Principal-Agent Problems in Search and Bargaining: Evidence from Dual Agency in Residential Real Estate Transactions*

Seung-Hyun Hong
Department of Economics
University of Illinois, Urbana-Champaign
hyunhong@illinois.edu

April 8, 2016

Abstract

In this paper, we quantify the extent of conflicts of interest between home sellers and listing agents, and examine the importance of such conflicts of interest in terms of their effects on transaction outcomes and welfare loss. Building upon a stationary search framework with Nash bargaining, we develop a structural housing search model that encompasses a mixture of two polar cases – fully aligned incentives versus fully misaligned incentives. This mixture model allows us to measure different degrees of the distortion in incentives without solving games between sellers and agents. We further incorporate the decision to withdraw a listing into the housing search decision, thereby integrating a search framework with a discrete choice dynamic framework to recover reservation values. Using listing data from a residential Multiple Listing Service, we first find that transaction prices are negatively correlated with agent-owned properties, which is inconsistent with findings in the literature and may incorrectly suggest the absence of incentive misalignment. We then estimate our structural model using the nested pseudo-likelihood algorithm, and recover reservation values of the seller and the agent. We find that agent-sellers have lower reservation values than client-sellers in our data, which explains the aforementioned negative correlation. We further find that the estimated mixture probability of the fully misaligned incentive case is about 0.7 on average, indicating a high degree of conflicts of interest between the seller and the listing agent. However, after the implementation of a legislation that requires agents to disclose their dual agency relationship to their clients, this probability has decreased by a 8 percentage point, which can be translated into an increase in transaction prices by 6.84% for dual agency and by 4.34% for non-dual agency.

Keywords: search, housing markets, conflicts of interest, dual agency, structural estimation

JEL classification: C41, C51, L14, L85, R31

*I am grateful to George Deltas, Lu Han, Guillermo Marshall, Yasutora Watanabe, and seminar participants at National University of Singapore and Hong Kong University of Science and Technology for their helpful comments. Sen Ma provided excellent research assistance in data processing. This paper is based on work supported by (while serving at) the National Science Foundation. Any opinion, findings, and conclusions or recommendations expressed in this paper are those of the author and do not necessarily reflect the views of the National Science Foundation.

1 Introduction

The majority of consumers rely on real estate agents to buy and sell their homes, since real estate agents are better informed about local housing markets and housing transactions than most consumers who infrequently engage in real estate transactions. This informational advantage, on the other hand, may also provide distorted incentives for agents to seek their own interest rather than their client's interest. For example, listing agents may prefer selling houses quickly, and thus recommend their client-sellers to accept a current offer, even if it might be better for clients to wait longer to receive higher offers. In fact, several recent studies have provided empirical evidence consistent with this possibility, suggesting a misalignment of incentives between home sellers and their agents (Rutherford, Springer, and Yavas, 2005; Levitt and Syverson, 2008a). In principle, however, real estate agents have fiduciary duties to their clients in most North American housing markets. Therefore, the presence of conflicts of interest between sellers and agents implies that agents might be breaching their fiduciary duties, exploiting their informational advantage.

In this paper, we aim to investigate to what extent such breaches of fiduciary duties might occur in real estate brokerage settings. In particular, we quantify the extent to which the agent's incentives might conflict with the client's incentives. We further evaluate the importance of such conflicts of interest, in terms of their effects on transaction outcomes and welfare loss. To this end, we develop a structural housing search model that encompasses both the home seller's problem and the listing agent's problem, allowing for different degrees of conflicts of interest between sellers and agents. We estimate the model using the data from the residential Multiple Listing Service (MLS) which includes both sold listings and withdrawn listings.

A growing literature has empirically examined conflicts of interest between home sellers and listing agents by comparing agent-owned vs. client-owned listings (Rutherford, et al, 2005; Levitt and Syverson, 2008a), MLS vs. for-sale-by-owner (Hendel, Nevo, and Ortalo-Magne, 2009), traditional agents vs. discounted agents (Levitt and Syverson, 2008b; Berheim and Meer, 2013), and dual agency vs. non-dual agency (Gardiner, et al, 2007; Kadiyali, et al, 2012). However, most studies in this literature have provided their evidence based on reduced-form regressions of transaction prices and time-on-market using data from completed transactions. The key difficulty in such regressions is that reservation prices of sellers and agents are not observed, but reservation prices are one of important determinants of transaction prices, since sellers or agents accept buyers' offers only if these offers are higher than their reservation prices. To see the issue clearly, suppose that both agent-sellers (i.e. listing agents who sell their houses) and client-sellers (i.e. home owners who hire agents to sell their houses) own houses with the same characteristics. However, if agent-sellers have higher reservation prices than client-sellers, observed transaction prices will be higher for agent-owned properties than client-owned properties, not necessarily because of misaligned incentives between sellers and agents, but simply due to lower reservation prices of clients.[1] As a result, such regressions may entail selection bias and omitted variable bias, in that the observed sample includes only transactions with prices higher than reservation prices, and unobserved reservation prices are likely correlated with the key regressor such as the agent-owned dummy.

These issues can be addressed in our approach, since the estimation of our housing search model allows us to recover reservation values of sellers and agents. Specifically, we build upon a stationary search framework with Nash bargaining in the labor search literature (e.g. Eckstein and Wolpin, 1995; Eckstein and van den Berg, 2007), in which a reservation wage is essentially the same as the expected value of being unemployed, that is, the continuation value from search. This continuation value is similar to the fixed point from the Bellman equation in dynamic discrete choice models. While the continuation value is often solved by the nested fixed point algorithm (Rust, 1987) or alternative approaches (e.g. Hotz and Miller, 1993; Hotz, et al, 1994; Aguirregabiria and Mira, 2002) in the dynamic discrete choice literature, it is normally estimated as a parameter in the labor search literature (e.g. Flinn, 2006; Eckstein and van den Berg, 2007) because the continuation value and the parameter for instantaneous utility during search are both unknowns in the reservation wage equation and thus not separately identified (Flinn and Heckman, 1982). We address this under-identification issue by exploiting the additional data on withdrawn listings and incorporating the seller's (or agent's) withdrawal decision into the housing search decision. We then recover the continuation value by employing the nested pseudo-likelihood (NPL) algorithm (Aguirregabiria and Mira, 2002).

To quantify different degrees of conflicts of interest between home sellers and listing agents,

[1] Note that the comparison is made between agent-sellers' reservation prices for their own houses and *client-sellers' reservation prices for their houses*, as opposed to *agents' reservation prices for their clients' houses*. Obviously, if an agent does not fully represent his client-seller's interest, his reservation price for his client's house would be lower than his client's reservation price, even though the agent would have the same reservation price as his client, had the house been owned by himself. In this case, the house would be sold at lower price than agent-owned properties with the same characteristics, which thus reflects misaligned incentives between sellers and agents. In contrast, our point is that the client could have lower reservation price than the agent, in which case sales price would be lower for the client's house than for the house owned by the agent, even if the agent fully represents the client's interest.

our housing search model also includes three types of home sellers: agent-seller, “sophisticated” seller, and “uninformed” seller. For agent-sellers, listing agents are also home sellers, and so there is no conflict of interest between the principal and the agent. Sophisticated sellers are as informed about local housing markets as listing agents. Accordingly, these sellers hire agents to use listing services but make home selling decisions on their own, which represents the polar case where the seller’s incentive is fully aligned with the agent’s incentive. By contrast, uninformed sellers rely entirely on their agents for home selling decisions, in which case agents recommend the decision that benefits themselves the most. Hence, this represents the other polar case of completely misaligned incentives. For most listings owned by clients, the reality is likely to be somewhere between the two polar cases above, but it is difficult to directly model such an intermediate case, since it involves potentially complex games between sellers and agents. As a result, we instead consider the mixture of these two polar cases, which allows us to measure different degrees of the distortion in incentives without solving games between sellers and agents.

The key factor to distinguish between the three types of sellers above is different commission incentives between these cases. In particular, we consider the variation in commissions from dual agency, for which the same agent represents both the buyer and the seller, thus receiving commissions from both sides.[2] Therefore, for non-dual agency, agent-sellers’ revenues are transaction prices minus commissions to cooperating agents, but for dual agency, agent-sellers receive the entire transaction price as their revenues. However, sophisticated sellers receive transaction prices with the subtraction of commissions to both agents, whether the transaction involves dual agency or not. In contrast, for the uninformed seller case in which the listing agent makes a decision, the agent receives a full commission for dual agency, and a half commission for non-dual agency. We thus identify our model by using these different commission structures as well as exclusion restrictions.[3] The estimation of our model then allows us to recover reservation values for each case.

We use the MLS data from downtown area in a large North American city during the period 1996-2006. Our data include over 100,000 listings with information on dual agency and agent-owned property as well as detailed housing characteristics. In addition, during our sample period, the state of this city implemented a new legislation requiring real estate agents in dual agency to inform their

[2]Typically, the listing agent receives a commission (5-6% of transaction price) from the seller, and then splits it with the cooperating agent (i.e. buyer’s agent). Hence, the listing agent keeps the full commission under dual agency.

[3]For instance, per-period listing costs are incurred by agents but unlikely to be considered by sophisticated sellers who pay commissions only if the transaction is completed.

clients about dual agency and the nature of their agency relationship. This legislation is similar to the disclosure policy considered in Han and Hong (2016). Using our data, we first estimate reduced-form regressions of transaction outcomes. The results show that dual agency is negatively correlated with both transaction prices and days on market, but these negative correlations have weakened after the disclosure policy was implemented. These correlations seem to be consistent with the possibility that double commissions from dual agency provide distorted incentives for agents to recommend accepting lower offers, but that the disclosure policy may mitigate the incentive distortion. However, we also find that agent-owned properties tend to be sold at lower prices than client-owned properties, which is opposite to the findings in Rutherford, et al (2005) and Levitt and Syverson (2008a). Nevertheless, these results do not necessarily suggest evidence for or against a misalignment of incentives between sellers and agents. They rather illustrate the aforementioned issues in identifying conflicts of interest based on reduced-form regressions of transaction outcomes.

Therefore, we estimate our housing search model and recover reservation values for sellers and agents. In particular, we find that recovered reservation values for agent-sellers are lower than those for client-sellers which are the mixture between those of sophisticated sellers and uninformed sellers. This result implies that lower transaction prices for agent-owned properties, relative to client-owned properties, are simply due to lower reservation values of agent-sellers in our data, as opposed to the absence of a misalignment of incentives between sellers and agents. We further find that the mixture probability for the sophisticated seller case is about one third. Note that this mixture probability can be considered as a measure of conflicts of interest, in that as this probability is lower, the extent of conflicts of interest is higher. Hence, our result suggests a sizable degree of incentive misalignment between sellers and agents in our data. In addition, we find that this probability has increased by a 8 percentage point after the implementation of the disclosure policy. To assess the importance of this increase, we compute changes in the expected transaction prices when the mixture probability is increased by a 8 percentage point. The results show that transaction prices would increase by 6.84% for dual agency and by 4.34% for non-dual agency, which suggests that the impact of the disclosure policy is not trivial, and also that dual agency entails more conflicts of interest than non-dual agency.

This paper contributes to three strands of the literature. First, our study contributes to the aforementioned empirical literature on incentive misalignment between home sellers and listing agents. As discussed, most studies in this literature are based on reduced-form analysis. Consequently, our paper complements these studies by providing a structural approach that allows us to quantify the degree of the distortion in incentives and evaluate welfare implications.[4] Second, this paper is also related to a small body of literature on dual agency and in-house transactions (e.g., Evans and Kolbe, 2005; Gardiner, et al, 2007; Kadiyali, et al, 2012; Han and Hong, 2015). Except for Han and Hong (2016), these studies use MLS data for completed transactions and rely on reduced-form regressions of transaction outcomes on dual agency. The findings from these studies are mixed, in that the treatment effects of dual agency vary across different studies. However, our paper shows that these regressions entail selection biases and omitted variable bias discussed earlier,[5] which may explain inconclusive results in these studies. By contrast, Han and Hong (2016) combine both reduced-form and structural approaches, in order to investigate the extent to which the buyer's agent strategically promotes in-house transactions to the buyer, and further evaluate welfare impacts of strategic in-house transactions.[6] Our paper thus complements Han and Hong (2016) by examining the interactions between sellers and listing agents, using a structural approach.

Lastly, this paper is related to a small body of literature on structural housing search models (Carrillo, 2012; Merlo, Ortalo-Magne, and Rust, 2014), as well as two large literatures on structural labor search models (e.g., Eckstein and Wolpin, 1995; Eckstein and van den Berg, 2007) and discrete choice dynamic models (e.g., Rust, 1987; Hotz and Miller, 1993). Carrillo (2012) uses the MLS data for completed transactions and develops a stationary equilibrium search model that allows for heterogeneity in trading motivation of both buyers and sellers as well as the directing role of the asking price. Our structural model complements his model, in that we allow for withdrawal decisions as well as incentive misalignment between sellers and listing agents. Merlo, et al (2014) use unique UK data containing the complete history of each listing, and develop a non-stationary dynamic model for the seller's home selling problem. They consider the seller's withdrawal decision as well, but they focus more on modeling the seller's decision to adjust asking price, and on explaining the

[4]Conflicts of interest between principals and agents are not unique to the real estate brokerage industry, and a large empirical literature has also documented the presence of conflicts of interest between principals and agents in various other settings. See, e.g., Gruber and Owings (1996), Mehran and Stulz (1997), Hubbard (1998), Garmaise and Moskowitz (2004), and Jiang, Nelson, and Vytlacil (2014). Because most studies in this literature are also based on reduced-form analysis, our paper provides similar contributions to this literature by providing an alternative approach based on a structural estimation.

[5]See Section 2.4 for more detailed discussion on these econometric issues.

