Understanding In-House Transactions in the Real Estate Brokerage Industry*

Lu Han
Rotman School of Management
University of Toronto
lu.han@rotman.utoronto.ca

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

August 2, 2015

Abstract

About 20% of residential real estate transactions in North America are in-house transactions, for which buyers and sellers are represented by the same brokerage. This paper examines to what extent in-house transactions are explained by agents' strategic incentives as opposed to matching efficiency. Using home transaction data, we find that agents are more likely to promote internal listings when they are financially rewarded and such effect becomes weaker when consumers are more aware of agents' incentives. We further develop a structural model and find that about one third of in-house transactions are explained by agents' strategic promotion, causing significant utility loss for homebuyers.

Keywords: incentive misalignment, real estate brokerage, in-house transaction, agent-intermediated search, structural estimation

JEL classification: C35, C51, L85, R31

*We thank the editor, two anonymous referees, conference participants at the Stanford Institute of Theoretical Economics, Pre-WFA Summer Real Estate Symposium, Rotman Real Estate Microstructure Conference, AREUEA Meeting, Israel Real Estate and Urban Economics Symposium, as well as seminar participants at the Atlanta Fed, Korea University, National Science Foundation, Penn State, Seoul National University, UC Berkeley, UIUC, University of Amsterdam, and University of Toronto for their helpful comments. Contact information: Lu Han, Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, Ontario, Canada M5S 3E6. Seung-Hyun Hong, 213 David Kinley Hall, 1407 West Gregory Drive, Urbana, IL 61801.

1 Introduction

Over 80% home buyers and sellers carry out their transactions with the assistance of licensed real estate agents. Yet concerns persist that the misalignment between the goals of agents and those of their clients can cause a loss in consumers' benefit. For example, recent studies have shown that the current commission structure leads real estate agents to leave their own homes on the market longer and sell at a higher price, compared to homes they sell for their clients (e.g., Rutherford, Springer, and Yavas, 2005; Levitt and Syverson, 2008a). A phenomenon that is equally important for understanding incentive issues and their implications for market efficiencies but has received much less attention is in-house transactions, that is, transactions for which buyers and sellers are represented by the same brokerage office.

In-house transactions account for about 20% of home transactions in North American markets. In theory, in-house transactions could create information efficiency and reduce transaction costs, leading to an efficient match between home buyers and sellers. However, it is also possible that agents may strategically promote in-house transactions for their own financial interest. In particular, since matching internal listings with internal buyers helps clear inventories faster and increase the chance of securing both ends of a transaction, brokerage firms often pay a higher commission to reward agents engaged in in-house transactions (Gardiner, Heisler, Kallberg, and Liu, 2007).

Strategic in-house transactions, if present, have at least two deleterious effects on home buyers and sellers. First, in the search stage, real estate agents may misguide buyers (sellers) by directing their interest to internal listings (buyers), resulting in a suboptimal choice for consumers. Second, in the negotiation stage, an apparent conflict of interest arises from having the same agency represent both buyers and sellers, making it impossible for the agency to help one party without hurting the other. For these reasons, many jurisdictions in North America have now introduced disclosure requirements for dual agency in order to help consumers avoid unintended dual agency relationship.[1]

[1] Massachusetts, for example, requires that real estate brokerages and agents involved in dual agency transactions

In this paper, we aim to contribute to the literature by empirically examining the extent to which strategic in-house transactions occur and further assessing their welfare impact on the efficiencies of home matching. To motivate our empirical strategy, we discuss two types of in-house transactions: *matching-based* and *strategically-promoted*. In the first case, a buyer may purchase an internal listing, because the listing yields the highest utility for the buyer. This could be due to the information advantages and transaction efficiencies associated with in-house transactions. In the second case, in-house transactions can occur for incentive reasons. When agents are financially rewarded by their brokerage firms for selling internal listings, the information advantage of agents relative to their clients may compound incentive conflicts, making it possible for cooperating agents (i.e., buyers' agents) to influence buyers' choice by directing their interest to internal listings first, resulting in a suboptimal match. However, the impact of strategic promotion should be weaker when clients are more informed about agents' incentives.

To test these implications, we utilize a rich dataset that covers one third of properties transacted through the Multiple Listing Service (MLS) in a large North American metropolitan area from 2001 to 2009. The dataset has three appealing features. First, it contains detailed information about house characteristics, neighborhood information, listing and transaction prices, as well as real estate brokerage firms on both sides of a transaction. Second, it includes properties that have been transacted multiple times, which allows us to control for unobserved house characteristics. Third, the sample period covers a natural experimental opportunity permitted by a new legislation (Real Estate and Business Brokerages Act – referred to as "REBBA" hereafter) that requires agents engaged in in-house sales to inform their clients about the dual agency relationship in writing.

Exploiting these unique data, we first estimate a reduced-form model to test the presence of strategic promotion. Our estimation strategy is akin to a difference-in-differences approach. We first

obtain informed written consent from both sellers and prospective buyers before completing a transaction (254 Code of Massachusetts Regulations 3.00 13.b). Similar laws have been implemented in other states including Wisconsin (Wisconsin Statutes 452.135) and Illinois (225 Illinois Compiled Statutes 454, Article 15).

exploit differences in commission structures. Specifically, agents in a traditional brokerage firm give their firm a predetermined ratio (usually 40-60%) of their commission revenues on the per-transaction basis. Full commission brokerage firms, on the other hand, allow their agents to retain 100% of commission revenues but require fixed amount of upfront fees instead (Munneke and Yavas, 2001). The revenues from the traditional brokerage firms strictly increase with the number of either end of transactions, and hence, these firms are more likely to offer their agents higher bonuses for promoting in-house sales (Conner, 2010). Such promotion bonus would be particularly attractive for cooperating agents if commission fees they receive from listing agents are lower than the market rate. Thus, we expect that cooperating agents working for traditional brokerage firms are more likely to engage in in-house transactions, and that this effect is stronger if the commission rate offered by the listing agent is lower. However, these effects alone can be problematic, as the commission structure/rate could endogenously vary with the degree of matching efficiency involved in in-house transactions.

To identify the presence of strategic in-house transactions, we then examine differences in the incidence of in-house transactions for agents with different commission incentives before and after the implementation of the REBBA. Presumably, by making consumers more informed about the agency relationship and related incentive issues, the REBBA policy constrains agents' ability to promote internal listings, while leaving other types of in-house transactions unaffected. Thus, the identification in our model does not require the commission rates or split structure to be exogenous. Instead, it relies on the assumption that no other commission related factors differentially affect the incidence of in-house transactions at the same time as the REBBA was implemented. To show this is indeed the case, we control for a large number of time-varying house and brokerage observable characteristics. We also include the interaction of house fixed effects with the REBBA as well as the interaction of brokerage fixed effects with the REBBA to control for possible time-variation in unobservable house and brokerage characteristics that may be correlated with commission variables.

Our reduced-form results show that cooperating agents are more likely to engage in in-house transactions when they split the commission fees with firms on the per-transaction basis. This effect is stronger when they receive less compensation from listing agents. More importantly, such effects are substantially weakened after the introduction of the REBBA. Together, these results are highly in line with the theoretical predictions, hence providing strong evidence for the presence of strategic in-house transactions. These results are robust to a rich set of controls and various fixed effects.

In light of the reduced-form evidence for strategic promotion, we further attempt to quantify the extent of strategic versus efficient in-house transactions, and evaluate the welfare consequence of strategic in-house transactions before and after the REBBA. This calls for structural estimation, because matching efficiencies are generally unobserved and hard to quantify. The key idea of our structural approach is as follows: for a given buyer, her decision on whether purchasing an internal listing reflects the difference between the net utility that she obtains from internal versus external listings and the net cost that she incurs when searching for internal versus external listings. If her cooperating agent strategically promote internal listings, such promotion would artificially increase the buyer's cost of searching for external listings. Thus, to the extent that the idiosyncratic matching values for internal and external listings can be estimated, we can recover the implicit costs that the agent may impose on the buyer for searching external listings.

To this end, we first use a nonparametric hedonic approach developed by Bajari and Benkard (2005) to recover the unobserved house characteristic and buyer-specific preferences for house characteristics. We then exploit econometric matching techniques (e.g., Heckman, et al. 1997, 1998) to recover the idiosyncratic match value that a buyer obtains from internal listings as well as from external listings. This enables us to estimate the implicit cost that buyers incur when shopping for external versus internal listings. To identify part of the cost that is due to agents' promotion, we again rely on the difference-in-differences strategy, exploiting variations generated by commission variables combined with the REBBA policy, both of which are well-motivated by the theory.

We find that about 64.3% of in-house transactions can be explained by buyers' own preference. In this case, agents' strategic promotion does not lead to a distortion in the home search process, because home buyers' *ex ante* preference for internal listings agrees with agents' interest. This is more likely to occur when agents work for a firm with a larger pool of listings, as larger firms tend to produce better internal matches. The remaining in-house transactions are likely due to agents' strategic promotion. For these transactions, we find that an agent's promotion of internal listings imposes a substantial cost when a buyer searches for external listings. In this case, even if a buyer's interest is best matched by an external listing, she ends up purchasing a house from internal listings because expected utility gain from purchasing externally is not sufficient to outweigh the associated cost imposed by the agent's promotion. Lastly, we find that the REBBA has helped homebuyers make more informed choices and constrained agents' ability to strategically promote, thereby increasing aggregate buyer welfare by $690 million in the sample market.

The rest of the paper is organized as follows. Section 2 discusses the related literature. Section 3 provides the institutional background and discusses theoretical predictions about strategic promotion. Section 4 describes our data, and Section 5 presents reduced-form evidence for strategic promotion. Section 6 further develops our structural model and presents the results to quantify the extent of strategic promotion and its associated welfare loss. Section 7 concludes the paper.

2 Related Literature

Broadly speaking, our paper is informed by an important literature on the distortion of agents' incentives (e.g., Gruber and Owings, 1996; Mehran and Stulz, 1997; Hubbard, 1998; Garmaise and Moskowitz, 2004). In light of the central role of housing markets in the recent economy, there has been substantial interest in examining incentive issues in different sectors in housing markets. One such sector is the mortgage industry. For example, Jiang, Nelson, and Vytlacil (2014) find that loans issued by mortgage brokers have higher delinquency rates, reflecting incentive structures that compensate brokers mainly based on origination volume rather than loan performance.

Another sector that plays a crucial role in housing markets is the brokerage industry. An extensive and growing literature has examined the consequence of the misalignment between goals of real estate agents and those of home sellers. For example, recent work has examined the effects on selling price and time on the market of agent-owned versus client-owned properties (Rutherford, Springer, and Yavas, 2005; Levitt and Syverson, 2008a), MLS-listed versus FSBO properties (Hendel, Nevo, and Ortalo-Magne, 2009), and properties sold by traditional agents versus discounted agents (Levitt and Syverson, 2008b; Berheim and Meer, 2008). One common thread between these papers is that the current commission arrangements have resulted in a distortion of agents' incentives, which in turn affects how much a house is sold for and how long it takes to sell.

Despite a significant interest in real estate agents' incentive issues, their importance in the specific context of in-house transactions has not been extensively studied. This seems surprising given the sheer magnitude of in-house transactions and obvious incentive issues that could arise from the dual agency representation. Gardiner, Heisler, Kallberg, and Liu (2007) are among the first to study the impact of dual agency in residential housing markets. They find that dual agency reduced the sales price and the time on the market and that both effects were weaker after a law change in Hawaii in 1984 which required full disclosure of dual agency. Using repeated sales properties, Evans and Kolbe (2005) examine the effect of dual agency on home price appreciation. In addition, Kadiyali, Prince, and Simon (2012) study the impact of dual agency on sales and listing price, as well as time on the market. However, like the previous literature on the real estate brokerage, these studies focus on transaction outcomes for home sellers. None of the existing work examines the consequences of agents' incentives on the quality of home match, which is the key transaction outcome for home buyers. The lack of such work is in large part due to the difficulty of determining the quality of a match between a buyer and a house.

In this paper, we marry the insights from the incentive distortion literature to the methodologies developed in the recent structural industrial organization literature (e.g., Bajari and Benkard, 2005;

Bajari and Kahn, 2005). Specifically, we develop a structural model of in-house transactions and propose an approach to recover the idiosyncratic match value in home transaction process. By linking our empirical work to agent-intermediated search theory, we are also able to distinguish between different sources of in-house transactions – ranging from strategic promotion to efficient matching. Doing so allows us to evaluate the economic harm that the incentive misalignment brings to homebuyers. Such evaluation contributes to a better understanding of market efficiency in this important industry. In this regard, our work also complements the recent literature that examines social inefficiencies resulted from free entry in the real estate brokerage industry (Hsieh and Moretti, 2003; Han and Hong, 2011; Jia Barwick and Pathak, 2014).

3 In-House Transactions in the Real Estate Brokerage Industry

3.1 Institutional background

If cooperating agents' interests are fully aligned with home buyers' interests, there should be no efficiency loss associated with in-house transactions. However, if agents have strategic interest to promote internal listings, buyers' benefits would be inevitably sacrificed, and a suboptimal match would be generated. Two characteristics of the residential real estate brokerage industry make the possible incentive issues particularly concerning for in-house transactions.

First, the agency relationship in real estate transactions does not encourage cooperating agents to represent the best interests of their buyers. In a typical multiple listing agreement for a real estate transaction, the listing agent has a contractual relationship with the seller, which explicitly defines his fiduciary obligations to the seller. The usual MLS agreement constitutes an offer of sub-agency to all other MLS members. The cooperating agent who brings the buyer to close the deal is deemed to have accepted the sub-agency offers and hence has fiduciary duties to the seller. Those duties effectively preclude the cooperating agent from adequately representing the buyer, even though the agent appears to work for the buyer.[2] While the conflicting loyalty by cooperating agents for buyers

[2]See Olazabal (2003) for detailed discussion on the agency relationship.

may seem obvious, many buyers are not aware of the agency relationship and rely heavily on their agents in searching for a home and negotiating the price of a home. The incentive misalignment problem is likely to worsen in in-house transactions, since agents from the same agency are more likely to share the information with each other and influence their clients' decisions from both ends.

Second, both academic researchers and market practitioners have noted that brokerage firms tend to offer a promotion bonus to agents who successfully sell in-house listings.[3] There are at least two motivations for such promotions. First, in-house transactions help the firm clear inventory faster, allowing agents to earn commissions from existing clients sooner and hence have more time/resources to compete for new clients. Second, by promoting in-house sales, brokerage can potentially influence clients' decision from both sides, making a transaction easier to go through and hence maximizing the chance of capturing commission income from both ends.[4] Consistent with this, a descriptive analysis of our data shows that brokerage firms with a higher fraction of in-house sales have a larger number of total transactions (separately counting each end of a transaction), even after controlling for the number of agents in each firm.

For these reasons, cooperating agents may strategically promote in-house transactions. For example, a cooperating agent may show her client internal listings before external listings.[5] Alternatively, a cooperating agent may take her client to visit externally listed houses before visiting the internally listed house, but these external listings would be selected to appear less attractive than the internal listings that the agent tries to promote. These efforts are strategic and may lead to an in-house transaction that is inconsistent with the interest of home buyers.

[3] For example, Gardiner, Heisler, Kallberg and Liu (2007) find that many brokerage firms give a financial reward to agents who successfully match internal clients with internal listings. Similarly, a popular industry practice book, *Buying a Home: The Missing Manual*, reports that some agencies pay agents a bonus for selling in-house listings because the agency makes more money in such transactions. In addition, a recent report by the *Consumer Advocates in American Real Estate* explicitly points out that agents who avoid in-house transactions may bear with some financial consequences, such as a less favorable commission split with the brokerage firm.

