R_s^*(\beta_s, s)$ should facilitate us the derivation of $\phi(p_s^*|s)$.

3.3   The buyer's two stage search problem

The buyer's optimal behavior can be described with a two stage search model. In the first stage, she samples an ad and uses the information contained in it to decide whether she should visit this home. In the second stage, given that she has visited this property, she obtains additional information about its value and the outcome of the bargaining game and decides between purchasing it or searching for a new ad next period. In this section, we formally describe the solution to such problem.

When a buyer picks an ad, she observes the home's features X, its posting price p_s, and b^o, a component of her buyer-home match value. From her point of view, a pair $\{p_s, s\}$ is an

---

[21]This condition says that an increase in s increases p_s^* in the sense of first-order stochastic dominance.

independent realization from the joint distribution of posting prices and home characteristics $\Gamma(p_s, s)$, while a specific value of b^o is an independent realization from the distribution G_o. Notice that $\Gamma(p_s, s)$ is well defined by the optimal sellers' pricing strategies $p_s^*(\beta_s, s)$ as well as the exogenous distributions of discount factors K_s and valuations Ψ.

When the buyer visits a property, she pays a known visiting cost c_b.[22] During the visit, the other component of her buyer-home match value b^u and the outcome of the bargaining game are revealed to her. At this stage, she chooses between buying the property and staying in the market for another period. We assume that the buyer chooses the optimal strategy that maximizes her expected value of search.

Before she decides to visit a home, her value of search W_b is the discounted expected utility of the maximum between visiting a property and waiting for another ad next period

$$W_b = \beta_b \int \int \int \max\left\{U^e(p_s^*, s, b^o, W_b), W_b\right\} d\Gamma(p_s^*, s)dG_o(b^o). \tag{5}$$

$U^e(p_s^*, s, b^o, W_b)$ is the buyer's expected value of having an opportunity to visit a property at the time she looks at the listing, before observing the realization of b^u, and before knowing who gets to make the take-it-or-leave-it offer

$$
\begin{aligned}
U^e(p_s^*, s, b^o, W_b) &= \theta \int \max\{b^o + b^u + \delta s - p_s^*, W_b\}dG_u(b^u) \\
&\quad + (1-\theta) \int \max\{b^o + b^u + \delta s - p_b, W_b\}dG_u(b^u) - c_b.
\end{aligned}
\tag{6}
$$

Because buyers know the optimal strategies $p_s^*(\beta_s, s)$ and $R_s^*(\beta_s, s)$, they are aware of a seller's reservation value $R_s^*(p_s^*, s)$ once they have observed p_s and s in an ad. Hence, if buyers

---

[22]Visiting costs include transportation and monetary opportunity costs (time costs), as well as non pecuniary (emotional) costs of touring a property.

have the opportunity to make a counter offer, it is optimal for them to ask $p_b = R_s^*(p_s^*, s)$

which does not depend on b^u. Using this result, we integrate the right hand side of equation

(6) by parts and obtain[23]

$$U^e(p_s^*, s, b^o, W_b) = W_b + j(p_s^*, s, b^o, W_b), \tag{7}$$

where

$$j(p_s^*, s, b^o, W_b) = \theta \int_A [1 - G_u(b^u)] db^u + (1 - \theta) \int_B [1 - G_u(b^u)] db^u - c_b,$$

$$A = \{b^u : b^u \geq W_b + p_s^* - \delta s - b^o\},$$

and

$$B = \{b^u : b^u \geq W_b + R_s^*(p_s^*, s) - \delta s - b^o\}.$$

Now, let us find the value of search W_b^* that solves equations (5) and (7). We substitute

equation (7) in (5), rearrange, and find

$$W_b^* = \frac{\beta_b}{1 - \beta_b} \int \int \int \max\{j(p_s^*, s, b^o, W_b^*), 0\} \, d\Gamma(p_s^*, s) dG_o(b^o). \tag{8}$$

It is easy to see that there is a unique W_b^* that solves equation (8).[24]

It is straightforward to show that U^e is decreasing with respect to the posting price.[25]

This fact implies that it is optimal for the buyer to follow a reservation strategy such that,

given a particular set of housing characteristics s and buyer-home match component b^o, she

---

[23]Details of the integration are provided in the appendix.

[24]The right hand side of equation (8) is no less than zero, and decreasing in W_b^* (since $j(p_s^*, s, b^o, W_b^*)$ is clearly decreasing in W_b^*). On the other hand, the right hand side of equation (8) crosses the origin and has a positive slope. Thus, a unique solution W_b^* exists.

[25]We use equation (7) -and Leibnitz rule-, to show that U^e is monotone

$$\frac{\partial U^e}{\partial p_s^*} = -\theta(1 - G_u(V_p)) - (1 - \theta)\frac{\partial R_s^*}{\partial p_s^*}(1 - G_u(V_r)) \leq 0,$$

where $V_p = W_b + p_s^* - \delta s - b^o$, and $V_r = W_b + R_s^*(p_s^*|s) - \delta s - b^o$.

visits the property if and only if the posting price p_s^* is below a reservation price $p_b^{r*}(s, b^o)$. Hence, for any s and b^o, the optimal buyer's reservation price must be such that her value of having an opportunity to visit a property equals her value of search

$$U^e(p_b^{r*}, s, b^o, W_b^*) = W_b^*.$$

To find the optimal $p_b^{r*}(s, b^o)$, we replace this optimality condition in (7) and solve

$$\theta \int_{A(p_b^{r*}, s, b^o)} [1 - G_u(b^u)]db^u + (1 - \theta) \int_{B(p_b^{r*}, s, b^o)} [1 - G_u(b^u)]db^u = c_b, \qquad (9)$$

where

$$A(p_b^{r*}, s, b^o) = \{b^u : b^u \geq W_b^* + p_b^{r*} - \delta s - b^o\},$$

and

$$B(p_b^{r*}, s, b^o) = \{b^u : b^u \geq W_b^* + R_s^*(p_b^{r*}, s) - \delta s - b^o\}.$$

Theorem 2: The solution to the buyer's two-step search is defined by a unique value W_b^* and a function $p_b^{r*}(s, b^o)$ that solve equations (8) and (9) respectively, along with the optimal strategies: (a) visit a property if, given a particular realization of s and b^o, an observed ad's posting price $p_s \leq p_b^{r*}(s, b^o)$; (b) if she has visited a home and does not have the opportunity to make a counter offer, she buys the property if and only if $\tilde{b} - p_s > W_b^*$; and (c) if she has visited a home and has the opportunity to make a counter offer, she should make a take-it-or-leave-it-offer (which will always be accepted) of $p_b = R_s^*(p_s, s)$ if and only if $\tilde{b} - R_s^*(p_s, s) > W_b^*$. In addition, W_b^* is increasing in β_b, while p_b^{r*} is decreasing in β_b. Proof in appendix.

It is useful to analyze the case when the distributions of s and b^o are degenerate. When this is the case, we are able to find analytical solutions for W_b^* and p_b^{r*} and provide an

intuitive interpretation of the buyer's optimal decisions.

Theorem 2a: When the distributions of s and b^o are degenerate, the solution to the buyer's two-step search model is defined by the unique pair $\{W_b^*, p_b^{r*}\}$ that solves equations (10) and (11) simultaneously, along with the optimal strategies described in (a), (b), and (c) in Theorem 2.

$$\int^{p_b^{r*}} \Gamma(p_s^*)(-\frac{\partial U^e}{\partial p_s^*})dp_s^* \quad = \quad \frac{1-\beta_b}{\beta_b}W_b^* \quad (10)$$

$$\theta \int_{W_b^*+p_b^{r*}-\delta s} [1 - G_u(b^u)]db^u + (1-\theta) \int_{W_b^*+R_s^*(p_b^{r*})-\delta s} [1 - G_u(b^u)]db^u \quad = \quad c_b \quad (11)$$

Proof in appendix.

Equation (10) states that the expected benefits from sampling a posting price lower than the buyer's reservation price p_b^{r*} (the left hand side) should be the same as the per-period expected return of staying in the market (since $\frac{1-\beta_b}{\beta_b}$ equals the per-period discount rate). Equation (11) implies that the optimal W_b^* and p_b^{r*} must be such that, the buyer's expected benefit from visiting a property equals her visiting cost.

3.4 Equilibrium

From the sellers' point of view, there is a distribution of heterogenous buyers in the market, each one of them with a different value of search. Because sellers are rational individuals who know the optimal buyers' strategies $p_b^{r*}(\beta_b, s, b^o)$, $W_b^*(\beta_b)$, and the relevant exogenous cumulative distribution functions, each is able to determine the probability that a buyer visits her property and is willing to trade.