[6]An in-house transaction refers to the transaction for which the same *brokerage* represents both the buyer and the seller, whereas under dual agency, the same *agent* represents both the buyer and the seller. Though it would be interesting to examine in-house transactions as well, incentive structures are more complicated in the case of in-house transactions – see Han and Hong (2016) for more details. Hence, this paper only focuses on dual agency.

high degree of stickiness of asking prices observed in their data. Our model also complements their model, in that we integrate a structural labor search model with a discrete choice dynamic model, and further incorporate the interactions between sellers and agents by using a mixture model.

The paper is organized as follows. In Section 2, we develop our housing search model with three types of sellers. This section also discusses the model's implications on transaction outcomes, and characterizes biases inherent in reduced-form regressions of transaction outcomes. Section 3 describes our MLS data, and presents descriptive statistics and reduced-form regression results. In Section 4, we discuss our estimation approach and econometric issues. Section 5 reports the estimation results and implications. Section 6 concludes the paper.

2 Model

To examine a misalignment of incentives between the listing agent and the seller, we develop the search and bargaining model for home selling problems with dual agency. We examine events after house j is listed on Multiple Listing Services (MLS). The timing of events is as follows. At the beginning of each day, the seller (or the listing agent) decides whether to withdraw the listing, or continue to list the house. If the listing is withdrawn, the seller exits the market.[7] However, if house j is listed on MLS at period t, a buyer may visit the house with probability λ_j. If a buyer visits the house, she may be a dual buyer (i.e. the buyer represented by the agent who has also listed house j) with probability θ_j, or a non-dual buyer (i.e. a buyer represented by other agent) with probability $1-\theta_j$. Once the buyer visits, her match value with the house is realized, and the buyer bargains with the listing agent. The resulting price is determined by the Nash bargaining solution. If the price is higher than the reservation price of the listing agent (or the seller), the agent accepts the price and the house is sold. However, if the house is not sold or no buyer visits the house, then the above events will repeat at period $t+1$.

We assume that λ_j and θ_j are given exogenously, in that they are determined by other factors, aside from the listing agent's strategic behavior. In other words, we allow the listing agent to influence the seller's problem, but not the buyer's problem which is captured by λ_j and θ_j. For λ_j

[7]Some sellers may delist their houses from MLS and list them again (but with new listing numbers). There are two reasons for doing so. One is to reset the days on market, in which case delisting should not be considered as withdrawal. The other is to withdraw from the current market and sell houses in a better housing market, in which case delisting can be considered as withdrawal, while re-listed houses can be essentially considered as new listings. For this reason, we assume that once the listing is withdrawn, the seller exits the market. In our data, we can link consecutive listings of the same house, and attempt to separate these two types of delisting.

and θ_j, we consider a reduced-form model described in Section 4.1. In reality, the buyer's problem may not be completely isolated from the listing agent's behavior, but this possibility is assumed away in this paper for two reasons. First, our focus is mostly on the interactions between sellers and listing agents, rather than those between buyers and their agents. Second, it is difficult to model strategic interactions between buyers and their agents, given limited data on buyers.

To model different degrees of misaligned agent incentives, we consider three cases corresponding to three types of home sellers. The first type includes agent-sellers who are both listing agents and home sellers at the same time. For the second and third types, sellers hire listing agents, so that the principal is different from the agent. The second type consists of "sophisticated" sellers who are informed about their housing market as much as listings agents. These sophisticated sellers hire agents to use listing services, but these sellers make home selling decisions on their own. In contrast, the third type is composed of "naive" or "uninformed" sellers who are not informed about their housing market. As a result, they need to rely fully on their agents for home selling decisions.

Note that the first type of agent-sellers represent the case in which there is no conflict of interest between the principal and the agent. We also observe agent-sellers in the data, and so we consider them as a benchmark case. For the second type with sophisticated sellers, the listing agent has to recommend the decision that maximizes his client's objective function, since any other recommendations will be rejected by sophisticated sellers. Hence, this case is equivalent to having fully aligned incentives between the seller and the agent, except that the seller's objective function in this case is not exactly the same as those of agent-sellers. By contrast, uninformed sellers will follow any advice from their agents, in which case the listing agent will recommend the decision that maximizes his own objective function. Accordingly, the third type reflects the case where the agent's incentive is completely misaligned with the seller.

Except for agent-sellers, the extent of incentive misalignment is likely to vary across different sellers and agents, but it is difficult to model all possible cases. For this reason, we consider the mixture of two polar cases represented by the second and third types of sellers, thereby capturing different degrees of conflicts of interest between sellers and listing agents. The following subsections describe these three cases in more detail.

2.1 Agent-Seller Case

Consider house j owned by a listing agent. Once the agent lists his own house on MLS, he will make two kinds of decisions each day: the first is whether to sell the house if a buyer visits the house, and the second is whether to continue listing the house if it is not sold. We use the following variables to denote these events: $s_{j,t}$ is the indicator variable for selling house j on day t; b_j is the dummy for the event that a buyer visits the house; $w_{j,t}$ is the indicator variable for withdrawing the listing from the market (i.e. $w_{j,t} = 0$ if the agent continues to list house j). We assume that the agent needs to incur per-period fixed cost c_j to continue to list the house on the market. The listing cost c_j includes any cost associated with selling the house, such as expenses or efforts spent on marketing the house.[8] In what follows, we first consider the selling decision and then model the withdrawal decision.

To model the selling decision, we need to consider the buyer's behavior because the selling decision occurs only if a buyer visits the house. However, partly due to limited data on buyers, but also due to our main focus on the seller and the listing agent, we do not model the buyer's search problem. Instead, we only consider a simple model for the buyer. Specifically, we assume that house j can be visited by at most one buyer each day.[9] Buyer i's match value for house j, M_{ij}, consists of two parts – a common value and a match value specific to each buyer. The common value is assumed to be $X_j\beta + \xi_j$, where X_j and ξ_j are observed and unobserved house characteristics, and β is a vector of parameters common to any buyer. The match value specific to buyer i, denoted by $\tilde{\mu}_{ij}$, is realized only after buyer i visits house j. Buyer i's utility from buying house j is then equal to $M_{ij} - \ln p_{ij}$. We denote the value of her outside option, i.e. not buying house j, by $\bar{U}_i$. Hence, her net utility is given by

$$M_{ij} - \bar{U}_i - \ln p_{ij} = X_j\beta + \xi_j + \mu_{ij} - \ln p_{ij}, \tag{1}$$

where $\mu_{ij} = \tilde{\mu}_{ij} - \bar{U}_i$. We further assume that μ_{ij} follows a known distribution $G(\mu)$.

If the agent-seller sells his house to a dual buyer, he will receive p_{ij}. If he sells the house to a non-dual buyer, he will receive $\left(1 - \frac{\tau}{2}\right) p_{ij}$, where p_{ij} is the transaction price and τ is the

[8] We assume that agents incur the same listing cost c_j, regardless of dual agency or non-dual agency. In addition, we assume away transaction costs that agents or sellers may incur during bargaining and closing stages. However, if dual agency reduces listing costs or transaction costs, it would be important to capture this impact of dual agency on agent costs, because it may affect welfare analysis. We will investigate this issue in the future version.

[9] It is possible that multiple buyers can visit the house on a given day, in which case a bidding war could occur. To simplify the analysis, however, we assume that only one buyer can visit the house each day.

commission rate. Typically, the seller pays a commission (5-6% of the transaction price) to the listing agent, and the listing agent evenly splits the commission with the buyer's agent. Hence, for non-dual agency, the agent-seller pays $\frac{\tau}{2}p_{ij}$ to the buyer's agent.

The transaction price p_{ij} is determined by Nash bargaining. If the bargaining between the agent-seller and the buyer does not reach an agreement, the agent-seller receives the disagreement value denoted by $\bar{D}_j^a$. The specific expression for $\bar{D}_j^a$ is derived below when we consider the withdrawal decision. The resulting price from the bargaining between buyer i and the agent-seller of house j is determined by Nash bargaining that solves

$$\max_{\ln p_{ij}}[\ln p_{ij} - \bar{D}_j^a]^\alpha [M_{ij} - \ln p_{ij} - \bar{U}_i]^{1-\alpha} \quad \text{if } b_j^d = 1,$$
$$\text{or } \max_{\ln p_{ij}} \left[\ln\left(1-\frac{\tau}{2}\right)p_{ij} - \bar{D}_j^a\right]^\alpha [M_{ij} - \ln p_{ij} - \bar{U}_i]^{1-\alpha} \quad \text{if } b_j^n = 1,$$

where α is the bargaining power of the agent-seller, and b_j^d (or b_j^n) is the indicator variable for a dual (or non-dual) buyer's visit. Note that we consider the problem above in terms of $\ln p$ instead of p, because $\ln p$ fits the data better, while choosing $\ln p$ is not different from choosing p. The solution to this problem is then given by

$$\ln p_{ij}^{a,d} = \alpha(M_{ij} - \bar{U}_i) + (1-\alpha)\bar{D}_j^a \quad \text{if } b_j^d = 1, \tag{2}$$
$$\ln p_{ij}^{a,n} = \alpha(M_{ij} - \bar{U}_i) + (1-\alpha)\left[\bar{D}_j^a - \ln(1-\tau/2)\right] \quad \text{if } b_j^n = 1, \tag{3}$$

where we use the superscript a, d (or a, n) to indicate that the price is the solution to Nash bargaining involving an agent-seller and a dual (or non-dual) buyer.

Therefore, when a dual buyer visits house j, the agent-seller agrees to sell his house if $\ln p_{ij}^{a,d} \geq \bar{D}_j^a$, which is equivalent to $M_{ij} - \bar{U}_i \geq \bar{D}_j^a$ by using the solution in (2). Similarly, when a non-dual buyer visits, the agent-seller sells his house if $\ln\left[\left(1-\frac{\tau}{2}\right)p_{ij}^{a,n}\right] \geq \bar{D}_j^a$, which is equivalent to $M_{ij} - \bar{U}_i \geq \bar{D}_j^a - \ln(1-\tau/2)$ by using the solution in (3). Using these inequalities and (1), we can write the probability of transaction with a dual buyer as

$$\begin{aligned}\varphi_j^{a,d} \equiv \Pr(s_j = 1 | a_j = 1, b_j^d = 1) &= \Pr\left[X_j\beta + \xi_j + \mu_{ij} \geq \bar{D}_j^a\right], \\ &= 1 - G\left(\bar{D}_j^a - X_j\beta - \xi_j\right),\end{aligned}$$

and the probability of transaction with a non-dual buyer as

$$\varphi_j^{a,n} \equiv \Pr(s_j = 1 | a_j = 1, b_j^n = 1) = 1 - G\left[\bar{D}_j^a - \ln\left(1-\frac{\tau}{2}\right) - X_j\beta - \xi_j\right],$$

where a_j is the dummy for agent-seller.

To model the decision to continue listing the house, let us begin with the agent-seller's value function, $V_j^a(s_{j,t-1})$, where $s_{j,t-1} = 1$ (or 0) if the house was sold at period $t-1$ (or not). We consider only $V_j^a(s_{j,t-1} = 0)$, since $s_{j,t} = 1$ is a terminal state, and the agent-seller's expected payoff for $s_{j,t} = 1$ is explicitly written below. Hence, we simply use V_j^a to denote $V_j^a(s_{j,t-1} = 0)$. The value function is then the unique solution of the Bellman equation:

$$V_j^a = \max_{w_{j,t} \in \{0,1\}} \begin{cases} v_j^a(0) + \epsilon_{j,t}^a(0) & \text{if } w_{j,t} = 0 \\ v_j^a(1) + \epsilon_{j,t}^a(1) & \text{if } w_{j,t} = 1 \end{cases} \tag{4}$$

where $\epsilon_{j,t}^a(w_{j,t} = 0)$, or simply $\epsilon_{j,t}^a(0)$, is the idiosyncratic error term capturing unobserved utility associated with continuing to list the house, and $\epsilon_{j,t}^a(1)$, is the idiosyncratic error reflecting the outside option value from withdrawing the listing. Note that once the listing is withdrawn, the seller exits the market. Accordingly, $v_j^a(w_{j,t} = 1)$, or simply $v_j^a(1)$, is assumed to be zero. However, if the agent-seller continues to list the house, he will incur listing cost c_j and sell the house to a dual buyer with probability $\lambda_j \theta_j \varphi_j^{a,d}$; or a non-dual buyer with probability $\lambda_j(1-\theta_j)\varphi_j^{a,n}$; or fail to sell the house with probability $1 - \lambda_j \theta_j \varphi_j^{a,d} - \lambda_j(1-\theta_j)\varphi_j^{a,n}$. Therefore, the agent-seller's expected value of continued search is written as

$$\begin{aligned} v_j^a(0) &= -c_j + \lambda_j \theta_j \varphi_j^{a,d} E_\mu(\ln p_{ij}^{a,d} | a_j = 1, b_j^d = 1, s_{j,t} = 1) \\ &\quad + \lambda_j(1-\theta_j)\varphi_j^{a,n} E_\mu \left[\ln \left(1 - \frac{\tau}{2}\right) p_{ij}^{a,n} \middle| a_j = 1, b_j^n = 1, s_{j,t} = 1 \right] \\ &\quad + \left[1 - \lambda_j \theta_j \varphi_j^{a,d} - \lambda_j(1-\theta_j)\varphi_j^{a,n}\right] \rho E V_j^a, \end{aligned} \tag{5}$$

where ρ is the discount factor. The second and third terms in (5) are the agent-seller's expected payoff from selling the house, where the expectation is taken over μ. The last term is the agent-seller's expected value when $s_{j,t} = 0$, and EV_j^a is the integrated value function given by

$$EV_j^a = \int V_j^a(s_{j,t} = 0) f(\epsilon^a) d\epsilon^a,$$

where $f(\epsilon^a)$ denotes the density function for $(\epsilon^a(0), \epsilon^a(1))$. We follow the literature on discrete choice dynamic models, and assume that $\epsilon^a(0)$ and $\epsilon^a(1)$ are extreme value random variables that are independent and identically distributed across listings and over time. Note that $\epsilon_{j,t+1}^a$ is not realized at period t, which is why EV_j^a, instead of V_j^a, enters (5). Note also that EV_j^a is the agent-seller's disagreement value, $\bar{D}_j^a$, since EV_j^a is the maximum expected value that the agent-seller can obtain when his bargaining with the buyer breaks down.