[4] To see this, note that signing a contract with a client does not provide a guarantee for an agent to receive any commission as the transaction may not occur during the agent's contract term. This is particularly a concern for cooperating agents as they tend to have less exclusive and shorter contracts (or even no contract) with buyers.

[5] Similarly, a listing agent may show his client's house to internal buyers before external buyers. In this paper, we focus our discussion on cooperating agents, but the logic can be easily extended to listing agents.

Of course, an in-house transaction could also occur due to spontaneous visits or information sharing. For example, a buyer may see a for-sale sign on a property and call the listing agent whose name is listed on the sign. Similarly, an agency may become a dual agency if a buyer who is represented by a cooperating agent independently discovers a house where the listing agent works for the same agency as the buyer's agent. It is not obvious whether these types of transactions would generate an efficient matching outcome or a suboptimal choice for consumers. However, their existence makes detecting strategic promotion empirically challenging. In what follows, we derive theoretical predictions that underpin our empirical approach to identify strategic promotion.

3.2 Strategic and Efficient In-House Transactions

In Appendix A, we present an agent-intermediated search model that applies search diversion theory developed in online shopping literature (Hagiu and Jullien, 2011) to the real estate brokerage industry. The model incorporates an important feature that real estate agents receive a share of *realized* sales revenues and this share is *larger* when a transaction occurs within the same brokerage office. The model yields the following intuitive result: the optimal amount of strategic promotion in equilibrium increases with the financial incentives an agent receives from promoting in-house transactions. Motivated by the practice in the real estate brokerage industry, we argue that such financial incentives are reflected by the amount of commission fees that agents receive in each transaction and how they split the fees with their affiliated brokerage offices.

In a residential real estate transaction, the commission rate for a cooperating agent is typically predetermined when the listing is posted on the MLS. While the commission rate is usually set at 2.5%, some listing agents would offer a higher or lower rate to cooperating agents. Intuitively, by rewarding cooperating agents a greater proportion of the commission, an external listing agent can effectively offset the promotion bonus that the cooperating agent receives from her own firm for promoting internal listings. Conversely, when the commission rate offered by a listing agent is low, the cooperating agent is more likely to respond to the financial incentives offered by the brokerage firm for promoting in-house transactions. The strategy of using substandard commission rates to artificially increase the frequency of dual-agency transactions is discussed in Yavas, *et al* (2013) and also evidenced by a recent industry report.[6] Thus, we expect that lower commission rates offered by listing agents are associated with a stronger presence of strategic in-house transactions.

In addition to commission fees, commission structure also matters. As noted earlier, real estate agents need to split commission fees with their affiliated brokerage offices, in return for the brand value and for the supporting services that brokerage offices provide. In practice, full commission brokerage firms, such as ReMax, let the agents keep all commission fees but require a fixed amount of upfront fees each month. More traditional brokerages firms, such as Royal LePage, split commission fees with their agents on the per-transaction basis. Naturally, the revenues in the latter type of brokerage firms strictly increase with the number of either end of transactions. Therefore, these brokerage firms are more likely to reward their agents for selling internal listings. Thus, we expect that the per-transaction split commission structure is associated with a stronger presence of strategic in-house transactions.

While agents may have financial incentives to promote in-house transactions, their ability to do so depends on whether buyers are aware of agents' incentives to strategically promote. In particular, our model shows that the strength of strategic promotion would be weaker if buyers are more aware of agents' financial incentives to promote. As discussed in Section 4, our sample covers a natural experimental opportunity permitted by a legislation that required real estate agents engaged in in-house transactions to disclose the possibility of strategic promotion to both buyers and sellers. This provides an opportunity for us to empirically test this prediction.

In-house transactions could also occur for efficiency rather than incentive reasons. We define an in-house transaction as "efficient" if buyer's net utility from internal listings is larger than her net utility from external listings, either *ex ante* or *ex post*. The model predicts different types of efficient

[6] For example, a recent report by the *Consumer Advocates in American Real Estate* states that "offering less than the going rate in your area will decrease the financial attractiveness of your home [to cooperating agents] and increases the likelihood that your broker will collect a double commission" (see an article titled "Dual Agency Schemes" in http://www.caare.org/ForBuyers, accessed August 1, 2014).

in-house transactions that result from a mixture of transaction efficiencies and information advantages.

While we do not attempt to fully model the sources of efficient in-house transactions, it can be shown that the efficiency loss associated with strategic in-house transactions depends on the difference in the matching quality that a given buyer obtains from internal and external listings. Empirically, we do not observe matching quality. However, we can proxy the difference in the matching quality by looking at how typical a house is and how many listings the brokerage possesses. Intuitively, if buyers are looking for more or less homogeneous houses (e.g., tract home), and if such homes are available both internally and externally, the potential loss of matching quality associated with purchasing an internal listing should be relatively small. In addition, matching in housing markets is typically characterized by increasing returns to scale (Ngai and Tereyro 2014; Genesove and Han 2012b). When a brokerage firm has a larger number of listings which a buyer can choose among, there should be less dispersion in the buyer's valuation of her most-preferred house from the market-wide pool and from the internal listings. Although the promoted listings may not match the buyer's preference best, the resulting efficiency loss should be smaller since these listings are closer to the buyer's preference.

In sum, there are a number of brokerage- and transaction-specific features that can be tied to predictions about in-house transactions. These range from sources of strategic promotion to sources of efficient matching. Specifically, the model laid out in Appendix A generates the following theoretical predictions.

Prediction 1: An agent who splits commissions with the affiliated brokerage firm on the per-transaction basis and/or receives lower commission fees is more likely to strategically promote internal listings.

Prediction 2: An agent's ability to promote internal listings is weaker when her buyer client is more aware of her strategic incentives.

Prediction 3: The efficiency loss associated with in-house transactions is smaller when buyers look for typical homes and/or when the cooperating brokerage offices have a larger number of listings.

Together, these predictions provide a basis for the difference-in-differences strategy used in our empirical analysis. The model also implies that a full control of efficient matching can be obtained by comparing a buyer's expected utility from internal and external listings, which further motivates the structural approach that we exploit in Section 6.

4 Data

The main source of our data is the Multiple Listing Service (MLS) in a large North American metropolitan area from January 1, 2001 to December 31, 2009. This market experienced a boom in the middle 2000s, with a peak in 2007 and the first half of 2008, followed by a temporary decline in the second half of 2008, and then an immediate strong rebound afterwards, with sales volume reaching the pre-crisis peak level in the first half of 2009. Our sample covers 28 districts which comprise a third of the metropolitan area. There are over 200,000 transactions and about 1,500 brokerage firms. The MLS data contain detailed information on house characteristics, including the number of bedrooms, the number of washrooms, lot size, the primary room size, dummy variables for basement, garage space and occupancy. In addition, the data provide neighborhood information, listing and transaction prices, as well as real estate brokerage firms on both sides of a transaction. Properties are identified in the MLS data by district, MLS number, address, unit number (if applicable). To avoid some extreme cases, we exclude the following transactions from the estimation sample: (1) transactions for which the sales price is less than $\$30,000$ or more than $\$3,000,000$; (2) transactions for which the cooperating commissions are less than 0.5% or more than 5%; (3) listings that stay on the market for less than one day or more than one year.[7]

We define in-house transactions as transactions for which the cooperating agent and the listing agent are associated with the same brokerage office. In our sample, about 20% of transactions occur

[7] We also estimated our model using somewhat different cutoffs (e.g. the cooperating commission rates are less than 1%; listings stay on the market for fewer than 2 days), our results are robust to these changes.

within the same brokerage office. Tables 1-2 report the fraction of in-house transactions by brokerage office size. In Table 1, we rank cooperating brokerages in order of their total market shares in our data, and group them by their rankings. In Table 2, we group cooperating brokerages by the number of real estate agents. Both tables show that larger brokerages tend to have relatively higher fractions of in-house transactions, as what one would expect.

One might wonder whether these in-house transactions can simply be a result of independent hiring decisions made separately by buyers and sellers. In that case, conditional on a given buyer working with brokerage j, the probability that the buyer purchases a house listed by the same brokerage should be equal to the market share of listing brokerage j. In other words,

$$\Pr(\text{listing} = j | \text{cooperating} = j) = \Pr(\text{listing} = j). \tag{1}$$

However, as shown in Figure 1, brokerage-office-level fractions of in-house transactions at any given district of the sample city are much higher than the dashed line which depicts the fractions predicted from (1). This suggests that a significant fraction of in-house transactions cannot be explained by independent interactions among brokerage firm, hence providing a key motivation for the empirical analysis in this paper.

As noted earlier, a legislation named the REBBA was implemented in the sample city in March 2006. According to the legislation, if an agent represents or provides services to both a buyer and a seller or more than one buyer, then the agent should "*in writing, at the earliest predictable opportunity and before any offer is made,* inform all buyers and sellers involved in that trade of the nature of the registrant's relationship to each buyer and seller." The agents need to disclose not only the fact that the listing and cooperating agents work for the same office but also the fact that the cooperating agent is a sub-agent of the listing agent and hence has fiduciary duties to the seller.[8] This means that a buyer could decide not to make an offer to the house of interest and continue searching if she

[8]The anecdotal evidence suggests that agents sometimes need to disclose the financial rewards they received from the brokerage firm for promoting in-house sales.

is concerned that the "conflict of the interests" might hurt the matching quality. By making clients more aware of the agency relationship and the possible incentive issues, the REBBA is most likely to affect the incidence of in-house transactions that occur for strategic reasons while leaving other types of in-house transactions unaffected.

Table 3 shows a slight downward trend for in-house transactions in our sample, with a discrete drop after 2006. As noted above, unlike many U.S. housing markets, the market under study did not experience a crash during 2006. Instead, it was in the midst of a boom that did not end until the second half of 2008. Thus, the discrete drop in the fraction of in-house transactions in 2006 is unlikely explained by a housing market downturn. Instead, it seems to be consistent with the implementation of the REBBA in 2006 that constrained agents' ability to strategically promote. We will investigate this possibility taking into account the changes in the market conditions in the next section. Note that the downward trend in in-house transactions further continued in the years after 2006. This is not surprising, as we expect that it takes time for the policy to be fully enforced and for consumers to fully understand the incentive issues behind dual agency.

5 Testing Strategic Promotions: A Reduced-Form Approach

5.1 Testing the Effects of Commission Incentives

To test the presence of strategic promotion, we estimate the following linear probability model

$$E(d_{ibt}|Z_{it}, X_{it}, W_{bt}, \eta_{ibt}) = Z_{it}\alpha + X_{it}\beta + W_{bt}\delta + \eta_{ibt}, \tag{2}$$

where d_{ibt} is the indicator variable for whether transaction i at period t is an in-house transaction carried out by brokerage b, and Z_{it} is a vector of firm- and transaction-specific variables related to commission structure/rate in transaction i. Specifically, $Z_{it} = (COMM_{it}, COMM_{it} \times REBBA_t)$, where $COMM_{it}$ is defined below, and $REBBA_t$ is a dummy variable for the REBBA. X_{it} refers to a vector of control variables including house lot size, number of bedrooms, number of washrooms, dummy variables for the basement, garage space, and occupancy status. W_{bt} refer to brokerage-level variables such as the number of internal listings by brokerage b in the same district during the month before the transaction.[9] In addition, η_{ibt} contains various fixed effects for location, time, brokerage, and house. Throughout the paper, we cluster the standard errors at the city block level to allow for the spatial and temporal dependence within neighborhood blocks.

The key variable of interest is $COMM_{it}$, which captures agents' commission incentives to promote in-house transactions. As described in Section 3.2, this is measured by two commission variables. The first is split, a firm-specific dummy variable that equals 1 if the cooperating agent splits commission fees with the brokerage firm on the per-transaction basis. The second is low.commission, a transaction-specific dummy variable that equals 1 if the commission fees received by the cooperating agent from the listing agent in a given transaction are lower than 2.5% of the house price. Note that the commission fees are determined at the beginning of the listing process and remain the same until a transaction is completed. As discussed earlier, agents who split commission fees with firms on the per-transaction basis are more likely to receive a promotion bonus; and a lower commission rate offered by the listing agents from other brokerages would make the in-house promotion bonus effectively more attractive to the cooperating agent. Following Prediction 1 in Section 3.2, we hypothesize that agents with stronger commission incentives are more likely to engage in in-house transactions, thus we would expect the coefficients on split, low.commission, and split×low.commission to be positive.

Though commission structures and fees are predetermined, we cannot infer strategic promotion directly from a straightforward comparison of transactions with different commission structures/fees, since brokerage firms may intentionally set their commission policies in an attempt to capture transaction cost savings resulted from in-house sales. Some of these savings may be passed onto the buyers and sellers, improving transaction efficiencies in general. If this is the case, then the higher probability of in-house transactions associated with the per-transaction split structure and/or lower commission fees cannot be interpreted as evidence for "strategic" promotion.

[9] If in-house transactions help enhance search efficiency, liquidity theory suggests that such benefits are bigger for firms with a larger number of listings. This effect would be captured by a positive δ in (2).

Hence, we take a difference-in-differences approach by including a term $COMM_{it} \times REBBA_t$. Specifically, we examine differences in the incidence of in-house transactions for agents with different commission structure/rates before and after the REBBA. Following Prediction 2 in section 3.2, we hypothesize that agents' ability to strategically promote will be weaken after buyers are more aware of agents' financial incentives. Thus, our empirical exercise is a joint test of the hypotheses that promotion of internal listings takes place under certain commission incentives and that the ability to promote is weakened after the REBBA. In other words, the identification in our model does not require the assumption that commission variables are exogenous; instead it relies on the assumption that no other commission related factors differentially affect the incidence of in-house transactions when the time as the REBBA was implemented.

There are a number of legitimate concerns with our approach. One concern is that we do not observe the actual promotion bonus directly but must infer it. Given the lack of information on the brokerage internal compensation scheme, this issue is inherent in doing research in this area. As argued above, we draw such inference based on two commission variables, both of which are motivated by the industry practice. In what follows, we will further deal with this issue by estimating a rich set of specifications to build a strong case that the commission effects we examine are due to strategic promotion and not due to other unobserved factors. Another concern is related to the identifying assumption described above. In particular, commission structures/rates may have changed after REBBA either because brokerage firms' characteristics have shifted over time or because the pool of houses that attract low commission rates have changed in a way that is not observed from the data. In Section 5.3, we discuss these concerns at length and address them by controlling for a rich set of fixed effects. The results show that changes in unobserved firm/house characteristics, contemporaneous with the REBBA, are unlikely to alter the interpretation of our key findings. To further strengthen the validity of our estimates, Section 5.4 also provides a direct test of the identification assumption.

5.2 Baseline Results

In the baseline estimation, four different versions are estimated, and the results are reported in Table 4. We begin with the simplest specification where $COMM_{it}$ is measured by a single commission variable, split. The related coefficient estimates in column 1 are statistically significant and consistent with what we expected. In particular, splitting commission fees with the firm on the per-transaction basis increases the probability of in-house transactions by 1.3 percentage point, while such effect disappears substantially after the implementation of the REBBA.

Despite our rich control of housing attributes, one might be concerned that unobserved house quality could vary so much to render the estimates imprecise. One plausible control variable for unobserved house quality is the listing price, because the listing price is likely to reflect not only observed but also unobserved house quality. Column 2 adds the listing price to the baseline specification. The coefficients on the split variables remain almost the same, both in magnitude and in significance.