The probability that a buyer visits a seller who has posted a price p_s and owns a

type s home is $q^*(p_s|s) = \Pr\{p_b^{r*}(\beta_b, s, b^o) > p_s\}$. In addition, $\gamma_s^*(p_s|s) = \Pr\{\tilde{b} - p_s > W_b^*(\beta_b)|p_b^{r*}(\beta_b, s, b^o) > p_s\}$ is the probability that the buyer is willing to buy the property given that she has visited the home and did not have the opportunity to make a counteroffer; and $\gamma_b^*(p_s|s) = \Pr\{\tilde{b} - R_s^*(p_s) > W_b^*(\beta_b)|p_b^{r*}(\beta_b, s, b^o) > p_s\}$ is the probability that the buyer buys the property given that she has visited the property and had the option to make a counteroffer.

Notice that $q^*(p_s|s), \gamma_s^*(p_s|s)$, and $\gamma_b^*(p_s|s)$ are well defined by the buyer's optimal strategies and the exogenous cumulative distribution functions G_o, G_u and K_b. Furthermore, we show in the appendix that $1 - \phi^*(p_s|s) = q^*(p_s|s)\gamma_s^*(p_s|s)$ is a well defined decreasing function.

Equilibrium conditions: The equilibrium of the model is determined by a fixed point in the following probability distributions:

$$q(p_s|s) = q^*(p_s|s) \quad ; \quad \gamma_s^*(p_s|s) = \gamma_s(p_s|s) \quad ; \quad \gamma_b^*(p_s|s) = \gamma_b(p_s|s)$$

Thus, the solution to the model is defined by a Bayesian Nash Equilibrium, where, given every buyers' and sellers' beliefs and optimal strategies, no one has an incentive to deviate from them.

Theorem 3: There exists probability functions $\{\phi^*, \gamma_s^*, \gamma_b^*\}$ such that the equilibrium conditions are satisfied. Proof: Work in progress.

The strategy to prove the existence of an equilibrium is to show that the space of functions $\{\phi^*, \gamma_s^*, \gamma_b^*\}$ maps continuously into itself. Once it is shown that the space spanned by these functions is a compact convex subset of a Banach space, existence follows from Schauder's

Fixed Point Theorem.[26]

3.5   Changes in the information technology

Developments in the information technology enhances the content of a home listing and provides buyers with additional information at the time they look at an ad. These additional information should change the agents' optimal strategies and the equilibrium of the market. In this section, we show how our theoretical model could be used to predict the effects of these improvements on the Real Estate Market.

A natural way to evaluate the effects of better information technology consists of assuming that this new technology allows buyers to learn a greater portion of their buyer-home match value at the time they look at an ad. That is, we should explore how the solution to the equilibrium search model changes when the variance of b^u decreases relative to the variance of b^0. These technological changes should affect the distribution of posting and transaction prices, the buyer's visiting rate, and the average time that a seller stays in the market.

Because in our equilibrium model the variables are related in complicated nonlinear ways, we cannot give precise theoretical insights about the size of these effects. For example, as more information becomes available, buyer's are able to make a more careful screening process before they decide to visit a property. Thus, the buyer's value of search increases, and the visiting rate diminishes. It is not clear how improved information may affect the probability of agreement between a buyer and a seller given that the buyer has decided to visit a particular property. In one hand, buyers visit only those properties with a relative high observed match value b^0; on the other hand, the value of their outside option (value

---

[26]The structure of Theorem 3's proof follows the existence proof in Stern (1990).

of search) has increased. Therefore, the final shift of the sellers' demand function $1 - \phi^*$ is uncertain. When buyers' optimal strategies change, sellers' optimal behavior is affected as well. It is straightforward to see that positive shifts in the demand function $1 - \phi^*$ makes sellers set higher posting prices and reservation values. However, we do not know the direction of the shift in demand and, thus, cannot make any theoretical predictions about the changes in optimal sellers' behavior nor about the changes in equilibrium.

To answer our question of interest, we should use, then, numerical methods to solve our equilibrium model and perform comparative statics exercises. However, before we attempt to do these tasks, we need to estimate the model to have a reliable benchmark for our comparisons.

4 Data and reduced form analysis

4.1 Study Area and Data

The area of our study includes Charlottesville City and Albemarle County. These are two adjacent locations that are part of the Charlottesville, VA Metropolitan Statistical Area. The City of Charlottesville is located in Central Virginia, approximately 100 miles southwest of Washington, D.C. and 70 miles northwest of Richmond, Virginia. Albemarle County surrounds Charlottesville City, and its north border lies approximately 80 miles southwest of Washington, D.C. Both areas occupy approximately 733 square miles (Charlottesville 10 and Albemarle 723). As one of the fastest growing areas in the state, the population increased by 16.7% between 1990 and 2003. According to the US Census, the combined population in these locations was 126,832 in 2003. In 2002, the total number of housing units in Charlottesville City and Albemarle County was 52,716, and, of those units, 58%

were owner-occupied.

The Charlottesville and Albemarle Association of Realtors (CAAR) has provided us with Multiple Listing Services (MLS) data for all completed real estate transactions in Charlottesville City and Albemarle County during the years 2000 through 2003. The property data consist of 3,910 individual transaction records with information on posting prices, transaction prices, number of days in the market, and detailed property characteristics that include the home address.

To avoid biases in our analysis produced by outliers, we exclude from our database 160 observations corresponding to properties that were sold for less than $45,000 or more than $450,000. In addition, to be consistent with our theoretical model, we also exclude 58 transactions (1.5%) where the transaction price was above the posting price. Then, using the individual addresses, we were able to match 2,876 observations with the US Census Block Codes and construct a matched dataset with both housing and neighborhood characteristics.[27] We include five variables from the US Census that we believe are important to explain neighborhood desirability; these are: population density, proportion of blacks, median age, household size, and household income. The first four variables were tabulated for each Census Block while the variable "median household income" was obtained for each Census Block Group only. Descriptive statistics for this matched dataset are presented in Table 1.

Based on our 2,876 records, the average transaction price was $196,400, with a minimum of $50,000 and a maximum of $449,300. The posting price was, on average, $4,200 higher than the transaction price. Despite this fact, the distribution of both posting and transaction

---

[27]To match our database with the US Census, we assigned a Census Block Code (CBC) to each of our records. However, in 816 cases, we were unable to link the reported addresses with the CBC. We dropped these unmatched observations from the sample.

Table 1: Descriptive Statistics. Completed Real Estate Transactions Albemarle and Charlottesville, VA 2000 - 2002

Variable	Mean	Median	Std.Dev.	Min.	Max.	Obs.
Posting price ($ thousands)	200.6	181.3	85.5	52.0	495.0	2876
Transaction price ($ thousands)	196.4	179.1	83.1	50.0	449.3	2876
Days in the market	43.3	26.0	46.1	1.0	199.0	2876
Home characteristics						
Finished square footage	1980.2	1910.1	701.9	261.0	6500.0	2876
Number of full bathrooms	2.03	2.00	0.67	0	5.00	2876
Number of bedrooms	3.35	3.00	0.80	0	9.00	2876
Total acres	0.99	0.26	2.34	0	36.20	2876
Age of the property	23.75	17.00	23.45	0	251.0	2876
Air conditioning	0.90	1.00	0.31	0	1	2876
Basement	0.55	1.00	0.50	0	1	2876
Central heat	0.97	1.00	0.18	0	1	2876
Sewer system	0.75	1.00	0.43	0	1	2876
Home Owner Association	0.51	1.00	0.50	0	1	2876
Pool	0.03	0.00	0.16	0	1	2876
One story only	0.31	0.00	0.46	0	1	2876
Detached	0.79	1.00	0.41	0	1	2876
Neighborhood characteristics						
Log of population per square mile in census block	7.02	7.50	1.54	1.42	10.17	2876
Proportion of blacks living in the census block	0.10	0.05	0.14	0.00	1.00	2876
Median age in the census block	38.51	37.70	7.48	18.50	80.10	2876
Mean household size in the census block	2.52	2.57	0.43	1.00	4.50	2876
Median household income in census block group	55.5	53.9	16.8	15.6	116.7	2876

Figure 1: Nonparametric density of posting and transaction prices.

The density has been evaluated at 100 equally-spaced points in the range of the data using a normal kernel function (with a bandwith of 16,5).
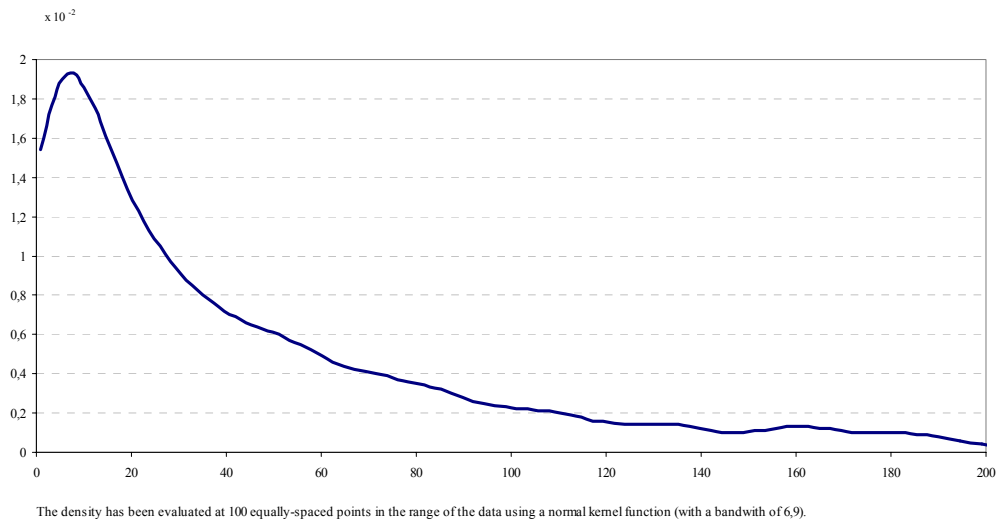
prices are quite similar. As shown in Figure 1, both distributions are unimodal and skewed to the right.