For the estimation problem, we integrate out ϵ^a in (5) and consider the smoothed Bellman equation given by

$$EV_j^a = \int \max_{w_{j,t}\in\{0,1\}} \left[v_j^a(0) + \epsilon_{j,t}^a(0),\ \epsilon_{j,t}^a(1)\right] f(\epsilon^a)d\epsilon^a. \tag{6}$$

We then define the agent-seller's conditional choice probability (CCP) for withdrawal as follows:

$$\Pr(w_{j,t} = 1|a_j = 1, s_{j,t-1} = 0) = \int I\left\{1 = \arg\max_{w_{j,t}} \left[v_j^a(0) + \epsilon_{j,t}^a(0),\ \epsilon_{j,t}^a(1)\right]\right\} f(\epsilon^a)d\epsilon^a,$$

where $I(\cdot)$ is the indicator function. The extreme value assumption implies

$$\begin{aligned}
\Pr(w_{j,t} = 0|a_j = 1, s_{j,t-1} = 0) &= \frac{\exp\left(v_j^a(0)\right)}{1+\exp\left(v_j^a(0)\right)}, \\
\Pr(w_{j,t} = 1|a_j = 1, s_{j,t-1} = 0) &= \frac{1}{1+\exp\left(v_j^a(0)\right)},
\end{aligned}$$

from which we can also obtain the following inversion of the CCP:

$$v_j^a(0) = \ln\left[\Pr(w_{j,t} = 0|a_j = 1, s_{j,t-1} = 0)\right] - \ln\left[\Pr(w_{j,t} = 1|a_j = 1, s_{j,t-1} = 0)\right]. \tag{7}$$

We further rewrite the smoothed Bellman equation in (6) as follows:

$$\begin{aligned}
EV_j^a &= \Pr(w_{j,t} = 0|a_j = 1, s_{j,t-1} = 0)\left\{v_j^a(w_{j,t} = 0) + E_\epsilon[\epsilon^a(0)|w_{j,t} = 0]\right\} \\
&\quad + \Pr(w_{j,t} = 1|a_j = 1, s_{j,t-1} = 0)\left\{E_\epsilon[\epsilon^a(1)|w_{j,t} = 0]\right\}.
\end{aligned} \tag{8}$$

Using the well-known results, we obtain

$$\begin{aligned}
E_\epsilon[\epsilon^a(0)|w_{j,t} = 0] &= \gamma - \ln\left[\Pr(w_{j,t} = 0|a_j = 1, s_{j,t-1} = 0)\right], \\
E_\epsilon[\epsilon^a(1)|w_{j,t} = 1] &= \gamma - \ln\left[\Pr(w_{j,t} = 1|a_j = 1, s_{j,t-1} = 0)\right],
\end{aligned}$$

where γ is the Euler's constant. Plugging these equations and (7) into (8) yields

$$EV_j^a = \gamma - \ln\left[\Pr(w_{j,t} = 1|a_j = 1, s_{j,t-1} = 0)\right]. \tag{9}$$

The derivation above suggests that we can use the approaches by Hotz and Miller (1993) or Aguirregabiria and Mira (2002) to estimate the model.

2.2 Sophisticated Seller Case

If sellers are client-sellers, that is, they are not real estate agents, they need to hire listing agents to sell their houses. Listing agents provide their clients with listing services (e.g. listing houses on the MLS, advertising and marketing houses, etc.) and home selling advice based on their expertise in local housing markets. As discussed above, we consider two types of client-sellers who hire listing agents – sophisticated sellers and uninformed sellers. We assume that a client-seller is sophisticated with the probability κ_j, or uninformed with the probability $1 - \kappa_j$. Sophisticated sellers are supposed to have the same expertise in local housing markets as listing agents, so that their decisions to sell houses are not influenced by agents' strategic behavior. As a result, the role of the agent is minimized for the case of sophisticated sellers, and it is sufficient to consider only the seller's problem, while ignoring the agent's problem.

Similar to the agent-seller case, sophisticated sellers make two decisions – whether to sell the house if a buyer visits, and whether to withdraw the listing if it is not sold – and face the same distribution of buyers in terms of visits and match values. However, there are two key differences from the agent-seller case. First, sophisticated sellers do not take into account listing cost c_j when they decide not to withdraw their houses, since c_j is the cost incurred by the listing agent and the seller pays a commission fee to the listing agent only after the transaction is completed. Hence, the sophisticated seller's expected value of continued search, which we denote by $v_j^s(0)$, does not include c_j. Second, the seller's payoff from selling the house is $(1-\tau)p_{ij}$ regardless of a dual buyer or a non-dual buyer.

Therefore, whether a dual buyer visits or a non-dual buyer visits the house, the sophisticated seller's negotiated price is determined by Nash bargaining that solves

$$\max_{\ln p_{ij}} \left[\ln(1-\tau)p_{ij} - EV_j^s\right]^{\alpha} \left[M_{ij} - \ln p_{ij} - \bar{U}_i\right]^{1-\alpha},$$

where EV_j^s is the integrated value function for the sophisticated seller. We then obtain

$$\ln p_{ij}^s = \alpha(M_{ij} - \bar{U}_i) + (1-\alpha)\left[EV_j^s - \ln(1-\tau)\right]. \tag{10}$$

Given this price, the seller agrees to sell his house if $\ln\left[(1-\tau)p_{ij}^s\right] \geq EV_j^s$, which is equivalent to $X_j\beta + \xi_j + \mu_{ij} \geq EV_j^s - \ln(1-\tau)$. The sophisticated seller's probability of transaction is given by

$$\varphi_j^s \equiv \Pr(s_j = 1 | a_j = 0, k_j = 1, b_j = 1) = 1 - G\left[EV_j^s - \ln(1-\tau) - X_j\beta - \xi_j\right],$$

where k_j equals 1 if the seller is sophisticated, or 0 if uninformed.

The withdrawal decision is modeled similar to the agent-seller's case. The sophisticated seller's value function is the unique solution of the Bellman equation:

$$V_j^s = \max_{w_{j,t}\in\{0,1\}} \left[v_j^s(0) + \epsilon_{j,t}^s(0),\ \epsilon_{j,t}^s(1)\right],$$

where $\epsilon_{j,t}^s(0)$ and $\epsilon_{j,t}^s(1)$ are the idiosyncratic error terms that respectively capture the sophisticated seller's unobserved utility associated with continuing to list the house and outside option value from withdrawing, and $v_j^s(0)$ is given by

$$v_j^s(0) \quad = \quad \lambda_j \varphi_j^s E_\mu \left[\ln(1-\tau)p_{ij}^s | a_j = 0, k_j = 1, b_j = 1, s_{j,t} = 1\right] + \left[1 - \lambda_j \varphi_j^s\right] \rho E V_j^s. \tag{11}$$

Assuming that $\epsilon^s(0)$ and $\epsilon^s(1)$ are i.i.d. extreme value random variables, we obtain the conditional choice probability as

$$\Pr(w_{j,t} = 0 | a_j = 0, k_j = 1, s_{j,t-1} = 0) \quad = \quad \frac{\exp\left(v_j^s(0)\right)}{1 + \exp\left(v_j^s(0)\right)},$$

from which we can also obtain the following inversion of the CCP:

$$v_j^s(0) = \ln\left[\frac{\Pr(w_{j,t} = 0 | a_j = 0, k_j = 1, s_{j,t-1} = 0)}{\Pr(w_{j,t} = 1 | a_j = 0, k_j = 1, s_{j,t-1} = 0)}\right].$$

Similar algebra as in the agent-seller case yields

$$EV_j^s = \gamma - \ln\left[\Pr(w_{j,t} = 1 | a_j = 0, k_j = 1, s_{j,t-1} = 0)\right].$$

2.3 Uninformed Seller Case

Uninformed sellers need to rely fully on their listing agents, so that their decisions to sell houses are completely influenced by their agents. Hence, for uninformed sellers, it is sufficient to consider only the agent's problem, while ignoring the seller's problem. Similar to the previous cases, the listing agent representing uninformed sellers makes the selling decision as well as the withdrawal decision, and face the same distribution of buyers. The key difference from the previous cases is that the agent receives $\tau p_{i,j}$ for a transaction with a dual buyer, or $\frac{\tau}{2} p_{i,j}$ for a transaction with a non-dual buyer. In addition, the agent takes into account listing cost c_j for the withdrawal decision. We also distinguish the agent's value function from the seller's value function (i.e. V_j^a and V_j^s) by using π_j^u to denote the value function of the listing agent representing an uninformed seller.

If a buyer visits the house, the listing agent bargains with the buyer, and the resulting price is determined by Nash bargaining that solves

$$\max_{\ln p_{ij}} \left[\ln\left(b_j^d \tau p_{ij} + b_j^n \frac{\tau}{2} p_{ij}\right) - E\pi_j^u\right]^{\alpha} \left[M_{ij} - p_{ij} - \bar{U}_i\right]^{1-\alpha},$$

where $E\pi_j^u$ is the integrated value function in the case of uninformed sellers. The solution to this problem is as follows:

$$\ln p_{ij}^{u,d} = \alpha(M_{ij} - \bar{U}_i) + (1-\alpha)\left[E\pi_j^u - \ln\tau\right] \quad \text{if } b_j^d = 1, \tag{12}$$

$$\ln p_{ij}^{u,n} = \alpha(M_{ij} - \bar{U}_i) + (1-\alpha)\left[E\pi_j^u - \ln(\tau/2)\right] \quad \text{if } b_j^n = 1. \tag{13}$$

Given these prices, the agent agrees to sell the house[10] if $\ln\left[\tau p_{ij}^{u,d}\right] \geq E\pi_j^u$ for a dual buyer, or if $\ln\left[\frac{\tau}{2} p_{ij}^{u,n}\right] \geq E\pi_j^u$ for a non-dual buyer. Using (12) and (13), these inequalities imply the uninformed seller's probability of transaction as follows:

$$\varphi_j^{u,d} \equiv \Pr(s_j = 1 | a_j = 0, k_j = 0, b_j^d = 1) = 1 - G\left[E\pi_j^u - \ln\tau - X_j\beta - \xi_j\right],$$

$$\varphi_j^{u,n} \equiv \Pr(s_j = 1 | a_j = 0, k_j = 0, b_j^n = 1) = 1 - G\left[E\pi_j^u - \ln(\tau/2) - X_j\beta - \xi_j\right].$$

Similar to the previous cases, the agent's value function for $s_{j,t-1} = 0$ is written as

$$\pi_{j,t}^u = \max_{w_{j,t} \in \{0,1\}} \left[v_j^u(0) + \epsilon_{j,t}^u(0),\ \epsilon_{j,t}^u(1)\right],$$

where $\epsilon^u(0)$ and $\epsilon^u(1)$ are i.i.d. extreme value random variables capturing the agent's unobserved utility from continuing to list the house and outside option value from withdrawing. The agent's expected value of continued search is written as

$$\begin{aligned} v_j^u(0) = & -c_j + \lambda_j \theta_j \varphi_j^{u,d} E_\mu \left[\ln(\tau p_{ij}^{u,d}) | a_j = 0, k_j = 0, b_j^d = 1, s_{j,t} = 1\right] \\ & + \lambda_j (1-\theta_j) \varphi_j^{u,n} E_\mu \left[\ln\left(\frac{\tau}{2} p_{ij}^{a,n}\right) \middle| a_j = 0, k_j = 0, b_j^n = 1, s_{j,t} = 1\right] \\ & + \left[1 - \lambda_j \theta_j \varphi_j^{u,d} - \lambda_j (1-\theta_j) \varphi_j^{u,n}\right] \rho E\pi_j^u. \end{aligned} \tag{14}$$

Given the extreme value assumption, we obtain the conditional choice probability as

$$\Pr(w_{j,t} = 0 | a_j = 0, k_j = 0, s_{j,t-1} = 0) = \frac{\exp\left(v_j^u(0)\right)}{1 + \exp\left(v_j^u(0)\right)},$$

[10] More precisely speaking, the agent recommends selling the house, and the uninformed seller follows the advice and decides to sell the house.

from which we can also obtain the following inversion of the CCP:

$$v_j^u(0) = \ln\left[\frac{\Pr(w_{j,t}=0|a_j=0,k_j=0,s_{j,t-1}=0)}{\Pr(w_{j,t}=1|a_j=0,k_j=0,s_{j,t-1}=0)}\right].$$

Similar algebra as before yields

$$E\pi_j^u = \gamma - \ln\left[\Pr(w_{j,t}=1|a_j=0,k_j=0,s_{j,t-1}=0)\right].$$

The estimation of our model in this section is essentially done by maximum likelihood estimation, but there are some complications that we need to address. See Section 4 for the likelihood function and the details on our estimation approach.

2.4 Implications on Transaction Outcomes

In this subsection, we first examine our model's implications on transaction outcomes, and then discuss potential biases inherent in reduced-form regressions.

