An Empirical Stationary Equilibrium Search Model of the Housing Market

Paul E. Carrillo*

First Version: December 2006

This version: August 2010

Abstract

We specify and estimate a computationally tractable stationary equilibrium model of the housing market. The model is rich and incorporates many of its unique features: buyers’ and sellers’ simultaneous search behavior, heterogeneity in their motivation to trade, transaction costs, a trading mechanism with posting prices and bargaining, and the availability of an exogenous advertising technology that induces endogenous matching. Estimation is conducted using Maximum Likelihood methods and Multiple Listing Services data. The estimated model is used to simulate housing market outcomes when a) the amount of information displayed on real estate listings increases and b) real estate agent’s commission rates change.

Keywords: Equilibrium Search Models, Internet, Housing Market.
JEL Codes: C51, D58, D80, D83.

* Assistant Professor, Department of Economics, The George Washington University. I am grateful to three referees for their valuable comments. In addition, I owe special thanks to Steven Stern, Maxim Engers and Edgar Olsen, for their help and guidance. I would like to thank David Albouy, Jim Albrecht, Naoko Akashi, Patrick Bajari, Christina Rennhoff, Andrey Pavlov, John Rust, Katja Seim, Ken Wilbur, Anthony Yezer, and seminar participants for helpful discussions and comments. I also gratefully acknowledge financial support from the Banco Central del Ecuador and from the University of Virginia's Bankard Fund for Political Economy. All remaining errors are my own.
1 Introduction

Buyers’ and sellers’ search behavior is an important feature of many markets. For example, in the housing market, both buyers and sellers actively participate and incur expensive costs during the search process to buy or to sell a home. For sellers, it is costly to keep their homes on the market, while buyers incur in high pecuniary and non-pecuniary expenses when visiting properties. Variation in these costs alone may explain a large portion of price dispersion and, also, why certain transactions occur while some others don’t.

With these considerations in mind, in this paper we specify and estimate a stationary equilibrium model of the housing market. Since we are particularly interested in modeling home buyers’ and sellers’ search behavior, we build our model based on the economic theory of search. Empirical and theoretical equilibrium search models have been widely used to explain many important features of labor markets but, surprisingly, there are considerably fewer applications in the housing market. While Wheaton (1990) and Albrecht et al. (2007) are examples of previous equilibrium search-theoretic work, we are not aware of any empirical study that has attempted to estimate a stationary equilibrium search model of the housing market. This is, precisely, one of the important contributions of this research.

Specifying and estimating a model that adequately describes the features of the housing market and that, at the same time, conforms the observed data reasonably well is challenging. To capture the unique nature of this market, we extend the framework of traditional equilibrium search models in the labor literature in several dimensions. In particular, we specify a stationary equilibrium model that incorporates: a) buyers’ and sellers’ simultaneous search behavior,2 b) heterogeneity in agents’ motivation to trade, c) transaction costs, d) a trading process that buyers engage in is rather different as they actively sample homes from a distribution of home sellers. One of the precise objectives of this paper is to point out the differences between home buyers’ and home sellers’ search strategies which are reflected in our modeling choices. Why do home buyers and sellers employ different search strategies? While this question is beyond the scope of our paper, the answer may be closely related to what middlemen (real estate

---

1 In these type of models, workers search sequentially, drawing price offers from a known distribution. On the demand side, firms react optimally to individual’s behavior and price dispersion arises as an outcome of the equilibrium. See, for example, Eckstein and Wolpin (1990), Albrecht and Axell (1984), and Mortensen (1990). A revision of this literature is presented in the next section.

2 One could argue that sellers in the housing market are not searching in the conventional way one thinks about search. Rather than actively searching and drawing buyers at random, they wait patiently for offers. Home sellers’ behavior, however, can be and has been described by search models (see for example, Howoritz 1992). They passively wait for random offers to arrive, incur high transaction costs and their behavior can be characterized by optimal reservation strategies. The search process that buyers engage in is rather different as they actively sample homes from a distribution of home sellers. One of the precise objectives of this paper is to point out the differences between home buyers’ and home sellers’ search strategies which are reflected in our modeling choices. Why do home buyers and sellers employ different search strategies? While this question is beyond the scope of our paper, the answer may be closely related to what middlemen (real estate
mechanism with posting prices and bargaining, and e) an exogenous (online) advertising technology that induces endogenous matching between agents. These features distinguish our theoretical model from the existing literature. In addition, given the structure of our model, we show that it is feasible to estimate with seller’s data only.

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 same advertising technology (the internet) to advertise her product (housing unit) displaying its posting price and features. Each period, buyers use this technology to sample price postings (listings), learn the characteristics of units, and decide whether 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 listing. 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. While real estate agents’ behavior is not modeled, we assume that sellers pay 6% of the transaction value to cover agents’ fees. 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. A stationary 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. The data used for estimation consists of Multiple Listing Services (MLS) data (posting prices, transaction prices, time agents) do for their clients in this market. How do real estate agents influence buyers’ and sellers’ search strategies? Do agents increase the surplus from trade? Our model cannot answer these interesting questions because we do not address agents’ strategic behavior. We discuss some of these issues in the last section of the paper.

3Strategic behavior of middlemen in markets with costly search has been characterized. For example, Rust and Hall (2003) develop a stationary equilibrium model with search frictions where middlemen’s and market makers’ behavior is explicitly modeled. Given the complexity of our model, we are forced to ignore these interesting issues.

4See, for example, Eckstein and Wolpin (1990), Kiefer and Neumann (1994) and Bunzel et al. (2001).
that a property stays on the market, and a detailed set of home characteristics) for real estate transactions in Charlottesville City and Albemarle County (VA) during the years 2000 through 2002. The estimated model is able to replicate the pricing data remarkably well. For instance, posting and transaction prices are predicted with almost the same accuracy of a linear regression model despite the variety of assumptions and the complexity of our structural model. Furthermore, parameter estimates have a reasonable interpretation and provide interesting insights about the housing market. In particular, the coefficients of our model can assess the importance of the information that listings provide to potential buyers. Our findings suggest that more than half of the buyer’s home valuation is gathered at the time when a listing is viewed. Conditional on the housing unit’s characteristics, however, only 3% of the idiosyncratic buyer-home match is obtained through the information provided by the listing. The rest of the match is revealed when the property is visited. This finding is consistent with the observation that most home buyers visit a unit before a purchase is made.

The estimated model is used to perform two counterfactual experiments. First, we let the model predict changes in the housing market when the amount of information displayed by the listings increase. Because most housing units are advertised through the internet, this experiment implicitly evaluates the effects of improvements in the amount of information displayed in online advertisements on market outcomes.\(^5\) Such evaluation is relevant given that online housing sites have notably enhanced the amount of information they display by incorporating pictures and virtual tours to the existing MLS and that improvements are forecasted for the near future. Furthermore, as technological constraints rapidly decrease, the amount of information transmitted online and displayed on internet based listings are likely to rise.\(^6\) The counterfactual experiment suggests that additional online information decreases market prices. For instance, if buyers could obtain all the relevant information they need to make a purchase decision from a listing, posting and transaction prices would decrease by 4.0% and 4.4%, respectively; on the other hand, the average time that a home

\(^5\)Hence, our findings are related to the literature that explains how the internet has influenced online and offline markets (see, for example, Brown and Goolsbee 2002, Goolsbee 2000, and D’urso 2002).

\(^6\)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.
stays on the market would increase by about seven days.

Our second counterfactual describes how market outcomes change when real estate agents’ commission rates vary. In the U.S. the standard commission rate for real estate agents is 6% of the sale price. While this commission is paid by the seller at the time of the transaction, in equilibrium, these costs should be shared by both home buyers and home sellers. Thus, as agents’ commission rates vary, the new equilibrium should reflect both supply and demand responses. We find that commission costs are about equally shared by buyers and sellers. That is, if commission rates decrease from 6 to 4%, list and transaction prices diminish by about 1%. Reducing commission rates would decrease list and transaction prices but the effects on marketing time are negligible.

In the next section, we compare our research to the existing literature. The third section presents the theoretical model. In the fourth section, we describe the data. Sections 5 and 6 discuss the estimation methods and results, respectively. The last section concludes and points out limitations and opportunities for future research. In particular, we highlight that in order to describe temporary market imbalances in this market, the stationarity assumption should be relaxed.

2 Literature review

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 housing market.\(^7\) 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.\(^8\)


\(^8\)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).
and Judd (1983), Albrecht and Axell (1984), and Mortensen (1990) have introduced theoretical models that determine non-degenerate price distributions and search behavior as an equilibrium outcome. These models, however, introduce search behavior of one side of the market only. We believe that in the housing market both buyers and sellers actively search for a match and, thus, a model with two-sided search should be considered.

Mortensen and Wright (2002), Rust and Hall (2003) and Albrecht et al. (2007) introduce equilibrium search models with two-sided search. Mortensen and Wright (2002) assume that buyers and sellers have heterogeneous preferences and 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. Rust and Hall (2003) extend a model developed by Spulber (1996) to analyze the microstructure of trade in a market with search frictions, middlemen (real estate agents in the housing context) and market makers. Buyers, sellers and middlemen are heterogeneous which drives price dispersion in equilibrium. In Albrecht et al. (2007)’s model, home buyers and home sellers have homogeneous preferences, but they transition from a relaxed to a desperate state at an exogenous rate. This assumption creates a non-degenerate equilibrium price distribution. We extend the previous literature by introducing two important features of the real estate market: 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. To introduce posting prices in our equilibrium search model we assume (as in Chen and Rosenthal 1996) 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

---

9 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.
phenomena, the estimation of such models has received considerable attention. Eckstein and Wolpin (1990) estimate a generalization of the Albrecht and Axell model. 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). 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 $\beta_b$ and $\beta_s$ as the buyers’ and sellers’ value of 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 continuous and differentiable distributions $K_b$ and $K_s$, respectively. Note that the lower a household’s $\beta$ 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 $s_t$ as the per-period utility flow that sellers obtain by owning this good, and let $s = \frac{s_t}{r} = X\gamma$ be the level of lifetime utility that sellers obtain by owning this good, where $r$ is the common discount factor and $\gamma$ is a vector of parameters. Notice that $s$ represents the quality of the home and

---

10 Canals and Stern (2001) provide a comprehensive survey of empirical search models.

11 Formally, each of these discount factors are formed by two components; that is, $\beta_b = \beta^\circ \beta_b^*$ and $\beta_s = \beta^\circ \beta_s^*$. The first component $\beta^\circ = \frac{1}{1+r}$ discounts the future using a discount rate $r$ that is common to all buyers and sellers. On the other hand, the second components, $\beta_b^*$ and $\beta_s^*$, are random variables that capture idiosyncratic differences in buyers’ and sellers’ motivation when buying or selling their home.
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 continuous and differentiable distribution $\Psi$, which is common knowledge to every agent in the market.