In this area, most homes sell relatively fast. While the mean time that a home stays on the market is 43 days, twenty two percent of the properties sold in less than one week and fifty percent sold in less than 26 days. On the other hand, a small number of homes (8.8%) stay for more than four months on the market. The density of the time that a home stays on the market is skewed to the right and unimodal (see Figure 2).

A typical home is about 24 years old, has 1,980 square feet, two bathrooms, and is located in a US Census block where 10% of its population is black. About 90% of these homes are equipped with air conditioning, while only 3% have a swimming pool.

Before analyzing the descriptive statistics of the neighborhood characteristics, notice that these statistics are weighted by the number of homes sold in each Census block and do not necessarily represent an accurate description of the whole population of Charlottesville

Figure 2: Nonparametric density of the number of days that a property stays on the market.

The density has been evaluated at 100 equally-spaced points in the range of the data using a normal kernel function (with a bandwith of 6,9).

City and Albemarle County. Instead, they describe only those locations where real estate transactions were made. For example, the median household income in a Census block group where a home was sold was, on average, $55,500. On the other hand, the median household income of Charlottesville City and Albemarle County was $31,007 and $50,749, respectively. With these considerations, a representative home in our sample is located in a US Census block where the median age of its inhabitants is 38.5 years and the mean household size is 2.5.

Finally, there is significant dispersion in the characteristics of the neighborhoods. For example, while there are many areas with no blacks living in them, there are several US Census blocks populated by blacks only.

4.2 The determinants of posting prices, transaction prices, and time in the market

Before estimating the structural model, it is important to understand what are the relevant factors that determine posting prices, transaction prices, and the time until a property is sold. In this section, we use simple hedonic Ordinary Leasts Squares (OLS) models to identify these factors, and later we utilize these results as a benchmark to evaluate the performance of the full structural model.

We let the posting price, the transaction price, and the log of time on the market be the dependent variables of three independent linear regression models: (A), (B) and (C), respectively. The explanatory variables include the property and neighborhood's characteristics from our matched database. Notice that we have not included time variant independent variables; that is, we have not specified year nor month dummy variables. We omit these variables to be consistent with the stationary nature of our theoretical model. Later, we shall test the validity of this assumption.

The results of the OLS regressions of models (A) and (B) are presented in Table 2. All the coefficients from (A) and (B) have the same sign, and the R^2 of both regressions is roughly 0.70, which is a typical level for housing price models (Mason and Quigley 1996).

Notice that we have specified linear regerssion models to explain posting and transaction prices. Instead, we could have used a log-linear or other type of non-parametric specification (see, for example, Bin 2004). To test the robustness of the linear specifications (A) and (B), we have estimated log-linear pricing models and found that the predictive power of the latter models was slightly higher (the Mean Square Error in A and B decreased in 2% and 2.1%, respectively). Nevertheless, we have chosen to use the linear models, since the

interpretation of the coefficients in a linear pricing equation is similar to the interpretation of some parameters in our structural model. In particular, the coefficients γ (in $s = X\gamma$) in our equilibrium model represent the marginal contribution of a home characteristic to the value of the home. The same interpretation have the coefficients of a linear hedonic pricing equation. These similarities facilitate the comparison between the structural and the reduced form models.

The coefficients' estimates of models (A) and (B) suggest that one additional square foot increases the posting and transaction price by $68 and $65 respectively. A surprising finding is that, after conditioning on square footage, one fewer bedroom adds over $4, 300 to the transaction price of a property. This fact suggests that agents prefer homes with larger bedrooms. Because we expect that the age of the property would affect its value in a nonlinear way, we include a quadratic and a cubic term in our specifications.[28] The estimates suggest that the transaction price of a new property declines rapidly at a decreasing rate for the first 40 years and slowly appreciates after that. For example, model (B) predicts that a ten year old property sells for $37,100 less than a new property, a twenty year old property sells for $9,000 less than the a ten year old home, and a 190 year old home has the same value as a new property. Buyers are willing to pay more for homes that are located in high-populated or high-income areas. As the mean age of the neighborhood increases, so does the price of the homes; the opposite is true for household size. It is difficult, however, to give a meaningful interpretation about the coefficients of these two variables because, most likely, they are capturing unobserved variation in income at the block level. Finally, there is

---

[28]Although the coefficient on the linear term is not statistically significant, we reject at the 1% significance level the joint null hypothesis that these three coefficients are equal to zero.

Table 2: Hedonic Price OLS Regression Models

	Dependent variable			
	Model (A) Posting price		Model (B) Transaction price	
Constant	-15.12	(15.74)	-17.01	(15.03)
Finished square footage	0.068*	(0.00)	0.065*	(0.00)
Number of full bathrooms	5.20*	(2.40)	4.36	(2.32)
Number of bedrooms	-5.74*	(2.08)	-5.16*	(2.05)
Log of acreage (zero if none)	7.25*	(1.01)	6.71*	(0.97)
Log of age of the property (zero if new)	-1.61	(6.33)	-5.49	(6.14)
Log of age of the property ^{2}	-10.89*	(3.70)	-9.25*	(3.57)
Log of age of the property ^{3}	2.19*	(0.58)	1.96*	(0.56)
Air conditioning	11.96*	(4.10)	12.93*	(4.00)
Basement	5.40*	(2.21)	5.38*	(2.16)
Central heat	12.63*	(6.38)	13.50*	(5.88)
Sewer system	-3.99	(3.72)	-4.13	(3.63)
Home Owner Association	14.54*	(2.62)	15.02*	(2.53)
Pool	15.64*	(7.44)	15.05*	(7.30)
One story only	-9.63*	(2.41)	-9.64*	(2.31)
Detached	42.18*	(2.75)	41.88*	(2.66)
Log of population per square mile in census block	3.06*	(1.02)	3.45*	(0.99)
Proportion of blacks living in the census block	-57.89*	(7.47)	-57.76*	(7.27)
Median age in the census block	1.00*	(0.19)	0.97*	(0.18)
Mean household size in the census block	-5.99*	(2.90)	-4.92	(2.78)
Median household income in census block group	0.48*	(0.08)	0.47*	(0.08)
R^2	0.705		0.707	
Number of observations	2876		2876	

Standard errors are in parentheses. Astericks indicate those parameters siginif-icant at the 5 percent significance level. The covariance matrix was calculated using the White-Heteroscedasticity-Consistent Method.

a significant "premium" that individuals pay for living in a non-black neighborhood which averages more than $57,000.

To explain the determinants of the time that a house stays on the market (Time on the Market, TOM), we present in Table 3 the OLS estimates of a log-linear duration model (Model C). Model (C)'s R^2 is significantly lower than (A)'s or (B)'s, but it is consistent with other findings in the literature (see for example, Horowitz 1992). We also find that the size and age of the house, the population density of the block, and the share of blacks living in the neighborhood, are statistically significant predictors of the time that a house stays on the market. Bigger properties stay longer on the market, that is, a 10% increase in square footage implies a 6.1% increase in the expected TOM. New and old properties sell quicker than middle age ones. Finally, our results suggest that properties located in black neighborhoods stay 58% longer on the market than properties in non-black vicinities. The rest of the coefficients are not statistically significant.