The main transaction outcomes are the listing's transaction price and time on market (TOM). To examine predictions on transaction prices, we consider the negotiated prices derived in (2), (3), (10), (12) and (13). For easy comparison, all these prices are rewritten below:

$$\begin{array}{rrcl}
\text{agent-seller:} & \ln p_{ij}^{a,d} & = & \alpha(X_j\beta+\xi_j+\mu_{i,j}) + (1-\alpha)EV_j^a, \\
 & \ln p_{ij}^{a,n} & = & \alpha(X_j\beta+\xi_j+\mu_{i,j}) + (1-\alpha)\left[EV_j^a - \ln\left(1-\frac{\tau}{2}\right)\right], \\
\text{sophisticated seller:} & \ln p_{ij}^{s} & = & \alpha(X_j\beta+\xi_j+\mu_{i,j}) + (1-\alpha)[EV_j^s - \ln(1-\tau)], \\
\text{uninformed seller:} & \ln p_{ij}^{u,d} & = & \alpha(X_j\beta+\xi_j+\mu_{i,j}) + (1-\alpha)[E\pi_j^u - \ln\tau], \\
 & \ln p_{ij}^{u,n} & = & \alpha(X_j\beta+\xi_j+\mu_{i,j}) + (1-\alpha)\left[E\pi_j^u - \ln\left(\frac{\tau}{2}\right)\right].
\end{array}$$

To examine predictions on TOM, we consider the listing's probability of transaction, because a higher probability of transaction implies a higher hazard rate, thus shortening TOM. For easy comparison, the probabilities of transaction for all cases are also rewritten as follows:

$$\begin{array}{rrcl}
\text{agent-seller:} & \varphi_j^{a,d} & = & 1 - G\left[EV_j^a - X_j\beta - \xi_j\right], \\
 & \varphi_j^{a,n} & = & 1 - G\left[EV_j^a - \ln\left(1-\frac{\tau}{2}\right) - X_j\beta - \xi_j\right], \\
\text{sophisticated seller:} & \varphi_j^{s} & = & 1 - G\left[EV_j^s - \ln(1-\tau) - X_j\beta - \xi_j\right], \\
\text{uninformed seller:} & \varphi_j^{u,d} & = & 1 - G\left[E\pi_j^u - \ln\tau - X_j\beta - \xi_j\right], \\
 & \varphi_j^{u,n} & = & 1 - G\left[E\pi_j^u - \ln\left(\frac{\tau}{2}\right) - X_j\beta - \xi_j\right].
\end{array}$$

Comparing transaction prices and probabilities above leads to the following predictions:

Prediction 1: All else being equal, transaction prices are lower for dual agency than for non-dual agency. Similarly, TOM will likely be shorter for dual agency than for non-dual agency.

Prediction 2: As there are more sophisticated sellers in the market, the difference in transaction prices and TOM between dual agency and non-dual agency will become smaller.

Prediction 3: The implications on transaction prices and TOM between agent-sellers and client-sellers are ambiguous.

Prediction 1 follows from $0 < -\ln\left(1 - \frac{\tau}{2}\right)$ and $-\ln\tau < -\ln(\tau/2)$, which implies that $\ln p_{i,j}^{a,d} < \ln p_{i,j}^{a,n}$ and $\ln p_{i,j}^{u,d} < \ln p_{i,j}^{u,n}$, as well as $\varphi_j^{a,d} > \varphi_j^{a,n}$ and $\varphi_j^{u,d} > \varphi_j^{u,n}$. In other words, dual agency generates higher commissions for agents than non-dual agency, *ceteris paribus*. Hence, if transactions involve dual agency, agents have stronger incentives to sell (or recommend to sell) houses cheaply and quickly. Prediction 2 is implied by the fact that transaction prices and probabilities are the same for sophisticated sellers, whether transactions involve dual agency or not. As a result, if more sellers are informed and sophisticated, the difference in transaction prices and TOM between dual agency and non-dual agency is likely to be smaller.

As stated in Prediction 3, however, our model does not provide unambiguous predictions regarding transaction prices and TOM between agent-sellers and client-sellers. For example, if we consider the difference in prices between agent-sellers and uninformed sellers under dual agency, we obtain that $\ln p_{i,j}^{a,d} - \ln p_{i,j}^{u,d} = (1-\alpha)[EV_j^a - E\pi_j^u + \ln\tau]$. Given that agent-sellers receive full transaction prices, while agents for uninformed client-sellers receive τ fraction of prices, EV_j^a is likely to be higher than $E\pi_j^u$. Hence, if we ignore $\ln\tau$, then $p_{i,j}^{a,d}$ is likely to be greater than $\ln p_{i,j}^{u,d}$, which is similar to the prediction from Levitt and Syverson (2008a). However, even if commission rate τ is small, $\ln\tau$ can be nontrivial. Hence, it is unclear whether $EV_j^a > E\pi_j^u - \ln\tau$. Moreover, if we also consider sophisticated sellers, the predicted price for client-sellers will be a convex combination of $\ln p_{i,j}^s$ and $\ln p_{i,j}^{u,d}$ (or $\ln p_{i,j}^{u,n}$) in the case of dual agency (or non-dual agency). Consequently, transaction prices for agent-sellers are not necessarily higher than those for client-sellers, even when everything else is equal.

To test these predictions, one may regress transaction outcomes on variables of interest. For example, to test Predictions 1 and 2, we could use a dummy variable for dual agency, and some proxy variables related to sophisticated sellers, such as a dummy variable for disclosure policy that requires agents to inform their clients of dual agency and the nature of their agency relationship. In such regressions, a negative coefficient on dual agency and a positive coefficient on an interaction between dual agency and disclosure policy would be consistent with Predictions 1 and 2.

However, one critical issue with these regressions is that transaction prices (or time on market until sale) are observed only if transactions are completed, that is, prices are higher than the seller's (or agent's) reservation value, which suggests that selection biases might arise in such regressions. To see this issue clearly, let us ignore ξ_j, and further assume that $\mu_{i,j}$ is independent of X_j and follows the standard normal distribution. As an illustration, consider observations with agent-sellers only. Suppose that we regress transaction prices on X_j and a dummy for dual agency, denoted by d_j. Hence, we consider the conditional expectation of prices as follows:

$$
\begin{aligned}
& E\left[\ln p_{i,j}|X_j, d_j, a_j = 1, s_j = 1\right] \\
= \; & d_j \times E\left[\ln p_{i,j}^{a,d}|X_j, d_j = 1, a_j = 1, s_j = 1\right] + (1-d_j) \times E\left[\ln p_{i,j}^{a,n}|X_j, d_j = 0, a_j = 1, s_j = 1\right] \\
= \; & d_j \times \left\{\alpha X_j\beta + (1-\alpha)EV_j^a + \alpha E[\mu_{i,j}|d_j = 1, a_j = 1, s_j = 1]\right\} \\
& +(1-d_j) \times \left\{\alpha X_j\beta + (1-\alpha)[EV_j^a - \ln(1-\tau/2)] + \alpha E[\mu_{i,j}|d_j = 0, a_j = 1, s_j = 1]\right\}.
\end{aligned}
$$

But $s_j = 1$ only if prices are higher than the seller's reservation value. Given the standard normal assumption on $\mu_{i,j}$, we can obtain inverse Mills ratio as follows: $E[\mu_{i,j}|d_j = 1, a_j = 1, s_j = 1] = \frac{\phi\left(EV_j^a - X_j\beta\right)}{1-\Phi\left(EV_j^a - X_j\beta\right)}$ and $E[\mu_{i,j}|d_j = 0, a_j = 1, s_j = 1] = \frac{\phi\left(EV_j^a - \ln(1-\frac{\tau}{2}) - X_j\beta\right)}{1-\Phi\left(EV_j^a - \ln(1-\frac{\tau}{2}) - X_j\beta\right)}$. Hence, it follows that

$$
\begin{aligned}
E\left[\ln p_{i,j}|X_j, d_j, a_j = 1, s_j = 1\right] \; = \; & \delta_0 + \delta_1 \, d_j + X_j\tilde{\beta} + (1-\alpha)EV_j^a + \frac{\alpha d_j \phi(EV_j^a - X_j\beta)}{1-\Phi(EV_j^a - X_j\beta)} \\
& + \frac{\alpha(1-d_j)\phi(EV_j^a - \ln(1-\frac{\tau}{2}) - X_j\beta)}{1-\Phi(EV_j^a - \ln(1-\frac{\tau}{2}) - X_j\beta)},
\end{aligned} \tag{15}
$$

where $\delta_0 = -(1-\alpha)\ln(1-\frac{\tau}{2})$, $\delta_1 = (1-\alpha)\ln(1-\frac{\tau}{2})$, and $\tilde{\beta} = \alpha\beta$.

If δ_1 could be estimated consistently in (15), we could test Prediction 1. However, the equation (15) illustrates two key issues. First, there is selection bias, as indicated by the inverse Mills ratio terms in (15). Second, omitted variable bias may arise for two reasons: (i) we do not directly observe EV_j^a, and (ii) d_j may be correlated with EV_j^a, because whether $d_j = 1$ or 0 depends on θ_j (probability of dual buyer's visit) and $\varphi_j^{a,d}$ (probability of transaction under dual agency), but $\varphi_j^{a,d}$ also depends on EV_j^a.

The example above only considers a reduced-form regression of transaction price on a dummy for dual agency among agent-sellers, simply because we can obtain slightly simpler exposition for this case. The issues discussed above, nonetheless, can be applied to any regressions of transaction prices, because transaction prices are observed only if houses are sold, and they may reflect the seller's reservation value which might be correlated with variables of interest. Therefore, reduced-form results are likely to be biased, unless these issues are addressed. The main difficulty of addressing these issues is that we do not directly observe the seller's (or the agent's) reservation values. In contrast, our structural model developed above suggests how to recover those reservation values. Section 4 provides details on our estimation approach.

3 Data

3.1 Data Description

The main source of our data is the Multiple Listing Service (MLS) in a large North American city. We obtained the residential listing data for downtown area in this city from January 1, 1996 to December 31, 2006. Hence, our data do not include the period for the recent housing market crisis. The data encompass over 100,000 listings: about 45% of listings were sold, while the remaining listings were not sold. However, unsold listings do not mean the withdrawal from the market, because listing agents or sellers may de-list their listings and then quickly list them again to reset the days on market. For this reason, we link consecutive listings of the same house by using addresses and property identification numbers. Specifically, if an unsold listing is de-listed and the house of this listing is not re-listed within 180 days, this listing is defined to be "withdrawn" from the market. Using this definition of withdrawal, we find that the fraction of withdrawn listings is about 23% in our data, and the remaining 32% of listings are temporarily de-listed and shortly re-listed. In our main analysis, we exclude the latter type of listings, and focus on two kinds of listings – one that is eventually sold, and the other that is eventually withdrawn. However, temporarily de-listed listings are still used to recompute the days on market which is equal to the sum of all listing periods of the same house on the market.

Our data contain information on listing agent's ID for all listings, and cooperating agent's ID for sold listings. We use this information to identify dual agency: dual is equal to 1 if listing agent's ID is the same as cooperating agent's ID; and 0 otherwise. In addition, our data include an indicator variable for agent-owned listing, which allows us to identify agent-sellers. As noted by many other papers that used MLS data, our MLS data also contain very rich information on most houses on the market,[11] including listing prices, transaction prices (for sold listings), and detailed housing characteristics, such as the number of bedrooms, the number of bathrooms, house age, and housing tenure.

Table 1 reports summary statistics for the key variables. The first column shows the mean of each variable for all sample which includes both sold listings and withdrawn listings. We also report these means separately for withdrawn listings in column 2, and for sold listings in column 3. In column 1, the average listing price is \$473,282, while the average transaction price is \$433,851.[12] The difference is \$39,431, which seems to be substantial. However, this is mostly due to higher listing prices of withdrawn listings. Once we condition on sold listings (see column 3), transaction prices are lower than listing prices only by \$4,473 on average. The days on market (DOM) are computed by summing across all of a house's listing periods, and its mean is 159 days for all sample. We also report last listing's days on market (LDOM) by excluding days on market of temporarily de-listed listings. Not surprisingly, LDOM is smaller than DOM.

The comparison of columns 2 and 3 shows that the difference in average listing prices between withdrawn listings and sold listings is \$103,496, and the difference in DOM (or LDOM) is 119 days (or 52 days). Higher listing prices and longer DOM for withdrawn listings might result from higher reservation values of these listings' sellers (or agents). As for housing characteristics, sold listings are on average two years older than withdrawn listings, and slightly smaller in terms of the number of bedrooms and bathrooms. Though most listings in our data are condominiums which are common housing tenure in downtown area, sold listings still include more condos than withdrawn listings. These differences between completed transactions and withdrawn listings suggest that using the sample of only completed transactions might entail potential selection issues.

In columns 4 and 5, we present summary statistics for the sample of sold listings with dual agency and that with non-dual agency. The comparison of these two columns indicates that dual agency tends to involve lower listing prices and transaction prices, as well as shorter DOM. This is consistent with commission-related incentives between dual agency and non-dual agency, though testing this is difficult because of selection issues, as discussed in Section 2.4. In addition, columns 4-5 show that houses from dual agency tend to be slightly younger and smaller that those from

[11] However, our data do not include houses that are for sale by owner (FSBO).
[12] We use the Consumer Price Index to deflate prices, and all prices in this paper are in 2010 dollar.

non-dual agency. Columns 6-7 report the mean of each variable for the sample of sold listings owned by agents, vs. clients. These two columns show that agent-owned houses tend to be listed at lower prices, and also sold at lower prices and more quickly than client-owned houses.

3.2 Reduced-Form Results

To further examine the patterns in data, we run regressions of transaction outcomes on the variables related to agent's incentives. We consider the following three variables related to agent's incentives. The first is dual, a dummy for dual agency; the second is $\mathsf{agent\text{-}owned}$, a dummy for listing owned by an agent; and the third is $\mathsf{disclosure}$, a dummy for the period after the requirement for disclosing dual agency was implemented. Note that around 2000, the state of the city considered in this paper implemented a new regulation which required real estate agents engaging in dual agency to inform their clients of dual agency and the nature of their agency relationship. This regulation is similar to the disclosure policy considered in Han and Hong (2016). Because agents under dual agency do not have to split their commissions with a buyer's agent, they are likely to accept lower prices compared to non-dual agency. But the disclosure policy might have informed more sellers about dual agency, which potentially increased the number of sophisticated sellers, thus weakening the commission-related effect of dual agency. As for agent-owned vs. client-owned, the standard prediction, as in Levitt and Syverson (2008a), is that agent-owned properties sell at higher prices and stay on the market longer than client-owned properties. Due to the issues discussed in Section 2.4, however, it is difficult to test these predictions using standard regressions. Therefore, the reduced-form results below provide only correlations between key variables, and do not necessarily provide evidence for or against conflicts of interest between sellers and agents.