A seller joins the housing market by placing a listing that informs all the other agents: a) that her home is for sale, b) of the posting price and, c) of 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.\(^{12}\) If a sale is made, sellers pay a fixed cost $T$ (fixed costs paid by the seller) and a percentage of the transaction price $k$ (real estate agent’s commission rate). Both $T$ and $k$ are exogenous.

In addition to the differences in motivation to trade, we add another source of heterogeneity. We let the lifetime utility that properties provide to buyers vary for each buyer-home combination. Specifically, we let $\tilde{b}$ be the random buyer-home match value that captures the lifetime utility a specific buyer derives from owning one particular property and let it depend on the home quality and a random component; that is, $\tilde{b} = \delta s + b$ where $\delta$ is a scalar parameter and $b$ is a draw from an i.i.d. mean zero random variable. The parameter $\delta$ captures average percentage differences in the properties’ valuations between buyers and sellers, and realizations of $b$ correspond to the specific value that one buyer assigns to one particular property.\(^{13}\) These assumptions imply that a unique home (fully described by $s$) could provide different levels of utility to different buyers or to the same buyer in two different time periods. While this may seem like a strong implication, this approach is common in models with stochastic matching where the potential transaction surplus realized when a buyer and a seller meet is pair-specific and random (see, for example, Pissarides 2000).

\(^{12}\)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.

\(^{13}\)Notice that all buyers agree about the unit’s mean valuation; that is $\delta$ is non-random and constant for all buyers.
In our model, however, the random-match component affects buyers only. Finally, let us assume that \( b \) is composed of two independent (from each other and for any home-and-buyer combination) mean zero random errors, \( b^o \) and \( b^u \). That is, \( b = b^o + b^u \), where \( b^o \) and \( b^u \) are i.i.d. mean zero random variables with variances \( \sigma_{b^o}^2 \) and \( \sigma_{b^u}^2 \) and exogenous continuous and differentiable cumulative distributions \( G_o \) and \( G_u \), respectively.

Let us now describe the process of how \( b \) is revealed to buyers. Buyers search “home for sale” listings sequentially and there is no recall. Every period, a buyer samples one listing that provides her information about the home’s characteristics \( \Phi \), its posting price \( p_s \), and \( b^o \), a fraction of her total random buyer-home match value. When a buyer observes a listing, 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. It is useful at this point to clarify the differences between \( s \), \( b^o \) and \( b^u \). At the time buyers observe a listing, they are able to infer their average valuation of the unit \( \delta_s \), because the listing displays all of its relevant characteristics \( X \). Conditional on \( X \), homes are identical. Given our assumptions, however, buyers’ valuation of this (identical) unit may vary since it depends on the realizations of the buyer-home match components. \( b^o \) and \( b^u \) provide the same type of information and do not depend on the home’s characteristics. The only difference is that the former is revealed at the time the listing is observed while the latter is observed only if a visit is made. What will matter in our application is the relative size of these errors. We return to this point later.

It may not be realistic to assume that buyers see all characteristics of the unit at the time the listing is viewed. Alternatively, one could assume that listings reveal only a subset of \( X \). The latter approach could be used to analyze interesting strategic behavior such as seller’s information disclosure choices and the buyer’s responses to the listing’s contents. Given the complexity of our model, however, we are forced to make the simplifying assumption that

\[ \text{\footnote{The reader may notice that, for a given set of housing characteristics } X, \text{ every seller receives the same lifetime utility } s = X \gamma. \text{ On the other hand, we assume that the lifetime utility } \tilde{b} \text{ that buyers receive varies for each buyer-home combination. Why is the value of a home different for buyers and sellers? Due to transaction costs, consumption of housing services does not adjust continuously. That is, home owner’s preferences should change by a large amount before they decide to move and become “sellers.” In our model, the parameters that explain this “transformation” are } \delta \text{ and the realization of } b. \text{ In particular, before former buyers become sellers, the buyer-home match value is lost and the mean utility that they get from the home decreases by } 1/\delta. \text{ A natural way to relax this assumption consists of including a random component in } s, \text{ which does not change the nature nor the main results of our model.}} \]
listings provide all relevant characteristics to buyers.\(^{15}\)

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 \(\theta\), 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 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).\(^{16}\) During the meeting, the buyer decides whether she buys the home (paying either \(p_s\) or \(p_b\)), or searches again for a new listing (posting price) next period. When a buyer (seller) buys (sells) a property, she exits the market forever.

### 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

---

\(^{15}\)We think this is not an unreasonable assumption because, in most markets, real estate agents are required to provide a large amount of information on their listings (such as the unit’s location, square footage, number of bathrooms, bedrooms, among others) and, presumably, these are the most important variables that influence buyer’s demand.

\(^{16}\)Notice that an implicit assumption of our bargaining model is that sellers do not have perfect information about buyers’ preferences at the time they meet. That is, if buyers get to make a counteroffer they will offer the seller’s reservation value. On the other hand, sellers never know the specific realization of the buyer-home match component \(b\) and cannot anticipate the gains from trade of a potential transaction. If it is assumed that sellers have perfect information about buyers’ preferences, 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.
\[
\Pi^e_s = q \theta [\gamma_s(N(p_s) - s) + (1 - \gamma_s)\beta_s \Pi^e_s] + q(1 - \theta) [\gamma_b \max\{N(R_s) - s, \beta_s \Pi^e_s\} + (1 - \gamma_b)\beta_s \Pi^e_s] + [1 - q]\beta_s \Pi^e_s,
\]
where \(N(p_s) = (1 - k)p_s - T\) is the cash proceeds net of real estate commissions. Here \(0 < k < 1\) is the agents’ commission rate and \(T\) represent closing costs (specific to the seller). The term \(\max\{N(R_s) - s, \beta_s \Pi^e_s\}\) represents the seller’s decision if she receives an offer less than the list 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 \(N(p_s) - s\) 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 \(N(R_s) - s\); finally, if trade does not happen, she returns to the market and keeps her discounted value of search \(\beta_s \Pi^e_s\).

The seller’s problem consists of choosing the optimal reservation value \(R^*_s\) and posting price \(p^*_s\) that maximize her expected profit. We have assumed that the buyer knows all the characteristics of the seller. Thus, if the buyer makes the take-it-or-leave-it offer (and if her private valuation exceeds \(R_s\)) she should offer the smallest value acceptable to the seller; that is, she should offer the value of \(R^*_s\) that solves the equation

\[
N(R^*_s) - s = \beta_s \Pi^e_s.
\]

We substitute this expression in equation (1) and obtain that for any optimal \(R^*_s\),

\[
\Pi^e_s = \theta q\gamma_s(N(p_s) - s) + (1 - \theta q\gamma_s)(N(R^*_s) - s).
\]

Differentiating this equation with respect to \(p_s\), we derive that the optimal seller’s posting price \(p^*_s\) solves

\[
N(p^*_s) - N(R^*_s) = (1 - k)\frac{1 - \phi(p^*_s|s)}{\phi(p^*_s|s)},
\]
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: \(1 - \phi(p_s|s) = q(p_s|s)\gamma_s(p_s|s)\).
To find the optimal reservation value, we substitute equation (2) for the expected profit in equation (3). After a bit manipulation, we find that the optimal reservation value is defined by

$$N(R^*_s) = \frac{\beta_s \theta(1 - \phi(p^*_s|s))N(p^*_s) + (1 - \beta_s)s}{1 - \beta_s[1 - \theta(1 - \phi(p^*_s|s))]}.$$ (5)

**Proposition 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) $\phi$ 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 (5) and (4) simultaneously. For any seller, and for any $\beta_s$ and $s$:

$$p^*_s(\beta_s, s) \geq R^*_s(\beta_s, s) \geq \frac{s + T}{1 - \theta};$$ in addition, these functions are increasing in both arguments. Proof in Appendix.

The two assumptions of Proposition 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. Condition (b) makes the reasonable conjecture that, everything else constant, the higher the quality of a home, the higher the probability it is sold.

The results of Proposition 1 are all intuitive. 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 the derivation of $\phi(p^*_s|s)$.

### 3.3 The buyer’s 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

---

17 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).

18 This condition says that an increase in $s$ increases $p^*_s$ in the sense of first-order stochastic dominance.
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 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 $p_s^*(\beta, s)$ as well as the exogenous distributions of discount factors $K_s$ and valuations $\Psi$.

When the buyer visits a property, she pays a known visiting cost $c_b$. These include transportation and monetary opportunity costs (time costs), as well as non pecuniary (emotional) costs of touring a property. 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

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

$U^e(p_s^*, s, b^o, W_b)$ 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

$$U^e(p_s^*, s, b^o, W_b) = \theta \int \max \{b^o + b^u + \delta s - p_s^*, W_b\} dG_u(b^u)$$

$$+(1-\theta) \int \max \{b^o + b^u + \delta s - p_b, W_b\} dG_u(b^u) - c_b. \quad (7)$$

Because buyers know the optimal strategies $p_s^*(\beta, s)$ and $R_s^*(\beta, s)$, 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
have the opportunity to make a counter offer, it is optimal for them to ask \( p_b = R^*_a(p^*_s, s) \) which does not depend on \( b^u \). Using this result, we integrate the right hand side of equation (7) by parts and obtain\(^{19}\)

\[
U^c(p^*_s, s, b^o, W_b) = W_b + j(p^*_s, s, b^o, W_b),
\]

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

\[
B = \{ b^u : b^u \geq W_b + R^*_a(p^*_s, s) - \delta s - b^o \}.
\]

Now, let us find the value of search \( W^*_b \) that solves equations (6) and (8). We substitute equation (8) into (6), rearrange, and find

\[
W^*_b = \frac{\beta_b}{1 - \beta_b} \int \int \max \{ j(p^*_s, s, b^o, W^*_b), 0 \} d\Gamma(p^*_s, s) dG_o(b^o).
\]

It is easy to see that there is a unique \( W^*_b \) that solves equation (9).\(^{20}\)

It is straightforward to show that \( U^c \) is decreasing with respect to the posting price.\(^{21}\) 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 visits the property if and only if the posting price \( p^*_s \) is below a reservation price \( p^*_b(s, b^o) \).

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

\[
U^c(p^*_b(s, b^o), W^*_b) = W^*_b.
\]

\(^{19}\)Details of the integration are provided in the appendix.

\(^{20}\)The right hand side of equation (9) 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 (9) crosses the origin and has a positive slope. Thus, a unique solution \( W^*_b \) exists.