In column 3, we add a different commission variable – low.commission – and its interaction with split. The coefficient estimates are consistent with what we expected. Splitting commission fees and receiving a lower commission fee increase the probability of in-house transactions. Moreover, the incentive effect is particularly strong when both commission variables are in effect. For example, splitting commission fees alone increases the probability of in-house transactions by 1.1 percentage point; this effect is further increased to 4.2 percentage points for agents receiving lower commission fees. The finding is consistent with our hypothesis that a lower commission payment makes the promotion bonus offered by the brokerage firm more attractive to the cooperating agent, and hence gives the latter stronger incentives to sell in-house listings. More importantly, these incentive effects are largely reduced after the implementation of the REBBA, as reflected by the negative coefficients on the REBBA interaction terms. The strong positive coefficients associated with the commission variables, and particularly their interactions, suggest that financial incentives at least partially explain in-house transactions, as predicted by the theory. The substantial weakening impact of the REBBA on the commission effects further suggests that the variations in in-house transactions caused by commission changes are an indicator of agents' strategic behavior rather than transaction efficiencies. The results are robust to the inclusion of the listing price, as shown in column 4. In what follows, we will treat columns 3-4 as primary specifications for more robustness checks.

5.3 Robustness Checks

The baseline specification results point to a strong presence of strategic promotions in explaining in-house transactions. However, one might be concerned that the estimated strategic effects could be due to a set of unobserved factors either at the house level or at the brokerage level that are correlated with commission policies. If this is the case, then we would not be able to convincingly interpret our finding as the evidence for strategic promotion. In this section, we provide a set of robustness checks to address this concern.

Table 5 presents a set of robustness checks that deal with time-invariant unobserved heterogeneity. Listing prices are controlled in the bottom panel (columns 6-10), but not in the top panel (columns 1-5). For the ease of comparison, columns 1 and 6 (in both Tables 5-6) respectively repeat the results presented in columns 3-4 in Table 4. As a further control for unobserved house attributes, we restrict the sample to houses that were sold multiple times in our sample period.[10] Doing so allows us to control for house fixed effects at the cost of dropping two thirds of observations. Columns 2 and 7 of Table 5 present the results. The estimates on commission and REBBA variables are qualitatively consistent with columns 1 and 6, although some of them lose statistical significance, probably due to the substantially reduced sample size.

In columns 3 and 8, we return to the full sample but control for the idiosyncratic brokerage fixed effects. If in-house transactions are more likely to occur for certain brokerage firms due to their specific policies or network size, its effect on our estimates should be controlled by including brokerage firm fixed effects. However, the key coefficient estimates on commission and REBBA variables continue to

[10]We drop houses that were sold multiple times within six months, because they are likely due to "flipping".

be strong and significant and have expected signs, suggesting that unobserved brokerage factors are unlikely to change the interpretation of our findings.

As our theoretical model implies, in-house transactions could occur for efficiency rather than incentive reasons. For example, when a brokerage has superior information about properties and buyers' demand curve in a specific housing market segment, an in-house transaction can lead to a better and quicker matching outcome. In the remaining columns, we control for brokerage specialization by adding the interactions of brokerage firm dummies with neighborhoods (columns 4 and 9) and with price ranges (columns 5 and 10). The former are intended to control for specialization based on geographical areas, while the latter for specialization based on certain price ranges. In both specifications, we find that agents with stronger common incentives, measured by receiving lower commission fees and split commission with the firm, are more likely to be engaged in in-house transactions. Moreover, these effects are much weakened after the REBBA. Together, these estimates suggest that our findings about the strategic promotion are quite robust.

So far we have shown that time-invariant unobserved house attributes or brokerage factors are unlikely to alter the interpretation of our key results. However, this does not rule out the possibility that there might be important differences in unobserved house or firm characteristics before and after the REBBA that are correlated with commission variables. If this is the case, the interpretation of our difference-in-differences estimates would still be questionable. In Table 6, we address this concern by explicitly controlling for time-varying unobserved heterogeneity particularly before and after the implementation of the REBBA.

Along this line, one possible story is that some properties might be easier to sell because of some unobserved attractive characteristics, and that these properties might be sold internally and carry lower commissions. If the fraction of such properties changed after the REBBA, then the finding of the weakened impact of commission effects could be due to the shift in the distribution of unobserved property characteristics, rather than the change in agents' strategic promotion. To address this concern, we restrict the sample to the repeated sales and include house fixed effects interacted with the REBBA dummy. As shown in columns 2 and 7 of Table 6, the commission variables continue to have a significant and positive effect and their interactions with REBBA continue to have a significant and negative effect. This provides reassuring evidence for our main finding.[11]

Similarly, one may argue that there might be differences in brokerage characteristics before and after the REBBA, which could affect both commission policies and frequency of in-house transactions. For example, the pool of brokers offering lower commission fees might have shifted after the REBBA for non-strategic reasons. Alternatively, full-commission and split-commission brokerage firms could have experienced different trends before and after the REBBA. While we cannot completely rule out this possibility, we find no evidence that such possibility would affect our key result. In particular, we return to the full sample and include brokerage fixed effects interacted with the REBBA dummy. As shown in columns 3 and 8 of Table 6, the key coefficient estimates on commission and REBBA variables are again consistent with what we expected, both in sign and in significance.

One could further postulate that the fraction of in-house transactions induced by brokerage specialization might have also changed after the implementation of the REBBA, which would affect the interpretation of our key estimates. To address this concern, we additionally include a triple interaction term *brokerage* × *region* × *REBBA* (in columns 4 and 9) and another triple interaction term *brokerage* × *price range* × *REBBA* (in columns 5 and 10) to control for the time variation in the fraction of in-house transactions due to different types of brokerage specialization. The resulting estimates are again consistent with what we have expected, confirming that time variation in brokerage specialization is unlikely to change the interpretation of our key findings.

[11] To include house fixed effects interacted with the REBBA dummy, we need to use only houses that were sold multiple times during the pre-REBBA (or post-REBBA) period. The number of houses that fit these criteria is 14,225 for the pre-REBBA and 5,700 during the post-REBBA, which results in 29,531 observations during the pre-REBBA period and 11,601 observations during the post-REBBA period.

5.4 Testing the Identification Assumption

Our identification assumption is that no other commission related factors differentially affect the incidence of in-house transactions at the same time as the REBBA was implemented. The before- and after-REBBA changes in the house- and brokerage-specific factors are fully absorbed in the robustness check specifications above. In this subsection, we provide a more direct test of the identification assumption by examining whether there is any systematic variation in characteristics of houses sold under different commission incentives around the time of the REBBA.

Specifically, we regress each observed house attribute on the commission fee and split structure variables as well as their interactions with REBBA, controlling for other house attributes and housing market conditions. The results are presented in Table 7. Reassuringly, we find that the coefficients on the commission variables, interacted with the REBBA, are small and statistically insignificant, suggesting that there were no systematic changes in observed attributes of houses sold under different commission structures before and after the implementation of the REBBA. Given that we find no systematic changes in observed attributes of houses sold under different commission structures before and after the REBBA, it is unlikely that there would be systematic changes in unobserved attributes that would bias our results. Together, we believe that the weight of the evidence points toward a strong support for our identification assumption.

5.5 Hot Markets and Atypical Homes

In this subsection, we examine how the estimated strategic in-house transactions vary with housing market cycles and distinctness of homes. One might expect that as the housing market heats up, the likelihood of bidding wars would increase and buyers would have increasing incentives to use the same brokerage that listed the house. They might be particularly so if buyers look for a home that is more or less similar to others. Note that in all specifications above, we have controlled for year×month fixed effects. Thus, the temporal component of local business effects is fully absorbed.

To investigate whether the phenomenon of bidding wars and the distinctness of houses complicate the interpretation of our results, we split the sample by the sales-to-list-price ratio and by a constructed house atypicality index. First, we use the sales-to-list ratio to capture the intensity of bidding wars: transactions with sales-to-list ratio exceeding 1 are considered to occur in a hot market; and others are considered to occur in a cold market. This is because the observation of a sales price greater than list price typically requires that there be multiple bidders, either active or potential (Han and Strange, 2014). Of course, there will be situations with multiple bidders who all bid below list price in which case our measure underestimates the bidding wars. However, the term "bidding war" seems to connote extreme bidding. Hence transactions with sales-to-list ratio exceeding 1 can be plausibly interpreted as being the instances of extreme bidding and so hot markets.

Second, within hot and cold markets, we further compare houses that look alike (e.g., "tract homes") with houses that look more distinct (e.g., "luxury houses"). To do so, we create an atypicality index for all houses in a standard way (Haurin, 1988). This index should be interpreted as the aggregate value of deviation of a property's characteristics from the sample mean in each neighborhood. We consider a house "typical" if its atypicality index is less than 0.1 (about the 50th percentile in our sample), and "distinct" otherwise. Splitting the sample by both sales-to-list ratio and the atypicality index yields four subsamples: hot market with typical houses, hot market with distinct houses, cold market with typical houses; cold market with distinct houses.

The results are reported in Table 8. Several findings emerge. First of all, most of the estimates on the commission variables and their interactions with the REBBA remain statistically significant and have expected signs.[12] This suggests that the estimated strategic promotion effects are robust to the controls on bidding intensity and house atypicality, providing additional support for our main findings. Second, we find that the magnitude of the strategic promotion effect is larger in hot markets, regardless of whether houses are typical or distinct. There are two reasons to expect a stronger strategic

[12]The coefficient on the REBBA interaction term in the hot market with typical homes remains strongly negative but less significant, possibly due to a much smaller number of observations in this case.

promotion effect in hot markets. One is that when housing markets are hot, agents are motivated to clear inventories faster so that they can spend more time competing for new clients/listings. Since dual agency helps speed up the transaction process, agents have stronger incentives to match internal listings with internal buyers. The other reason is that in hot markets where bidding wars are prevalent, it is difficult for buyers to find an ideal home on their own, making it easier for agents to influence their buyer clients in their search and bargain process and hence promote internal listings. Third, we find that within hot or cold markets, the differences in the estimated strategic effect between typical and distinct houses are statistically insignificant. This suggests that the degree of atypicality alone does not affect in-house transactions caused by strategic promotion. On the other hand, we expect the atypicality factor might affect efficient in-house transactions. We investigate this possibility in Section 6.2.

6 Quantifying Strategic Promotions: A Structural Approach

So far we have obtained a set of findings that are consistent with our theoretical predictions. While these results are strongly supportive of the presence for the strategic promotions, they cannot tell us the extent of the strategic promotion in explaining in-house transactions. In addition, they cannot help us evaluate the welfare impact of strategic promotion and the associated disclosure requirement. The challenge to conducting such exercise in a reduced-form way is that matching efficiencies are generally unobserved and hard to quantify. To address these issues, we develop a structural model. Our model is described in Section 6.1, and the estimation results are presented in Section 6.2.

6.1 Structural Model

The key idea of our structural approach is as follows: a buyer's decision on whether to purchase an internal listing reflects the difference between the net utility that she obtains from internal versus external listings and the net cost that she incurs when searching for internal versus external listings. If her cooperating agent strategically promote internal listings, such promotion would artificially increase the buyer's cost of searching for external listings. Thus, to the extent that the idiosyncratic matching values for internal and external listings can be estimated, we can recover the implicit costs that the agent may impose on the buyer for searching external listings.

To implement this idea, we develop our model in three steps. In the first step, we build on and modify the hedonic framework developed by Bajari and Benkard (2005) and Bajari and Kahn (2005), which allows us to locally recover buyer-specific preferences for house characteristics.[13] Next, using the recovered preference, we construct idiosyncratic match values that a buyer obtains from internal and external listings. In the third step, using the observed decisions on in-house transactions and the recovered match values, we estimate the implicit costs associated with strategic promotion. In what follows, we begin with the modified version of the hedonic framework, and then describe the second and third steps. The estimation details are provided in Appendix B.

6.1.1 Modified Hedonic Framework

To describe the model, let us consider market $t \in T$, where there are $i = 1, \ldots, I_t$ home buyers who are looking for houses, and $j = 1, \ldots, J_t$ housing units that sellers put on the market. The interactions of a large number of buyers and sellers will lead to hedonic equilibrium in which buyers match to houses, and the resulting equilibrium prices are determined by the hedonic price function that maps housing characteristics to prices as follows: $p_j = \mathbf{p}_t(X_j, \xi_j)$, where p_j is the sales price of house j, X_j is a $1 \times m$ vector of observed attributes of house j, ξ_j is the unobserved house characteristic, and $\mathbf{p}_t$ is the price function in market t that varies across markets, reflecting different equilibria. In our application, we consider the price function given by

$$\log\left(\mathbf{p}_t(X_j, \xi_j)\right) = \alpha_{j,0} + \sum_{k=1}^{m} \alpha_{j,k} x_{j,k} + \eta_t + \xi_j,$$

[13]Bajari and Benkard (2005) and Bajari and Kahn (2005) improve upon the hedonic two-step approach of Rosen (1974) and Epple (1987) by incorporating a nonparametric estimation for the hedonic price function and by proposing an approach to recover the unobserved product characteristic. Note that their framework is not originally intended for constructing counterfactual match values for internal listings versus external listings, nor recovering implicit costs associated with strategic promotion. Therefore, we modify the hedonic framework in order to incorporate strategic promotion and rationalize our approach to construct counterfactual match values.

where $\alpha_j = (\alpha_{j,0}, \ldots, \alpha_{j,m})$ is a vector of the hedonic coefficients that represent the implicit prices faced by each buyer who has chosen house j, and η_t captures market fixed effects.[14]

Because our goal is to recover a buyer's preferences, we focus on the buyer's problem. We posit that buyers' utility functions are defined over house characteristics X_j and ξ_j, as well as the composite commodity denoted by e. The buyer's problem is to maximize her utility $u_i(X_j, \xi_j, e)$ subject to the budget constraint. Following Bajari and Benkard (2005) and Bajari and Kahn (2005), we impose a functional form assumption for identification of the utility function. In particular, we assume the linear utility function[15] given by

$$u_i(X_j, \xi_j, e) = \sum_{k=1}^{m} \beta_{i,k} x_{j,k} + \beta_{i,0} \xi_j + e \tag{3}$$

where $\beta_i = (\beta_{i,0}, \ldots, \beta_{i,m})$ is a vector of buyer-specific random coefficients capturing buyer i's preferences for housing characteristics.

To allow for a decision that leads to an in-house transaction, we modify the standard hedonic framework by assuming that the buyer's decisions are determined in two stages: in the first stage, the buyer chooses X_j and ξ_j; in the second stage, the buyer decides on d_j, where d_j is the indicator variable for an in-house transaction. To incorporate the second stage while maintaining the hedonic framework, we specifically use two modifications as follows.

First, we assume that the budget constraint is given by $e + p_j + g_i(d_j) = y_i$, where the price of the composite commodity is normalized to one, y_i is buyer i's income, and $g_i(d_j)$ is assumed to reflect the implicit costs and benefits associated with d_j. In particular, we assume that $g_i(d_j) = c_i(1 - d_j) - \gamma_i d_j$, where c_i is a random coefficient representing extra search costs for external listings if buyer i's agent strategically promotes internal listings; while γ_i is a random coefficient capturing potential transaction cost savings if buyer i purchases a house from internal listings. Intuitively, strategic promotion occurs

[14] This price function is a linear approximation of $\mathbf{p}_t(X_j, \xi_j)$ in a local neighborhood of house j's characteristics. Hence, α_j varies across different houses, and so they are estimated nonparametrically by using the approach described in Appendix B. In this price function, η_t includes district fixed effects as well as year×month fixed effects. Note also that different markets (based on either location or time) lead to different equilibria, so that houses with the same characteristics can have different prices if they are in different markets.

[15] We experimented with other functional forms such as log-linear utility, but find that our results are still robust.

when the agent introduces noise into the search process by making it more costly for the buyer to shop for external listings, which is captured by c_i. On the other hand, an in-house transaction could generate transaction efficiencies that implicitly benefit buyer i. While these benefits do not affect the idiosyncratic match value, they still affect buyer i's decision by generating transaction cost savings. We capture these benefits by γ_i.