In Table 2, we use the sample of sold listings, and regress transaction prices on dual dummy, an interaction between dual and disclosure, agent-owned dummy, as well as various control variables, including housing characteristics, listing prices, and time fixed effects. Column 1 reports the baseline results, while columns 2-4 present the results using different fixed effects. In all columns, we find that the coefficient on dual is negative and statistically significant, while the coefficient on dual$\times$disclosure is positive and statistically significant, which is consistent with the commission-related effects of dual agency. However, the coefficient on agent-owned is negative in columns 1-2, and 4, while it is estimated to be positive in column 3. These results are inconsistent with findings from Hendel, et al. (2009), Levitt and Syverson (2008a), and Rutherford, et al. (2005).

Table 3 reports the results from using two variables related to the probability of transaction. Panel A uses DOM as the dependent variable, while Panel B considers the dummy variable for sold listings as the dependent variable. Both variables are related to the probability of completing transaction, but they do not exactly capture the probability of transaction. In Panel A, we use the sample of sold listings, while in Panel B, we use the sample of both sold and withdrawn listings. Because dual agency is observed only if the listing is sold, dual and dual×disclosure are not included in the regressions in Panel B. The regressions of DOM in Panel A show that listings tend to be sold faster if it involves dual agency, rather than non-dual agency, though this effect becomes weaker once listing agent fixed effects are included in column 4. The disclosure policy seems to have weaken this effect, but all coefficients on dual×disclosure are not estimated precisely. The coefficient on agent-owned is mostly negative, but most coefficients are not estimated precisely, which is also inconsistent with findings in the literature.

In contrast, the regressions of the sold dummy in Panel B show that the coefficient on agent-owned is negative and statistically significant in all columns. Hence, agent-owned listings are more likely to be withdrawn from the market than client-owned listings, which might be due to higher reservation values of agents than those of clients. Of course, this seems to be inconsistent with the preceding results that agent-owned houses are sold at lower prices and more quickly than client-owned houses, because faster sales at lower prices could result from lower reservation prices.

For at least three reasons, however, these reduced-form results do not necessarily provide evidence for or against conflicts of interest between sellers and listing agents. First, houses owned by agents tend to be listed at lower prices in our data, suggesting that agents in this market might have owned cheaper houses than clients. Second, the average transaction prices in Hendel, et al. (2009) and Rutherford, et al. (2005) are lower than \$200,000, and those in Levitt and Syverson (2008) are lower than \$300,000. Accordingly, our data include more expensive houses than the data analyzed by these papers, and home ownership decisions by agents in the downtown market analyzed in this paper might be different from those studied in other papers. Third, regressions of housing prices are likely biased, unless the issues discussed in Section 2.4 are addressed. Hence, by using only reduced-form results from regressions of transaction prices, it is difficult to provide evidence related to conflicts of interest between home sellers and agents. This further motivates our structural approach.

4 Econometric Model

In this section, we first derive the likelihood function to estimate the model developed in Section 2. We then describe details on our estimation approach, and discuss identification of the key parameters in the model.

4.1 Likelihood Function

Our data consist of listings that are withdrawn (i.e., $w_j = 1$) or sold (i.e., $w_j = 0$). For each listing, we observe whether the house is owned by an agent (i.e., $a_j = 1$) or a client (i.e., $a_j = 0$). Accordingly, we first construct the individual likelihood function for agent-owned listings, and then derive the likelihood function for client-owned listings which is a mixture of the sophisticated seller case and the uninformed seller case. The key error terms in our model are $(\epsilon^a, \epsilon^s, \epsilon^u)$ and μ. The former is the error terms in the value function of the seller (or agent), and the latter is the error term in the buyer's utility that captures the unobserved match value specific to each buyer. As discussed in Section 2, we assume that ϵ's follow the extreme value distribution. As for μ, we assume that it is distributed $N(0, \sigma_\mu)$, that is, $G(\mu) = \Phi\left(\frac{\mu}{\sigma_\mu}\right)$, where $\Phi(\cdot)$ is the standard normal distribution function. Because our analysis uses the logarithm of transaction prices, this assumption is essentially the same as assuming the lognormal distribution which is a common distributional assumption in the literature on the labor search model.[13]

The individual likelihood function for agent-owned listing, denoted by ℓ_j^a, is composed of three components: the first part for the events before the listing was withdrawn or sold; the second for withdrawal; and the third component for transaction either through dual agency or non-dual agency. If listing j was withdrawn or sold after T_j days on market, it must be that the agent-seller decided to continue listing the house for each day during the previous $T_j - 1$ days, and for each day, either no buyer visited, or the agent-seller decided not to sell the house to a buyer who visited. Therefore, the first component of ℓ_j^a is written as

$$\left\{ \frac{\exp\left(v_j^a(0)\right)}{1+\exp\left(v_j^a(0)\right)} \left(1 - \lambda_j \theta_j \varphi_j^{a,d} - \lambda_j (1-\theta_j) \varphi_j^{a,n}\right) \right\}^{T_j - 1},$$

where $v_j^a(0)$ is given by (5), and the probability of transaction (i.e., the agent-seller accepting the

[13]Note also that this assumption satisfies the "recoverability condition" in Flinn and Heckman (1982), since the normal distribution can be recovered from what is truncated from below.

buyer's offer) is rewritten as $\varphi_j^{a,d} = 1 - \Phi\left(\frac{EV_j^a - X_j\beta - \xi_j}{\sigma_\mu}\right)$ and $\varphi_j^{a,n} = 1 - \Phi\left(\frac{EV_j^a - \ln\left(1-\frac{\tau}{2}\right) - X_j\beta - \xi_j}{\sigma_\mu}\right)$ respectively for a dual buyer and for a non-dual buyer.

As for λ_j, the buyer's arrival rate, and θ_j, the probability of a dual buyer's visit (conditional on a buyer's visit), we assume that these probabilities take the following functional forms:

$$\lambda_j = \frac{\exp(W_j\delta)}{1+\exp(W_j\delta)}, \quad \theta_j = \frac{\exp(Y_j\eta)}{1+\exp(Y_j\eta)},$$

where W_j is a vector of variables related to a buyer's visit in general (e.g. overall housing market condition) and Y_j is a vector of variables related to a dual buyer's visit, while δ and η are corresponding parameters. As mentioned before, we assume that λ_j and θ_j are not determined by the listing agent's strategic behavior, and consider only a reduced-form model for λ_j and θ_j because of our main focus on strategic interactions between sellers and listing agents. However, one may be still concerned that λ_j and θ_j could be also influenced by the agent's strategic behavior. To examine this issue, note that a dual buyer's visit may arise from three possibilities: first, the buyer who visits the house is not represented by any agent, in which case the listing agent becomes the buyer's agent; second, the buyer is represented by the same listing agent, but is shown the house because the buyer is the best match for the house; third, the same listing agent represents the buyer and shows the house for strategic reasons. Among these three cases, only the third case implies a potential issue.[14] Though it would be ideal to incorporate this case into our model, we acknowledge that it is difficult to model such strategic interactions between agents and buyers, given that our data contain limited information on buyers.[15]

The second component of ℓ_j^a only includes the probability that the agent-seller decides to withdraw, and so it is given by

$$\left\{\frac{1}{1+\exp\left(v_j^a(0)\right)}\right\}^{w_j}.$$

[14]Though the mechanism in the third case cannot be directly examined without modeling interactions between the agent and the buyer, it might be still reflected in our mixture model indirectly. For example, if listing agents' strategic behavior against their client-sellers is further facilitated by their influence on buyers, this strategic effect might further reduce the mixture probability of sophisticated sellers. Though we cannot separate different mechanisms, we might still estimate the overall impact of strategic behavior.

[15]Han and Hong (2016) undertook such modeling work to quantify the extent of strategic promotions by the buyer's agent. However, their structural model is based on hedonic framework which is incompatible with search/bargaining framework in this paper. We hope that future research in this area will improve on these approaches by considering a full equilibrium model for both sides.

Note that $v_j^a(0)$ in (5) includes the following conditional expectation of transaction prices:

$$\begin{aligned} E\left[\ln p_j | a_j = 1, b_j^d = 1, s_j = 1\right] &= \alpha(X_j\beta + \xi_j) + (1-\alpha)EV_j^a + \alpha E(\mu | a_j = 1, b_j^d = 1, s_j = 1), \\ E\left[\ln p_j | \, a_j = 1, b_j^n = 1, s_j = 1\right] &= \alpha(X_j\beta + \xi_j) + (1-\alpha)\left[EV_j^a - \ln(1-\tau/2)\right] \\ &\quad + \alpha E(\mu | a_j = 1, b_j^d = 1, s_j = 1). \end{aligned}$$

Given the normal distribution assumption on μ, it follows that

$$\begin{aligned} E(\mu | a_j = 1, b_j^d = 1, s_j = 1) &= E(\mu | \mu \geq EV_j^a - X_j\beta - \xi_j) = \sigma_\mu \frac{\phi\left(\frac{EV_j^a - X_j\beta - \xi_j}{\sigma_\mu}\right)}{1 - \Phi\left(\frac{EV_j^a - X_j\beta - \xi_j}{\sigma_\mu}\right)}, \\ E(\mu | a_j = 1, b_j^n = 1, s_j = 1) &= \sigma_\mu \frac{\phi\left(\frac{EV_j^a - \ln\left(1-\frac{\tau}{2}\right) - X_j\beta - \xi_j}{\sigma_\mu}\right)}{1 - \Phi\left(\frac{EV_j^a - \ln\left(1-\frac{\tau}{2}\right) - X_j\beta - \xi_j}{\sigma_\mu}\right)}. \end{aligned}$$

The expected value of continued search, $v_j^a(0)$, also includes two parameters ρ and α that cannot be identified given our data. We thus follow the literature and fix these two parameters.[16] To simplify the expression below, we present the likelihood function with $\alpha = 1$. In our actual estimation, however, we experiment with different values for α.

The third component of ℓ_j^a includes the probability of continuing to list the house at period T_j, as well as the probability of a dual (or non-dual) buyer's visit and the probability density function to observe the transaction price. Because reported transaction prices may be measured with errors, we allow for measurement errors. Following the labor search literature (see, e.g., Eckstein and van den Berg, 2007), we assume that $\ln p_j^o = \ln p_j + u_j$, where p_j^o is observed transaction price, and u_j is distributed $N(0, \sigma_u)$ and independent of p_j. Hence, we obtain the following joint probability that the agent-seller agrees to sell the house to a dual buyer and that $\ln p_j^o$ is realized:

$$\left[1 - \Phi\left(\frac{EV_j^a - (1-\varrho^2)(X_j\beta + \xi_j) - \varrho^2 \ln p_j^o}{\sigma_\mu\sqrt{1-\varrho^2}}\right)\right] \times \frac{1}{\sqrt{\sigma_\mu^2 + \sigma_u^2}}\phi\left(\frac{\ln p_j^o - X_j\beta - \xi_j}{\sqrt{\sigma_\mu^2 + \sigma_u^2}}\right),$$

where $\varrho^2 = \frac{\sigma_\mu^2}{\sigma_\mu^2 + \sigma_u^2}$, and $\phi(\cdot)$ is the standard normal probability density function.[17] We can obtain

[16] In the dynamic estimation literature, the discount factor ρ is typically fixed. In the labor search literature, it has been shown that the bargaining parameter α is not identified without additional structure and information. See, e.g., Eckstein and Wolpin (1995), Eckstein and van den Berg (2007), and Flinn (2006). For example, Eckstein and Wolpin (1995) assume that $\alpha = 0.5$. In contrast, Flinn (2006) adds buyer-side (i.e., firm-side) data to his main data on sellers (i.e., workers), and further incorporates additional structure to identify the bargaining parameter. Such information is not available in our data, and so we fix α in our estimation.

[17] See also Eckstein and Wolpin (1995) for a similar expression for this joint probability.

a similar expression for the non-dual agency case. Combining all these components of ℓ_j^a yields

$$\begin{aligned}
\ell_j^a \;=\; & \left\{ \frac{\exp\left(v_j^a(0)\right)}{1+\exp\left(v_j^a(0)\right)} \left(1-\lambda_j\theta_j\varphi_j^{a,d}-\lambda_j(1-\theta_j)\varphi_j^{a,n}\right)\right\}^{T_j-1} \times \left\{\frac{1}{1+\exp\left(v_j^a(0)\right)}\right\}^{w_j} \\
& \times \left\{ \frac{\exp\left(v_j^a(0)\right)}{1+\exp\left(v_j^a(0)\right)} \lambda_j \left(\theta_j \left[1-\Phi\left(\frac{EV_j^a-(1-\varrho^2)(X_j\beta+\xi_j)-\varrho^2\ln p_j^o}{\sigma_\mu\sqrt{1-\varrho^2}}\right)\right]\right)^{d_j}\right. \\
& \times \left((1-\theta_j)\left[1-\Phi\left(\frac{EV_j^a-\ln\left(1-\frac{\tau}{2}\right)-(1-\varrho^2)(X_j\beta+\xi_j)-\varrho^2\ln p_j^o}{\sigma_\mu\sqrt{1-\varrho^2}}\right)\right]\right)^{1-d_j} \\
& \left.\times \frac{1}{\sqrt{\sigma_\mu^2+\sigma_u^2}}\phi\left(\frac{\ln p_j^o-X_j\beta-\xi_j}{\sqrt{\sigma_\mu^2+\sigma_u^2}}\right)\right\}^{1-w_j}
\end{aligned} \tag{16}$$

where d_j is a dummy variable for dual agency.