\(^{21}\)We use equation (8) -and Leibnitz rule- to show that \( U^c \) is monotone

\[
\frac{\partial U^c}{\partial p^*_s} = -\theta (1 - G_u(V_p)) - (1 - \theta) \frac{\partial R^*_a}{\partial p^*_s} (1 - G_u(V_r)) \leq 0,
\]

where \( V_p = W_b + p^*_s - \delta s - b^o \), and \( V_r = W_b + R^*_a(p^*_s|s) - \delta s - b^o \).
To find the optimal \( p_b^* (s, b^o) \), we replace this optimality condition in (8) and solve

\[
\theta \int_{A(p_b^*, s, b^o)} [1 - G_u(b^u)] db^u + (1 - \theta) \int_{B(p_b^*, s, b^o)} [1 - G_u(b^u)] db^u = c_b, \tag{10}
\]

where

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

and

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

**Proposition 2** The solution to the buyer’s two-step search is defined by a unique value \( W_b^* \) and a function \( p_b^* (s, b^o) \) that solve equations (9) and (10) 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 \( p_s \leq p_b^* (s, b^o) \); (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 \( \bar{b} - p_s > W_b^* \); 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 \( \bar{b} - R_s(p_s, s) > W_b^* \). In addition, \( W_b^* \) is increasing in \( \beta_b \), while \( p_b^* \) is decreasing in \( \beta_b \). 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 \( p_b^* \) and provide an intuitive interpretation of the buyer’s optimal decisions.

**Proposition 3** 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 \( \{ W_b^*, p_b^* \} \) that solves equations (11) and (12) simultaneously, along with the optimal strategies described in (a), (b), and (c) in Proposition 2. Proof in appendix.

\[
\frac{\beta_b}{1 - \beta_b} \int \Gamma(p_s^*) (-\frac{\partial U_s}{\partial p_s^*}) dp_s^* = W_b^* \tag{11}
\]

\[
\theta \int_{\frac{W_b^* + p_b^* - \delta s}{W_b^* + R_s(p_b^*) - \delta s}} [1 - G_u(b^u)] db^u + (1 - \theta) \int_{\frac{W_b^* + R_s(p_b^*) - \delta s}{W_b^* + p_b^* - \delta s}} [1 - G_u(b^u)] db^u = c_b \tag{12}
\]

Equation (11) states that the expected benefits from sampling a posting price lower than the buyer’s reservation price \( p_b^* \) (the left hand side) should be the same as the per-period
expected return of staying in the market (since $\frac{1-\beta}{\beta}$ equals the per-period discount rate). Equation (12) implies that the optimal $W_b^*$ and $p_b^*$ must be such that, the buyer’s expected benefit from visiting a property equals her visiting cost.

### 3.4 Equilibrium

**Definition**

From the sellers’ point of view, there is a distribution of heterogeneous 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, $p_b^*$, $W_b^*$, and the relevant exogenous cumulative distribution functions, each seller is able to determine the probability that a buyer visits her property and is willing to trade. For instance, the rate at which buyers visit a seller who has posted a price $p_s$ and owns a type $s$ home, $q^*(p_s|s)$, may be computed using the optimal buyer behavior rules described in Proposition 2 since only those properties for which $p_s < p_b^*(\beta_b, s, b^p)$ will be visited. Similarly, buyer’s strategies may be used to infer $\gamma_s^*(p_s|s)$, the probability that the buyer is willing to buy the housing unit given that she has visited the home and did not have the opportunity to make a counteroffer. This is evaluated by computing the (conditional) probability that the buyer’s gains from trade $\bar{b} - p_s$ exceed her outside option $W_b^*(\beta_b)$ provided a home visit has occurred. The probability that trade occurs given that a buyer has visited the home and had the opportunity to make a counteroffer, $\gamma_b^*(p_s|s)$ may be evaluated by computing the (conditional) probability that the buyer’s gains $\bar{b} - R_s$ exceed her outside option $W_b^*(\beta_b)$. Hence, $1 - \phi^*(p_s|s) = q^*(p_s|s)\gamma_s^*(p_s|s)$ is also well defined. In the appendix, we provide more insights about these functions and show that $1 - \phi^*(p_s|s)$ is decreasing.

Before equilibrium can be defined, notice that without entry of new buyers and sellers, the distribution of buyer’s and seller’s types cannot be stationary. In our model, those who are more motivated to trade would exit the market sooner which would shift the distribution of buyer’s and seller’s discount factors as time increases. Some assumptions are thus needed to guarantee stationarity. In particular, we assume that new entrants exactly match the type of those buyers and sellers who exited after trading.22

---

22 A similar approach was taken by Rust and Hall (2003). In their model, however, the rate at which
Let \( \omega_b \) be the rate at which buyers exit the market. This rate depends on the buyer’s type (discount factor)

\[
\omega_b(\beta_b) = k_b(\beta_b) \chi_b(\beta_b),
\]

where \( k_b \) is the buyer’s discount factor density and \( \chi_b \) is the (unconditional) probability that a type \( \beta_b \) buyer exits the market in any period. Notice that \( \chi_b \) is well defined by optimal buyer’s strategies.\(^{23}\) To guarantee that \( K_b \) is constant over time, the rate at which type \( \beta_b \) buyers exit the market should match the rate at which new type \( \beta_b \) buyers enter. This can be achieved if we assume that new buyers are sampled at random from the probability density

\[
\tilde{k}_b(\beta_b) = \omega_b(\beta_b) / \tilde{\omega}_b,
\]

where \( \tilde{\omega}_b = \int_0^1 \omega_b(x) dx \).

A similar analysis applies to sellers. Let \( \omega_s \) be the rate at which sellers exit the market, which depends on both the seller’s discount factor \( \beta_s \) and home’s characteristics \( s \)

\[
\omega_s(\beta_s, s) = k_s(\beta_s) \tau(s) \chi_s(\beta_s, s).
\]

Here \( k_s \) is the seller’s discount factor density, \( \tau \) is the density of \( s \), and \( \chi_s \) measures the probability that a seller trades her home in any period.\(^{24}\) The rate at which a type \((\beta_s, s)\) seller leaves should match the rate at which a seller of the same type enters the market. This is achieved if we assume that new sellers are randomly sampled from the probability density

\[
\tilde{k}_s(\beta_s, s) = \omega_s(\beta_s, s) / \tilde{\omega}_s,
\]

agents enter the market is exogenous and the types of active buyers and sellers in the market is endogenously determined. In our model, the distribution of active buyer’s and seller’s types is exogenous \((K_s, K_b, \text{and } \Psi)\), but the rate at which newcomers enter is endogenous.

\(^{23}\)To be specific, the rate at which a type \( \beta_b \) exits the market is

\[
\chi_b(\beta_b) = E \left\{ \Pr \{ p_b^*(\beta_b, s, b^o) > p_s \} \left[ \theta \Pr \{ b^o + b^u + \delta s - p_s > W^*_b(\beta_b)p_b^*(\beta_b, s, b^o) > p_s \} + (1 - \theta) \Pr \{ b^o + b^u + \delta s - R_s > W^*_b(\beta_b)p_b^*(\beta_b, s, b^o) > p_s \} \right] \right\}.
\]

Here, the first term is the probability that a buyer visits a property. The terms inside the brackets compute the probability of trade given that a visit has been made. Notice that, from a type \( \beta_b \) buyer’s point of view, \( s, b^o, b^u \) and \( p_s \) are random.

\(^{24}\)To be specific,

\[
\chi_s(\beta_s, s) = q^*(p_s(\beta_s)|s) \left\{ \theta \gamma^*_s(p_s(\beta_s)|s) + (1 - \theta) \gamma^*_b(p_s(\beta_s)|s) \right\}.
\]

Here, the first term is the probability that a buyer visits a property. The terms inside the brackets compute the probability of trade given that a visit has been made.
where }\bar{\omega}_b = \int_0^\infty \int_0^1 \omega_s(x, y) dx dy.

**Definition 4** A stationary equilibrium is defined by a set of exogenous distributions \((\Psi, G_o, G_u, K_s, K_b)\) and a function }\phi^*(p_s|s) such that a) home sellers’ and home buyers’ search behavior is consistent with Propositions 1 and 2, respectively, and b) new entrants exactly match the type of those buyers and sellers who exit after trading.

**Computation**

We are not able to provide analytical insights about the equilibrium properties of our model. Instead, numerical methods are used to solve it both for estimation purposes and for counterfactual analysis. Before we can start solving the model, the exogenous distributions’ functional forms \((\Psi, G_o, G_u, K_s, K_b)\) as well as other parameter values need to be specified. We will provide details about these issues later but, for now, assume they are appropriate.

To solve the model and find the equilibrium, we use an iterative approach. We begin by picking a starting value for the function }\phi which is denoted by }\phi_0(p_s|s). Given this function, sellers’ optimal list prices }p^*_s(s, }\beta_s) and reservation values }R^*_s(s, }\beta_s) are computed as described in Proposition 1. These previous calculations as well as Proposition 2 are used when computing the buyers’ value functions and solving their two-step search problem. Optimal buyers’ strategies can be summarized by the functions }p^*_b(}\beta_b, s, }\beta_b) and }W^*_b(}\beta_b), which in turn can be used to find a new estimate of the function }\phi: }\phi_1(p_s|s). Given }\phi_1, the seller problem may be solved again. This procedure continues until convergence, that is, until we reach a number }n such that }||\phi_n - }\phi_{n-1}|| < }\varepsilon, where }\varepsilon is a “small” non-negative value.

It is appropriate at this point to discuss some of the convergence properties of our iterative method. Unless otherwise noted, the model is always solved using the same starting values. In particular, we let }\phi_0 = 1 - \exp\{\alpha(p_s - s)\}, where }p_s > s and }\alpha = 0.08. We call this the “baseline” function. To facilitate the convergence of our numerical routine, we smooth the convergence process, that is the estimate of }\phi after }n iterations is computed as a weighted average of }\phi_{n-1} and }\phi_n.\textsuperscript{25} Given these adjustments and a value of }\varepsilon of 5E-04, it takes about

\textsuperscript{25}The weight given to }\phi_n is computed as }w = \max\{\min\{0.1 \times T, 0.4\}, 0.05\}, where }T is the maximum absolute difference between }\phi_{n-1} and }\phi_{n-2}. These numbers have been set by trial-and-error to avoid non-convergence problems. If the procedure iterates more than 20 times without improvements, we decrease the upper bound of the minimum operator (0.4) by 10%.
70 iterations before our program converges.

The choice of the baseline function \( \phi_0 \) may seem arbitrary and readers may wonder if results are sensitive to this choice. To alleviate these concerns, equilibrium has been computed using alternative starting values and results are displayed in Table 1. We first use the same type of exponential specification used in the baseline function but change the value of \( \alpha \) to 0.05 and 0.40. Then, we let \( \phi_0 = \eta \frac{z - \bar{z}}{s} \), where \( 0 < \frac{z - \bar{z}}{s} < 1 \) and compute the equilibrium when \( \eta \) equals 0.80 and 0.95. We refer to this case as the “linear” function. Finally, we compute a “normal” specification where \( \phi_0 = \Phi\left(\frac{z - \bar{z}}{\tau}\right) \), \( z = \frac{p_s - s}{\bar{z}} \), and \( \Phi \) is the standard normal cumulative distribution function. For our purposes, we fix \( \bar{z} = 0.20 \) and let \( \tau \) take the values of 0.20 and 0.40. Results shown in Table 1 suggest that the choice of \( \phi_0 \) may affect the speed of convergence, but does not seem to have an effect in equilibrium. It is clear that the average and maximum absolute differences between the function \( \phi^* \) computed with the baseline starting values and the same function computed using alternative starting values are negligible. This exercise is not by any means an attempt to provide a proof of equilibrium uniqueness but to show some empirical evidence that multiple equilibria may not be a serious concern in our application.