A common approach in the literature when estimating TOM hedonic equations -as in Jansen and Jabson (1980), Kang and Gardner (1989) and Yavas and Yang (1995)- consists of including the percent difference between the actual posting price and the predicted posting price from an OLS model, as an explanatory variable in (C). These models conjecture that, after controlling for the property's characteristics, higher than average posting prices should lead to longer TOM. However, if there are any unobserved (for the econometrician) housing characteristics that affect both the posting price and the TOM, there may be important endogeneity biases with this approach. In fact, we expect that unobserved home features will be negatively correlated with TOM and positively correlated with the posting price, causing this coefficient to be biased downwards. Thus, we choose not to include this variable

Table 3: Linear Duration Regression Models

	Dependent variable Model (C) Log days in the market	
Constant	-0.33	(0.90)
Log of finished square footage	0.61*	(0.13)
Number of full bathrooms	0.02	(0.05)
Number of bedrooms	-0.004	(0.05)
Log of acreage (zero if none)	-0.001	(0.03)
Log of age of the property (zero if new)	-1.38*	(0.14)
Log of age of the property ^{2}	0.51*	(0.08)
Log of age of the property ^{3}	-0.06*	(0.01)
Air conditioning	-0.17	(0.10)
Basement	-0.20	(0.06)
Central heat	-0.02	(0.16)
Sewer system	-0.15	(0.10)
Home Owner Association	-0.07	(0.07)
Pool	0.08	(0.15)
One story only	-0.003	(0.06)
Detached	0.05	(0.08)
Log of population per square mile in census block	-0.059*	(0.03)
Proportion of blacks living in the census block	0.58*	(0.21)
Median age in the census block	0.006	(0.00)
Mean household size in the census block	0.084	(0.07)
Median household income in census block group	0.000	(0.00)
R^2	0.133	
Number of observations	2876	

Standard errors are in parentheses. Astericks indicate those parameters siginificant at the 5 percent significance level. The covariance matrix was calculated using the White-Heteroscedasticity-Consistent Method.

in our specification. The structural model we propose solves this problem by explicitly modeling and controlling for unobserved housing heterogeneity.

5 Estimation

5.1 The Likelihood Function

The set of parameters of our structural model Θ can be estimated using Maximum Likelihood.[29] As we described in the previous Chapter, for each transaction i, we observe the posted price p_{si}, the transaction price p_{mi}, the number of days the property stays in the market t_i, and a set of home characteristics X_i^o. In this section, we use the information in our dataset to specify the relevant likelihood function.

Before we estimate the equilibrium model specified in the previous section, we need to solve it. That is, given all the exogenous variables and parameters of the model, we shall find the function ϕ^* (or equivalently q^*, γ_s^* and γ_b^*) such that both buyers and sellers have no incentives to deviate from their optimal endogenous strategies. In the appendix, we provide details about the numerical methods used to solve the model.

It is important to recognize that certain features of the property, which are displayed in pictures and/or detailed comments in a Multiple Listing Service (MLS) ad, can be observed only by the agents in the market and not by the econometrician. Thus, when estimating the model, we need to control for unobserved housing characteristics. We model this unobserved heterogeneity by letting $s_i = X_i^o \gamma + u_i$ and assuming that u_i is an i.i.d. mean zero error with density f_u. In addition, we assume that u is independent of the seller's discount factor β_s.

---

[29]In what follows, we condition on a set of parameter values Θ. But, for expositional purposes, we omit it from our notation.

First, let d_i equal one if the posting price equals the transaction price, and zero otherwise, and let us derive the likelihood contribution of an observation where the transaction price was below the posting price. Notice that, given a posting and a transaction price, we can use the structure of the model (equations 4 and 3 and the optimal ϕ^*) to calculate u_i and β_{si}. That is, with information on $\{p_{si}, p_{mi}, X_i^o\}$ we compute u_i and β_{si} as the unique values that simultaneously solve (12) and (13).

$$p_{mi} = p_{si} - \frac{1 - \phi^*(p_{si}|X_i^o\gamma + u_i)}{\phi^{*\prime}(p_{si}|X_i^o\gamma + u_i)} \tag{12}$$

$$\beta_{si} = \frac{p_{mi} - X_i^o\gamma - u_i}{p_{mi} - X_i^o\gamma - u_i + (p_{si} - p_{mi})\theta[1 - \phi^*(p_s|X_i^o\gamma + u_i)]} \tag{13}$$

For any given u_i, we can determine the corresponding $s_i = X_i^o\gamma + u_i$; and, using q_s^*, γ_s^*, and γ_b^*, the well defined functions determined by the solution of the equilibrium model, we find the probability that trade occurs at a transaction price below the posting price

$$l_m(p_{si}, p_{mi}, X_i^o, u_i) = (1 - \theta)q_s^*(p_{si}|s_i)\gamma_b^*(p_{mi}|s_i).$$

In a similar way, notice that

$$l_p(p_{si}, X_i^o, u_i) = \theta q_s^*(p_{si}|s_i)\gamma_s^*(p_{si}|s_i)$$

is the probability that in a period a buyer visits a seller and purchases the property for the posting price. Thus, the unconditional probability that a property does not sell in one period is defined by

$$l_o(p_{si}, p_{mi}, X_i^o, u_i) = 1 - l_p(p_{si}, X_i^o, u_i) - l_m(p_{si}, p_{mi}, X_i^o, u_i),$$

and the probability of observing a property staying t_i periods in the market is

$$l_{om}(t_i|p_{si}, p_{mi}, X_i^o, u_i) = l_o(p_{si}, p_{mi}, X_i^o, u_i)^{t_i - 1} l_m(p_{si}, p_{mi}, X_i^o, u_i).$$

32

Using the previous definitions, we construct the likelihood contribution of an observation $\{t_i, p_{si}, p_{mi}, X_i^o, d_i = 0\}$ as

$$l_{tpm}(t_i, p_{si}, p_{mi}, X_i^o) = l_{om}(t_i | p_{si}, p_{mi}, X_i^o, u_i) k_s(\beta_{si}) f_u(u_i) |J_i|,$$

where k_s is the density of the seller's βs, and $|J_i|$ is the absolute value of the determinant of the Jacobian of the transformation implied by equations (12) and (13).

Now, let us focus on evaluating the likelihood contribution of an observation where the home sold for the posting price. Notice that, because we do not observe the seller's reservation value, we cannot recover the value of the unobserved home characteristics u_i as we previously did. However, if we condition on a particular value of u, we can use equation (12) to estimate the seller's reservation value $p_m(u) = p_m(p_{si}, X_i^o, u)$. With this information, and using equation (13), we manage to recover the value of $\beta_s(u) = \beta_s(p_{si}, p_m(u), X_i^o, u)$, and compute the probability of observing a property staying t_i periods in the market

$$l_{op}(t_i | p_{si}, p_{mi}, X_i^o, u) = l_o(p_{si}, p_m(u), X_i^o, u)^{t_i-1} l_p(p_{si}, X_i^o, u).$$

Thus, the likelihood contribution of observation $\{t_i, p_{si}, p_{mi}, X_i^o, d_i = 1\}$ is

$$l_{tp}(t_i, p_{si}, p_{mi}, X_i^o) = \int l_{op}(u) k_s(\beta_s(u)) |J_i(u)| f_u(u | p_m(u) > X_i^o \gamma + u) du.$$

Note that we integrate the likelihood contributions for all values of u that are consistent with the model. That is, only those values of u that satisfy the condition that $p_m(u) > s$ are considered.

The log-likelihood contribution L_i of observing $\{t_i, p_{si}, p_{mi}, X_i, d_i\}$ is

$$L_i(t_i, p_{si}, p_{mi}, X_i^o, d_i; \Theta) = (1 - d_i) \log(l_{tpm}) + d_i \log(l_{tp}),$$

and the MLE parameter estimates are the ones that maximize the log-likelihood of observing the sample

$$\Theta^{MLE} = \arg\max_{\Theta} \left\{ \sum_{i=1}^{n} L_i(p_{si}, p_{mi}, t_i, X_i^o, d_i; \Theta) \right\}. \tag{14}$$

5.2 Functional form assumptions and identification

Before we solve the model, we need to specify functional forms for all the relevant exogenous distributions. We let both β_b and β_s be a transformation of a normally distributed random variable. That is,

$$\log\left(\frac{\beta_j}{1 - \beta_j}\right) \sim N(u_{\beta_j}, \sigma_{\beta_j}^2); \qquad j = b, s.$$

In addition, we let b^o, b^u, and u be mean zero normal random variables with standard deviations σ_{b^o}, σ_{b^u} and σ_u respectively. Given these functional form assumptions and a set of parameters values, we use the numerical methods described in the appendix to solve the model.

To make estimation feasible, we make the additional identifying assumptions that the distributions of β_b and β_s are the same, that is $K_s = K_b = K$. We need to make this assumption because, in our sample, we do not observe buyer's behavior, and thus, cannot infer anything about the buyer's motivation to trade. This assumption is reasonable to the extent that both buyers and sellers in the Real Estate market are equally influenced, on average, by the same kind of factors that could affect their motivation to trade (for example, changes in life style, income, and employment). Thus, when making comparative statics analysis, we need to be aware of this normalization to provide a correct interpretation of our results.

With the previous assumption, we use MLE to estimate the model's parameter vector

$$\Theta = \left\{ \mu_\beta, \sigma_\beta, \sigma_{b^u}, \sigma_{b^o}, \sigma_u, \gamma, \delta, c_b, \theta \right\}.$$

It is important to provide some intuition about how we identify Θ in our likelihood function. Variation in prices and home characteristics identify the vector γ, in a similar manner as OLS does in a linear regression. Differences between posting prices and transaction prices and the structure of the model identify μ_β, σ_β, and σ_u. The joint distribution of posting prices and time in the market and the observed agent's behavior identify σ_{b^u}, σ_{b^o} and c_b. Finally, the share of transactions where the transaction price was the posting price identifies θ.