Second, we assume that the first stage is determined separately from the second stage. Note that after substituting the budget constraint into $u_i(X_j, \xi_j, e)$, we can write the buyer's problem as

$$\max_{(X_j,\xi_j),d_j} u_i(X_j, \xi_j, y_i - \mathbf{p}_t(X_j, \xi_j) - g_i(d_j)).$$

Our assumption implies that despite the presence of d_j, the buyer's choice with respect to X_j and ξ_j is still optimal, in that the following first order conditions hold and they do not depend on d_j.

$$\beta_{i,k} = \frac{\partial \mathbf{p}_t(X_{j^*}, \xi_{j^*})}{\partial x_{j,k}}, \text{ and } \beta_{i,0} = \frac{\partial \mathbf{p}_t(X_{j^*}, \xi_{j^*})}{\partial \xi_j}, \tag{4}$$

where j^* denotes house j chosen by buyer i. Therefore, if we recover the slope of the price function locally, then we can also locally recover buyer i's random coefficient β_i, that is, buyer-specific preferences for house characteristics.[16] Moreover, we can recover ξ_j by using the approach in Bajari and Benkard (2005) and Bajari and Kahn (2005).[17] This is very useful for our purpose, since recovered β_i and ξ_j help us to construct buyer-specific match values for internal versus external listings.

Notwithstanding these advantages, there are a few potential concerns with our approach. As for the first assumption, one may think that it would be natural to model strategic promotion by restricting the buyer's choice set, as agents-steered buyers face different choice sets than do non-steered buyers. However, this would lead to the discrete choice model with random choice sets, since actual choice sets are not observed. Because there are a large number of houses in each market for

[16] We do so by following Bajari and Benkard (2005) and Bajari and Kahn (2005) who propose an approach based on a nonparametric estimation of the price function. See Appendix B for the description of our estimation method.

[17] Bajari and Benkard (2005) show that if we assume that ξ_j represents a composite of all unobserved features of the house, that u_i is strictly increasing in ξ_j, and that X_j is independent of ξ_j, then ξ_j can be recovered during the first-stage estimation of the price function. This approach is described in Appendix B.

a buyer to choose among, such a model would require heavy computations, which is why we do not use this alternative approach. Nevertheless, as long as the alternative approach also assumes that commission structure/rate and REBBA determine agents' incentives to restrict the buyer's choice set, both approaches essentially lead to the same empirical predictions.[18]

Regarding the second assumption, a legitimate concern is that the presence of strategic promotion implies that the buyer may make potentially suboptimal choices with respect to X_j and ξ_j, in which case the first order conditions in (4) may not hold. We allow for this possibility by introducing optimization errors as follows.

$$\beta_{i,k} = \nu_i \frac{\partial \mathbf{p}_t(X_{j^*}, \xi_{j^*})}{\partial x_{j,k}}, \text{ and } \beta_{i,0} = \nu_i \frac{\partial \mathbf{p}_t(X_{j^*}, \xi_{j^*})}{\partial \xi_j}, \tag{5}$$

where ν_i's are random variables with positive support and unit mean.[19] With the presence of ν_i, the choice of (X_{j^*}, ξ_{j^*}) may not lead to the highest utility for buyer i. However, as long as ν_i's are not correlated with X_j and ξ_j, we can still recover buyer-specific preferences. To see this, note that the price function in the hedonic framework captures the equilibrium prices, rather than some absolute function which is fixed across markets. Therefore, if buyers' choices are determined by (5) instead of (4), the resulting equilibrium prices will be different from the equilibrium prices associated with (4), in which case the price function we can recover is not $\mathbf{p}_t(X_{j^*}, \xi_{j^*})$, but $\widetilde{\mathbf{p}}_t(X_{j^*}, \xi_{j^*})$, where $\widetilde{\mathbf{p}}_t(X_{j^*}, \xi_{j^*}) = \nu_{j^*} \mathbf{p}_t(X_{j^*}, \xi_{j^*})$ and $\nu_{j^*} = \nu_i$ for buyer i who has chosen j^*. Nevertheless, this does not prevent us from recovering buyer-specific preferences, because once we locally identify the slope of $\widetilde{\mathbf{p}}_t(X_{j^*}, \xi_{j^*})$, we can locally recover β_i by

$$\beta_{i,k} = \frac{\partial \widetilde{\mathbf{p}}_t(X_{j^*}, \xi_{j^*})}{\partial x_{j,k}}, \text{ and } \beta_{i,0} = \frac{\partial \widetilde{\mathbf{p}}_t(X_{j^*}, \xi_{j^*})}{\partial \xi_j}.$$

Thus, as long as potential suboptimal choices with respect to X_j and ξ_j are represented by optimization

[18] Under our approach, buyers know all housing options available to them. However, due to agents' strategic promotion, buyers face a higher search cost when buying external houses, which implies that these external houses are less likely to be included in the buyers' *effective* choice sets. To the extent that agents under our approach and agents under the alternative approach face the same commission incentives to promote internal listings, both approaches will generate qualitatively the same predictions about strategic in-house transactions.

[19] This type of optimization error is considered in Reiss and Wolak (2007).

errors as above, our hedonic framework still allows us to recover buyer-specific preferences.

Finally, our model assumes that the housing attributes X are continuous, so the first order conditions hold with equality. In reality, some housing attributes (e.g. lot size) take continuous values, other housing attributes (e.g. # bedrooms) take only discrete values – integer values in our case. In the latter case, the first order conditions will hold with inequality. Nevertheless, our approach to estimate the preference parameters from the equality conditions can be valid under the following conditions. If the price function is convex in a given discrete attribute $x_{j,k}$ (as in our case),[20] $\beta^*_{i,k}$ is partially identified by the inequality first order conditions, in which case our approach to set $\hat{\beta}_{i,k} = \frac{\partial \mathbf{p}_t(X_{j*}, \xi_{j*})}{\partial x_{j,k}}$ is similar to using the mid point for an interval variable. If the divergence between the estimated parameter from our approach and the true parameter has zero mean and is conditionally mean independent of $x_{j,k}$, our approach generates an unbiased estimator for the preference parameter. See Appendix C for the proof.

To the best of our knowledge, this paper presents the first attempt to recover buyer-specific preferences and idiosyncratic match values for internal versus external listings in the home matching process intermediated by real estate agents. Though a full-equilibrium model of brokerage choice interacted with house choice is more desirable, this is beyond the scope of the current paper, given our focus on recovering the extent of strategic promotion and the significant computational complexity that a full-equilibrium model may entail. We hope that future research in this area will improve upon our approach by developing more generalized models with less restrictive assumptions.

6.1.2 Constructing Counterfactual Match Values

Once we locally recover buyer-specific preferences, β_{ik}, and unobserved house characteristic, ξ_j, buyer i's match value for the purchased house j can be computed as follows:

$$U_j(\beta_i) = \sum_{k=1}^{m} \beta_{i,k} x_{j,k} + \beta_{i,0} \xi_j - p_j. \tag{6}$$

[20] If the price function is linear in $x_{j,k}$, the inequality first order condition becomes equality.

Let $V^1(\beta_i)$ and $V^0(\beta_i)$ respectively denote buyer i's match values for internal listings and external listings. The calculation in (6) then allows us to recover $V^1(\beta_i)$ for buyers in in-house transactions and $V^0(\beta_i)$ for buyers in cross-house transactions. However, to construct counterfactual $V^0(\beta_i)$ (or $V^1(\beta_i)$) for buyers in in-house (or cross-house) transactions, we need to know what other external (or internal) listings these buyers considered when searching for houses.

From the model constructed above, it follows that the buyer's choice of a house with (X_{j^*}, ξ_{j^*}) must be either optimal or close to being optimal (with optimization error ν_i). Thus, for a buyer who bought an internal listing, we can construct her counterfactual match value for external listings by using econometric matching techniques (e.g. Heckman, et al. 1997, 1998) that put higher weights on the houses with similar characteristics (in terms of both X_j and ξ_j) as house j^* and lower weights on those with different characteristics. Specifically, for a given buyer i who bought house j^* through an in-house transaction, we first compute $U_s(\beta_i) = \sum_{k=1}^{m} \beta_{i,k} x_{s,k} + \beta_{i,0} \xi_s - p_s$ for $s \in D_i^0$, where D_i^0 denotes a set of external listings in the same market. We then compute the weighted average of $U_s(\beta_i)$ using econometric matching techniques, where weights are computed by multivariate kernels as detailed in Appendix B. Similarly, we also construct counterfactual match values of internal listings for buyers who bought external listings.

6.1.3 Recovering Implicit Costs Associated with Strategic Promotion

In the second stage, the buyer makes a decision on whether purchase an in-house listing ($d_j = 1$). Four possible cases can occur in the second stage. The first is that houses with characteristics (X_{j^*}, ξ_{j^*}) or similar characteristics are available only among internal listings, in which case the buyer will choose $d_j = 1$.[21] The second is that such houses are available only among external listings, in which case the buyer will choose $d_j = 0$. In the remaining two cases, such houses are available in both internal listings and external listings, so that we can compute $V^1(\beta_i)$ and $V^0(\beta_i)$. If $V^1(\beta_i) + \gamma_i \geq V^0(\beta_i) - c_i$,

[21] Note that counterfactual match values are essentially weighted means, where weights are given by multivariate kernels as in Appendix B. Hence, for the first case, all externally listed houses are sufficiently different from the desired house, so that their weights are approximately zero. Given the lack of alternatives in the external listings, an in-house transaction would be optimal.

the buyer will choose $d_j = 1$, whereas if $V^1(\beta_i) + \gamma_i < V^0(\beta_i) - c_i$, the buyer will choose $d_j = 0$.

Clearly the first two cases entail efficient transactions. Combining the remaining two cases yields

$$d_{j^*} = 1 \text{ (or } = 0) \quad \Rightarrow \quad V^1(\beta_i) - V^0(\beta_i) + \gamma_i + c_i \geq 0 \text{ (or } < 0). \tag{7}$$

The inequality in (7) illustrates three key sources behind in-house transactions. The first is $V^1(\beta_i) - V^0(\beta_i)$, reflecting the positive utility gain that a buyer obtains from internal versus external listings. The second is γ_i, capturing transaction efficiencies from in-house transactions other than the utility gain reflected by $V^1(\beta_i) - V^0(\beta_i)$. The third is c_i, which is the extra cost of searching for external listings under agents' strategic promotion. Thus, efficient in-house transactions can be computed as the sum of in-house transactions for which $V^1(\beta_i) \geq V^0(\beta_i)$ as well as transactions for which buyers' optimal houses can only be found among internal listings. Accordingly, the remaining in-house transactions with $V^1(\beta_i) < V^0(\beta_i)$ provides an upper bound on the magnitude of strategic promotion in explaining in-house transactions. This upper bound provides useful information, but to obtain more information on the extent of strategic promotion, we need to estimate the distribution of γ_i and c_i. To this end, we follow Bajari and Kahn (2005) and impose a parametric assumption to estimate a discrete choice model based on the inequality in (7). Note that we do not attempt to fully separate c_i from γ_i, since it would require strong and arbitrary assumptions. We instead focus on the marginal effect of strategic promotion by further exploiting the difference-in-differences strategy discussed in Section 5.

6.2 Results from Structural Estimation

As discussed in the previous subsection, once we recover V^1 and V^0, we can quantify the extent of efficient matching by comparing V^1 and V^0. Panel A of Table 9 reports that among all the in-house transactions, the fraction of efficient matching is 0.643 for our sample period, indicating that 64.3% of buyers purchased houses from internal listings because they derive higher utility from internal listings than external listings.[22] This percentage is 62% for the pre-REBBA sample and 68.1% for the post-

[22]Panel B of Table 9 reports the fraction of efficient matching among all the cross-house transactions. This includes cross-house transactions with $V^1 < V^0$ as well as transactions for which buyers' optimal houses can only be found

REBBA sample, suggesting that the policy has improved efficient matching, possibly by discouraging strategic behavior to some extent. The estimate of 64.3% also implies that 35.7% of buyers end up purchasing a house from internal listings even though their interest is best matched by a house listed by other brokerages, indicating the extent of strategic promotion.

Note that this definition of efficient matching is an indication of the matching *outcome* rather than agents' *intention*. That is, even if an agent strategically promotes internal listings, the resulting in-house transaction is still considered efficient as long as the internal listing provides better match value for the buyer than external listings, in which case the buyer's "trust for the agent" is justified by the matching efficiency. In this sense, 35.7% should be considered as a lower bound for agents' strategic behavior in the sample market. On the other hand, 35.7% should also be considered as an upper bound for the actual suboptimal outcome resulted from strategic promotions. As we discussed earlier, when a buyer purchases an internal listing that does not match her preference best, it could be either because the buyer is constrained in her choice set due to agents' strategic promotion or because the buyer values transaction cost savings generated from an in-house transaction.

To further examine the determinants of efficiencies generated from in-house transactions, we use the sample of in-house transactions and consider the regression of the utility gain from internal versus external listings, measured by $V^1 - V^0$ (in 100,000). The results are reported in Table 10. A few findings emerge from column 1. First, the coefficients on ln(#listings) – #listings of the buyer's brokerage – and the top 10 franchise dummy are positive and statistically significant, indicating that larger firms tend to produce better internal matches. In addition, more typical houses are associated with smaller expected utility losses from in-house transactions. Intuitively, a potential buyer is more likely to find a better match from internal listings, if her brokerage owns a large number of listings spanning a large set of attributes and neighborhoods, or if the buyer is looking for tract homes and such

externally. In theory, all cross-house transactions should be optimal, but in practice, our approach may also yield suboptimal cross-house transactions in that $V^1 > V^0$. Table 9 shows that 95.3% of cross-house transactions are efficient, which assures that our approach is reasonable.

homes are available both internally and externally. Together, these findings are strongly in line with Prediction 3 in Section 3.2. Moreover, the coefficient on the REBBA is statistically significant and is about 0.06 in both columns, indicating that the average utility gain from purchasing an internal listing (relative to an external listing) is increased by about $6,000. This is consistent with the hypothesis that the legislation has reduced the extent of strategic promotions. Moving to column 2, we further control for housing characteristics. The key coefficients are largely similar to those in column 1. The coefficient on the atypicality index remains negative and statistically significant, although smaller in magnitude due to the fact that the atypicality index is essentially a function of house characteristics.

Now that we recovered $V^1 - V^0$, we further attempt to quantify the extent of strategic promotion controlling for possible transaction efficiencies. To this end, we use the step 3 estimation approach described in Appendix B and estimate the logit model based on the inequality in (7). To identify the effect of strategic promotion, we again apply the difference-in-differences approach discussed in the reduced-form analysis. Let Z_i denote a vector of variables related to search cost (c_i) but not related to transaction savings (γ_i), and W_i denote a vector of variables related to transaction savings (γ_i) but possibly related to search cost (c_i) as well. We thus include commission variables and their interactions with the REBBA in Z_i, while including #listings and days on the market in W_i.

Table 11 presents the results from estimating the logit model. In all specifications, we set the coefficient on $V^1 - V^0$ equal to 1, consistent with the inequality model in (7). Column 1 presents the baseline results. The coefficients on the split dummy and the low commission dummy are significantly positive, suggesting that agents are more likely to promote internal listings (and thereby increasing buyers' cost of searching for external listings) when they split commissions on the per-transaction basis. Such effect is much stronger when agents receive lower commission fees from listing agents. Moreover, the commission effects are reduced substantially after the implementation of the REBBA.