3.5 Changes in the information technology and real estate commission rates

As it was discussed in the introduction, our model can be used to perform several counterfactual exercises. In particular, we use it to explore the effects on the housing market of a) changes in the amount of information displayed by the listings and b) changes in real estate commission rates.

Developments in information technology enhances the content of home listings and provides buyers with additional information at the time they look at an ad. This additional information should change the agents’ optimal strategies and the equilibrium of the 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 when they view ads. 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. To evaluate these changes, we analyze how market outcomes vary as we modify \( \lambda = \frac{\sigma_b^2}{\sigma_b^2} \), the ratio between the variance of \( b^0 \) and the total variance of the buyer-home match value \( \sigma^2_b = \sigma^2_{b^0} + \sigma^2_{b^u} \).

Because the variables in our equilibrium model 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, buyers are able to perform a more careful screening process before they decide to visit a property. Thus, buyers’ 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 once the buyer has decided to visit a particular property. On the one hand, buyers visit only those properties with a relatively high observed match value \( b^0 \); on the other hand, the value of their outside option (value 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 \( 1 - \phi^* \) make sellers set higher posting prices and higher 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.

In the U.S., the standard commission rate for real estate agents is 6% of the sale price. While this commission is paid by the seller at the time of the transaction, in equilibrium, these costs should be shared by both home buyers and home sellers. From a partial equilibrium point of view, seller’s list prices and reservation values should increase with commission rates. However, changes in sellers’ posting and reservation prices have a direct effect on buyers’ optimal search strategies, shifting the demand function as well. Thus, as agent’s commission rates vary, the new equilibrium should reflect both supply and demand responses. Again, due to the complex nature of the model we cannot make any theoretical predictions about the size of these effects on equilibrium.

To answer these interesting questions several comparative statics exercises will be performed in the following sections. In particular, we use numerical methods to compute counterfactual experiments that simulate market outcomes when a) the amount of information
displayed on listings increases and b) agent’s commission rates change. Before we attempt to do these tasks, however, we need to estimate the model to have a reliable benchmark for our comparisons.

Before estimation is implemented, it is useful to analyze how changes in parameter values affect the model’s predictions. As before, we make assumptions about the exogenous distributions’ functional forms and pick appropriate parameter values for a baseline model. The baseline model is used to predict mean list prices, mean transaction prices and mean marketing time. These predictions are shown in the first row of Table 2. We then predict mean market outcomes when one parameter of the baseline model is modified at a time. For example, results in Table 2 suggest that if \( \lambda \) rises from 0.027 to 0.5 (and every other parameter remains constant), list and transaction prices would decrease by about 3% while marketing time will increase by about 3 days. These results provide some initial insights about our first counterfactual of interest. We will elaborate on this point later. In addition, we show how market outcomes change when other parameters of the demand function, \( \sigma_b \), \( \delta \) and \( c_b \), vary. We modify these parameters by relatively large amounts to magnify their effects and show that these modifications affect the joint distribution of list prices, transaction prices and time on the market.

4 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. These 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, their combined population was 126,832 in 2003. In 2002, the total number of housing units in Charlottesville

26Specific functional form assumptions are detailed in section 5.2. Parameter values are shown in Tables 4 and 5.
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 2002. 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 black, 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 3.

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. In about 58% of transactions, the list price was higher than the transaction price. On average, the difference between posting prices and transaction prices was close to $4,200. Despite these facts, the distribution of both posting and transaction prices are quite similar (both distributions are unimodal and skewed to the right).

In this area, most homes sell relatively quickly. 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 on the market for more than four months.

27To 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.
A typical home is about 24 years old, has 1,980 square feet, two bathrooms, and is located in a 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 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 Census blocks populated by blacks only.

5 Estimation

Before the model can be estimated, we need to provide specific functional form assumptions for the exogenous distributions \( K_s(\xi_s) \), \( K_b(\xi_b) \), \( G_o(\sigma_o) \), and \( G_u(\sigma_u) \), where \( \xi_s, \xi_b, \sigma_o, \) and \( \sigma_u \) are parameters that characterize them. Then, for a given set of parameter values \( \Theta = \{\gamma, \theta, c_b, \delta, \xi_s, \xi_b, \sigma_o, \sigma_u\} \) and exogenous housing characteristics \( X = \{X_i\}_{i=1}^N \), we use the numerical algorithm described in section 3 to solve our model. Once equilibrium is found, the joint likelihood of observing the sample may be computed. The set of parameters of our structural model \( \Theta \) can be estimated using Maximum Likelihood. Details are provided below.

5.1 The Likelihood Function

\(^{28}\text{In what follows, we condition on a set of parameter values } \Theta. \text{ But, for expositional purposes, we omit it from our notation.}\)
As we described in the previous section, 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.

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 \(\beta_s\).

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 5 and the optimal \(\phi^*\)) to calculate \(u_i\) and \(\beta_{si}\). That is, with information on \(\{p_{si}, p_{mi}, X_i^o\}\) we compute \(u_i\) and \(\beta_{si}\) as the unique values that simultaneously solve

\[
N(p_{mi}) = N(p_{si}) - (1 - k) \frac{1 - \phi^*(p_{si} | X_i^o \gamma + u_i)}{\phi^*(p_{si} | X_i^o \gamma + u_i)} \frac{N(p_{mi}) - X_i^o \gamma - u_i}{N(p_{mi}) - X_i^o \gamma - u_i + (N(p_{si}) - N(p_{mi})) \theta [1 - \phi^*(p_{si} | X_i^o \gamma + u_i)]}.
\]

For any given \(u_i\), we can determine the corresponding \(s_i = X_i^o \gamma + u_i\); and, using \(q_s^*, \gamma_s^*, \) and \(\gamma_b^*\), 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

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

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

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

where $k_s$ is the density of the seller’s $\beta$s, and $|J_i|$ is the absolute value of the determinant of the Jacobian of the transformation implied by equations (15) and (16).

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 (15) to estimate the seller’s reservation value $p_m(u) = p_m(p_{si}, X_i^o, u)$. With this information, and using equation (16), 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

$$l_{tp}(t_i, p_{si}, p_{mi}, X_i^o) = \int l_{op}(u)k_s(\beta_s(u))|J_i(u)|f_u(u)p_m(u)\frac{X_i^o\gamma + u + T}{1 - k}du.$$ 

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 $p_m(u) > \frac{s + T}{1 - k}$ are considered.

The log-likelihood contribution $L_i$ of observing

$$\{t_i, p_{si}, p_{mi}, X_i^o, d_i\}$$

is

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

24
and the MLE parameter estimates are the ones that maximize the log-likelihood of observing
the sample

\[ \Theta^{MLE} = \arg \max_{\Theta} \left\{ \sum_{i=1}^{n} L_i(p_{si}, p_{mi}, t_i, X^o_i, d_i; \Theta) \right\} . \]  \hspace{1cm} (17)

### 5.2 Functional form assumptions and identification

Before we solve the model, we need to specify functional forms for all of the relevant exo-
genous distributions. We let both \( \beta_b \) and \( \beta_s \) be a transformation of a normally distributed
random variable \( \tilde{\beta} \). That is,

\[ \tilde{\beta}_j = \log \left( \frac{\beta_j}{1 - \beta_j} \right) \sim N(u_{\beta_j}, \sigma_{\beta_j}^2); \quad j = b, s. \]

In addition, we let \( b^o, b^u, \) and \( u \) be mean zero normal random variables with standard de-
viations \( \sigma_{b^o}, \sigma_{b^u} \) and \( \sigma_u \) respectively. The total buyer home component \( b \) is thus a mean
zero random normal random variable with standard deviation \( \sigma_b = \sigma_{b^o} + \sigma_{b^u} \). Given these
functional form assumptions and a set of parameters values, we use the numerical methods
described earlier to solve the model.

To make estimation feasible, we make the additional identifying assumptions that the
distributions of \( \beta_b \) and \( \beta_s \) 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 in
fluenced, on average, by the same kind of factors that could affect their motivation to trade (for example,
changes in life style, income, and employment). When making comparative statics analysis,
we need to be aware of this normalization to provide a correct interpretation of our results.
We also use our knowledge of the industry to calibrate agent’s commission rates \( k \) to 6% of
the transaction value. In addition, we assume that seller’s do not pay any relevant closing
costs \( (T = 0) \).\(^{29}\)

With the previous assumptions, we use MLE to estimate the model’s parameter vector

\[ \Theta = \{ \mu_{\beta}, \sigma_{\beta}, \sigma_b, \lambda, \sigma_u, \gamma, \delta, c_b, \theta \} . \]

\(^{29}\)In the U.S., most closing costs are paid by buyers.
Notice that, rather than estimating $\sigma_{y^c}$ and $\sigma_{y^v}$, in our estimation routine, we identify $\sigma_b$ and $\lambda$.

It is important to provide some intuition about how we identify $\Theta$ in our likelihood function. Notice first that variation in prices and home characteristics identify the vector $\gamma$ in a similar manner as OLS in a linear hedonic regression. The rest of the parameters are identified by relying on the parametric structure of the model. In our model, sellers of identical homes may have different posting prices and reservation values if their motivation to trade differs. That is, for a given $s$, higher posting prices should point out sellers with higher discount factors. Given a specific discount factor, however, our model makes precise predictions about the gap between posting prices and reservation values. Differences between the model’s predictions ($P_s - R_s$) and what we observe in the data ($p_{si} - p_{mi}$) can then be attributed to unobserved characteristics of the unit. In sum, differences between posting prices and transaction prices together with the structure of the model identify $\mu_\beta, \sigma_\beta$, and $\sigma_u$.

The share of transactions where the transaction price equaled the posting price identifies $\theta$. To see this, let $\tilde{\theta}$ be the probability that the posting price is the transaction price given that trade occurs; that is $\tilde{\theta} = \Pr\{p = p^*_s|\text{trade}\}$. Notice that $\tilde{\theta}$ may be easily estimated with our data (it is 0.42) and that using Bayes’ rule we obtain that $\tilde{\theta} = \frac{\theta \Pr\{\text{trade}|p = p^*_s\}}{\Pr\{\text{trade}\}}$. Given the model’s estimates for the per-period likelihood of trade, it is easy to see then how $\theta$ is identified from the data. Notice furthermore that $\theta$ is not simply equal to $\tilde{\theta}$ and that the structure of the model gives us extra insights into bargaining.