6 Results

6.1 Parameter estimates

In Tables 4 and 5, we present the Maximum Likelihood estimates of the structural model parameters, and, in the rest of this section, we discuss the implications of these results. The estimate of θ implies that there is a 0.47 probability that a home sells for the posting price, and, because θ is close to one half, buyers and sellers have a similar level of market power.

Since we are measuring periods in calendar days, we expect the mean discount factor μ_β to be close to one. As we anticipated, the estimated daily mean discount factor is 0.9997, which implies a monthly mean discount factor of 0.989 and a yearly one of 0.88. These results suggest that, on average, agents value an opportunity to sell their home one year from today 12% less than having the same opportunity now.

The estimate of σ_β is small. However, small variation in the agents' discount-factors

35

Table 4: Parameter Estimates of the Structural Model: Baseline parameters

Variable	Baseline model's parameters	
θ	0.47*	(0.009)
μ_β	0.99967*	(0.00004)
σ_β	0.0053*	(0.000001)
c_b	0.506*	(0.0226)
δ	1.08*	(0.0035)
σ_u	45.694*	(0.750)
$\sigma_{\tilde{b}}$	80.50*	(0.029)
λ	0.031*	(0.00063)
Number of observations	2876	

Standard errors are in parentheses. Aster-
icks indicate those parameters significant at
the 5 percent significance level.

implies important differences in both buyer's and seller's optimal behavior. To illustrate

this statement, we present in Figure 3 the seller's optimal posting prices and reservation

values for different values of β_s (given an estimated mean home quality $\bar{s}$ of \$160,929). The

smallest (largest) discount factor we picked is 0.99963 (0.9999), which corresponds to the 0.2

th (99.8th) percentile of the distribution K_s. Within this interval, there exists significant

variation in both posting prices and reservation values, implying that the transaction price

of such a property could be anywhere between \$170,000 and \$220,000. As in Glower et al

(1998), we find that more motivated agents post cheaper prices and sell their homes at lower

prices.[30]

According to the Charlottesville and Albemarle Real Estate Association's Director, on

average, potential buyers visit homes-for-sale in pairs (couples) for approximately 45 minutes

per visit. Thus, assuming that the median visitor has an hourly income of \$19.2 per hour,[31]

---

[30]However, our results are not fully comparable with those found by Glower et al (1998). While these
authors assess the effect of five (observed) variables -related to the seller's motivation to trade- in posting and
transaction prices, we use a structural model to evaluate how equilibrium prices change as the (unobserved)
discount rate varies.

[31]To construct this estimate, we assume that the median household supplies an average of 60 hours per

Figure 3: Reservation Value and Posting Price as a Funtion of the Discount Factor (evaluated at the mean home quality $\bar{s}$)
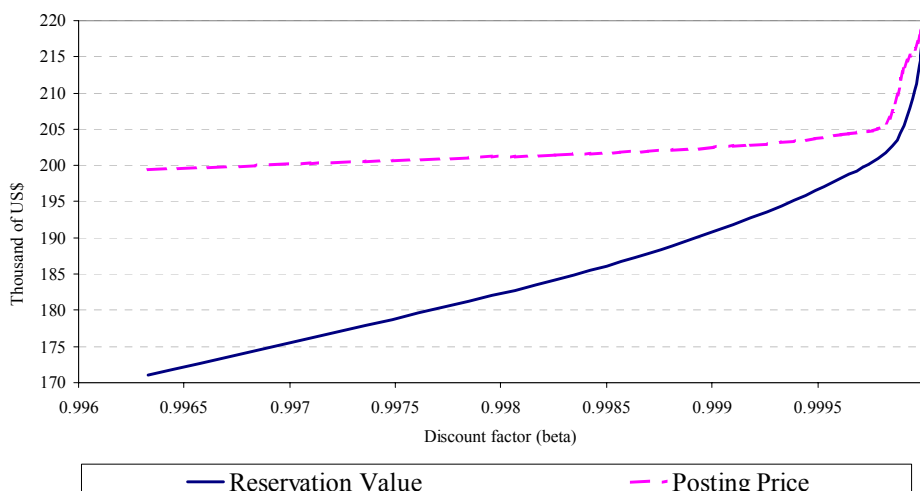
the related couple's touring time-opportunity costs add up to approximately $29. Our buyer's visiting cost c_b estimate is far greater ($506) meaning that other costs, such as transportation and non pecuniary (emotional) costs, account for the vast majority (more than 90%) of it.

The estimate of δ implies that buyers value homes 8% more than sellers.

σ_u represents the dispersion of all unobserved factors that affect the value of a home. The errors' standard deviation of the OLS hedonic models in Table 2 have a quite similar interpretation and, thus, we expect σ_u to be close to them. In fact, the estimate of σ_u is $45.7, and the differences with the OLS errors' standard deviation are less than 2%.

The estimate of $\sigma_{\tilde{b}}$, the standard deviation of the idiosyncratic buyer-home match value $\tilde{b}$, is relatively large ($80.5). This means that there is significant heterogeneity in buyer's tastes in the housing market.

Recall from the theoretical model that $\tilde{b}$ has two independent components: a) b^o, the

week. Then, we divide the area's yearly median household income ($55.500) by the total number of hours per year worked by this median household.

match value learned when the buyer observes an online ad, and b) b^u, the match value learned when the buyer visits and tours the home. To evaluate how much relevant information do online ads provide to real estate customers, we estimate $\lambda = \frac{\sigma_{b^o}^2}{\sigma_{\tilde{b}}^2}$, the ratio between the variance of b^o and the variance of $\tilde{b}$. The estimate of λ is 0.03 suggesting that, on average, only 3% of the relevant information that buyers collect before making a purchase decision is obtained through online ads. To a certain extent, this is a surprising result. However, notice that many of the current online advertising features, such as virtual tours, for example, were not available during the first years of our sample.

The estimates of γ (Table 5) have the same sign and are close in magnitude to the coefficients of the OLS price hedonic regressions. Our results suggest that one additional square foot increases the value of a home by $69, while an extra bathroom raises it by $2,176. As was the case with the hedonic regressions, an additional bedroom reduces the value of a property, suggesting that sellers dislike small bedrooms. Homes that have a yard are more valuable, and, on average, a property that has one additional acre than the median is worth an extra $13,400. The coefficients on age suggest that the value of a new property declines rapidly at a decreasing rate for the first 31 years and slowly appreciates after that. In particular, the model predicts that a new property is $48,000, $58,000, and $0 more valuable than a ten, twenty and 220-year-old property, respectively. Having air conditioning and central heat adds approximately $8,600 and $19,000 to the value of a home, respectively. Finally, all coefficients on the neighborhood variables have the expected signs.

Table 5: Parameter Estimates of the Structural Model: Seller's home value

Variable	Seller's home value	
	$s = X\gamma$	
Constant	-53.802*	(12.4)
Finished square footage	0.069*	(0.0013)
Number of full bathrooms	2.176	(1.63)
Number of bedrooms	-9.119*	(1.18)
Log of acreage (zero if none)	8.492*	(1.06)
Log of age of the property (zero if new)	-11.30*	(4.34)
Log of age of the property ^{2}	-8.735*	(2.16)
Log of age of the property ^{3}	2.006*	(0.30)
Air conditioning	8.529*	(3.05)
Basement	1.311	(1.99)
Central heat	18.969*	(4.40)
Sewer system	-5.154	(3.49)
Home Owner Association	11.348*	(2.70)
Pool	12.555*	(4.95)
One story only	-10.371*	(2.17)
Detached	36.876*	(3.26)
Log of population per square mile in census block	3.318*	(0.82)
Proportion of blacks living in the census block	-60.442*	(7.46)
Median age in the census block	0.906*	(0.12)
Mean household size in the census block	-13.707*	(2.68)
Median household income in census block group	0.436*	(0.06)
Number of observations	2876	

Standard errors are in parentheses. Astericks indicate those parameters siginificant at the 5 percent significance level.