In columns 1-2, we attempt to control for potential savings in transaction costs. Two natural controls for cost savings are the number of listings by the buyer's brokerage and the buyer's time on the market. While the former is observed, the latter is not. However, to the extent that buyers' and sellers' time on the market are strongly positively correlated (Genesove and Han, 2012a), sellers' time on the market should serve as a good control for transaction efficiencies. Column 1 includes only the number of internal listings, while column 2 includes both control variables. The estimates are similar in both columns, suggesting that transaction cost savings associated with in-house transactions are unlikely to bias our estimates for strategic promotion. In column 3, we further include the total number of district-level monthly listings of all brokerages to control for overall market conditions. The results again remain robust. To further control for the the possible changes in house or brokerage characteristics that could be correlated with potential commission changes at the time of REBBA, we include the recovered unobserved house attribute, ξ_j, and its interaction with the REBBA dummy in column 4; and franchise fixed effects interacted with the REBBA dummy in column 5.[23] In both cases, the estimates remain close to those in column 1, both in magnitude and in precision.

The estimates reported in Table 11 also allow us to quantify the implicit cost of searching for external listings due to agents' strategic promotion. To this end, we use the coefficient estimates on commission variables, which are hypothesized to affect the extent of in-house transactions due to agents' incentives but not due to transaction efficiencies. The computed cost and utility gain/loss from in-house transactions are reported in Table 12, where we use the sample of in-house transactions before REBBA. To provide a benchmark, we first compute the median of $V^1 - V^0$ as a measure of utility gain/loss that a buyer derives from purchasing an internal listing versus an external listing. Panel A of Table 12 shows that the median of $V^1 - V^0$ is $25,501[24] for in-house transactions with $V^1 > V^0$, indicating that these transactions are indeed matching efficient, as they deliver a utility gain of $25,501 for an average home buyer (compared to the case where the buyer purchases an external listing). To assess aggregate welfare implications, we further add up from the distribution of house-

[23]Ideally, we wish to include brokerage fixed effects and their interactions with the REBBA dummy, but given that there are over 1,500 brokerage firms, including brokerage fixed effects in the discrete choice model is not tractable. As a result, we use franchise fixed effects instead.

[24]Note that this is interpreted in dollar values, since $U_j(\beta_i)$ is linear in house price.

specific values, and find that the magnitude of total utility gains is equivalent to 10.9% of the total volume of internal sales for the sample. Panel B, on the other hand, shows that the median utility loss is -$18,440 for in-house transactions with $V^1 < V^0$, indicating a substantial mismatch between buyers and the internal listings they bought. We further find that the magnitude of total utility losses in this case is equivalent to 4.1% of the total volume of internal sales.

In Panel C, we use in-house transactions with $V^1 < V^0$, and compute implicit costs associated with agents' promotion. We find that the median value is $19,467, and aggregated costs are equivalent to 4.3% of the total volume of internal sales. Note that this implicit cost should be interpreted not only as the extra cost that a buyer incurs when looking for a house from the pool of external listings, but also as a shadow price that her agent virtually adds to the price of external listings.[25] For in-house transactions with $V^1 > V^0$, agents' strategic interest does not lead to a distortion in the home search process, because home buyers' *ex ante* preference for internal listings agrees with agents' incentive to strategically promote. However, this is not the case for in-house transactions with $V^1 < V^0$. In the latter case, even though a buyer's preference is best matched by an external listing, the expected utility gain from purchasing externally is not sufficient to outweigh the associated cost imposed by the agent's strategic promotion, and hence, the resulting in-house matching is suboptimal.

Lastly, to investigate the effect of the REBBA on strategic in-house transactions, we first compute the predicted probabilities of in-house transactions using our samples.[26] Panel A of Table 13 reports the mean of these probabilities for the period before and after REBBA. It shows that the fraction of in-house transactions has declined from 19.4% to 17.4%. We then compute the counterfactual probabilities of in-house transactions in the absence of REBBA, and find that the mean of these probabilities is 18.8% as reported in Panel B. Hence, in the absence of REBBA, the fraction of in-house transactions would have been 18.8%, instead of 17.4%. These results indicate that REBBA has

[25]The latter concept is similar to the "virtual price" of unavailable goods, which is introduced in the literature that analyzes rationing (Hicks 1940; Rothbarth 1941) and new goods (Hausman 1996, 1999).

[26]For in-house (cross-house) transactions that we cannot match similar external (internal) listings, cross-house (in-house) transactions are impossible, and so their predicted probabilities of in-house sales are set equal to one (zero).

weakened the impact of strategic promotion on buyers' home search process, which accounts for 70% of a decrease in in-house transactions before and after REBBA. In Panel D, we compute the difference between the buyer's welfare with REBBA and the counterfactual buyer's welfare without REBBA.[27] We find that the removal of REBBA would decrease aggregate buyer welfare by $690 million, which is equivalent to 2.23% of the total volume of sales after REBBA.

7 Conclusion

About 20% of residential real estate transactions in North America occur within the same brokerage office. In this paper, we examine the causes and implications of in-house transactions for home buyers. We find that real estate agents are more likely to be engaged in in-house transactions when they are financially motivated, and this effect is weakened after the implementation of the REBBA. These findings are consistent with an agent-intermediated search model and provide strong evidence for the presence of strategic promotions.

To quantify the extent of in-house transactions that occur for strategic reasons, we propose a structural approach to recover the match values that a home buyer obtains from internal listings and external listings, which allows us to explicitly control for possible matching efficiencies. Our estimates suggest that about 64.3% of in-house transactions provide an efficient matching outcome, while the remaining in-house transactions are likely caused by strategic promotion. The latter cause a utility loss for home buyers, indicated by a substantial increase in search cost imposed by agents when these buyers search for external listings. Lastly, the REBBA has weakened the impact of agents' strategic promotion on the home matching process, significantly accounting for the decrease in in-house transactions before and after the regulatory change.

These findings are particularly relevant in the current housing markets as most states in the

[27] To compute the counterfactual buyer's welfare without REBBA, we again set the coefficients on the interactions between REBBA and commission variables equal zero, and compute counterfactual net utility for the buyer. Because the absence of REBBA can increase implicit costs associated with external listings, the buyer who bought her house from external listings might purchase from internal listings, in which case her counterfactual utility should include V^1 instead of V^0. Hence, we take into account this possibility by making use of counterfactual V^1 estimated from our second step.

U.S. have now required agency disclosure, indicating a regulatory reliance on disclosure to reduce inefficiency resulted from in-house sales. Our result suggests that the legislation does have desired effects by helping homebuyers make more informed choices and by constraining agents' ability to strategically promote. However, it cannot completely eliminate information inefficiencies, possibly due to the difficulty of monitoring and enforcing the required disclosure.

References

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, (2008), "Do Sellers' Brokers Raise or Reduce Home Prices? Evidence of Agency Costs from an Unusual Real Estate Market," NBER Working Paper Series 13796.

Conner, N. (2010), *Buying a Home: The Missing Manual*, Sebastopol: O'Reilly Media.

Epple, D. (1987), "Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products," *Journal of Political Economy*, *95(1)*, pp. 59-80.

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.

Fan, J., and Gijbels, I. (1996), *Local Polynomial Modelling and Its Applications*, London: Chapman & Hall.

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.

Genesove, D. and Han, L. (2012a), "Search and Matching in the Housing Market," *Journal of Urban Economics*, *72(1)*, pp. 31-45.

Genesove, D. and Han, L. (2012b), "Measuring the Thinness of Real Estate Markets," Working Paper. University of Toronto.

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

Hagiu, A. and Jullien, B. (2011), "Why Do Intermediaries Divert Search?" *RAND Journal of Economics*, *42 (2)*, pp. 337-62.

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 Strange, W. (2014), "Bidding Wars for Houses," *Real Estate Economics*, *41(3).*

Hausman, J. (1996), "Valuation of New Goods under Perfect and Imperfect Competition," in *The Economics of New Products*, eds. T. Bresnahan and R. Gordon, Chicago: University of Chicago Press, pp. 209-237.

Hausman, J. (1999), "Cellular Telephone, New Products, and the CPI," *Journal of Business and Economic Statistics*, *17(2)*, pp. 188-194.

Heckman, J., Ichimura, H., and Todd, P. (1998), "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme," *Review of Economic Studies*, *64*, pp. 605-654.

Heckman, J., Ichimura, H., and Todd, P. (1997), "Matching as an Econometric Evaluation Estimator," *Review of Economic Studies*, *65*, pp. 261-294.

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.

Hicks, J. (1940), "The Valuation of the Social Income," *Economica*, *42*, pp. 105-124.

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.

Hong, S.-H. (2013), "Measuring the Effect of Napster on Recorded Music Sales: Difference-in-Differences Estimates Under Compositional Changes," *Journal of Applied Econometrics*, *28(2)*, pp. 297-324.

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

Jia Barwick, P. and Pathag, P. (2014), "The Costs of Free Entry: An Empirical Study of Real Estate Agents in Greater Boston," forthcoming at RAND Journal of Economics.

Jiang, W., Nelson, A., and Vytlacil, E. (2014), "Liar's Loan? Effects of Origination Channel and Information Falsification on Mortgage Deliquency," *Review of Economics and Statistics*, *96(1)*, pp. 1-18.

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.

Munneke, H.J and Yavas A. (2001),"Incentives and Performance in Real Estate Brokerage," *The Journal of Real Estate Finance and Economics* 22 (1), 5-21.

Ngai, L. and S. Tenreyro (2014),"Hot and cold seasons in the housing market," *American Economic Review*, forthcoming.

Olazabal, A. (2003), "Redefining Realtor Relationships and Responsibilities: The Failure of State Regulatory Responses," *Harvard Journal on Legislation*, *40*, pp. 65-132.

Reiss, P. and Wolak, F. (2007), "Structural Econometric Modeling: Rationales and Examples from Industrial Organization," in *Handbook of Econometrics, Vol. 6A*, eds. Heckman, J. and Leamer, E, Amsterdam: Elsevier , pp. 4277-4415.

Rosen, S. (1974), "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, *82(1)*, pp. 34-55.

Rothbarth, E. (1941), "The Measurement of Changes in Real Income Under Conditions of Rationing," *Review of Economic Studies*, *8*, pp. 100-107.

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.

Yavas, A., Waller, B, and Bian, X. (2013), "Commission Splits in Real Estate Transactions," *the 48 th AREUEA Annual Conference Paper.*

Figure 1: Cooperating Brokerage's Fraction of In-House Transactions at District-level

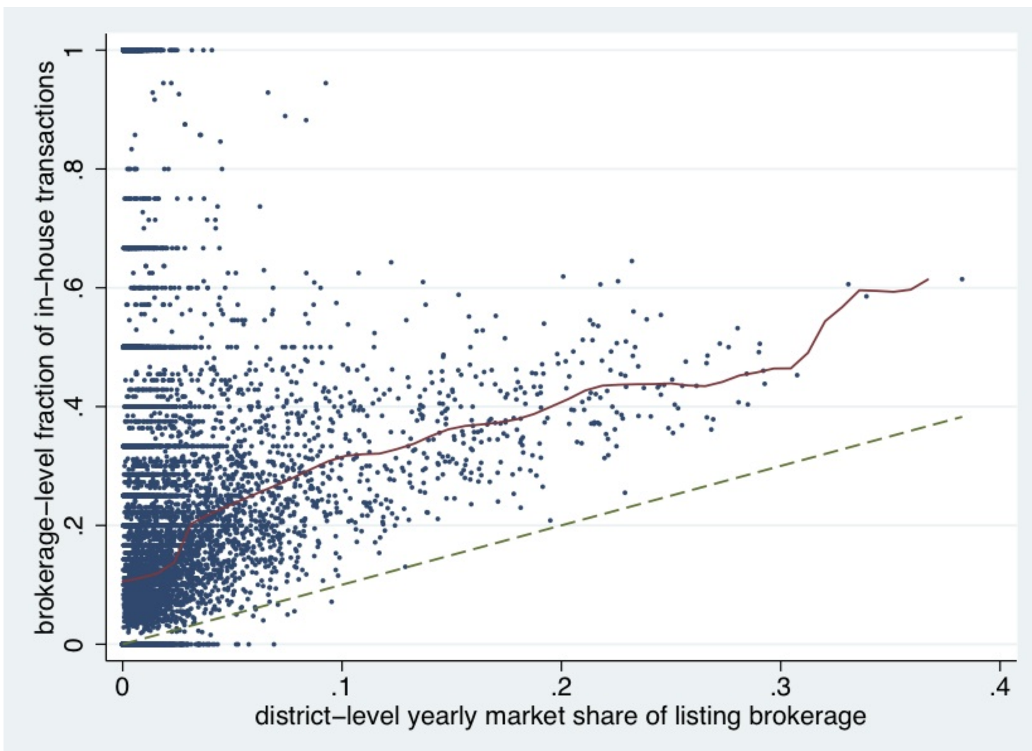

Table 1: Fraction of In-House Transactions by Brokerage Ranking[a]

	fraction of in-house	total share by brokerage ranking
ranking.1-10	0.220	0.214
ranking.11-50	0.195	0.333
ranking.51-100	0.161	0.182
ranking.101-200	0.166	0.150
ranking.201-	0.142	0.121
All	0.183	1.000

[a]We rank cooperating brokerages in order of their total market shares in our data. We then group them by their rankings. The "fraction of in-house" means the fraction of in-house transactions among total transactions carried out by brokerages in each ranking group. The "total share by brokerage ranking" refers to the share of each ranking group's total transactions among all transactions in our data.

Table 2: Fraction of In-House Transactions by #Agents[a]

	fraction of in-house	total share by #agents
#agent.501-	0.220	0.093
#agent.201-500	0.182	0.242
#agent.101-200	0.177	0.235
#agent.51-100	0.180	0.151
#agent.11-50	0.189	0.126
#agent.1-10	0.161	0.053
#agent.unknown	0.173	0.100
All	0.183	1.000

[a]We collect information on the number of agents for each brokerage, and group brokerages by the number of agents. The "fraction of in-house" means the fraction of in-house transactions among total transactions carried out by brokerages in each group. The "total share by #agents" refers to the share of each group's total transactions among all transactions in our data.

Table 3: Fraction of In-House Transactions by Year[a]

	fraction of in-house	total share by year
2001	0.202	0.099
2002	0.189	0.115
2003	0.189	0.113
2004	0.186	0.119
2005	0.192	0.117
2006	0.193	0.110
2007	0.177	0.123
2008	0.169	0.093
2009	0.153	0.111
All	0.183	1.000

[a]The "fraction of in-house" means the fraction of in-house transactions among total transactions in each year. The "total share by year" refers to the share of each year's total transactions among all transactions in our data.

Table 4: Baseline Results for "Difference-in-differences" using REBBA[a]

	(1)	(2)	(3)	(4)
split	0.013***	0.012***	0.011***	0.010***
	(0.002)	(0.002)	(0.002)	(0.002)
split×rebba	-0.014***	-0.014***	-0.011***	-0.011***
	(0.004)	(0.004)	(0.004)	(0.004)
low.commission			0.037***	0.037***
			(0.007)	(0.007)
low.commission×rebba			0.005	0.006
			(0.010)	(0.010)
split×low.commission			0.031**	0.031**
			(0.013)	(0.013)
split×low.commission×rebba			-0.040**	-0.038**
			(0.016)	(0.016)
rebba	-0.030**	-0.067***	-0.033**	-0.071***
	(0.015)	(0.015)	(0.015)	(0.015)
#listings	0.009***	0.008***	0.009***	0.008***
	(0.000)	(0.000)	(0.000)	(0.000)
house characteristics	yes	yes	yes	yes
occupancy types	yes	yes	yes	yes
listing price	no	yes	no	yes
year×month fixed effects	yes	yes	yes	yes
district fixed effects	yes	yes	yes	yes
$\bar{R}^2$	0.069	0.070	0.069	0.071
observations	206956	206956	206956	206956

[a]The dependent variable is the indicator variable for whether the transaction is in-house. House characteristics include ln(lot.front), ln(lot.depth), dummy variables for #bedrooms, #washrooms, and #garages. Occupancy types are the indicator variables for different types of occupants. #listings is the district-level monthly number of listings by the same brokerage as the buyer's cooperating brokerage. For listing price, we use the logarithm of prices. Robust standard errors clustered at the city block level in parentheses. * denotes significance at a 10% level, ** denotes significance at a 5% level, and *** denotes significance at 1% level.