The individual likelihood function for client-owned listings, denoted by $\ell_j^{s,u}$, is composed of similar components as above, except that it is a mixture of the sophisticated seller's listing and the uninformed seller's listing. In addition, the sophisticated seller's problem remains the same, regardless of dual agency or non-dual agency, which simplifies the likelihood component for the sophisticated seller is simpler than those for other cases. Recall that κ_j denotes the probability that the client-seller of listing j is sophisticated. We assume that κ_j is given by $\kappa_j = \frac{\exp(Z_j\zeta)}{1+\exp(Z_j\zeta)}$, where Z_j is a vector of variables related to the seller's information, and ζ is a vector of parameters. A similar derivation as above yields the following individual likelihood for client-owned listing j:

$$\begin{aligned}
\ell_j^{s,u} \;=\; & \left[\kappa_j \left\{\frac{\exp\left(v_j^s(0)\right)}{1+\exp\left(v_j^s(0)\right)}\left(1-\lambda_j\varphi_j^s\right)\right\}^{T_j-1} \times \left\{\frac{1}{1+\exp\left(v_j^s(0)\right)}\right\}^{w_j}\right. \\
& \times \left\{\frac{\exp\left(v_j^s(0)\right)}{1+\exp\left(v_j^s(0)\right)}\lambda_j\left[1-\Phi\left(\frac{EV_j^s-\ln(1-\tau)-(1-\varrho^2)(X_j\beta+\xi_j)-\varrho^2\ln p_j^o}{\sigma_\mu\sqrt{1-\varrho^2}}\right)\right]\right\}^{1-w_j} \\
& +(1-\kappa_j)\left\{\frac{\exp\left(v_j^u(0)\right)}{1+\exp\left(v_j^u(0)\right)}\left(1-\lambda_j\theta_j\varphi_j^{u,d}-\lambda_j(1-\theta_j)\varphi_j^{u,n}\right)\right\}^{T_j-1} \times \left\{\frac{1}{1+\exp\left(v_j^u(0)\right)}\right\}^{w_j} \\
& \times \left\{\frac{\exp\left(v_j^u(0)\right)}{1+\exp\left(v_j^u(0)\right)}\lambda_j\left(\theta_j\left[1-\Phi\left(\frac{E\pi_j^u-\ln(\tau)-(1-\varrho^2)(X_j\beta+\xi_j)-\varrho^2\ln p_j^o}{\sigma_\mu\sqrt{1-\varrho^2}}\right)\right]\right)^{d_j}\right. \\
& \left.\left.\times\left((1-\theta_j)\left[1-\Phi\left(\frac{E\pi_j^u-\ln\left(\frac{\tau}{2}\right)-(1-\varrho^2)(X_j\beta+\xi_j)-\varrho^2\ln p_j^o}{\sigma_\mu\sqrt{1-\varrho^2}}\right)\right]\right)^{1-d_j}\right\}^{1-w_j}\right] \\
& \times\left[\frac{1}{\sqrt{\sigma_\mu^2+\sigma_u^2}}\phi\left(\frac{\ln p_j^o-X_j\beta-\xi_j}{\sqrt{\sigma_\mu^2+\sigma_u^2}}\right)\right]^{1-w_j}
\end{aligned} \tag{17}$$

where $v_j^s(0)$ and $v_j^u(0)$ are respectively given in (11) and (14), $\varphi_j^s = 1 - \Phi\left(\frac{EV_j^s - \ln(1-\tau) - X_j\beta - \xi_j}{\sigma_\mu}\right)$, $\varphi_j^{u,d} = 1 - \Phi\left(\frac{E\pi_j^u - \ln(\tau) - X_j\beta - \xi_j}{\sigma_\mu}\right)$, and $\varphi_j^{u,n} = 1 - \Phi\left(\frac{E\pi_j^u - \ln(\tau/2) - X_j\beta - \xi_j}{\sigma_\mu}\right)$. Similar to $v_j^a(0)$, $v_j^s(0)$ and $v_j^u(0)$ include the conditional expectation of prices as follows:

$$\begin{aligned}
E\left[\ln p_j | a_j = 0, k_j = 1, b_j = 1, s_j = 1\right] &= X_j\beta + \xi_j + \sigma_\mu \frac{\phi\left(\frac{EV_j^s - \ln(1-\tau) - X_j\beta - \xi_j}{\sigma_\mu}\right)}{1 - \Phi\left(\frac{EV_j^s - \ln(1-\tau) - X_j\beta - \xi_j}{\sigma_\mu}\right)}, \\
E\left[\ln p_j | a_j = 0, k_j = 0, b_j^d = 1, s_j = 1\right] &= X_j\beta + \xi_j + \sigma_\mu \frac{\phi\left(\frac{E\pi_j^u - \ln(\tau) - X_j\beta - \xi_j}{\sigma_\mu}\right)}{1 - \Phi\left(\frac{E\pi_j^u - \ln(\tau) - X_j\beta - \xi_j}{\sigma_\mu}\right)}, \\
E\left[\ln p_j | a_j = 0, k_j = 0, b_j^n = 1, s_j = 1\right] &= X_j\beta + \xi_j + \sigma_\mu \frac{\phi\left(\frac{E\pi_j^u - \ln(\tau/2) - X_j\beta - \xi_j}{\sigma_\mu}\right)}{1 - \Phi\left(\frac{E\pi_j^u - \ln(\tau/2) - X_j\beta - \xi_j}{\sigma_\mu}\right)}.
\end{aligned}$$

Combining (16) and (17), we finally obtain the log likelihood function as

$$\ln L = \sum_{j=1}^{N} a_j \ln \ell_j^a + (1 - a_j) \ln \ell_j^{s,u}, \tag{18}$$

where N is the number of observations.

4.2 Estimation

The estimation of our model is essentially done by maximum likelihood (ML) estimation. However, a crucial difficulty in the ML approach is that we do not directly observe EV_j^a, EV_j^s, and $E\pi_j^u$. A typical approach in the labor search literature is to treat such a reservation value as a parameter, and estimate it along with other parameters (see, e.g., Eckstein and van den Berg, 2007). In contrast, our approach is to incorporate estimation methods from the literature on discrete choice dynamic models. In particular, we exploit additional information on withdrawal decisions in our data, and employ the conditional choice probabilities (CCP) and the inversion approach developed by Hotz and Miller (1993). Because our model contains a mixture of the sophisticated seller case and the uninformed seller case, however, we cannot estimate CCP for sophisticated sellers or uninformed sellers directly. For this reason, we use the iterated version of the Hotz and Miller's approach, that is, the nested pseudo-likelihood (NPL) algorithm developed by Aguirregabiria and Mira (2002).

Specifically, our NPL algorithm proceeds as follows. We begin by estimating the probability of withdrawal using the sample of both withdrawn and sold listings. We then use these initial withdrawal probabilities to compute the initial values for EV_j^a using (9), and similarly for EV_j^s and $E\pi_j^u$. Next, we treat EV_j^a, EV_j^s, and $E\pi_j^u$ as known values and estimate the full model by maximizing the likelihood function in (18). However, our initial withdrawal probabilities will not be correct estimates of CCP, because they do not account for our mixture model, and the sample includes only the withdrawal decisions at the last period. As a result, we use the estimated parameters to recompute the withdrawal probabilities and update continuation values, that is, EV_j^a, EV_j^s, and $E\pi_j^u$. We then re-estimate the full model. We repeat this procedure until it converges.[18]

4.3 Identification

The key primitives of our model include continuation values, listing cost c_j, and parameters in λ_j, θ_j, and κ_j. This section discusses identification of these parameters. As mentioned in the previous section, we do not directly observe continuation values, and a typical approach in the labor search literature is to estimate them as model parameters, in which case the parameter for search cost or instantaneous utility during search is recovered using the estimated continuation value and the equation for the value of continued search (see, e.g., Flinn, 2006; Eckstein and van den Berg, 2007). This approach has been used in the labor search literature, because of under-identification in the basic search model (see Flinn and Heckman, 1982).

In contrast, we overcome this under-identification issue by exploiting the additional information on withdrawal decisions, and further employ the NPL algorithm to recover continuation values. Moreover, the variation in commission incentives between dual agency and non-dual agency for different types of sellers and agents allows us to construct different value functions for each type of seller/agent, which further enables us to recover EV_j^a, EV_j^s, and $E\pi_j^u$ separately. Note also that recovered continuation values are critical in addressing selection biases and omitted biases discussed in Section 2.4. As is clear from the equation (15), these biases arise because we do not directly observe reservation values.

With regard to other key parameters, their identification relies mainly on exclusion restrictions, in that some variables enter only one equation but not the other equations. We assume that $c_j = K_j\omega$, where K_j includes variables related to listing costs, and ω contains corresponding

[18] Pesendorfer and Schmidt-Dengler (2010) point out that the sequential NPL method for estimating dynamic games can lead to inconsistent estimates. Note, however, that this criticism is not about the NPL method developed in Aguirregabiria and Mira (2002), but about the NPL method for estimating dynamic games (Aguirregabiria and Mira, 2007) which may have difficulties converging to all equilibria in multi-player games.

parameters. As for excluded variables in K_j, we consider variables related to the competition between listing agents. Hsieh and Moretti (2003) and Han and Hong (2011) show that an increase in the number of real estate agents would lead to more intense non-price competition between agents, thus increasing their costs in selling houses. Accordingly, K_j includes the number of agents in each local market. However, local competition between listing agents, or the number of agents in each local market, is unlikely to affect λ_j, θ_j, and κ_j directly.

As for parameters in λ_j, θ_j, and κ_j, recall that $\lambda_j = \frac{\exp(W_j\delta)}{1+\exp(W_j\delta)}$, $\theta_j = \frac{\exp(Y_j\eta)}{1+\exp(Y_j\eta)}$, and $\kappa_j = \frac{\exp(Z_j\zeta)}{1+\exp(Z_j\zeta)}$, where W_j is a vector of variables related to a buyer's visit in general, Y_j is a vector of variables related to a dual buyer's visit, and Z_j is a vector of variables related to the seller's information. Regarding W_j, we consider variables capturing overall housing market conditions and buyer's search costs that are not directly related to listing costs, or θ_j, or κ_j. For example, these variables include some city-level housing price indices or variables that indicate hot market or buyer's market (e.g. local average DOM; winter vs. summer). Because θ_j is the conditional probability that a buyer is represented by the same listing agent, conditional on a buyer's visit, Y_j does not necessarily include variables from W_j. For a listing represented by a given listing agent, θ_j is likely to be lower if there are more buyers represented by other listing agents (different from the given listing agent). For this reason, we include the number of such buyers in Y_j. As for κ_j, the client-seller of listing j is more likely to be sophisticated if more information is available to the seller. For instance, the introduction of disclosure requirement for dual agency could inform sellers more about the agency relationship in real estate transactions. Advances in the Internet have also made sellers acquire more information on housing markets, thus allowing them to become more sophisticated. Hence, the local adoption rate of the Internet could be included in Z_j.

Lastly, note that the model contains ξ_j, unobserved house characteristics. To address this issue, we use detailed house characteristics as control variables. In addition, we include listing prices as another control variable. Note that listing prices are likely to reflect not only observed house quality, but also unobserved house quality. We might also consider alternative approaches such as those used by Bajari and Benkard (2005) and Bajari and Kahn (2005), in which they recover ξ_j from nonparametric regressions of housing prices. However, their approaches build upon the hedonic framework that assumes competitive housing markets, which is inconsistent with the search framework considered in this paper. Therefore, using their approach requires considerable modifications of the hedonic framework, which is beyond the scope of this paper.

5 Estimation Results

5.1 Parameter Estimates

This section presents the results from the estimation of our model. Table 4 reports the parameter estimates for each component of our model.[19] Panel A includes the parameter estimates for house characteristics, $X_j\beta$, which enter the buyer's utility, capturing a common value (excluding a buyer-specific match value μ_{ij}). In a way to account for unobserved house characteristics, we also include listing prices as a control variable. In Panel B, we report the parameter estimates for λ_j, the arrival rate of a buyer. Because the buyer's arrival rate is related to housing market conditions, we consider the following variables that reflect local housing markets. We first include the monthly S&P/Case-Shiller Home Price Index for the city studied in this paper which captures overall housing market conditions. Because housing market conditions vary across different geographic areas, we also use average days on market (DOM) for each census tract, where the average DOM is computed for all listings (excluding listing j) in the same census tract as listing j. Note that the census tract-level average DOM is computed for each year, but not for year-month, because some census tracts contain a very small number of listings in some year-months. To capture seasonal changes at detailed geographic areas, we use average DOM for each zip code and every year-month, where the average is computed by excluding listing j. All coefficients for these variables are estimated precisely, and their signs are reasonable: a higher housing market index is related to a housing boom, in which more buyers enter the market; a shorter DOM is also correlated with a housing boom, which is consistent with negative coefficients on both measures of average DOM.

In Panel C of Table 4, we report the parameter estimates for θ_j, the conditional probability that the buyer who visited the house is represented by the same agent who lists the house. We include the ratio of buyers represented by agents other than the listing agent of listing j, relative to total number of buyers who purchased houses in a given year. The coefficient on this variable is negative, as expected, and it is estimated precisely. Panel D reports the parameter estimates for listing cost c_j. We consider the number of listing agents who listed any house in the MLS each year. The coefficient estimate for this variable is positive, suggesting that more agents in the market lead to more intense competition between agents, thus increasing listing costs. In Panel

[19] We report the preliminary estimates with a rather simple specification. However, this specification might be too parsimonious, given complexity of housing markets. We will consider more specifications and check the robustness of the estimates in our future version of the draft.