Figure 4: Nonparametric density of actual and predicted posting prices.
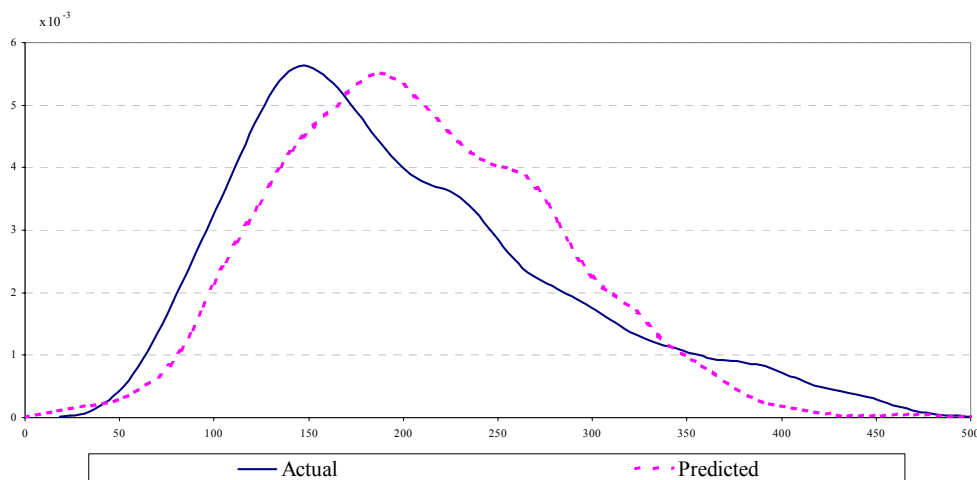
Both densities has been evaluated at 100 equally-spaced points in the range of the data using a normal kernel function.
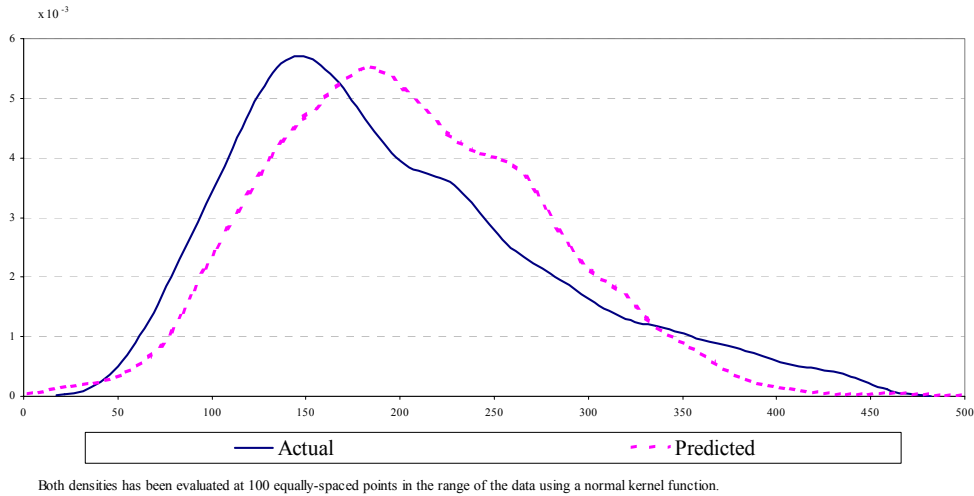
6.2 Within sample fit

Overall, the model seems to fit the pricing data reasonably well. Within our sample, we have used the structural model to predict posting and transaction prices and constructed pseudo-R^2 indexes.[32] For each of these variables, the pseudo-R^2 were about 0.69, which is slightly below the R^2s derived from the OLS models. Despite the variety of assumptions and the complexity of our structural model, posting and transaction prices are predicted almost with the same accuracy of a linear regression model.

In Figures 4 and 5 we present the densities of the predicted posting and transaction prices, respectively. In both cases, the model displays a reasonably good representation of our sample data. It seems, however, that our model is unable to capture the skewness of the actual data.

---

[32]To construct the within sample fit, we use the solution to our estimated equilibrium model and use simulation methods to predict posting prices, transaction prices, and time on the market.

Figure 5: Nonparametric density of actual and predicted transaction prices.

Both densities has been evaluated at 100 equally-spaced points in the range of the data using a normal kernel function.

As in Horowitz (1992), our model fails to predict the time that a property stays on the market (TOM). It can be seen from Figure 6 that there is little variation in the predicted TOM. All the predictions are from 42 to 80 days, while the observed times are from 1 to 200 days. Thus, our model explains very little of the variation of TOM.

6.3 Changes in online information

One of the primary benefits of estimating a structural model is that the parameter estimates are said to be invariant to policy changes and, thus, the estimated model may be used to conduct counterfactual policy experiments. In this section, we use the estimated model to evaluate how equilibrium market outcomes change when the amount of online information increases.

Our experiment raises the amount of the relevant information that buyers obtain when looking at an online ad. That is, we let λ increase from 0.03 to 0.50 and to 1.00, and evaluate the effect of these changes on posting prices, transaction prices, and time on the market.

Figure 6: Nonparametric density of actual and predicted time on the market.
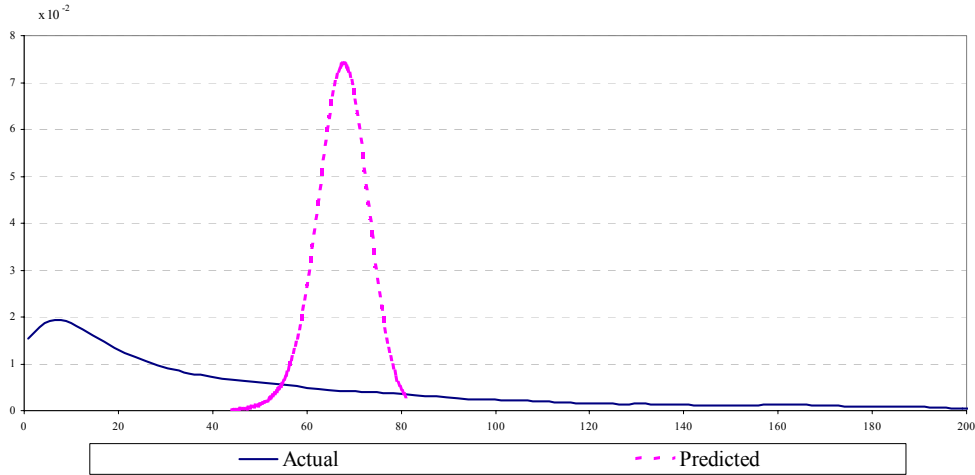
Both densities has been evaluated at 100 equally-spaced points in the range of the data using a normal kernel function.

Table 6: Effects of an increase in the amount of online information

Variable	Scenarios		
	$\lambda = 0.03$	$\lambda = 0.50$	$\lambda = 1.00$
Mean posting price ($ thousands)	207.8	204.5	201.8
Mean transaction price ($ thousands)	204.0	199.1	197.6
Mean time on the market (days)	67.2	69.2	71.3

The results of this exercise are displayed in Table 6.

Our results suggest that additional online information decreases market prices. For instance, if buyers could obtain from an online ad all the relevant information they need to make a purchase decision ($\lambda = 1$), posting and transaction prices would decrease by 2.9% and 3.1%, respectively. These results provide evidence that the internet increases market competition and are consistent with other findings in the literature.[33]

We also find evidence that, as more information is displayed in real estate online ads, the time that a property stays on the market raises. In particular, when λ equals one, the average TOM increases in about four days compared to our benchmark case ($\lambda = 0.03$). As

[33]See, for example, Brown and Goolsbee (2002).

online information increases, buyers have more options to choose from and take more time to make a final purchase decision. This result is consistent with that found by D'urso (2002).

7 Conclusions

We have specified an equilibrium two-sided search model that depicts many of the Real Estate market's real-life features. The theoretical model incorporates, in an equilibrium context, five very important characteristics of the REM: a) buyers' and sellers' search behavior, b) heterogeneity in agents' motivation to trade, c) transaction costs, d) a trading mechanism with posting prices and bargaining, and e) the availability of an online advertising technology. We estimate the theoretical model using Maximum Likelihood methods and Multiple Listing Services data.

We use the estimated model to predict how improvements in the information technology affect the average posting price, transaction price, and time until a property is sold. Our results suggests that improvements in online information displayed by real estate ads, decrease market prices but increase the time that a property stays on the market. Furthermore, we find that, on average, only 3% of the relevant information that home-buyers collect before making a purchase decision is obtained through online ads.

A Theorem's proofs

Proof of Theorem 1: Substituting (4) in (3) and rearranging, we obtain the following expression:

$$h(p_s^*|s)(p_s^* - s) = \frac{1 - \beta_s[1 - \theta(1 - \phi(p_s^*|s))]}{1 - \beta_s}. \tag{15}$$

Furthermore, assume that the hazard function is non-decreasing in p_s^*. Note that the left hand side of equation (15) evaluated at $p_s^* = s$ is zero, and that it is increasing in p_s^*; the right hand side of equation (15) evaluated at $p_s^* = s$ is a positive value greater or equal than one, and it is decreasing in p_s^*. Thus, there exists a unique solution for the posting price. Replacing p_s^* in (4) we solve for the optimal reservation value R_s^*, and, since the hazard function is nonnegative, it must be the case that $p_s^* \geq R_s^*$ (for any value of β_s and s). Using equation (15) and the implicit function theorem, we take the derivative of p_s^* w.r.t. β_s and obtain

$$\frac{\partial p_s^*}{\partial \beta_s} = \frac{\frac{\theta(1-\phi)}{(1-\beta_s)^2}}{(p_s^* - s)\frac{\partial h}{\partial p_s^*} + h + \frac{\beta_s}{1-\beta_s}\theta\frac{\partial \phi}{\partial p_s^*}} > 0.$$

It is also easy to see from equation (4) that $\frac{\partial R_s^*}{\partial p_s^*} = 1 + \frac{\frac{\partial h}{\partial p_s^*}}{h^2} > 0$. Thus, $\frac{\partial R_s^*}{\partial \beta_s} = \frac{\partial R_s^*}{\partial p_s^*}\frac{\partial p_s^*}{\partial \beta_s} > 0$.