Table 5: Robustness Checks to Time-Invariant Unobserved Heterogeneity[a]

	(1)	(2)	(3)	(4)	(5)
split×low.commission	0.031**	0.035*	0.024**	0.022**	0.024**
	(0.013)	(0.021)	(0.011)	(0.011)	(0.011)
split×low.commission×rebba	-0.040**	-0.032	-0.036**	-0.034**	-0.035**
	(0.016)	(0.032)	(0.015)	(0.015)	(0.015)
house characteristics	yes	no	yes	yes	yes
occupancy types	yes	yes	yes	yes	yes
listing price	no	no	no	no	no
year×month fixed effects	yes	yes	yes	yes	yes
district fixed effects	yes	yes	yes	yes	yes
house fixed effects	no	yes	no	no	no
brokerage fixed effects	no	no	yes	no	no
brokerage×region	no	no	no	yes	no
brokerage×price range	no	no	no	no	yes
$\bar{R}^2$	0.069	0.079	0.084	0.087	0.086
observations	206956	81662	206956	206956	206956
	(6)	(7)	(8)	(9)	(10)
split×low.commission	0.031**	0.035*	0.024**	0.021*	0.024**
	(0.013)	(0.021)	(0.011)	(0.011)	(0.011)
split×low.commission×rebba	-0.038**	-0.031	-0.034**	-0.033**	-0.034**
	(0.016)	(0.032)	(0.015)	(0.015)	(0.015)
house characteristics	yes	no	yes	yes	yes
occupancy types	yes	yes	yes	yes	yes
listing price	yes	yes	yes	yes	yes
year×month fixed effects	yes	yes	yes	yes	yes
district fixed effects	yes	yes	yes	yes	yes
house fixed effects	no	yes	no	no	no
brokerage fixed effects	no	no	yes	no	no
brokerage×region	no	no	no	yes	no
brokerage×price range	no	no	no	no	yes
$\bar{R}^2$	0.071	0.079	0.086	0.089	0.087
observations	206956	81662	206956	206956	206956

[a]The dependent variable is the indicator variable for whether the transaction is in-house. House characteristics include ln(lot.front), ln(lot.depth), dummy variables for #bedrooms, #washrooms, and #garages. Occupancy types are the indicator variables for different types of occupants. #listings is the district-level monthly number of listings by the same brokerage as the buyer's cooperating brokerage. For listing price, we use the logarithm of prices. Robust standard errors clustered at the city block level in parentheses. * denotes significance at a 10% level, ** denotes significance at a 5% level, and *** denotes significance at 1% level.

Table 6: Robustness Checks to Time-Varying Unobserved Heterogeneity[a]

	(1)	(2)	(3)	(4)	(5)
split×low.commission	0.031**	0.063*	0.020*	0.019*	0.021*
	(0.013)	(0.033)	(0.011)	(0.010)	(0.011)
split×low.commission×rebba	-0.040**	-0.089*	-0.032**	-0.031**	-0.033**
	(0.016)	(0.052)	(0.015)	(0.015)	(0.015)
house characteristics	yes	no	yes	yes	yes
occupancy types	yes	yes	yes	yes	yes
listing price	no	no	no	no	no
year×month fixed effects	yes	yes	yes	yes	yes
district fixed effects	yes	yes	yes	yes	yes
house×rebba fixed effects	no	yes	no	no	no
brokerage×rebba fixed effects	no	no	yes	no	no
brokerage×region×rebba	no	no	no	yes	no
brokerage×price range×rebba	no	no	no	no	yes
$\bar{R}^2$	0.069	0.086	0.085	0.088	0.087
observations	206956	41132	206956	206956	206956
	(6)	(7)	(8)	(9)	(10)
split×low.commission	0.031**	0.063*	0.020*	0.019*	0.021*
	(0.013)	(0.033)	(0.011)	(0.010)	(0.011)
split×low.commission×rebba	-0.038**	-0.089*	-0.031**	-0.030**	-0.032**
	(0.016)	(0.052)	(0.015)	(0.015)	(0.015)
house characteristics	yes	no	yes	yes	yes
occupancy types	yes	yes	yes	yes	yes
listing price	yes	yes	yes	yes	yes
year×month fixed effects	yes	yes	yes	yes	yes
district fixed effects	yes	yes	yes	yes	yes
house×rebba fixed effects	no	yes	no	no	no
brokerage×rebba fixed effects	no	no	yes	no	no
brokerage×region×rebba	no	no	no	yes	no
brokerage×price range×rebba	no	no	no	no	yes
$\bar{R}^2$	0.071	0.087	0.086	0.090	0.088
observations	206956	41132	206956	206956	206956

[a] The dependent variable is the indicator variable for whether the transaction is in-house. House characteristics include ln(lot.front), ln(lot.depth), dummy variables for #bedrooms, #washrooms, and #garages. Occupancy types are the indicator variables for different types of occupants. #listings is the district-level monthly number of listings by the same brokerage as the buyer's cooperating brokerage. For listing price, we use the logarithm of prices. Robust standard errors clustered at the city block level in parentheses. * denotes significance at a 10% level, ** denotes significance at a 5% level, and *** denotes significance at 1% level.

Table 7: Robustness Checks: Regressions of House Characteristics[a]

	Dependent Variable:				
	#bedrooms	#washrooms	#garages	ln(lot.front)	ln(lot.depth)
	(1)	(2)	(3)	(4)	(5)
split×rebba	-0.002	-0.001	-0.003	-0.004	0.002
	(0.001)	(0.001)	(0.003)	(0.004)	(0.003)
low.commission×rebba	-0.004	0.006	0.008	0.008	0.010
	(0.006)	(0.006)	(0.010)	(0.009)	(0.006)
split×low.commission×rebba	0.008	-0.005	0.029	-0.002	0.004
	(0.009)	(0.008)	(0.038)	(0.013)	(0.010)
other house characteristics	yes	yes	yes	yes	yes
occupancy types	yes	yes	yes	yes	yes
year×month fixed effects	yes	yes	yes	yes	yes
district fixed effects	yes	yes	yes	yes	yes
$\bar{R}^2$	0.327	0.362	0.267	0.421	0.168
observations	206956	206956	206956	206956	206956

[a]Each column uses a different house characteristic as the dependent variable. These regressions are similar to the regression reported in column 6 of Table 4, except that the dependent variable is different, and control variables include other housing characteristics not used in the dependent variable. Robust standard errors clustered at the city block level in parentheses. * denotes significance at a 10% level, ** denotes significance at a 5% level, and *** denotes significance at 1% level.

Table 8: Robustness Checks: Hot/Cold Markets and Typical/Distinct Houses [a]

	hot market typical homes (1)	cold market typical homes (2)	hot market distinct homes (3)	cold market distinct homes (4)
split×low.commission	0.082*	0.026*	0.090**	0.031**
	(0.042)	(0.014)	(0.039)	(0.015)
split×low.commission×rebba	-0.071	-0.034*	-0.091*	-0.048**
	(0.051)	(0.020)	(0.048)	(0.023)
house characteristics	yes	yes	yes	yes
occupancy types	yes	yes	yes	yes
listing price	yes	yes	yes	yes
year×month fixed effects	yes	yes	yes	yes
district fixed effects	yes	yes	yes	yes
$\bar{R}^2$	0.059	0.060	0.067	0.073
observations	14009	82367	15167	86240

[a]A "typical" market is defined to be houses with atypicality index less than 0.1 (about 50th percentile), and a "heterogeneous" market includes those with atypicality index larger than 0.1. We follow Haurin (1988) to compute atypicality index which intends to capture how far each house is from the average house in terms of characteristics. Note also that we compute atypicality index for each census tract separately. For census tract markets with a small number of listings, atypicality index is not meaningful, so we use only census tract markets with more than 100 listings. A "hot" market is defined to be houses with sales-list price ratio larger than 1, and a "cold" market is defined to be those with sales-list price ratio less than 1. These regressions are similar to the regression reported in column 4 of Table 4, except for the sample used in each regression. Robust standard errors clustered at the city block level in parentheses. * denotes significance at a 10% level, ** denotes significance at a 5% level, and *** denotes significance at 1% level.

Table 9: Fraction of Efficient Transactions [a]

A. among in-house transactions	
before the REBBA was implemented	0.620
after the REBBA was implemented	0.681
both periods	0.643
B. among cross-house transactions	
before REBBA was implemented	0.956
after REBBA was implemented	0.949
both periods	0.953

[a]Efficient in-house transactions include in-house transactions with $V^1 > V^0$ as well as in-house transactions for which similar external listings cannot be found. Similarly, efficient cross-house transactions include cross-house transactions with $V^1 < V^0$ as well as cross-house transactions for which similar internal listings cannot be found.

Table 10: Regressions of $V^1 - V^0$ for In-House Transactions [a]

	(1)	(2)
rebba	0.0577*	0.0632*
	(0.0340)	(0.0334)
ln(#listings)	0.0533***	0.0525***
	(0.0038)	(0.0037)
top10.coop.franchise	0.0951***	0.0896***
	(0.0175)	(0.0170)
atypicality index	-0.5059***	-0.1856***
	(0.0596)	(0.0535)
year×month	yes	yes
district	yes	yes
house characteristics	no	yes
$\bar{R}^2$	0.227	0.263
observations	33853	33853

[a] The dependent variable is $V^1 - V^0$ (in 100,000). The regressions are estimated by using only in-house transactions, so that $V^1 - V^0$ in these regressions measures the efficiency gain for internal listings relative to external listings. We use observations for which we can estimate V^1 and V^0. #listings is the district-level monthly number of listings by the same brokerage as the buyer's cooperating brokerage. top10.coop.franchise is the dummy variable for whether the cooperating brokerage of the transaction belongs to top 10 cooperating brokerage firms. Robust standard errors clustered at the city block level in parentheses. * denotes significance at a 5% level, and ** denotes significance at a 1% level.

Table 11: Step 3 Estimation Results[a]

	(1)	(2)	(3)	(4)	(5)
split×low.commission	0.384***	0.380***	0.380***	0.371***	0.371***
	(0.135)	(0.135)	(0.135)	(0.133)	(0.137)
split×low.commission×rebba	-0.501***	-0.499***	-0.499***	-0.471***	-0.424**
	(0.188)	(0.188)	(0.188)	(0.187)	(0.191)
split	0.196***	0.195***	0.195***	0.176***	0.176***
	(0.032)	(0.032)	(0.032)	(0.032)	(0.054)
split×rebba	-0.716***	-0.716***	-0.716***	-0.720***	-0.943***
	(0.050)	(0.050)	(0.050)	(0.050)	(0.079)
low.commission	0.181**	0.178**	0.178**	0.172**	0.169**
	(0.092)	(0.092)	(0.092)	(0.091)	(0.089)
low.commission×rebba	-0.116	-0.115	-0.115	-0.054	-0.030
	(0.128)	(0.128)	(0.128)	(0.127)	(0.128)
rebba	-1.698***	-1.693***	-1.692***	3.906***	3.270***
	(0.197)	(0.197)	(0.197)	(1.327)	(1.356)
#listings	0.024***	0.024***	0.024***	0.024***	0.023***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
days.on.market		0.001***	0.001***	0.001***	0.001***
		(0.000)	(0.000)	(0.000)	(0.000)
#total.listings			-0.000	-0.000	-0.000
			(0.000)	(0.000)	(0.000)
ξ_j				1.002***	1.006***
				(0.086)	(0.086)
ξ_j×rebba				0.503***	0.467***
				(0.121)	(0.122)
house characteristics	yes	yes	yes	yes	yes
occupancy types	yes	yes	yes	yes	yes
year×month	yes	yes	yes	yes	yes
district	yes	yes	yes	yes	yes
franchise×rebba	no	no	no	no	yes
observations	50412	50412	50412	50412	50409

[a]The dependent variable is the indicator variable for whether the transaction is in-house or not. We use observations for which we can estimate V^1 and V^0. The coefficient on $V^1 - V^0$ (in 100,000) is set equal to 1. #listings is the district-level monthly number of listings by the same brokerage as the buyer's cooperating brokerage. #total.listings is the district-level monthly number of listings of all brokerages. The results are estimated by logit. Column 4 is estimated by a conditional fixed-effect logit model, so that it uses only houses which alternate in-house transactions and cross-house transactions in different transactions. Robust standard errors clustered at the city block level in parentheses. * denotes significance at a 10% level, ** denotes significance at a 5% level, and *** denotes significance at 1% level.

Table 12: Utility Gain and Implicit Costs of In-House Transactions Before REBBA[a]

A. utility gain from in-house transactions with $V^1 > V^0$	
median of $V^1 - V^0$	25,501
aggregate $V^1 - V^0$	284 million
B. utility loss from in-house transactions with $V^1 < V^0$	
median of $V^0 - V^1$	18,440
aggregate $V^0 - V^1$	106 million
C. implicit costs associated with strategic promotion	
median costs	19,467
aggregate costs	112 million

[a]Utility gains are computed by $V^1 - V^0$, while the implicit costs are calculated by using the coefficient estimates from column 3 in Table 11. We use the sample of in-house transactions before the REBBA with non-zero costs. Panel C reports implicit costs for in-house transactions with $V^1 < V^0$.

Table 13: REBBA Effect on Strategic In-House Transactions[a]

A. mean predicted prob. of in-house transactions	
before REBBA	0.194
after REBBA	0.174
B. mean predicted prob. of in-house transactions without REBBA	
for the sample after REBBA	0.188
C. % reduction in in-house transactions attributable to REBBA reducing strategic promotion	
	= (0.188-0.174)/(0.194-0.174)×100% = 70%
D. an increase in aggregate buyer welfare due to REBBA, relative to total sales prices after REBBA	
	= \$690 million/\$31 billion×100% = 2.23%

[a]Panels A-B report the mean of predicted probabilities of in-house transactions. In Panel B, to compute the mean probability of in-house transactions without REBBA, we use only the samples after REBBA and compute the predicted probability of in-house transactions by setting the coefficients on the interactions between REBBA and commission variables equal zero. In Panel D, we compute the difference between the buyer's welfare with REBBA and the counterfactual buyer's welfare without REBBA. The table reports the aggregate value for these differences.

Web Appendices

Appendix A: An Agent-Intermediated Search Model

Our motivating theoretical framework follows closely Hagiu and Jullien (2011), who provide an inspiring economics analysis of search diversion in an online shopping setting. We apply their search diversion theory to the real estate brokerage industry and show that agents can misguide homebuyers by introducing noise in the home search process. Unlike their model that assumes an intermediary receives a fixed amount of revenues from each store visit by buyers regardless of actual sales, we make a different assumption to reflect the key compensation feature of the real estate brokerage industry. That is, agents receive a fixed percentage of *realized* sales revenues and this percentage is *larger* when a transaction occurs within the same brokerage firm. As shown later, such compensation feature is the driving source of agents' strategic promotion.

To simplify the analysis of the search process in the housing market, let us consider a setup where there are two types of buyers (buyer 1 and buyer 2), two types of houses (house 1 and house 2), and one cooperating agent.