All previous parameters have been identified without information about time on the market. The joint distribution of posting prices and marketing time allow us to identify $\delta$ and $c_b$. To see this, notice that the expected time on the market depends upon the list price and the unit’s characteristics. In a partial equilibrium context, if $c_b$ diminishes, buyers increase their search efforts and the unconditional mean time that a seller stays on the market should rise. In the same context, if $\delta$ changes, then both the gap between the list price and the buyers’ average valuation of the home ($p^*_s - \delta s$) as well as the per-period likelihood of trade are affected. Thus, mean time on the market and the covariation between list prices and marketing time identify $c_b$ and $\delta$, respectively.
Equilibrium effects help us identify the two remaining parameters of the buyer’s demand function, $\sigma_b$ and $\lambda$. As these two parameters change, both the buyer’s value of search and the seller’s optimal pricing strategies are affected in equilibrium. How are these two parameters separately identified? While it is not easy to see how this occurs given the complicated nonlinear nature of the equilibrium, we argue that changes in each of these parameters have different effects in the joint distribution of posting prices, transaction prices and time on the market. For example, our model predicts that marketing time increases in equilibrium when $\sigma_b$ or $\lambda$ rise (see Table 2). However, the effects on transaction prices are rather different. Equilibrium transaction prices rise with $\sigma_b$ but diminish with $\lambda$.

6 Results

6.1 Parameter estimates

In Tables 4 and 5, we present maximum likelihood estimates of the structural model parameters, and, in the rest of the section, we discuss the implications of these results. The estimate of $\theta$ implies that there is a 0.47 probability that the list price is a take-it-or-leave-it offer to the buyer. Because $\theta$ is close to one half, buyers and sellers have a similar level of market power.

The mean (and median) of the transformed discount factor $\tilde{\beta}$ is about 8.6. This implies daily, monthly and yearly median discount factors of 0.9998, 0.994 and 0.93, respectively. These results suggest that, the median agent values an opportunity to sell their home one year from today 7% less than having the same opportunity now. The estimate of $\sigma_\beta$ is somewhat small (1.125). However, small variation in the agents’ discount-factors implies important differences in both buyer’s and seller’s optimal behavior. To illustrate this point, we present in Figure 1 the seller’s optimal posting prices and reservation values for different values of $\beta_s$ (for the median home quality). The smallest (largest) annual discount factor we picked is 0.25 (0.999), which corresponds to the 0.2th (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
Our results suggest that buyers spend $506 each time they visit and consider purchasing a real estate asset. This result is consistent with Shimizu et al. (2003) who find that, in the Tokyo housing market, buyers spend about $1,700 (207,900 Japanese Yen) each time they visit and consider buying a home. Thus, this estimate provides further evidence that search costs in the housing market are significantly large and may be an important source of market inefficiency. The visiting cost includes time, transportation, and other nonpecuniary (emotional) costs that are spent during the process of visiting, touring, and deciding to buy a home. A simple calculation reveals that time and transportation costs may add up to $213 which is far lower than our visiting cost estimate of $506. This means that other nonpecuniary (emotional) costs account for the majority of it (58%).

If exogenous shocks (such as mobility of fertility shocks, for example) explain the entry of home buyers and sellers into the housing market, it may be expected that buyers would value homes more than sellers. Our estimate of $\delta$ suggest this is the case: on average buyers value homes about 8% more than sellers do. $\sigma_u$ represents the dispersion of all unobserved factors (to the econometrician) that affect the value of a home. Notice that the error’s standard deviation of a OLS hedonic model has a similar interpretation. Thus, we would expect $\sigma_u$ to be close in magnitude to it. We return to this point later.

The estimate of $\sigma_b$, the standard deviation of the idiosyncratic buyer-home match value $b$, is relatively large ($81,400). This means that there is significant heterogeneity in buyer’s tastes in the housing market. Recall from the theoretical model that $b$ has two independent

---

30 This is consistent with Glower et al. (1998) who find that more motivated agents post cheaper prices and sell their homes at lower prices. Our results, however, 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 The per-search time cost is the value of the time spent in one visit. Based on our interview with the Charlottesville and Albemarle Real Estate Association’s Director and on Shimizu et. al (2003), we postulate that the visiting time that the average household uses in one search adds up to 9 hours: a) one hour devoted to traveling time; b) two hours for the visit (on average, buyers visit homes-for-sale in pairs for approximately one hour per visit); and c) six hours for a post-search consideration period. As in Shimizu et al.(2003), we let the buyer’s wage rate be the per-hour opportunity cost. Assuming that the median visitor has an hourly income of $19.2 per hour (area’s yearly median household income of $55,500 divided by 2,880, the total number of hours per year worked by this median household), the related buyer’s time-opportunity costs add up to approximately $173. To estimate the buyer’s transportation costs, we have assumed that the average (one way) transportation time in the area is 20 minutes and the corresponding taxi-cab rate is approximately $20. Thus, the per-visit transportation cost is $40.
components: a) $b^\gamma$, the 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 listings provide to buyers, we estimate $\lambda = \frac{\sigma^2_\gamma}{\sigma^2_\gamma + \sigma^2_u}$ and $\kappa = \frac{V(\delta + b^\gamma)}{V(\delta + b^\gamma + b^u)}$.

As was previously discussed, $\lambda$ measures how much information about the idiosyncratic buyer-home match listings provide to real estate customers. Its estimate is 0.027 suggesting that, on average, about 3% of the information about the buyer-home match value is obtained through the listing. To a certain extent, the low value of $\lambda$ is surprising. However, notice that many of the current online advertising features, such as multiple pictures and virtual tours, were not available during the first years of our sample. Aside from revealing part of the buyer-home match, listings also provide information about the units’ characteristics ($X^0$ and $u$). Thus, to measure how much overall information listings reveal to buyers, we compute $\kappa$, the variance of the total buyer’s home valuation obtained when the listing is viewed relative to the variance of the total buyer’s valuation. Using our parameter estimates we find that $\kappa = 0.55$ which suggests that more than half of the buyer’s home valuation is obtained at the time when a listing is viewed. The rest is revealed when the housing unit is visited. This finding is consistent with the observation that most home buyers visit a unit before a purchase is made.

As we stated in the introduction, the amount of information displayed in online ads has increased over time, meaning that the value of $\lambda$ should have increased accordingly. To test this hypothesis, we divide our sample in three groups that correspond to the year when the transaction was made. Then, for each year, we estimate $\lambda$ while keeping constant the rest of the previously estimated parameters. The results of this exercise are displayed in Table 6 and provide evidence that the importance of online information has risen over time.

The estimates of $\gamma$ ($s = X\gamma$) are shown in the first column of Table 5. These coefficients represent the marginal contribution of a home characteristic to the value of the home. For instance, our estimates suggest that one additional square foot increases the value of a home by about $70 while an extra bathroom raises it by $2,160. 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 $13,450. An interesting finding is that, after conditioning on square footage, one extra bedroom decreases the value of a home by about $9,200, suggesting that agents prefer
homes with larger bedrooms.\textsuperscript{32} Because we expect the age of the property affects its value in a nonlinear way, we added quadratic and cubic terms to our specification. The coefficients on age suggest that the value of a new property declines rapidly at a decreasing rate for the first 35 years and slowly increases thereafter. In particular, the model predicts that a new property is $48,000, $59,000, and $0 more valuable than a ten, twenty and 230-year-old property, respectively. Having air conditioning and central heat adds approximately $8,600 and $19,000 to the value of a home, respectively.

All coefficients on the neighborhood variables have the expected signs. Agents value more homes that are located in high-population or high-income areas. As the mean age of the people in the neighborhood increases, so does its intrinsic valuation; the opposite is true for household size. However, it is somewhat difficult 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.

It is useful to compare our structural estimates with coefficients from a hedonic price regression. The second and third column of Table 5 show estimates of a linear regression where the dependent variable is the posting price and the transaction price, respectively. Explanatory variables include the same set of variables used in our structural model.\textsuperscript{33} The $R^2$ of both regressions is roughly 0.70, which is a typical level for housing price models (Mason and Quigley 1996). The coefficients of the hedonic pricing equations measure the implicit price of each housing attribute. Thus, one would expect the hedonic and structural coefficients to be somewhat similar. As expected, with the exception of the constant term, all coefficients from the hedonic models have the same sign and are close in magnitude to the structural parameters. Finally, notice that, as expected, the differences between the OLS errors’ standard deviations and our estimate of $\sigma_u$ are very small (less than 2%).

\textsuperscript{32}A similar finding was reported by Meese and Wallace (1997).

\textsuperscript{33}Notice that we have specified linear regression 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, we have estimated log-linear pricing models and found that the predictive power of the latter models was slightly higher (the mean square error decreases by about 2%). Nonetheless, the linear specification was chosen to facilitate comparison with the structural estimates. Similarly, notice that we have not included time variant independent variables in the hedonic specification such as year or month dummy variables. We omit these variables to be consistent with the stationary nature of our theoretical model.
The estimated model also provides insights about the relationship between list prices and time on the market. Previous studies have used real estate transaction data to measure the effect of list prices on time on the market (Kang and Gardner 1989, Yavas and Yang 1995, Knight 2002, and Anglin, Rutherford and Springer 2003, among others). The common reduced-form empirical approach is intuitive and straightforward. In a duration model, an explanatory variable that measures the seller’s “markup,” the (percentage) difference between the posting price and the true value of the home, is included. Since the true value of the unit is unobserved, it is usually replaced by the expected price estimated using a hedonic equation.\(^{34}\) The coefficient on the markup variable estimates the effects of misspricing on time on the market. While this method is useful to control for seller’s strategic behavior in duration models, it may underestimate the effect of overpricing on marketing time. For instance, it is likely that the unexplained residual in a hedonic equation is negatively correlated with the unexplained portion of a time-on-the-market model. That is, a home that has desirable “unobserved” features may sell at a higher price and, other things equal, faster. Since list prices and transaction prices are highly correlated, the markup variable may be negatively correlated with the error term of the duration equation as well. Thus, the coefficient on the markup variable could have a negative bias understating the effects of overpricing. The structural model we estimate solves this problem by explicitly modeling and controlling for unobserved housing heterogeneity.

We use the estimates of the model to compute the effects of changing the list price on time on the market. To perform this exercise, we pick a representative unit with a posting price of $200,000 that expects to be sold in 58 days. We vary the posting price and use the structural model to calculate the expected marketing time. Results are shown in Figure 2 and compared with a similar exercise where the conventional reduced-form approach is used.\(^{35}\) Our structural estimates suggest that there is a substantially large effect of overpricing on

\(^{34}\)This is the approach suggested by Yavas and Yang (1995) where the expected price (at the time the unit is listed) is computed with an auxiliary hedonic regression. In other applications, the true value of the home is proxied by the actual transaction price.