E, we present the parameter estimates for κ_j, the probability that a client-seller is sophisticated. We first include census tract-level information on higher education (BA degree and master degree) obtained from 2000 Census data. We also include the dummy variable for the implementation of a legislation that requires agents to disclose their dual agency relationship to their clients. The coefficient estimate for this variable is positive and statistically significant, suggesting that the introduction of the disclosure policy might have informed more sellers about dual agency as well as agency relationship, thus increasing κ_j.[20] Lastly, we consider the number of listings sold by the agent for listing j, excluding listing j. The reason behind this variable is that as agents are more experienced in carrying out transactions, they are likely to be much more informed than clients, thereby reducing the probability that the client is more informed. Confirming this reason, the coefficient estimate for this variable is negative and is also estimated precisely.

5.2 Implications of Estimated Model

The estimation of our model involves the computation of reservation values, i.e. continuation values from continuing to list the house. For listings by agent-sellers, we recover EV_j^a, and for listings by client-sellers, we recover both EV_j^s for the sophisticated seller case and $E\pi_j^u$ for the uninformed seller case. Since we use the logarithm of prices, we can consider these reservation values relative to the logarithm of prices. In Table 5, we present summary statistics of these recovered reservation values. The table shows that the mean of EV_j^a is about 11.49, while the mean of EV_j^s is 13.42. The mean of $E\pi_j^u$ is about 9.16. To interpret the value of $E\pi_j^u$, note that the agent for an uninformed seller agrees to sell the house if $\ln(\tau p_j) \geq E\pi_j^u$, i.e., $\ln p_j \geq E\pi_j^u - \ln\tau$ for dual agency, and if $\ln(\frac{\tau}{2}p_j) \geq E\pi_j^u$, i.e., $\ln p_j \geq E\pi_j^u - \ln\frac{\tau}{2}$ for non-dual agency. Hence, the reservation value of an uninformed seller who completely follows the agent's advice will be given by $E\pi_j^u - \ln\tau$ for dual agency and $E\pi_j^u - \ln\frac{\tau}{2}$ for non-dual agency. Accordingly, the mean value of 9.16 for $E\pi_j^u$ would imply that the uninformed seller's reservation value is about 12.16 for dual agency, and 12.85 for non-dual agency.[21]

Note that the mean of EV_j^a is lower than the mean of EV_j^s as well as the mean of the uninformed seller's reservation value. Because the reservation value of a client-seller is a mixture between

[20] Of course, this variable might also capture other time effects, but we do not attempt to isolate the effect of the disclosure policy in the current version. Nevertheless, the estimate indicates that the disclosure policy and other factors such as the advance in the Internet, which coincided with the introduction of the disclosure policy, have increased the probability of sophisticated sellers.

[21] We assume that $\tau = 0.05$, and so $\ln\tau = -2.996$, while $\ln\frac{\tau}{2} = -3.69$.

those of a sophisticated seller and an uninformed seller, this result implies that the agent-seller's reservation value is lower than the client-seller's reservation value in our data. Recall that the patterns reported in Section 3 show negative correlations between transaction prices and agent-owned properties, that is, agent-owned properties tend to be sold at lower prices than client-owned properties. If we had relied only on these reduced-form results, we might have incorrectly concluded that these correlations provided evidence against a misalignment of incentives between the agent and the seller. However, the estimates from our structural estimation indicate that the aforementioned negative correlations are simply due to lower reservation values of the agent-seller relative to those of the client-seller.

In Tables 6-7, we report summary statistics of the probabilities in the model recovered from the estimation. Table 6 shows that the buyer's arrival rate λ_j is 0.046 on average, whereas the conditional probability for a dual buyer θ_j (the probability that a buyer who visits the house is represented by the same agent who lists the house, conditional that a buyer visits the house) is 0.14 on average. Note that the fraction of dual agency observed in our data is about 0.2, that is, two out of 10 completed transactions are dual agency in our data. Hence, to the extent that the recovered θ_j captures exogenous factors for dual agency, the remaining fraction, 0.06, or 30% of dual agency, might have resulted from strategic behavior or the distortion in incentives. Note that the listing agent's lower reservation value under dual agency (relative to that under non-dual agency) is likely to result in a lower transaction price, thereby increasing the probability of completing the transaction for dual agency (relative to non-dual agency). In other words, if the listing agent's incentive had been fully aligned with the seller, the remaining 30% of dual agency might have occurred.

In Table 7, we find that κ_j, that is, the probability that a client-seller is sophisticated, is about 0.32 on average. As discussed before, we use a mixture of two polar cases – the sophisticated seller case for fully aligned incentives, and the uninformed seller case for fully misaligned incentives – in order to measure different degrees of conflicts of interest between sellers and agents, without solving potentially complex games between them. Hence, κ_j can be considered as a measure for the degree of the distortion in incentives of the listing agent, in that a lower κ_j indicates a higher degree of conflicts of interest. Therefore, 0.32 means that with the probability 0.32, listing agents are likely to breach their fiduciary duties to their clients by not fully representing the best interest of their clients. We also compute summary statistics of the recovered κ_j for the periods before and after the implementation of a legislation that requires agents to disclose their dual agency and the nature of their agency relationship to their clients. The table shows that the mean of κ_j before the disclosure policy is about 0.26, while that after the policy change is about 0.34. Thus, after the disclosure policy was introduced, κ_j has increased by a 8 percentage point on average.

5.3 Welfare Implications

Using the estimated model, we examine the importance of conflicts of interest in terms of their effects on transaction outcomes and welfare loss. In this version, we focus on transaction prices, since they are the key transaction outcome as well as the key factor to determine the seller's welfare. In that regard, we begin by computing expected log prices for different cases. Because conflicts of interest occur to the relationship between client-sellers and their agents, we use the sample of listings by client-sellers, and compute the expected $\ln p_j$, conditional on sales, as well as dual (or non-dual) agency. In Table 8, we consider two polar cases of the sophisticated seller case and the uninformed seller case, and report the expected $\ln p_j$ for both cases. In the table, $E[\ln p|k = 1, d = 1, s = 1, a = 0]$ means the expected log price, conditional on conditional on sophisticated seller ($k = 1$), dual agency ($d = 1$), transaction ($s = 1$), and client-seller ($a = 0$). These expected log prices are computed by accounting for selection biases (i.e. prices are observed only if they are higher than reservation values). Panel A (or Panel B) presents summary statistics of these expected log prices for dual agency (or non-dual agency). The table shows that the expected log prices are lower for the uninformed seller case than for the sophisticated seller case. The table also shows that the difference between these two polar cases is larger under dual agency than under non-dual agency, which is mainly because the expected log prices for the sophisticated seller cases are very similar, regardless of dual agency or non-dual agency, whereas those for the uninformed seller case are lower under dual agency than under non-dual agency. These results are consistent with the implications of our model.

Because the comparison of the two polar cases above might not be useful by itself, however, we further consider mixture cases under different values of κ_j in order to evaluate the impacts of different degrees of incentive misalignment on transaction prices. We consider three cases. The first is the benchmark case, using the current estimates for κ_j. For the second case, we increase κ_j by 0.08, that is, the value by which κ_j has increased after the implementation of the disclosure policy. The third case increases κ_j even more by 0.2. Both the second and third cases reflect more aligned incentives between sellers and agents. For each case, we use the sample of client-owned listings, and compute the expected $\ln p$ by $\hat{\kappa} \times E[\ln p|k = 1, d, s = 1, a = 0] + (1-\hat{\kappa}) \times E[\ln p|k = 0, d, s = 1, a = 0]$, where $\hat{\kappa}$ is κ used for each case. The results are presented in Table 9. Panel A (or Panel B) reports the results for dual agency (or non-dual agency). In Panel A, we find that increasing κ by 0.08 increases the average log price from 12.828 to 12.896, or by 0.068. Since we are comparing the logarithm of prices, 0.068 can be interpreted as an increase in prices by 6.8%. Similarly, we find that increasing κ by 0.2 results in an increase in transaction prices by 17.1%. These results clearly illustrate that reducing conflicts of interest (by increasing κ) would lead to a higher transaction price, thus benefiting the seller more. Panel B reports similar findings, but the magnitude of the change is lower compared to Panel A. In terms of percentage changes, increasing κ by 0.08 (or 0.2) would lead to an increase in transaction price by 4.3% (or 10.8%) when the transaction is done through non-dual agency. Note that the impacts on transaction prices are larger for dual agency than for non-dual agency, which also indicates that the degree of conflicts of interest is higher for dual agency than non-dual agency. To sum up, the results above show that conflicts of interest have sizable impacts on transaction prices.

6 Conclusion

In this paper, we investigate to what extent real estate agents might be breaching their fiduciary duties to their clients. To this end, we develop a structural housing search model that encompasses a mixture of two polar cases – fully aligned incentives and fully misaligned incentives. We estimate the model using the listing data that contain both sold listings and withdrawn listings. Using the estimated model, we find a high degree of conflicts of interest between sellers and agents. We also find that the distortion in incentives can be mitigated by the disclosure policy or any factor that could make sellers more informed.

In addition, our model is useful in characterizing potential biases in reduced-form regressions of transaction outcomes using the data from completed transactions only. Our results illustrate these issues, in that the patterns in our data show lower transaction prices of agent-owned properties compared to client-owned properties, from which one may incorrectly infer that there is no misalignment of incentives between sellers and agents in the market studied in this paper. Our approach addresses these issues, because we can recover reservation values of sellers and agents. Using the estimated model, we find that agent-sellers tend to have lower reservation values than client-sellers in our data. This difference in reservation values is more likely to explain the negative correlation between transaction prices and agent-owned properties in our data, rather than the absence of incentive misalignment, particularly given that we find a high degree of conflicts of interest from our structural estimation. These results therefore suggest that reduced-form regressions using transaction data may entail biases unless we account for selection bias and omitted variable bias stemming from unobserved reservation values. The approach developed in this paper can provide a solution to address these issues.

References

Aguirregabiria, V. and Mira, P. (2002), "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models," *Econometrica*, *70(4)*, pp. 1519-1543.

Aguirregabiria, V. and Mira, P. (2007), "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, *75(1)*, pp. 1-53.

Bajari, P. and Benkard, C. L. (2005), "Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach," *Journal of Political Economy*, *113(6)*, pp. 1239-1276.

Bajari, P. and Kahn, M. (2005), "Estimating Housing Demand with an Application to Explaining Racial Segregation in Cities," *Journal of Business and Economic Statistics*, *23(1)*, pp. 20-33.

Bernheim, D. and Meer, J. (2012), "Do Real Estate Brokers Add Value When Listing Services are Unbundled?," *Economic Inquiry*, *51*, pp. 1166-1182.

Carillo, P. (2012), "An Empirical Stationary Equilibrium Search Model of the Housing Market," *International Economic Review*, *53*, pp. 203-234.

Eckstein, Z. and van den Berg, G. (2007), "Empirical Labor Search: A Survey," *Journal of Econometrics*, 136, pp. 531-564.

Eckstein, Z., and Wolpin, K. (1995), "Duration to First Job and the Return to Schooling: Estimates from a Search-Matching Model," *Review of Economic Studies*, 62, pp. 263-286.

Evans, R. and Kolbe, P. (2005), "Homeowners' Repeat Sales Gains, Dual Agency and Repeated Use of the Same Agent," *Journal of Real Estate Research*, *27(3)*, pp 267-92.

Flinn, C. (2006), "Minimum Wage Effects on Labor Market Outcomes under Search, Matching, and Endogenous Contact Rates," *Econometrica*, 74, pp. 1013-1062.

Flinn, C. and Heckman, J. (1982), "New Methods for Analyzing Structural Models of Labor Force Dynamics," *Journal of Econometrics*, *18*, pp. 115-168.

Garmaise, M. and Moskowitz, T. (2004), "Confronting Information Asymmetries: Evidence from Real Estate Markets," *Review of Financial Studies*, *17(2)*, pp. 405-37.

Gardiner, J., Heisler, J., Kallberg, J. and Liu, C. (2007), "The Impact of Dual Agency," *Journal of Real Estate Finance and Economics*, *35*, pp. 39-55.

Gruber, J. and Owings, M. (1996), "Physician Financial Incentives and Cesarean Section Delivery," *RAND Journal of Economics*, *27(1)*, pp. 99-123.

Han, L. and Hong, S.-H. (2011), "Testing Cost Inefficiency Under Free Entry in the Real Estate Brokerage Industry," *Journal of Business and Economic Statistics*, *29(4)*, pp. 564-578.

Han, L. and Hong, S.-H. (2016), "Understanding In-House Transactions in the Real Estate Brokerage Industry," *RAND Journal of Economics*, forthcoming.

Hendel, I., Nevo, A., and Ortalo-Magne, F. (2009), "The Relative Performance of Real Estate Marketing Platforms: MLS versus FSBOMadison.com." *American Economic Review*, *99*, pp.1878-98.

Hotz, J. and Miller, R. (1993), "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, *60*, pp. 497-529.

Hotz, J., Miller, R, Sanders, S., and Smith, J. (1994), "A Simulation Estimator for Dynamic Models of Discrete Choice," *Review of Economic Studies*, *61*, pp. 265-289.

Hsieh, C. and Moretti, E. (2003), "Can Free Entry Be Inefficient? Fixed Commissions and Social Waste in the Real Estate industry," *Journal of Political Economy*, *111(5)*, pp. 1076-1122.

Hubbard, T. (1998), "An Empirical Examination of Moral Hazard in the Vehicle Inspection Market," *RAND Journal of Economics*, *29*, pp. 406-426.

Kadiyali, V., Prince J., and Simon D. (2012). "Is Dual Agency in Real Estate Transactions a Cause for Concern?" *Journal of Real Estate Finance and Economics*, *22*, pp. 521.

Levitt, S., and Syverson, C. (2008a), "Market Distortions When Agents Are Better Informed: The Value of Information in Real Estate," *Review of Economics and Statistics*, *90(4)*, 599-611.