Buyers: Buyers differ along two dimensions: preferences for houses and search costs. Along the first dimension, there are two types of buyers: type 1 buyers make up a fraction α of the population and derive net utilities u^H from visiting house 1 and u^L from visiting house 2; type 2 buyers make up a fraction $1-\alpha$ of the population and derive net utilities u^H from visiting house 2 and u^L from visiting house 1. We assume that $u^H > u^L$, which implies that *ex ante* type 1 buyers prefer house 1 over house 2, and that type 2 buyers prefer house 2 over house 1 in the sense that will be described below. Along the second dimension, buyers are differentiated by the search cost c they incur each time they visit a house. We use $F(c)$ to denote the cumulative distribution of c. They can only visit at most two houses sequentially.

More specifically, take buyer 1 as an example. Her valuation of a specific house h, v_h^1, is unknown prior to the visit but is learnt upon inspection of the house, so that the expected utility prior to visiting the house is $u_h^1 = \int_{p_h}^{k} (v_h^1 - p_h) dG(v_h^1)$, where $G(v)$ denote the cumulative distribution of v,[28] $k = H$ if $h = 1$ and $k = L$ if $h = 2$. Assuming that $0 < L < H$, it follows that $u_1^1 > u_2^1$, and that $u^H = u_1^1$ and $u^L = u_2^1$. In other words, *ex ante* house 1 is a better match for buyer 1 than house 2. Note that u_h^1 should be interpreted as encompassing the utility of just "looking around" the house plus the expected utility of actually buying the house, net of the price paid. Upon visiting a house, a buyer observes the realized value of being matched with a specific house, v_h^1, and then decides to whether to buy the house.

Houses: Houses also differ along two dimensions: matching quality and the listing brokerage firm. Along the first dimension, as described above, type 1 house stands for houses that *ex ante* match the buyer 1's preference best, whereas house 2 stands for houses that *ex ante* match the buyer 2's preference best. Along the second dimension, house 1 is listed by a firm that is different from the cooperating brokerage firm, whereas house 2 is listed by an agent affiliated with the same brokerage firm.

For simplicity, we assume that prices of houses are exogenously given at p_1 and p_2. This is because house prices are typically determined by general market conditions, which is much broader than the choice of intermediaries. In addition, the listing price of a house is publicly advertised before the cooperating agents and their buyers are engaged in the search process. To the extent that the sales and listing prices are highly positively correlated, the exogeneity assumption is justified.

Cooperating Agent: The cooperating agent observes each buyer's type (1 or 2) but not her search cost c. Since the agent is assumed to have superior information about houses available in the market, he immediately knows which house *ex ante* fits the buyer's preference best. Following Hagiu and Jullien (2011), we denote by q_1 the probability that the agent takes buyer 1 to house 1 for the *first* visit. If the cooperating agent always optimize the match process between buyers and houses,

[28] For any given buyer, v_1^1 and v_2^1 are assumed to be independently distributed.

then we should expect $q_1 = 1$. In contrast, we say that the cooperating agent "strategically" promotes his own firm's listings (i.e., house 2) whenever $q_1 < 1$.

The cooperating agent receives a fixed percentage of actual sales price as commission income when a transaction is completed. This income is then split with the agent's affiliated brokerage firm. In net, the cooperating agent obtains a fixed share of transaction price, τ_1 (or τ_2), from the sale of house 1 (or house 2). If an agent receives a bonus for promoting internal listings, then all else equal, $\tau_2 > \tau_1$. As a result, the cooperating agent for buyer 1 may sometimes find it profitable to recommend house 2 which generates the highest revenue, rather than house 1 which matches buyer 1's preference best. The incidence of $q_1 < 1$ captures precisely the inefficiency resulted from the commission structure described above.

Timing: The timing of decisions is as follows: (i) the agent publicly announces q_1; (ii) buyers observe q_1; (iii) buyers decide whether to follow agent's guidance, engage in the search process, and make their purchase decisions after visiting the house(s).

Solving the Model

Without loss of generality, let us focus our analysis on type 1 buyers. First, consider a type 1 buyer with high search costs, i.e., $c > u^H(p_1)$. In this case, the buyer would not visit any of the two houses, and as a result, the agent receives zero commission income.

Next, consider a type 1 buyer with low search costs, i.e., $c \leq u^L(p_2)$. She will visit both houses irrespective of where the agent directs her for her first visit. Upon visiting both houses, the buyer compares two houses and purchases the one that gives her the largest net realized utility, $\max\{v_1 - p_1, v_2 - p_2\}$. Accordingly, the probability of buyer 1 purchasing house 2, ρ_2^1, is given by:

$$\begin{aligned} \rho_2^1 &\equiv \Pr(v_2^1 - p_2 > v_1^1 - p_1) \\ &= \int_{p_1}^{H} \int_{v_1^1 - p_1 + p_2}^{L} dG_L(v_2^1) dG_H(v_1^1) \\ &= \int_{p_1}^{H} \left(1 - G_L(v_1^1 - p_1 + p_2)\right) dG_H(v_1^1) \end{aligned} \tag{8}$$

Thus the cooperating agent receives commission income $\tau_2 p_2$ with probability ρ_2^1 and $\tau_1 p_1$ with probability $1-\rho_2^1$.

Finally, consider a type 1 buyer with intermediate search costs, i.e., $u^L(p_2) \leq c \leq u^H(p_1)$. In this case, if the buyer is first sent to house 1 (which happens with probability q_1), she would make a purchase and stop visiting another house if the net realized value from buying house 1, $(v_1 - p_1)$, is greater than the expected utility of continuing visiting house 2, $\max\{u^L(p_2) - c, 0\}$. Since $u^L(p_2) \leq c$, $\max\{u^L(p_2) - c, 0\} = 0$, so that she will not visit house 2 with probability 1. If she is first sent to house 2 (which happens with probability $1 - q_1$), she would stop searching if and only if the net realized utility, $(v_2 - p_2)$, is greater than the expected utility of continuing visiting house 1, that is, $\max\{u^H(p_1) - c, 0\} = u^H(p_1) - c$. In the event when buyer 1 visits both houses, she will purchase house 1 with probability $1 - \rho_2^1$ and house 2 with probability ρ_2^1.

Knowing the probability q_1, a type 1 buyer follows the agent's guidance if her search cost is above $u^L(p_2)$ and below some critical value u_1, where $u_1 = c$ is implicitly defined by

$$q_1 u^H(p_1) + (1-q_1)\int \max\left(v_2 - p_2, u^H(p_1) - c\right) g_L(v_2)dv_2 - c = 0$$

Note that when $q_1 = 1$, we have $u_1 = u^H(p_1)$ and $\frac{du_1}{dq_1} = u^H(p_1) - u^L(p_2)$.

Turning to the agent's side, the revenue he derives from type 1 buyers is then:

$$\begin{aligned}\Pi_1 \quad = \quad & \left(\tau_1 p_1 (1-\rho_2^1) + \tau_2 p_2 \rho_2^1\right) F(u_L) + q_1 \tau_1 p_1 \left(F(u_1) - F(u_L)\right) \\ & +(1-q_1)\left[\left(\tau_1 p_1(1-\rho_2^1) + \tau_2 p_2 \rho_2^1\right)\int_{u^L}^{u_1} G_L(p_2 + u^H - c) f(c) dc \right. \\ & \left. + \tau_2 p_2 \int_{u^L}^{u_1} (1 - G_L(p_2 + u^H - c)) f(c) dc\right] \end{aligned} \tag{9}$$

The first term represents the revenue that the agent receives from type 1 with low search costs, i.e., with $c \leq u^L(p_2)$. The second term represents the revenue that the agent receives from type 1 buyers who have intermediate search costs, i.e., with $u^L(p_2) \leq c \leq u_1$, and have been efficiently matched to house 1 on their first visit. The third term represents the revenue that the agent receives from type 1 buyers who have intermediate search costs but have been strategically directed to house 2 first. Note that the first integrant term is the probability that the buyer decides to continue searching conditional on having visited house 2 in the first round of search. In this case, the agent receives $\tau_1 p_1$ with probability $1-\rho_2^1$ and $\tau_2 p_2$ with probability ρ_2^1.

Maximizing (2) over q_1 yields the following proposition, which contains our baseline results:

Proposition 1 *The cooperating agent "strategically" promotes in-house transactions (i.e., $q_1 < 1$) if and only if*

$$\frac{\tau_2 p_2}{\tau_1 p_1} > \frac{F(u^H) - F(u^L) - (1-\rho_2^1)\int_{u^L}^{u_1} G_L(p_2 + u^H - c) f(c) dc + f(u^H)(u^H - u^L)}{F(u^H) - F(u^L) - (1-\rho_2^1)\int_{u^L}^{u_1} G_L(p_2 + u^H - c) f(c) dc} \tag{10}$$

Proof: The cooperating agent maximizes (9) over q_1. Using the fact $u_1(q_1 = 1) = u^H(p_1)$ and $\frac{du_1}{dq_1}(q_1 = 1) = u^H(p_1) - u^L(p_2)$, it is straightforward to show that $\frac{\partial \Pi_1}{\partial q_1}(q_1 = 1) < 0$ if and only if (10) holds. ■

Strategic In-House Transactions

Condition (10) is central to understanding of agents' incentives to strategically promote in-house transactions. In particular, at $q_1 = 1$, all type 1 buyers with intermediate search costs will be first directed to houses that match their preference best, leading to an efficient matching outcome. By laying out conditions under which the cooperating agent lowers q_1 below 1, condition (10) immediately delivers several predictions of strategic in-house transactions that can be taken to the data.

Prediction 1: The commission structure matters. It is clear from condition (10) that the optimal amount of strategic promotion increases with the ratio $\frac{\tau_2 p_2}{\tau_1 p_1}$. If the prices of two houses are not too different from each other (which is not too unreasonable given that buyers usually specify a price range for houses they search for), the larger is the ratio $\frac{\tau_2}{\tau_1}$, the more likely condition (10) will hold, and the stronger is the agent's incentive to promote her own firm's listings. In the brokerage industry, agents need to split commission fees with their affiliated brokerage offices, in return for the brand value and for the supporting services that brokerage offices provide. In practice, full commission brokerage firms, such as ReMax, let the agents keep all commission fees but require a fixed amount of upfront fees each month. More traditional brokerages firms, such as Royal LePage, split commission fees with their agents on the per-transaction basis. Naturally, the revenues in the latter type of brokerage firms strictly increase with the number of either end of transactions. Therefore, these brokerage firms are more likely to reward their agents for selling internal listings, making $\tau_2 > \tau_1$.[29] Thus, we expect that the per-transaction split commission structure is associated with a stronger presence of strategic in-house transactions.

In addition to commission structure, commission rate also matters. Note that the commission rate for a cooperating agent is typically predetermined when the listing is posted on the MLS. While the commission rate is usually set at 2.5%, some listing agents would offer a higher or lower rate to cooperating agents. Intuitively, by rewarding cooperating agents a greater proportion of the commission, an external listing agent can effectively increase τ_1 in condition (10), and this helps offset the promotion bonus that the cooperating agent receives from her own firm for promoting internal listings. Conversely, when the commission rate offered by a listing agent is low, the cooperating agent is more likely to respond to the financial incentives offered by the brokerage firm for promoting in-house transactions. The strategy of using substandard commission rates to artificially increase the frequency of dual-agency transactions is discussed in Yavas, *et al* (2013) and also evidenced by a recent industry report.[30] Thus, we expect that lower commission rates offered by listing agents are associated with a stronger presence of strategic in-house transactions.

Prediction 2: The extent to which cooperating agents can promote in-house transactions depends on the difference in the matching quality that a given buyer obtains from internal and external listings. As we can see from condition (10), the bigger is $(u^H - u^L)$ and/or ρ_2^1, the smaller is the likelihood of strategic promotion ($q_1 < 1$). Intuitively, if the best house that a buyer can find from external listings is far better than the one she can find from internal listings, either *ex ante* (reflected by $u^H - u^L$) or *ex post* (reflected by ρ_2^1), then it becomes difficult for her agent to promote internal listings. Empirically,

[29]See the footnote 3 for discussion on the related industry practice.

[30]For example, a recent report by the *Consumer Advocates in American Real Estate* states that "offering less than the going rate in your area will decrease the financial attractiveness of your home [to cooperating agents] and increases the likelihood that your broker will collect a double commission" (see an article titled "Dual Agency Schemes" in http://www.caare.org/ForBuyers, accessed August 1, 2014).

we do not observe matching quality. However, matching in housing markets is typically characterized by increasing returns to scale (Ngai and Tereyro 2014; Genesove and Han 2012b). When a brokerage firm has a larger number of listings which a buyer can choose among, there should be less dispersion in the buyer's valuation of her most-preferred house from the market-wide pool and from the internal listings. As a result, the brokerage firm will find it easier to promote its own listings. Although the promoted listings may not match the buyer's preference best, the resulting efficiency loss should be smaller since these listings are closer to the buyer's preference.

Prediction 3: The brokerage firm's ability to strategically promote in-house transactions also depends on whether buyers are aware of agents' incentives to strategically promote. So far, the model has assumed that buyers faced with a known probability of q_1. If buyers are not aware of agents' strategic incentives, this would remove the dependence of $\frac{du_1}{dq_1}$ in deriving the derivative of Π_1 with respect to q_1. As a result, the right-hand-side of condition (10) is reduced to 1. In this case, the agent's incentive to promote in-house transactions would purely depend on the financial reward ($\frac{\tau_2}{\tau_1}$) and search cost. The quality difference would no longer matter, since buyers believe that agents always match them to their first best house and hence would not be sensitive to the difference between the first and second best houses ($u^H - u^L$). As discussed later, our sample covers a natural experimental opportunity permitted by a legislation that required real estate agents engaged in in-house transactions to disclose the possibility of strategic promotion to both buyers and sellers. This provides an opportunity for us to empirically test this prediction.

Efficient In-House Transactions

In-house transactions could also occur for efficiency rather than incentive reasons. We define an in-house transaction as "efficient" if the buyer's net utility from internal listings is larger than her net utility from external listings, either *ex ante* or *ex post.* In our model, the probability of efficient in-house transactions is given by:

$$P = \left(\alpha\rho_2^1 + (1-\alpha)(1-\rho_1^2)\right) F(u_L) + (1-\alpha)(F(u_2) - F(u_L)) \tag{11}$$

The first term in (11) refers to the probability of in-house transactions by type 1 and type 2 buyers with low search costs. With probability ρ_2^1, a type 1 buyer purchases house 2 because house 2 delivers larger net realized utility than house 1. Similarly, with probability $1-\rho_1^2$, a type 2 buyer purchases house 2. In both cases, transactions occur within the same brokerage firm, and these in-house transactions represent the outcome of buyers' own choices rather than agents' promotional efforts. In particular, the low search cost removes the reliance of buyers on agents in looking for ideal homes, resulting in an efficient match between buyers and houses, regardless of whether the transaction occurs within the same brokerage firm or not.

The second term in (11) refers to the probability of in-house transactions by type 2 buyers with intermediate search costs. It is straightforward to show that with probability $q_2 = 1$, all type 2 buyers will be first directed to house 2.[31] Moreover, these buyers would end up purchasing house 2, since the expected utility of visiting house 1 is less than the search cost. In this case, the agents' incentive to promote their own listings is consistent with the buyers' interest, because these listings match the buyers' *ex ante* preference best. This type of in-house transactions, although promoted by cooperating agents, represents an efficient matching outcome.