\(^{35}\)To compute the reduced-form model, we estimate a log-linear regression where the dependent variable is the log of marketing time. Independent variables include housing characteristics as well as the percentage difference between the list price and the expected transaction price. Expected transaction prices are estimated using an auxiliary hedonic regression. Given space constraints we do not report coefficients but these may be requested from the author.
marketing time. This effect is non-linear and increases exponentially as the markup rises. For instance, if the list price raises by 1%, the expected time on the market rises by about 7 days; if the list price increases by 5%, however, marketing time is expected to increase by approximately 370 days. As expected, the reduced-form estimates are substantially lower. This confirms that failure to control for unobserved heterogeneity can lead to misleading estimates of the effect of list prices on marketing time.

6.2 Goodness of fit

Within our sample, we use the structural model and simulation methods to predict posting and transaction prices. Overall the model fits the pricing data reasonably well. The mean posting price in the data was $200,600, and our model’s predictions are only 2.6%, or $5,200, higher. Furthermore, the model predicts that the mean transaction price is approximately $202,100, which is only about 2.9% higher than the mean transaction price of our sample data.

To have a measure of how well our model fits the observed data, we construct pseudo-$\hat{R}^2$ indexes. This indexes are computed in a similar way as the $R^2$ of an OLS regression. For example, let $E[p_{si}^*|X_i]$ be the expected posting price predicted by our model of a property with observed characteristics $X_i$, and let $p_{si}$ be the observed posting price. Then, the corresponding pseudo-$\hat{R}^2$ index is defined as follows

$$\hat{R}^2 = 1 - \frac{\sum_{i=1}^{n} (p_{si} - E[p_{si}^*|X_i])^2}{\sum_{i=1}^{n} (p_{si} - \bar{p}_{si})^2},$$

where $\bar{p}_{si}$ is the posting price’s sample mean, and $n$ is the sample size.

For both the posting price and the transaction price, the pseudo-$R^2$s were about 0.69, which is slightly below the $R^2$s 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 3 and 4 we present the densities of the actual versus the predicted posting and transaction prices, respectively. In both cases, the model displays a reasonably good representation of our sample data which is not a feature of most empirical equilibrium search
models in the literature (see, for example, Eckstein and Wolpin 1990). It seems, however, that our model is unable to replicate the skewness of the actual pricing data. We may be predicting a distribution of prices that is too symmetric around the mean for one particular reason. We have assumed that all the exogenous distributions in our model are directly related to the normal distribution, and this assumption is imposing a certain degree of symmetry in the distribution of prices. These normality assumptions were taken for computational convenience and could be relaxed by assuming a more flexible functional form of the errors’ densities.

As in Horowitz (1992), our model does not perform well when predicting the duration data. Figure 5 compares the density of the duration that a house stays on the market predicted by the model with the density of the observed data. Our model overpredicts both the median and the mean duration and inflates the right tail of the simulated distribution by overpredicting the number of properties that stay in the market for more than 200 days.36 There may be several sources of misspecification that may lead the model to inflate the average and median time on the market. We discuss these as well as other limitations on the last section of the paper.

6.3 Counterfactual experiments

One of the primary benefits of estimating a structural model is that it may be used to conduct counterfactual experiments. In this section, we use the estimated model to evaluate how equilibrium market outcomes change when a) the amount of information displayed in listings increases and b) agent’s commission rates vary.

The first experiment raises the amount of the relevant information that buyers obtain when looking at an online listing, and it can be described in terms of one parameter change. That is, we let \( \lambda \) 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. The results of this exercise are displayed in Table 7 and suggest that additional information decreases market

\[ \hat{R}^2 \]

\[ -0.18 \]

\[ \text{pseudo-R}^2 \]

\[ -0.18 \]

The median and the mean of the observed duration data is 26 and 43 days, respectively. Our model overpredicts both of these variables. According to the model’s predictions, the median and the mean time that a property stays on the market is 73 days and 165 days, respectively. Furthermore, our estimate of the pseudo-\( R^2 \) is negative (-0.18), meaning that, on average, our model does even worst than the sample mean.
prices. For instance, if buyers could obtain from a listing all the relevant information they need to make a purchase decision ($\lambda = 1$), posting and transaction prices would decrease by 4.0% and 4.4%, respectively. We interpret this findings as evidence that the availability of listings in the internet may increase market competition which is consistent with other findings in the literature.\footnote{See, for example, Brown and Goolsbee (2002).} We also find evidence that, as more information is displayed in real estate listings, the time that a property stays on the market increases. In particular, when $\lambda$ equals one, the average time on the market increases in seven days compared to our benchmark case ($\lambda = 0.03$). Presumably, as information on the listing 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).

In the second counterfactual, we let commission rates vary from 0% to 10% and evaluate how these changes affect housing market outcomes. Results are shown in Table 8. They suggest that the costs of real estate agents are about equally shared by both buyers and sellers despite the fact that the full rate is paid by the seller at the time of the transaction. For example, if rates go down from 6% to 4% and to 2%, our model predicts that transaction prices would decrease by about 2% and 1%, respectively. The model also predicts that as commission rates increase, time on the market diminishes. For instance, when rates go up from 6% to 10%, average time on the market decreases by about 6 days. Presumably, as commission rates and prices increase, buyers’ value of search decreases, they become less selective and choose to trade faster. When rates decrease from 6% to smaller levels, however, changes on time on the market are almost negligible.

7 Conclusions and avenues for future research

In this research we have specified and estimated a computationally tractable stationary equilibrium model of the housing market. This work contributes to the existing literature in several ways. From a theoretical perspective, the model explains posting prices, transaction prices, and time on the market in an integrated equilibrium context. From an empirical point of view, it shows that a stationary equilibrium search model can provide a reasonable
Parameter estimates and counterfactual analysis provide interesting insights about the housing market. Parameter estimates suggest that more than half of the buyer’s home valuation is gathered at the time when a listing is viewed. Conditional on the housing unit’s characteristics, however, about only 3% of the idiosyncratic buyer-home match is obtained through the information provided by the listing. The rest of the match is revealed when the property is visited. As more information is available on the listing and buyers are able to learn a greater share of this buyer-home match without having to visit the unit, our model predicts that posting and transaction prices would decrease. To a certain extent, this is not a surprising prediction because it is expected that as search costs decrease, markets would become more competitive. On the other hand, we find evidence that buyers would also become more selective and there would be a small increase in marketing time. Because most of the real estate listings are posted on the internet, our experiment implicitly evaluates the effects of improvements in the amount of information displayed in online advertisements on market outcomes. Hence, our findings may be of interest to the literature that seeks to explain how the use of the internet has influenced online and offline markets. A second counterfactual assesses the effects of changes in real estate agents’ commission rates on marketing outcomes. We find that commission costs are about equally shared by buyers and sellers despite the fact that sellers pay all of these costs. While reducing commission rates would decrease list and transaction prices, the effects on duration are very small.

Our modeling choices have several shortcomings. Perhaps the main limitation of our model is its stationary nature. In a stationary model such as ours, sellers’ posting and reservation prices as well as buyers’ optimal search strategies do not change with time. These implications are most likely rejected by the data in most housing markets. Presumably, sellers drop list and reservation prices and buyers increase their search efforts as a moving date approaches or when the listing-contract period is about to expire. Similarly, the fact that some sellers in the data withdraw their properties from the market cannot be accounted for in our model. The stationarity assumption (along with other assumptions about the matching technology) also implies that the time a property stays on the market follows a geometric distribution and that it has no duration dependence. Thus, the time-invariant approximation to the data.
nature of our framework precludes our model from picking the negative correlation between prices and marketing time that is generally observed in housing sales data. These restrictions may explain why our model is not able to fit marketing time data very well. Finally, in a stationary framework it is not possible to characterize any dynamic feature of the housing market. In particular, we are not able to describe temporary market imbalances, that is, periods when the number of buyers exceeds the number of sellers or vice versa. These imbalances probably explain a significant portion of changes in housing prices over time. Our stationary model may predict outcomes in markets where home buyers are, on average, more motivated to trade than sellers and vice versa. This exercise provides some (indirect) insights about how the ratio of the "number" of buyers per seller affects equilibrium but it cannot explain any dynamic issues.

While we acknowledge that the stationary nature of our model imposes important restrictions, we think it provides a useful benchmark for future research. Without too much effort, some of the stationary assumptions may be relaxed in a partial-equilibrium search framework. For example, optimal list prices and reservation values would change with time if there is a particular date when sellers need to move or when the listing-contract period is about to expire.\textsuperscript{38} Similarly, buyers’ value of search and optimal strategies would change when market conditions vary or as a particular moving-in date approaches. It is not clear to us, however, how equilibrium could be characterized (or if it even exists) in such circumstances. Extensions to the model that allow buyers and sellers to change their optimal behavior with time are likely to improve the overall fit of the model to the data (in particular, with respect to time on the market).

To provide a computationally tractable solution, we made several simplifying assumptions about the search behavior of home buyers and sellers. Some of these assumptions may be too restrictive. In particular, assumptions about sequential search, search without recall, and the independence between buyers’ search-intensity and their motivation to trade may not fully reflect the reality of this market. Similarly, the assumption that buyers observe \textit{all} characteristics of a housing unit at the time a listing is viewed may be unappealing for some readers. The implications of these assumptions are also likely to be rejected by

\textsuperscript{38}A similar approach is taken by Merlo, Ortalo-Magne and Rust (2006) and Salant (1991).
the data. For instance, it is clear that buyers sample and consider more than one listing at a time, that motivated buyers are more eager to search than less motivated ones, and that buyers learn new information about the unit’s features when the visit occurs. Some of these assumptions may be relaxed at a relatively low cost. For instance, our model could be easily modified to allow buyers to review multiple listings within a period that may be recalled in the future. This is a potentially useful extension that, combined with detailed data about the buyer’s search process (such as number of visits, number of listings viewed, and buyer’s demographics), could significantly improve the characterization of the buyer’s search mechanism. In our view, other extensions are not straightforward. For example, one could let buyers choose how many listings to pick and how many properties to visit (using Stern’s 1989 framework, for example). One could also assume that agents observe only a subset of the unit’s characteristics at the time they look at an online listing and that the rest of its characteristics are revealed when the visit is made. Then contents of the listing become endogenous (chosen by sellers) and based on what the listing reveals, buyers should update their beliefs about the value of the home and about the seller’s motivation to trade. These non-trivial extensions could be explored in future research.

For simplification and identification purposes, we introduce an innovative yet restrictive bargaining model. The bargaining model predicts that transaction prices occur either at the seller’s posting price or reservation value and that buyers make at most one counteroffer to sellers. These implications are likely rejected by the data because some properties do sell at prices above the list price and several rejected offers made by unsuccessful buyers are likely to occur before a seller accepts an offer (Merlo and Ortalo-Magne 2004). This contradicts our assumption that buyers have perfect information about sellers’ preferences. Besides computational constraints, we are forced to use these assumptions because we lack data about buyers’ offers. When appropriate data are available, explaining the dynamic process of how buyers and sellers bargain in the housing market can be quite challenging. Merlo, Ortalo-Magne and Rust (2006) show how this can be done using a rich dataset and stochastic dynamic programming.