Levitt, S., and Syverson, C. (2008b), "Antitrust Implications of Home Seller Outcomes when using Flat-Fee Real Estate Agents" *Brookings-Wharton Papers on Urban Affairs*.

Mehran, H. and Stulz R. (2007), "The Economics of Conflicts of Interest in Financial Institutions," *Journal of Financial Economics*, 85 (2), 267-96.

Merlo, A. and Ortalo-Magne, F. (2004), "Bargaining over Residential Real Estate: Evidence from England," *Journal of Urban Economics*, 56 (2), 192-216.

Merlo, A., Ortalo-Magne, F., and Rust, J. (2015), "The Home Selling Problem: Theory and Evidence," *International Economic Review*, *56*, pp. 457-484.

Pesendorfer, M. and Schmidt-Dengler, P. (2010), "Sequential Estimation of Dynamic Discrete Games: A Comment," *Econometrica*, *78(2)*, pp. 833-842.

Rust, J. (1987), "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, *55*, pp. 999-1033.

Rutherford, R., Springer, T., and Yavas, A. (2005), "Conflicts between Principals and Agents: Evidence from Residential Brokerage," *Journal of Financial Economics*, 76 (3), 627-65.

Table 1: Summary Statistics[a]

	all (1)	withdrawn (2)	sold (3)	dual (4)	non-dual (5)	agent-owned (6)	client-owned (7)
listing price	473,282	541,820	438,324	421,723	442,512	393,602	441,931
transaction price	433,851	N/A	433,851	418,561	437,707	387,866	437,559
DOM (all listings)	159	238	119	110	121	101	121
LDOM (last listing)	90	124	72	64	74	56	73
number of bedrooms	1.75	1.84	1.71	1.66	1.72	1.62	1.71
number of baths	1.63	1.68	1.60	1.56	1.61	1.50	1.61
house age	18.19	16.56	18.99	18.01	19.24	18.56	19.03
condo	0.88	0.84	0.91	0.87	0.91	0.89	0.91

[a]The table reports the mean of each variable. Each column contains summary statistics for the given sample. All sample includes both sold listings and withdrawn listings. Dual sample consists of sold listings with the same agent representing both seller and buyer. Agent-owned (or client-owned) sample includes sold listings owned by agent-sellers (or client-sellers). Both listing prices and transaction prices are deflated by the Consumer Price Index, and their values are in 2010 dollar. DOM (all listings) means days on market for all listing periods (summed across all of a house's listing periods), and LDOM (last listing) means last listing's days on market, thus excluding days on market of listings temporarily de-listed. Condo is the indicator variable for whether housing tenure is condo.

Table 2: Regression Results for Transaction Prices[a]

	dependent variable: ln(transaction price)			
	(1)	(2)	(3)	(4)
dual	-0.039***	-0.025***	-0.009*	-0.031***
	(0.004)	(0.004)	(0.004)	(0.004)
dual×disclosure	0.048***	0.034***	0.014**	0.034***
	(0.005)	(0.005)	(0.005)	(0.005)
agent-owned	-0.023***	-0.020***	0.025***	-0.007+
	(0.004)	(0.004)	(0.004)	(0.005)
house characteristics	yes	yes	yes	yes
listing price	yes	yes	yes	yes
year×month fixed effects	yes	yes	yes	yes
census track fixed effects	no	yes	yes	yes
house fixed effects	no	no	yes	no
listing agent fixed effects	no	no	no	yes
observations	53256	53230	32982	48891
R^2	0.850	0.879	0.990	0.894

[a]The sample includes only sold listings. Dual is a dummy for whether the same agent represented both buyer and seller in a transaction. Disclosure is a dummy for the period after disclosure requirement for dual agency was implemented. Agent-owned is a dummy for whether the listing is owned by an agent. House characteristics include dummy variables for condo and attached house, as well as dummy variables for #bedrooms, #rooms, #bathrooms, #garages, basement types, and house age. Column 2 excludes listings for which census track numbers cannot be assigned. Column 3 excludes properties that were sold only once during our sample period. Column 4 excludes listings handled by a listing agent who sold a house only once in our data. * denotes significance at a 10% level, ** denotes significance at a 5% level, and *** denotes significance at 1% level.

Table 3: Regression Results for Transaction Probability [a]

	A. dependent variable: DOM			
	(1)	(2)	(3)	(4)
dual	-14.81***	-12.70***	-10.09+	-6.67+
	(3.56)	(3.56)	(6.00)	(3.83)
dual×disclosure	0.53	1.07	7.34	3.42
	(4.34)	(4.34)	(7.93)	(4.65)
agent-owned	-4.37	-5.54+	0.80	-3.17
	(3.11)	(3.12)	(5.69)	(4.01)
house characteristics	yes	yes	yes	yes
listing price	yes	yes	yes	yes
year×month fixed effects	yes	yes	yes	yes
census track fixed effects	no	yes	yes	yes
house fixed effects	no	no	yes	no
listing agent fixed effects	no	no	no	yes
observations	53256	53230	32982	48891
R^2	0.102	0.120	0.823	0.236
	B. dependent variable: dummy for sold			
	(1)	(2)	(3)	(4)
agent-owned	-0.101***	-0.101***	-0.051***	-0.030***
	(0.006)	(0.006)	(0.012)	(0.007)
house characteristics	yes	yes	yes	yes
listing price	yes	yes	yes	yes
year×month fixed effects	yes	yes	yes	yes
census track fixed effects	no	yes	yes	yes
house fixed effects	no	no	yes	no
listing agent fixed effects	no	no	no	yes
observations	79889	79829	49803	69998
R^2	0.131	0.146	0.611	0.247

[a]Panel A includes the sample of sold listings, while Panel B includes the sample of both sold and withdrawn listings. DOM is days on market for all listing periods (summed across all of a house's listing periods). The dummy for sold is equal to 1 if the listing is sold; 0 if the listing is withdrawn. Dual is a dummy for whether the same agent represented both buyer and seller in a transaction. Disclosure is a dummy for the period after disclosure requirement for dual agency was implemented. Agent-owned is a dummy for whether the listing is owned by an agent. House characteristics include dummy variables for condo and attached house, as well as dummy variables for #bedrooms, #rooms, #bathrooms, #garages, basement types, and house age. Column 2 excludes listings for which census track numbers cannot be assigned. Column 3 in Panel A (or Panel B) excludes properties that were sold (or listed) only once during our sample period. Column 4 in Panel A (or Panel B) excludes listings handled by a listing agent who sold (or listed) a house only once in our data. * denotes significance at a 10% level, ** denotes significance at a 5% level, and *** denotes significance at 1% level.

Table 4: Parameter Estimates for Structural Model[a]

A. house characteristics		C. θ: prob. of dual buyer's visit	
# beds	0.0849***	ratio of buyers represented	-32.0443***
	(0.0107)	by other agents each year	(0.6857)
# rooms	-0.0171**	constant	30.1001***
	(0.0053)		(0.6801)
# baths	-0.0792***	D. c: listing cost	
	(0.0125)	# agents who listed	0.0015*
# garages	-0.0096	any house (by year)	(0.0006)
	(0.0080)	constant	-1.6489***
condo	-0.0479**		(0.2490)
	(0.0165)	E. κ: prob. of sophisticated seller	
ln(listing price)	0.5591***	# people w/BA degree	-0.0016***
	(0.0140)	(census tract-level)	(0.0001)
constant	5.5348***	# people w/master degree	0.0001*
	(0.1702)	(census tract-level)	(0.0001)
B. λ: arrival rate of buyer		#total population	0.0003***
Case-Shiller monthly index	0.0140***	(census tract-level)	(0.0000)
	(0.0002)	disclosure	0.2092***
zip code-level average DOM	-0.0132***		(0.0096)
(by year-month)	(0.0002)	agent-level # listings sold	-0.0005***
census tract-level average DOM	-0.0108***		(0.0001)
(by year)	(0.0002)	constant	-0.1886***
constant	-2.9480***		(0.0098)
	(0.0328)		

[a]Each panel in the table reports parameter estimates for each component in the likelihood function in (16) and (17). The model is estimated by using the nested pseudo-likelihood algorithm. In Panel A, Case-Shiller monthly index is the S&P/Case-Shiller Home Price Index for the city studied in this paper. Zip code-level (or census tract-level) average DOM is the average DOM of listings in the same zip code (or census tract) as a given listing, excluding the given listing. In Panel C, the ratio of buyers represented by other agents (i.e. agents other than the listing agent of a given listing) is the ratio against total number of buyers who purchased houses each year. In Panel D, # agents who listed any house is the number of listing agents who listed any house in the MLS each year. In Panel E, # people w/BA degree or master degree is the number of people with a given degree surveyed in 2000 Census. Disclosure is a dummy for the period after disclosure requirement for dual agency was implemented. Agent-level # listings sold is the number of listings sold by the listing agent of the given listing, excluding the given listing (if sold). * denotes significance at a 10% level, ** denotes significance at a 5% level, and *** denotes significance at 1% level.

Table 5: Recovered Reservation Values[a]

	mean (1)	25th pct (2)	median (3)	75th pct (4)
agent-seller case: EV^a	11.486	11.043	11.417	11.861
sophisticated seller case: EV^s	13.423	13.023	13.342	13.747
uninformed seller case: $E\pi^u$	9.164	8.561	9.153	9.816

[a]The table presents summary statistics of reservation values (i.e. continuation values) recovered from our estimation. 25th pct means 25th percentile. Both the agent-seller case and the sophisticated seller case report the seller's reservation value, whereas the uninformed seller case reports the reservation value of the listing agent who represents an uninformed seller. All values are relative to the logarithm of prices.

Table 6: Recovered Probabilities [a]

	mean (1)	25th pct (2)	median (3)	75th pct (4)
λ: arrival rate of buyer	0.046	0.023	0.039	0.059
θ: prob. of dual buyer's visit	0.141	0.126	0.130	0.142

[a]The table reports summary statistics of the recovered probabilities from the estimation. λ is the arrival rate of buyer, that is, the probability that a buyer visits a house each day. θ is the probability that a buyer who visits the house is represented by the same agent who lists the house (conditional that a buyer visits the house).

Table 7: Recovered κ: Prob. of Sophisticated Seller[a]

	mean (1)	25th pct (2)	median (3)	75th pct (4)
all sample	0.317	0.220	0.312	0.436
before disclosure policy	0.256	0.163	0.261	0.301
after disclosure policy	0.335	0.229	0.324	0.446

[a]The table reports summary statistics of recovered κ, that is, the probability that a client-seller is sophisticated. All sample includes both the periods before and after the introduction of new legislation including the disclosure policy. Before disclosure policy means the sample before the policy change, while after disclosure policy means the same after the policy change.

Table 8: Welfare Implications of Incentive Misalignment: Comparing Two Polar Cases[a]

	mean (1)	25th pct (2)	median (3)	75th pct (4)
	A. dual agency			
$E[\ln p \mid k=0, d=1, s=1, a=0]$	12.567	12.214	12.558	12.918
$E[\ln p \mid k=1, d=1, s=1, a=0]$	13.422	13.069	13.369	13.709
$E[\ln p \mid k=1, d=1, \ldots] - E[\ln p \mid k=0, d=1, \ldots]$	0.865	0.512	0.861	1.185
	B. non-dual agency			
$E[\ln p \mid k=0, d=0, s=1, a=0]$	12.932	12.511	12.903	13.341
$E[\ln p \mid k=1, d=0, s=1, a=0]$	13.474	13.129	13.404	13.740
$E[\ln p \mid k=1, d=0, \ldots] - E[\ln p \mid k=0, d=0, \ldots]$	0.542	0.114	0.567	0.944

[a]The table reports summary statistics of the expected $\ln p$ for client-sellers ($a = 0$), conditional on transaction ($s = 1$), sophisticated seller ($k = 1$) vs. uninformed seller ($k = 0$), as well as dual agency ($d = 1$) vs. non-dual agency ($d = 0$). The expected $\ln p$ is computed for all client-sellers, accounting for selection biases (i.e. prices are observed only if they are higher than reservation values). The difference, $E[\ln p|k = 1, \ldots] - E[\ln p|k = 0, \ldots]$, is computed for all observations, and the table reports summary statistics of these differences.

Table 9: Welfare Impacts of Aligning Incentives by Increasing κ[a]

	mean (1)	25th pct (2)	median (3)	75th pct (4)
	A. dual agency			
expected $\ln p$ w/current κ	12.828	12.502	12.867	13.146
expected $\ln p$ w/counterfactual κ (0.08 higher)	12.896	12.585	12.927	13.205
expected $\ln p$ w/counterfactual κ (0.2 higher)	12.999	12.699	13.014	13.296
% increase in expected $\ln p$ due to an increase in κ by 0.08	6.8	4.1	6.9	9.5
% increase in expected $\ln p$ due to an increase in κ by 0.2	17.1	10.2	17.2	23.7
	B. non-dual agency			
expected $\ln p$ w/current κ	13.068	12.722	13.101	13.408
expected $\ln p$ w/counterfactual κ (0.08 higher)	13.111	12.785	13.142	13.435
expected $\ln p$ w/counterfactual κ (0.2 higher)	13.176	12.873	13.200	13.477
% increase in expected $\ln p$ due to an increase in κ by 0.08	4.3	0.9	4.5	7.6
% increase in expected $\ln p$ due to an increase in κ by 0.2	10.8	2.3	11.3	18.9

[a]The table reports summary statistics of the expected $\ln p$ for client-sellers. The expected $\ln p$ with current κ is equal to $\kappa \times E[\ln p|k = 1, d, s = 1, a = 0] + (1 - \kappa) \times E[\ln p_j|k_j = 0, d_j, s_j = 1, a_j = 0]$, where κ is the recovered κ for listing j. The expected $\ln p$ with counterfactual κ (0.2 higher) is similar, except that we use $\kappa + 0.2$, instead of κ. The difference is computed for all observations, and the table reports summary statistics of these differences.