We let $\beta_s = 0$ in equation (3) and find that $R_s^* = s$. Because $0 \leq \beta_s \leq 1$ and $\frac{\partial R_s^*}{\partial \beta_s} \geq 0$, it must be the case that $p_s^* \geq R_s^* \geq s$.

To prove Theorem 1, we still need to show that p_s^* and R_s^* are increasing in s. To show this, we use equation (15) and the implicit function theorem to find

$$\frac{\partial p_s^*}{\partial s} = \frac{h - \left(\frac{\partial h}{\partial s}(p - s) + \frac{\partial \phi}{\partial s}\right)}{(p_s^* - s)\frac{\partial h}{\partial p_s^*} + h + \frac{\beta_s}{1-\beta_s}\theta\frac{\partial \phi}{\partial p_s^*}} > 0.$$

Notice that the expression in parenthesis in the numerator is always negative, because both $\frac{\partial \phi}{\partial s}$ and $\frac{\partial h}{\partial s}$ are negative. Furthermore, $\frac{\partial R_s^*}{\partial s} = \frac{\partial R_s^*}{\partial p_s^*}\frac{\partial p_s^*}{\partial s} > 0$ QED.

Proof of Theorem 2

To prove Theorem 2, we need to show the details of integration of equation (6), and the monotonicity of W_b^* and p_b^{r*}.

1) **Details of the integration of equation (6)**

Let $v_p = b^o + \delta s - p_s^*$ and integrate by parts the first term of **(6)**

$$
\begin{aligned}
E[\max\{\tilde{b} - p_s^*, W_b\}|p_s^*, s, b^o] &= \int \max\{b^u + b^o + \delta s - p_s^*, W_b\}dG_u(b^u)\\
&= \int_{W_b - v_p} (b^u + v_p)dG_u(b^u) + W_b G_u(W_b - v_p)\\
&= -[(b^u + v_p)(1 - G_u(b^u))]_{W_b - v_p}^{\infty}\\
&\quad + \int_{W_b - v_p} [1 - G_u(b^u)]db^u + W_b G_u(W_b - v_p)\\
&= W_b[1 - G_u(W_b - v_p)]\\
&\quad + \int_{W_b - v_p} [1 - G_u(b^u)]db^u + W_b G_u(W_b - v_p)\\
&= W_b + \int_{W_b + p_s^* - \delta s - b^o} [1 - G_u(b^u)]db^u.
\end{aligned}
$$

Using the same procedure, the second term of $U^e(p_s^*, s, b^o, W_b)$ becomes

$$
E[\max\{b + \delta s - R_s^*(p_s^*, s), W_b\}|p_s^*] = W_b + \int_{W_b + R_s^*(p_s^*, s) - \delta s - b^o} [1 - G_u(b^u)]db^u.
$$

2) **Monotonicity of W_b^* and p_b^{r*}Wb**

Define $g(W_b^*) = \int \int \int \max\{j(p_s^*, s, b^o, W_b^*), 0\} d\Gamma(p_s^*, s)dG_o(b^o)$ and rewrite equation (8) as $W_b^* = \frac{\beta_b}{1 - \beta_b} g(W_b^*)$. Clearly, $g(W_b^*)$ is a non-negative decreasing function. Then, we use the implicit function theorem to show that

$$
\frac{\partial W_b^*}{\partial \beta_b} = \frac{\frac{g(W_b^*)}{1 - \beta_b}}{\frac{\beta_b}{1 - \beta_b}[-g'(W_b^*)] + 1} > 0.
$$

Furthermore, we use equation (**??**) -and Leibnitz rule- to show that

$$\frac{\partial p_b^{r*}}{\partial W_b^*} = -\frac{\theta(1 - G_u(V_p)) + (1 - \theta)(1 - G_u(V_r))}{\theta(1 - G_u(V_p)) + (1 - \theta)\frac{\partial R_s^*}{\partial p_b^{r*}}(1 - G_u(V_r))} < 0,$$

where $V_p = W_b + p_b^{r*} - \delta s - b^o$, and $V_r = W_b + R_s^*(p_b^{r*}|s) - \delta s - b^o$. Thus, $\frac{\partial p_b^{r*}}{\partial \beta_b} = \frac{\partial p_b^{r*}}{\partial W_b^*}\frac{\partial W_b^*}{\partial \beta_b} < 0$.

Steps one and two complete the proof. QED.

Proof of Theorem 2a:

First, note that because s is non random and $b^o = 0$, equation (5) becomes

$$
\begin{aligned}
W_b &= \beta_b \int \max\{U^e(p_s^*), W_b\} d\Gamma(p_s^*) \\
\frac{W_b}{\beta_b} &= \int_{p_s^*:\{U^e(p_s^*)>W_b\}} U^e(p_s^*)d\Gamma(p_s^*) + W_b \int_{p_s^*:\{U^e(p_s^*)<W_b\}} d\Gamma(p_s^*) \\
&= \int_{p_s^*:\{p_s^*<U^{e-1}(W_b)\}} U^e(p_s^*)d\Gamma(p_s^*) + W_b \int_{p_s^*:\{p_s^*>U^{e-1}(W_b)\}} d\Gamma(p_s^*)
\end{aligned}
$$

where $U^{e-1}(.)$ is the inverse image of $U^e(.)$, which is well defined since we have already shown

that U^e is monotone. Let us integrate by parts and rearrange to obtain

$$
\begin{aligned}
\frac{W_b}{\beta_b} &= \int^{U^{e-1}(W_b)} U^e(p_s^*)d\Gamma(p_s^*) + W_b \int_{U^{e-1}(W_b)} d\Gamma(p_s^*) \\
&= [\Gamma(p_s^*)\{U^e(p_s^*)\}]_\infty^{U^{e-1}(W_b)} - \int^{U^{e-1}(W_b)} \Gamma(p_s^*)\frac{\partial U^e(p_s^*)}{\partial p}dp_s + W_b[1 - \Gamma(U^{e-1}(W_b))] \\
&= W_b\Gamma(U^{e-1}(W_b)) + \int^{U^{e-1}(W_b)} \Gamma(p_s^*)(-\frac{\partial U^e(p_s^*)}{\partial p})dp_s + W_b[1 - \Gamma(U^{e-1}(W_b))] \\
&= W_b + \int^{U^{e-1}(W_b)} \Gamma(p_s^*)(-\frac{\partial U^e(p_s^*)}{\partial p})dp_s.
\end{aligned}
$$

It is optimal for buyers to set a reservation strategy such that $U^e(p_b^{r*}) = W_b^*$. Then, we

replace this optimality condition in the former equation and obtain

$$\frac{\beta_b}{1-\beta_b} \int^{p_b^{r*}} \Gamma(p_s^*)\left(-\frac{\partial U^e(p_s^*)}{\partial p}\right) dp_s^* = W_b^*.$$

46

Finally, we let $U^e(p_b^{r*}) = W_b^*$ in equation (7) to find

$$\theta \int_{W_b^* + p_b^{r*} - \delta s} [1 - G_u(b^u)]db^u + (1 - \theta) \int_{W_b^* + R_s^*(p_b^{r*}) - \delta s} [1 - G_u(b^u)]db^u = c_b.$$

Thus, the solution to the search problem is equivalent to finding the pair $\{W_b, p_b^r\}$ that solves

the previous equation system. QED.

Proof that $\phi^*(p_s|s)$ is well defined and decreasing.

The function $q(p_s|s, b^o)$ is well defined by the exogenous distribution K_b and the optimal

buyers reservation price function p_s^{r*}.

$$q(p_s|s, b^o) = \Pr\{p_b^{r*}(\beta_b, s, b^o) > p_s|s, b^o\} = K_b(p_b^{r*-1}(p_s|s, b^o))$$

where $p_b^{r*-1}(p_s)$ is the function that determines β_b for any value p_s (which exists because we

have shown that $p_b^{r*}(\beta_b)$ is monotone). Then

$$\begin{aligned}
\gamma_s(p_s|s, b^o) &= \Pr\{b^o + b^u + \delta s - p_s > W_b^*(\beta_b)|p_b^{r*}(\beta_b, s, b^o) > p_s\} \\
&= \frac{\int^{p_b^{r*-1}(p_s|s,b^o)} (1 - G_u[W_b^*(\beta_b) + p_s - \delta s - b^o)\, dK_b(\beta_b)}{K_b(p_b^{r*-1}(p_s|s, b^o))},
\end{aligned}$$

$$\gamma_b(p_s|s, b^o) = \frac{\int^{p_b^{r*-1}(p_s|s,b^o)} (1 - G_u[W_b^*(\beta_b) + R_s^*(p_s) - \delta s - b^o)\, dK_b(\beta_b)}{K_b(p_b^{r*-1}(p_s|s, b^o))},$$

and

$$\begin{aligned}
\phi^*(p_s|s) &= 1 - \int q(p_s|s, b^o) * \gamma_s(p_s|s, b^o)dG_o(b^o) \\
&= 1 - \int \int^{p_b^{r*-1}(p_s|s,b^o)} (1 - G_u[W_b^*(\beta_b) + p_s - \delta s])\, dK_b(\beta_b)dG_o(b^o)
\end{aligned}$$

which is clearly non-increasing in p_s.