Despite high commission rates, the majority of transactions in the U.S. housing market are facilitated by real estate agents. Revealed preference suggests then that agents increase
the average gains from trade by a large amount. How is this achieved? Presumably, agents increase the rate at which potential buyers visit home sellers and also improve the quality of the matches. Our model cannot shed light on these interesting issues because it ignores the role of agents as middlemen. Agents’ strategic behavior and agents’ heterogeneity are omitted while commission rates are exogenously set. Given our data (all transactions in our sample were facilitated by real estate agents) and the model’s complexity, we are forced to make these simplifications. As demonstrated by Rust and Hall (2003) and Spence (1996), however, the strategic behavior of middlemen in a market with two-sided search can be characterized. How do these middlemen “search” in the housing market? How do they set list prices? These are interesting questions that deserve further research.

Finally, rather than working with a finite number of buyers and sellers, we assume there is a continuum of them. This assumption is convenient since it allows us to find some analytical answers to the partial equilibrium search problems. However, it is not realistic and may be too restrictive to describe an actual housing market where there is only a finite number of buyers and sellers. Modeling a market with a finite number of agents in a non-stationary environment could potentially be very useful to analyze the effects of temporary market imbalances. When markets are finite, however, the entire formulation of agent’s search strategies has to be modified to account for the fact that sampling occurs without replacement (presumably, buyers do not draw the same property more than one time). This is not an easy task and imposes a significant computational burden.

In sum, this research provides a theoretical and empirical contribution to the literature and gives interesting insights about the housing market. As it is the case with most economic models, some of our assumptions fail to conform reality. Nonetheless, we think it provides a reasonably good approximation to the data and hope it serves a benchmark for future studies.
8 Appendix

Proof of Proposition 1: Substituting (5) into (4) and rearranging, we obtain the following expression:

\[ h(p^*_s s)(p^*_s - \frac{s + T}{1 - k}) = \frac{1 - \beta_s [1 - \theta (1 - \phi (p^*_s s))]}{1 - \beta_s} \]  

(19)

Furthermore, assume that the hazard function is non-decreasing in \( p^*_s \). Note that the left hand side of equation (19) evaluated at \( p^*_s = \frac{s + T}{1 - k} \) is zero, and that it is increasing in \( p^*_s \); the right hand side of equation (19) evaluated at \( p^*_s = \frac{s + T}{1 - k} \) 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 \( \beta_s \) and \( s \)). Using equation (19) and the implicit function theorem, we take the derivative of \( p^*_s \) w.r.t. \( \beta_s \) and obtain

\[ \frac{\partial p^*_s}{\partial \beta_s} = \frac{\theta (1 - \phi)}{(1 - \beta_s)^2} \frac{\partial h}{\partial p^*_s} + h + \frac{\beta_s}{1 - \beta_s} \theta \frac{\partial \phi}{\partial p^*_s} > 0. \]

It is also easy to see from equation (4) that \( \frac{\partial R^*_s}{\partial \beta_s} = 1 + \frac{\partial h}{\partial p^*_s} > 0 \). 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 (5) and find that \( R^*_s = \frac{s + T}{1 - k} \). Because \( 0 \leq \beta_s \leq 1 \) and \( \frac{\partial R^*_s}{\partial \beta_s} \geq 0 \), it must be the case that \( p^*_s \geq R^*_s \geq \frac{s + T}{1 - k} \).

To prove Proposition 1, we still need to show that \( p^*_s \) and \( R^*_s \) are increasing in \( s \). To show this, we use equation (19) and the implicit function theorem to find

\[ \frac{\partial p^*_s}{\partial s} = \frac{h}{1 - k} \frac{\partial h}{\partial \beta_s} \frac{\partial \phi}{\partial p^*_s} + h + \frac{\beta_s}{1 - \beta_s} \theta \frac{\partial \phi}{\partial p^*_s} > 0. \]

Notice that the expression in parenthesis in the numerator is always negative, because both \( \frac{\partial \phi}{\partial s} \) and \( \frac{\partial h}{\partial \beta_s} \) are negative. Furthermore, \( \frac{\partial R^*_s}{\partial \beta_s} = \frac{\partial R^*_s}{\partial p^*_s} \frac{\partial p^*_s}{\partial \beta_s} > 0 \). QED.

Proof of Proposition 2

To prove Proposition 2, we need to show the details of integration of equation (7), and the monotonicity of \( W^*_b \) and \( p^*_l \).

1) Details of the integration of equation (7)
Let \( v_p = b^o + \delta s - p^*_s \) and integrate by parts the first term of (7)

\[
E[\max\{b-p^*_s, W_b\}|p^*_s, s, b^o] = \int max\{b^u + b^o + \delta s - p^*_s, W_b\}dG_u(b^u) = \int_{W_b-v_p}^{} (b^u + v_p)dG_u(b^u) + W_bG_u(W_b - v_p) = W_b[1 - G_u(W_b - v_p)] + \int_{W_b-v_p}^{} [1 - G_u(b^u)]db^u + W_bG_u(W_b - v_p)
\]

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 \( p^r_b^* \)

Define \( g(W^*_b) = \int \int \max\{j(p^*_s, s, b^o, W^*_b), 0\} \ d\Gamma(p^*_s, s)dG_o(b^o) \) and rewrite equation (9) 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

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

Furthermore, we use equation (10) -and Leibnitz rule- to show that

\[
\frac{\partial p^r_b^*}{\partial W_b^*} = -\frac{\theta(1 - G_u(V_p)) + (1 - \theta)(1 - G_u(V_r))}{\theta(1 - G_u(V_p)) + (1 - \theta)\frac{\partial R^*_s}{\partial p^*_b}(1 - G_u(V_r))} < 0,
\]

where \( V_p = W_b + p^r_b^* - \delta s - b^o \), and \( V_r = W_b + R^*_s(p^r_b^*|s) - \delta s - b^o \). Thus, \( \frac{\partial p^r_b^*}{\partial \beta_b} = \frac{\partial p^r_b^*}{\partial W_b^*} \frac{\partial W_b^*}{\partial \beta_b} < 0 \).

Steps one and two complete the proof. QED.

Proof of Proposition 3:
First, note that because \( s \) is non random and \( b^o = 0 \), equation (6) becomes

\[
W_b = \beta_b \int \max \{ U^e(p_s^*), W_b \} d\Gamma(p_s^*)
\]

\[
\frac{W_b}{\beta_b} = \int_{p_s^* : \{ U^e(p_s^*) > W_b \}} U^e(p_s^*) d\Gamma(p_s^*) + W_b \int_{p_s^* : \{ U^e(p_s^*) < W_b \}} d\Gamma(p_s^*)
\]

\[
= \int_{p_s^* : \{ p_s^* < U^{e^{-1}}(W_b) \}} U^e(p_s^*) d\Gamma(p_s^*) + W_b \int_{p_s^* : \{ p_s^* > U^{e^{-1}}(W_b) \}} d\Gamma(p_s^*)
\]

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

\[
\frac{W_b}{\beta_b} = \int_{U^{e^{-1}}(W_b)} U^e(p_s^*) d\Gamma(p_s^*) + W_b \int_{U^{e^{-1}}(W_b)} d\Gamma(p_s^*)
\]

\[
= W_b \Gamma(U^{e^{-1}}(W_b)) + \int_{U^{e^{-1}}(W_b)} \Gamma(p_s^*)(-\frac{\partial U^e(p_s^*)}{\partial p}) dp_s + W_b[1 - \Gamma(U^{e^{-1}}(W_b))]
\]

\[
= W_b + \int_{U^{e^{-1}}(W_b)} \Gamma(p_s^*)(-\frac{\partial U^e(p_s^*)}{\partial p}) dp_s.
\]

It is optimal for buyers to set a reservation strategy such that \( U^e(p_b^*) = W_b^* \). Then, we replace this optimality condition in the former equation and obtain

\[
\frac{\beta_b}{1-\beta_b} \int_{p_b^*} \Gamma(p_s^*)(-\frac{\partial U^e(p_s^*)}{\partial p}) dp_s = W_b^*.
\]

Finally, we let \( U^e(p_b^*) = W_b^* \) in equation (8) to find

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

Thus, the solution to the search problem is equivalent to finding the pair \( \{ W_b, p_b^* \} \) that solves the previous equation system. QED.

**Additional insights about** \( q \) **and** \( \gamma_s * \).

The function \( q(p_s|s, b^o) \) is well defined by the exogenous distribution \( K_b \) and the optimal buyers reservation price function \( p_b^{*} \).

\[
q(p_s|s, b^o) = \text{Pr}\{ p_b^{*}((\beta_b, s, b^o) > p_s|s, b^o) = K_b(p_b^{*}-1(p_s|s, b^o)) \}.
\]
where \( p^*_b(p_s) \) is the function that determines \( \beta_b \) for any value \( p_s \) (which exists because we have shown that \( p^*_b(\beta_b) \) is monotone). Then

\[
\gamma_s(p_s|s,b^o) = \Pr\{b^o + b^u + \delta s - p_s > W_b^*(\beta_b)|p^*_b(\beta_b, s, b^o) > p_s\} = \frac{\int_{p^*_b(p_s|s,b^o)} (1 - G_u[W_b^*(\beta_b) + p_s - \delta s - b^o]) dK_b(\beta_b)}{K_b(p^*_b(p_s|s,b^o))},
\]

and

\[
\phi^*(p_s|s) = 1 - \int q(p_s|s,b^o) \gamma_s(p_s|s,b^o) dG_o(b^o) = 1 - \int \int (1 - G_u[W_b^*(\beta_b) + p_s - \delta s]) dK_b(\beta_b) dG_o(b^o) \quad (20)
\]

which is clearly non-increasing in \( p_s \).