Thus, the model predicts two types of efficient in-house transactions. Under the first type, a buyer receives the largest *ex post* utility from an internal listing through her own comparison of all available listings. Under the second type, a buyer is directed by an agent to an internal listing that matches her *ex ante* preference best. Since the first type of in-house transactions are not driven by agents, we focus our discussion on the second type of in-house transactions, which is driven by mutual interests of buyers and their agents. Note that our model takes the buyer's choice of the cooperating brokerage

[31] To see this, note that for type 2 buyers, a similar condition as condition (10) can be obtained by changing only the left-hand side in (10) to $\frac{\tau_1 p_1}{\tau_2 p_2}$. Assuming that houses 1 and 2 are in the same price range, with the in-house promotion bonus, the left-hand side is less than 1, while the right-hand side is greater than 1. This implies that the threshold condition will never be met, and hence $q_2 = 1$.

as given (reflected by an exogenous α), hence we are unable to explicitly model the sources of efficient in-house transactions. In practice, buyers' *ex ante* preference for an internal listing may agree with the agent's self-promotion interest for various reasons. For example, an in-house transaction may lower transaction costs and improve the efficiency in the bargaining and closing stage, making buyers more likely to favor transactions within the same brokerage house. Alternatively, a buyer may choose a cooperating agent simply because the agent's affiliated firm specializes in listing houses that fit the buyer's interest best. In both cases, in-house transactions are caused by a mixture of transaction efficiencies and information advantages.

Appendix B: Details on Three-Step Estimation Approach

To estimate our model, we follow and modify the estimation approach used by Bajari and Kahn (2005) which involves three steps. The first step estimates the hedonic price function using nonparametric methods, and recovers buyer-specific utility parameters β_i. In the second step, we estimate $V^1(\beta_i)$ and $V^0(\beta_i)$. In the third step, we estimate the distribution of γ_i and c_i. In what follows, we describe our approach in detail. We also provide discussion on identification where applicable.

Step 1: estimating the price function and recovering ξ_j and β_i

In the first step, we recover the slope of the price function in a local neighborhood of the characteristics of house j^* chosen by buyer i. To this end, we use a nonparametric estimation of the hedonic price function, and in particular, we use the local linear regression.[32]

Following Bajari and Kahn (2005), we consider a linear approximation of $\mathbf{p}_t(X_j, \xi_j)$ in a local neighborhood of house j^*'s observed characteristics. Specifically, we consider

$$\log\left(\mathbf{p}_t(X_j, \xi_j)\right) = \alpha_{j,0} + \sum_{k=1}^{m} \alpha_{j,k} x_{j,k} + \eta_t + \xi_j, \tag{12}$$

[32]Fan and Gijbels (1996) provide detailed treatment on local linear (or polynomial) regression. We could instead use other nonparametric methods such as a kernel estimator (e.g. Nadaraya-Watson estimator or Gasser-Müeller estimator) or a series estimator. However, Bajari and Benkard (2005) found that a local linear kernel estimator as in Fan and Gijbels (1996) worked best. For this reason, we also use the local linear regression.

where $\alpha_j = (\alpha_{j,0}, \ldots, \alpha_{j,m})$ is a vector of the hedonic coefficients, η_t captures market fixed effects, and we use a logarithm of the price function instead of its level in order to improve the fitting of the price function. In our estimation, we first regress $\log(p_j)$ on X_j, district fixed effects, and year×month fixed effects. We then use the demeaned prices, $\widetilde{\log(p_j)} \equiv \log(p_j) - \hat{\eta}_t$, and estimate α_{j^*} for each value of $j = 1, \ldots, J_t$ by using local fitting methods which solve

$$\min_{\alpha} \sum_{j=1}^{J_t} \left\{ \widetilde{\log(p_j)} - \alpha_0 - \sum_{k=1}^{m} \alpha_k x_{j,k} \right\}^2 K_B(X_j - X_{j^*}), \tag{13}$$

where $K_B(\mathbf{v})$ is the kernel function. Given $K_B(X_j - X_{j*})$, α_{j^*} can be estimated by weighted least squares for each j^*. As for $K_B(\mathbf{v})$, we use the product of univariate Gaussian kernel, following Bajari and Kahn (2005) who used $K_B(\mathbf{v}) = \prod_{k=1}^{m} \frac{1}{b} N(\frac{1}{b} \frac{v_k}{\hat{\sigma}_k})$, where b is a scalar bandwidth, $N(\cdot)$ is the univariate Gaussian kernel, and $\hat{\sigma}_k$ is the sample standard deviation of v_k.

Once we estimate α_{j^*}, we can recover an estimate of ξ_{j^*}. Following Bajari and Benkard (2005) and Bajari and Kahn (2005), we recover an estimate of ξ_{j^*} from the residual in (12), which yields

$$\xi_{j^*} = \log(p_{j^*}) - \alpha_{j^*,0} - \sum_{k=1}^{m} \alpha_{j^*,k} x_{j^*,k} - \eta_t.$$

We then use (4) to recover $\beta_{i,k}$ as follows.

$$\hat{\beta}_{i,k} = \frac{\partial \mathbf{p}_t(X_{j^*}, \xi_{j^*})}{\partial x_{j,k}} = \frac{\partial \log(\mathbf{p}_t(X_{j^*}, \xi_{j^*}))}{\partial x_{j,k}} \times p_{j^*} = \hat{\alpha}_{j^*,k} \times p_{j^*}, \quad \forall k = 1, \ldots, m.$$

To recover $\beta_{i,0}$, the coefficient on ξ_j in (3), we use a similar equation as above. Since $\frac{\partial \log(\mathbf{p}_t(X_{j^*}, \xi_{j^*}))}{\partial \xi_j} = 1$ in (12), we can easily recover $\beta_{i,0}$ by $\hat{\beta}_{i,0} = \frac{\partial \mathbf{p}_t(X_{j^*}, \xi_{j^*})}{\partial \xi_j} = \frac{\partial \log(\mathbf{p}_t(X_{j^*}, \xi_{j^*}))}{\partial x_{j,k}} \times p_{j^*} = p_{j^*}$.

Step 2: estimating $V^1(\beta_i)$ and $V^0(\beta_i)$

For in-house transactions, we compute $V^1(\beta_i)$ by plugging the recovered β_i and ξ_j into (6). Similarly, we compute $V^0(\beta_i)$ for cross-house transactions. To estimate $V^0(\hat{\beta}_i)$ for buyer i with $d_{j^*} = 1$ and $V^1(\hat{\beta}_i)$ for buyer i with $d_{j^*} = 0$, we need to compute the weighted mean of $U_s(\beta_i)$ by putting higher weights on houses with similar characteristics as house j^*, while putting lower or no weights on houses with different characteristics. For this reason, we use a local linear matching method[33] to estimate

[33]See, e.g., Heckman, et al. (1997, 1998) and Hong (2013) for more details on a local linear matching method.

$E[U_s(\beta_i)|s \in D_i^0]$ for buyer i with $d_{j^*} = 1$ and $E[U_s(\beta_i)|s \in D_i^1]$ for buyer i with $d_{j^*} = 0$. Specifically, the local linear weighted mean is given by the intercept μ_0 in the minimization problem

$$\min_{\mu_0,\mu_1} \sum_{s \in D_i^{1-d_{j^*}}} \left\{U_s(\beta_i) - \mu_0 - (X_s - X_{j^*})'\mu_1\right\}^2 K_B(X_s - X_{j^*}) \times K_b(\xi_s - \xi_{j^*}),$$

where $K_B(\mathbf{v})$ is defined above, $K_b(v) = \frac{1}{b}N(\frac{1}{b}\frac{v}{\hat{\sigma}_k})$, and D_i^1 (or D_i^0) denotes a set of internal (or external) listings in the same market, so that if $d_{j^*} = 1$, we compute the local linear weighted mean by using houses in $D_i^{1-d_{j^*}} = D_i^0$.

Step 3: estimating the distribution of γ_i and c_i

To obtain more information on the extent of strategic promotion, we need to estimate the distribution of γ_i and c_i. To this end, we use (7) and impose a parametric assumption on the distribution of $\delta_i = \gamma_i + c_i$. Hence, we do not attempt to fully separate c_i from γ_i, but instead focus on the marginal effect of strategic promotion by using exclusion restrictions and a natural experiment from a policy change. Let us begin by considering the following specifications for γ_i and c_i:

$$\begin{aligned} \gamma_i = &\quad \gamma_0 + W_{1,i}\lambda_1 + W_{2,i}\lambda_2 + \epsilon_i, \\ c_i = &\quad c_0 + Z_i\theta_1 + W_{2,i}\theta_2 + \omega_i, \end{aligned} \tag{14}$$

where γ_0 and c_0 are the intercepts, ϵ_i and ω_i are the error terms, and W_i is a vector of variables related to transaction costs, but $W_{1,i}$ is only related to transaction costs, while $W_{2,i}$ is related to both transaction costs and strategic promotion. In (14), Z_i is a vector of variables related to strategic promotion but not related to transaction costs. Though we use excluded variables Z_i that only affect strategic promotion, we cannot separately identify γ_i and c_i, because we cannot distinguish γ_0 from c_0 without further restrictions, and $W_{2,i}$ enters both γ_i and c_i.

Therefore, our main approach for the step 3 considers $\delta_i = \gamma_i + c_i$ as follows:

$$\delta_i = \delta_0 + Z_i\theta_1 + W_{1,i}\delta_1 + W_{2,i}\delta_2 + \eta_i, \tag{15}$$

where $\delta_0 = \gamma_0 + c_0$, $\delta_1 = \lambda_1$, $\delta_2 = \lambda_2 + \theta_2$, and $\eta_i = \epsilon_i + \omega_i$. Hence, as long as we have excluded variables Z_i, we can identify and estimate the marginal effect of strategic promotion due to changes in Z_i. If we do not impose any assumption on η_i, we can obtain bounds on θ_1. To obtain point identification, we follow Bajari and Kahn (2005) and impose a parametric distribution on η_i. However, note that the identification of θ_1 does not rely on a particular parametric assumption for η_i. In our application, we assume a logistic distribution, simply because it is straightforward to estimate the model. We thus estimate the parameters using the following likelihood function based on (7):

$$\begin{aligned} L(\theta_1, \delta) \quad = \quad & \prod_i F\left(V^1(\beta_i) - V^0(\beta_i) + \delta_0 + Z_i\theta_1 + W_{1,i}\delta_1 + W_{2,i}\delta_2\right)^{d_j} \\ & \times \left[1 - F\left(V^1(\beta_i) - V^0(\beta_i) + \delta_0 + Z_i\theta_1 + W_{1,i}\delta_1 + W_{2,i}\delta_2\right)\right]^{1-d_j}, \end{aligned}$$

where $F(\cdot) = \exp(\cdot)/(1 + \exp(\cdot))$.

Appendix C: Estimation of Preferences for Discrete Attributes from First Order Conditions

If a housing attribute, $x_{j,k}$, only takes integer values, the first order conditions will hold with inequality. Any estimated $\beta_{i,k}$ that satisfies the inequality should be consistent with the optimal choice, so that the true $\beta^*_{i,k}$ will be only partially identified. Suppose that $x_{j,1}$ is the number of bedrooms. If the optimal $x_{j,1}$ is, say, 2, then consumer maximization implies

$$U_j(x_{j,1} = 2, p_j(x_{j,1} = 2)|\beta_i) \geq U_j(x_{j,1} = n, p_j(x_{j,1} = n)|\beta_i), \quad \forall n \neq 2,$$

where we fix all $x_{j,k}$'s ($k \neq 1$) and they are not included above to simplify the exposition. Given our linear utility assumption, the inequality above can be written as

$$\beta_{i,1} \times 2 + \sum_{k=2}^{m} \beta_{i,k} x_{j,k} + \beta_{i,0}\xi_j - p_j(x_{j,1} = 2) \geq \beta_{i,1} \times n + \sum_{k=2}^{m} \beta_{i,k} x_{j,k} + \beta_{i,0}\xi_j - p_j(x_{j,1} = n), \quad \forall n \neq 2,$$

which can be rewritten as

$$\beta_{i,1} \times (2 - n) \geq p_j(x_{j,1} = 2) - p_j(x_{j,1} = n), \quad \forall n \neq 2,$$

These inequalities can be reduced to

$$p_j(x_{j,1} = 2) - p_j(x_{j,1} = 1) \leq \beta_{i,1} \leq p_j(x_{j,1} = 3) - p_j(x_{j,1} = 2), \tag{16}$$

because $p_j(x_{j,1} = 3) - p_j(x_{j,1} = 2) \leq \frac{p_j(x_{j,1}=n) - p_j(x_{j,1}=2)}{n-2}$ for $n \geq 4$, as long as $p_j(x_{j,1})$ is convex or linear in $x_{j,1}$. In our application, we consider

$$\log\left(\mathbf{p}_t(X_j, \xi_j)\right) = \alpha_{j,0} + \sum_{k=1}^{m} \alpha_{j,k} x_{j,k} + \eta_t + \xi_j,$$

where $\alpha_j = (\alpha_{j,0}, \ldots, \alpha_{j,m})$ is a vector of random coefficients. Therefore, $p_j(x_{j,1}) = \exp(\alpha_{j,0} + \sum_{k=1}^{m} \alpha_{j,k} x_{j,k} + \eta_t + \xi_j)$, which is convex in $x_{j,1}$.

The discussion above suggests two results on identification of $\beta^*_{i,1}$. First, if the price function is linear in $x_{j,1}$ (e.g. $\mathbf{p}(X_j) = \alpha_{j,0} + \sum_{k=1}^{m} \alpha_{j,k} x_{j,k} + \xi_j,$), the inequality (16) becomes equality, since $p_j(x_{j,1} = 2) - p_j(x_{j,1} = 1) = p_j(x_{j,1} = 3) - p_j(x_{j,1} = 2)$. Thus, even though $x_{j,1}$ is not continuous, we can recover $\beta^*_{i,1}$ from the equality FOC. Second, if the price function is convex in $x_{j,1}$ as in our application, $\beta^*_{i,1}$ is partially identified by the inequality (16). Because $\frac{\partial p_j(x_{j,1}=2)}{\partial x_{j,1}}$ lies between $p_j(x_{j,1} = 2) - p_j(x_{j,1} = 1)$ and $p_j(x_{j,1} = 3) - p_j(x_{j,1} = 2)$, if we set $\beta_{i,1} = \frac{\partial p_j(x_{j,1}=2)}{\partial x_{j,1}}$, this is similar to using the mid point value for an interval variable. In this case, the issue is not that the buyer has chosen a particular attribute violating the equality FOC (this is the case of the suboptimal choice, which can be captured by the optimization error), but rather that the buyer has chosen an optimal amount of a given attribute, but this optimal choice can be rationalized by any values within the interval given by (16).

We further examine conditions under which our approach of setting $\hat{\beta}_{i,1} = \frac{\partial p_j}{\partial x_{j,1}}$ does not result in any bias. To this end, suppose that $\beta^*_{i,1}$ denotes the true value of $\beta_{i,1}$, and so the inequality (16) is satisfied for $\beta^*_{i,1}$. We can then consider a prediction error ϵ_i, so that $\hat{\beta}_{i,1} = \beta^*_{i,1} + \epsilon_i$. Note that our matching estimator ($V^1(\hat{\beta}_i)$ or $V^0(\hat{\beta}_i)$) is constructed from $U_j(\hat{\beta}_i) = \sum_{k=1}^{m} \hat{\beta}_{i,k} x_{j,k} + \hat{\beta}_{i,0}\xi_j - p_j$. Hence, for our matching estimator to be consistent, we need that $E(\beta_{i,1} x_{j,1})$ be equal to $\beta^*_{i,1} x_{j,1}$. This is satisfied when $E(\epsilon_i) = 0$ and $E(\epsilon_i | x_{j,1}) = 0$, since $E(\beta_{i,1} x_{j,1}) = E(\beta^*_{i,1} x_{j,1} + \epsilon_i x_{j,1}) = \beta^*_{i,1} x_{j,1} + E(\epsilon_i x_{j,1})$. We believe that these conditions are not too strong in our case.