QED.

B Sketch of the numerical methods used to solve the empirical model

Step	Description
1)	Using X and γ, determine Ψ (the distribution of s).
2)	Create a grid of size n_2 for values of s in equally spaced points that depends on the distribution Ψ
3)	Create grids of size n_1 for β_s and β_b of not necessarily equally spaced points.
4)	Specify an initial guess for the function $\phi(p_s|s)$ and store it in a (n_1 x n_2 x 2) array.
5)	For each combination of β_s and s, use equations (3), (4) and the array $\phi(p_s|s)$ to solve for $R_s^*(\beta_s, s)$ and $p_s^*(\beta_s, s)$.
6)	Use equation (8) to solve for $W_b^*(\beta_b)$, and store it in a (n_1 x 1) array.
7)	Use equation (9) and $W_b^*(\beta_b)$ to find $p_b^{r*}(\beta_b, s|b^o)$, and store it in a (n_1 x n_2 x n_3) array.
8)	Given $p_b^{r*}(\beta_b, s|b^o)$ and $W_b^*(\beta_b)$, compute $\phi^*(p_s|s, b^o)$, and store it in a (n_1 x n_2 x 2 x n_3) array.
9)	Using a discrete approximation, evaluate $\phi^*(p_s|s)$, and store it in a (n_1 x n_2 x 2) array.
10)	Stop if $\phi^*(p_s|s) = \phi(p_s|s)$; otherwise go back to 5.

References

[1] Anderson, S., and R. Renault, (2004), "Advertising Content", University of Virginia Working Paper.

[2] Albrecht, J., and B. Axell, (1984), "An Equilibrium Model of Search Unemployment" *Journal of Political Economy*, (92) , 824-40.

[3] Arnold, M., (1999), "Search, Bargaining and Optimal Asking Prices", *Real Estate Economics*, (27), 453-481.

[4] Belkin, J, D. Hempel, and D. Mcleavey, (1976), "An Empirical Study on Time market Using Multidimensional Segmentation of Housing Markets", *Journal of the American Real Estate and Urban Economics Association*, (4), 57-79.

[5] Bagnoli, M., and Bergstrom, T., (1989), "Log-Concave Probability and Its Applications", *University of Michigan Center for Economic and Social Theory*, Working Paper no. 89-32.

[6] Berndt, E., B. Hall, R. Hall, and J. Hausman, (1974), "Estimation and Inference in Nonlinear Structural Models", *Annals of Economic and Social Measurement* 3/4, 653-665.

[7] Bin, O., (2004), "A prediction comparison of housing sales prices by parametric versus semi-parametric regressions", *Journal of Housing Economics*, (13), 68–84.

[8] Brown J., and A. Goolsbee, (2002), "Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry", *Journal of Political Economy*, (110), 481-507.

[9] Bunzel, H., J. Christensen, P. Jensen, N. Kiefer, L. Korsholm, L. Muus, G. Neumann, and M. Rosholm, (2001), "Specification and Estimation of Equilibrium Search Models", *Review of Economic Dynamics*, (4) , 90-126.

[10] Burdett, K. and K. Judd, (1983), "Equilibrium Price Dispersion", *Econometrica*, (5), 955-70.

[11] Burdett, K. and D. Mortensen, (1998), "Wage differentials, Employer size, and Unemployment", *International Economic Review*, (39), 257-73.

[12] Burdett, K., R. Wright, and S. Shi, (2001), "Pricing and Matching with frictions", *Journal of Political Economy*, (109), 1060-85.

[13] Canals, J. and S. Stern, (2001), "Empirical Search Models", University of Virginia Working Paper.

[14] Chen, Y., and R. Rosenthal, (1996a), "Asking Prices as Commitment Devices", *International Economic Review*, (37), 129-55.

[15] Chen, Y., and R. Rosenthal, (1996b), "On the use of ceiling-price commitments by monopolists", *RAND Journal of Economics*, (27), 207-20.

[16] Christensen, B., and N. Kiefer, (1991). "The Exact Likelihood Function for an Empirical Job Search Model.", *Econometric Theory*, (7), 464-486.

[17] Diamond, P., (1971), "A Model of price Adjustment", *Journal of Economic Theory*, (3), 156-68.

[18] D'urso, V., (2002), "Home Buyer Search Duration and the Internet", MIT Sloan School of Management Working Paper 4271-02.

[19] Eckstein, Z., and K. Wolpin, (1990), "Estimating a Market Equilibrium Search Model from Panel Data on Individuals", *Econometrica*, (58), 783-808.

[20] Glower, M., D. Haurin, and P. Hendershott, (1998), "Selling Time and Selling Price: The Influence of Seller Motivation", *Journal of the American Real Estate and Urban Economics Association*, (26), 719-740.

[21] Goolsbee, A., (2000), "In a World Without Borders: the Impact of Taxes on Internet Commerce", *The Quarterly Journal of Economics*, (115), 561-76.

[22] Heckman, J., (1979), "Sample Selection Bias as a Specification Error", *Econometrica*, (47), 153-162.

[23] Horowitz, J., (1992), "The Role of the List Price in Housing Markets: Theory and an Econometric Model", *Journal of Applied Econometrics*", (2), 115-129.

[24] Julien, B., J. Kennes, and I. King (2000), "Bidding for labor", *Review of Economic Dynamics* , (3), 619-40.

[25] Kang, H., and M. Gardner, (1989) "Selling Price and Marketing Time in Residential Real Estate Market", *The Journal of Real Estate Research,* (4), 21-35.

[26] Kennan, J. and R. Wilson, (1993), "Bargaining with Private Information", *Journal of Economic Literature*, (31), 45-104.

[27] Kiefer, N., and G. Neumann, (1994), "Wage Dispersion with Homogeneity: The Empirical Equilibrium Search Model" in *Panel Data and Labor Market Dynamics* (Westergard -Nielsen N., Jensen P), North-Holland, New York, 57-74.

[28] Levitt, S., and C. Syverson, (2005), "Market Distortions when Agents are Better Informed: The Value of Information in Real Estate Transactions", NBER Working Paper No. 11053.

[29] Lippman, S., and J. McCall, (1976), "The Economics of Job Search: A Survey." Parts I and II. *Economic Inquiry*, (14), 155-189 and 347-368.

[30] Mackie-Mason, J., and H. Varian, (1995), "Pricing the Internet," in *Public Access to the Internet*, Brian Kahin and James Keller, eds., Cambridge, MA, MIT Press.

[31] Mason, C., Quigley, J., (1996). Non-parametric hedonic housing prices. *Housing Studies* 11, 373–385.

[32] McKnight, L., and J., (1997), Internet Economics, Cambridge, MA, MIT Press.

[33] Mortensen, D., (1990), "Equilibrium Wage Distributions: A Synthesis" , in *Panel Data and Labor Market Studies* (Hartog J., Ridder G.,Theeuwes J.) North-Holland, New York, 279-96.

[34] Mortensen, D., and R. Wright, (2002), "Competitive pricing and efficiency in search equilibrium", *International Economic Review*, (43), 1-20.

[35] Palmquist, R., (1980), "Alternative Techniques for Developing Real Estate Price Indexes", *The Review of Economics and Statistics*, (62), 442-448.

[36] Rosen, S., (1974), "Hedonic Prices and Implicit Prices: Product Differentiation in Pure Competition", *The Journal of Political Economy*, (82), 34-55.

[37] Rogerson, R., and R. Wright, (2004), "Search-Theoretic Models of the Labor Market: A Survey", Penn Institute for Economic Research, Working Paper 02-041.

[38] Rubenstein, A., (1982), "Perfect Equilibrium in a Bargaining Model", *Econometrica*, (50), 97-109.

[39] Stern, S., (1990), "The effects of optimizing behavior in matching models", *Review of Economic Studies,* (57), 647-660.

[40] Yavas, A., (1992), "A Simple Search and bargaining Model of Real Estate Markets," *Journal of the American Real Estate and Urban Economics Association*, (20), 533-548.

[41] Yavas, A., and S. Yang, (1995), "The Strategic Role of Listing Price in Marketing Real Estate: Theory and Evidence," *Real Estate Economics*, (23), 347-368.

[42] Yinger, J., (1981), "A Search Model of Real Estate Broker Behavior", *American Economic Review*, (71), 591-605.

[43] Zumpano, L., K. Johnson, and R. Anderson, (2003), "Internet use and real estate brokerage market intermediation", *Journal of Housing Economics,* (12), 134–150.