References


### Table 1
Equilibrium Convergence of Function $\phi$

<table>
<thead>
<tr>
<th>Starting value</th>
<th>Number of iterations before convergence</th>
<th>Absolute differences in $\phi$ compared to baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Average</td>
</tr>
<tr>
<td>Baseline (Exponential, $\alpha=0.08$)</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>Exponential</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha=0.05$</td>
<td>67</td>
<td>1.51E-04</td>
</tr>
<tr>
<td>$\alpha=0.40$</td>
<td>88</td>
<td>1.02E-04</td>
</tr>
<tr>
<td>Linear</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\eta=0.80$</td>
<td>49</td>
<td>1.17E-04</td>
</tr>
<tr>
<td>$\eta=0.95$</td>
<td>60</td>
<td>1.07E-04</td>
</tr>
<tr>
<td>Normal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau=0.20$</td>
<td>63</td>
<td>1.82E-04</td>
</tr>
<tr>
<td>$\tau=0.40$</td>
<td>68</td>
<td>1.18E-04</td>
</tr>
</tbody>
</table>
Table 2  
Equilibrium Comparative Statics  
Changes in Selected Parameters

<table>
<thead>
<tr>
<th>Model specification</th>
<th>Mean posting price ($ thousands)</th>
<th>Mean transaction price ($)</th>
<th>Mean days on the market</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline parameters</td>
<td>205.8</td>
<td>202.1</td>
<td>64.9</td>
</tr>
<tr>
<td>Changes in parameter values</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>Alternative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda = 0.027$</td>
<td>$\lambda = 0.50$</td>
<td>200.1</td>
<td>195.9</td>
</tr>
<tr>
<td>$\sigma_b = 81.4$</td>
<td>$\sigma_b = 120.0$</td>
<td>243.1</td>
<td>235.9</td>
</tr>
<tr>
<td>$c = 0.506$</td>
<td>$c = 0.20$</td>
<td>307.7</td>
<td>280.43</td>
</tr>
<tr>
<td>$\delta = 1.08$</td>
<td>$\delta = 1.04$</td>
<td>204.1</td>
<td>200.7</td>
</tr>
</tbody>
</table>
Table 3
Descriptive Statistics of Completed Real Estate Transactions in
Charlottesville, VA 2000-2002
Number of observations: 2,876

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>St. Dev.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posting price ($ thousands)</td>
<td>200.6</td>
<td>181.3</td>
<td>85.5</td>
<td>52</td>
<td>495</td>
</tr>
<tr>
<td>Transaction price ($ thousands)</td>
<td>196.4</td>
<td>179.1</td>
<td>83.1</td>
<td>50</td>
<td>449.3</td>
</tr>
<tr>
<td>Days on the market</td>
<td>43.3</td>
<td>26</td>
<td>46.1</td>
<td>1</td>
<td>199</td>
</tr>
<tr>
<td>Equals one if Ps &gt; Pm</td>
<td>0.579</td>
<td>1</td>
<td>0.494</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

*Home characteristics*

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>St. Dev.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finished square footage</td>
<td>1980.2</td>
<td>1910.1</td>
<td>701.9</td>
<td>261</td>
<td>6500</td>
</tr>
<tr>
<td>Number of full bathrooms</td>
<td>2.03</td>
<td>2</td>
<td>0.67</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Number of bedrooms</td>
<td>3.35</td>
<td>3</td>
<td>0.8</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Acres</td>
<td>0.99</td>
<td>0.26</td>
<td>2.34</td>
<td>0</td>
<td>36.2</td>
</tr>
<tr>
<td>Age of property</td>
<td>23.75</td>
<td>17</td>
<td>23.45</td>
<td>0</td>
<td>251</td>
</tr>
<tr>
<td>Air conditioning</td>
<td>0.9</td>
<td>1</td>
<td>0.31</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Basement</td>
<td>0.55</td>
<td>1</td>
<td>0.5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Central heat</td>
<td>0.97</td>
<td>1</td>
<td>0.18</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Sewer system</td>
<td>0.75</td>
<td>1</td>
<td>0.43</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Home Owner Association</td>
<td>0.51</td>
<td>1</td>
<td>0.5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Pool</td>
<td>0.03</td>
<td>0</td>
<td>0.16</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>One story only</td>
<td>0.31</td>
<td>0</td>
<td>0.46</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Detached</td>
<td>0.79</td>
<td>1</td>
<td>0.41</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

*Neighborhood characteristics*

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>St. Dev.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log of Population per square mile in census block</td>
<td>7.02</td>
<td>7.5</td>
<td>1.54</td>
<td>1.42</td>
<td>10.17</td>
</tr>
<tr>
<td>Proportion black in census block</td>
<td>0.1</td>
<td>0.05</td>
<td>0.14</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Median age in census block</td>
<td>38.51</td>
<td>37.7</td>
<td>7.48</td>
<td>18.5</td>
<td>80.1</td>
</tr>
<tr>
<td>Mean household size in census block</td>
<td>2.52</td>
<td>2.57</td>
<td>0.43</td>
<td>1</td>
<td>4.5</td>
</tr>
<tr>
<td>Median household income in census block group</td>
<td>55.5</td>
<td>53.9</td>
<td>16.8</td>
<td>15.6</td>
<td>116.7</td>
</tr>
</tbody>
</table>
Table 4
Parameter Estimates Baseline Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Baseline model's parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>θ</td>
<td>0.465 (0.009) *</td>
</tr>
<tr>
<td>μ_β</td>
<td>8.580 (0.012) *</td>
</tr>
<tr>
<td>σ_β</td>
<td>1.125 (0.003) *</td>
</tr>
<tr>
<td>cb</td>
<td>0.5059 (0.007) *</td>
</tr>
<tr>
<td>δ</td>
<td>1.084 (0.001) *</td>
</tr>
<tr>
<td>σ_μ</td>
<td>45.01 (0.350) *</td>
</tr>
<tr>
<td>σ_b</td>
<td>81.36 (0.799) *</td>
</tr>
<tr>
<td>λ</td>
<td>0.027 (0.005) *</td>
</tr>
</tbody>
</table>

Number of observations 2876

Standard errors are in parentheses. Asterisks indicate those parameters significant at the 5 percent significance level.
### Table 5
Structural Estimates Seller Home Value

<table>
<thead>
<tr>
<th>Structural model</th>
<th>Seller's home value</th>
<th>Linear OLS Models</th>
<th>Dependent variable:</th>
<th>Posting price</th>
<th>Transaction price</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>s = X * γ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-60.12 (2.76) *</td>
<td>-15.12 (15.74) *</td>
<td>-17.01 (15.03) *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finished square footage</td>
<td>0.069 (0.00) *</td>
<td>0.068 (0.00) *</td>
<td>0.065 (0.00) *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of full bathrooms</td>
<td>2.16 (1.57)</td>
<td>5.20 (2.40) *</td>
<td>4.36 (2.32) *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of bedrooms</td>
<td>-9.23 (0.99) *</td>
<td>-5.74 (2.08) *</td>
<td>-5.16 (2.05) *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log of (acreage +1)</td>
<td>8.54 (0.81) *</td>
<td>7.25 (1.01) *</td>
<td>6.71 (0.97) *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log of age of the property (zero if new)</td>
<td>-11.36 (0.50) *</td>
<td>-1.61 (6.33)</td>
<td>-5.48 (6.14) *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log of age of the property ^2</td>
<td>-8.79 (0.75) *</td>
<td>-10.90 (3.70) *</td>
<td>-9.25 (3.57) *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log of age of the property ^3</td>
<td>2.00 (0.16) *</td>
<td>2.19 (0.58) *</td>
<td>1.96 (0.56) *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Air conditioning</td>
<td>8.54 (2.83) *</td>
<td>11.96 (4.10) *</td>
<td>12.93 (4.00) *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basement</td>
<td>1.35 (1.97)</td>
<td>5.40 (2.21) *</td>
<td>5.38 (2.16) *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Central heat</td>
<td>18.67 (2.98) *</td>
<td>12.63 (6.38) *</td>
<td>13.50 (5.88) *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sewer system</td>
<td>-4.94 (2.63)</td>
<td>-4.00 (3.72)</td>
<td>-4.13 (3.63)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home Owner Association</td>
<td>11.27 (2.41) *</td>
<td>14.45 (2.62) *</td>
<td>15.01 (2.53) *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pool</td>
<td>12.41 (4.70) *</td>
<td>15.64 (7.44) *</td>
<td>15.05 (7.30) *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>One story only</td>
<td>-10.66 (0.23) *</td>
<td>-9.63 (2.41) *</td>
<td>-9.64 (2.31) *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detached</td>
<td>36.99 (2.86) *</td>
<td>42.18 (2.75) *</td>
<td>41.88 (2.66) *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log of Population per square mile in census block</td>
<td>3.28 (0.07) *</td>
<td>3.06 (1.02) *</td>
<td>3.45 (0.99) *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion of blacks living in the census block</td>
<td>-60.06 (4.78) *</td>
<td>-57.89 (7.47) *</td>
<td>-57.77 (7.27) *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median age in the census block</td>
<td>0.88 (0.05) *</td>
<td>1.00 (0.19) *</td>
<td>0.97 (0.18) *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean household size in the census block</td>
<td>-14.21 (0.14) *</td>
<td>-6.00 (2.90) *</td>
<td>-4.92 (2.78)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median household income in census block group</td>
<td>0.44 (0.04) *</td>
<td>0.49 (0.08) *</td>
<td>0.47 (0.08) *</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R2: 0.705 0.707
Standard error of regression: 46.6 45.2
Number of observations: 2876 2876

Standard errors are in parentheses. Asterisks indicate those parameters significant at the 5 percent significance level.
### Table 6

**Estimates of $\lambda$ Across Time**

<table>
<thead>
<tr>
<th>Year</th>
<th>$\lambda$</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>0.018</td>
<td>(0.016)</td>
</tr>
<tr>
<td>2001</td>
<td>0.027</td>
<td>(0.012) *</td>
</tr>
<tr>
<td>2002</td>
<td>0.035</td>
<td>(0.006) *</td>
</tr>
</tbody>
</table>

Standard errors are in parentheses. Asterisks indicate those parameters significant at the 5 percent significance level.
Table 7
Counterfactual Analysis
Effects of an Increase in the Amount of Information on the Listing

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\lambda = 0.027$</th>
<th>$\lambda = 0.50$</th>
<th>$\lambda = 1.00$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean posting price ($\text{th}\text{ousands}$)</td>
<td>205.8</td>
<td>200.1</td>
<td>197.5</td>
</tr>
<tr>
<td>Mean transaction price ($\text{th}\text{ousands}$)</td>
<td>202.1</td>
<td>195.9</td>
<td>193.1</td>
</tr>
<tr>
<td>Mean days on the market</td>
<td>64.9</td>
<td>68.2</td>
<td>71.8</td>
</tr>
</tbody>
</table>
Table 8  
Counterfactual Analysis  
Effects of a Change in Agent's Commission Rates

<table>
<thead>
<tr>
<th>Variable</th>
<th>0%</th>
<th>2%</th>
<th>4%</th>
<th>6%</th>
<th>8%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean posting price ($ thousands)</td>
<td>199.7</td>
<td>201.8</td>
<td>203.8</td>
<td>205.8</td>
<td>207.6</td>
<td>209.2</td>
</tr>
<tr>
<td>Mean transaction price ($ thousands)</td>
<td>195.9</td>
<td>198.1</td>
<td>200.1</td>
<td>202.1</td>
<td>203.9</td>
<td>205.1</td>
</tr>
<tr>
<td>Mean days on the market</td>
<td>66.1</td>
<td>66.6</td>
<td>67.1</td>
<td>64.9</td>
<td>64.0</td>
<td>58.6</td>
</tr>
</tbody>
</table>
Figure 1
Reservation Value and Posting Price as a Function of the Discount Factor
(Evaluated at the Median Home Quality)

Figure 2
Posting Price and Marketing Time
Both densities have been evaluated at 100 equally-spaced points in the range of the data using a normal kernel function.
Figure 5

Nonparametric Density of Actual and Predicted Marketing Time

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