THREE ESSAYS ON THE EFFICIENCY OF REAL ESTATE MARKETS

A Dissertation

by

XI ZHAO

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Li Gan
Co-Chair of Committee,	Qi Li
Committee Members,	Steven Puller
	Haipeng Chen
Head of Department,	Timothy Gronberg

August 2015

Major Subject: Economics

ABSTRACT

The U.S. real estate markets have undergone substantial fluctuations in recent years. This dissertation attempts to understand the effects of some market fundamentals on residential real estate market outcomes and efficiency from both theoretical perspective and empirical evidence. This research contains three research projects.

First, a number of papers have identified the positive return of market size on matching outcomes when the market exhibits frictions. In the second chapter, I develop a novel directed search model that connects home list price, reservation price with the sale outcomes and empirically test the thick market effects on trading efficiency in housing market using home transaction data in Dallas metropolitan area during 2006 to 2008. The results present strong and robust market size effects that houses on thicker market are listed and sold at higher prices and significantly faster speed.

The third chapter studies principal-agent problems in real estate markets, where the brokerage service is often used to facilitate home sales. The seller agent gets percentage commissions from the home owner and splits with the buyer agent in reward for producing the buyer. A seller agent sometimes serves as dual agent that represents both the seller and buyer sides and gets all commissions. I introduce a theoretical model and present evidence from Dallas metropolitan housing market that the agency structure may create principal-agent problems. I find that the dual-agent-assisted home sales on average give 2.6% more discount on final price than home sales that are assisted by two agents. Competition among home buyers may reduce the severity of principal-agent problems.

The fourth chapter deviates from the rational agent assumption and investigates the behavioral impacts of price endings on home sales. Recent literature in behavioral economics suggests that price endings have some psychological impacts on buyer's purchasing decision. In real estate markets, both round price and precise price (or nine ending price) strategies are used in home sales. From the panel data and regression discontinuity analysis of Dallas housing market transactions, I find homes listed with precise price are on average sold at 4.6% higher price than homes listed at round price only when prices are less than their nearby round prices, favoring the nine ending price literature.

DEDICATION

To my grandma, R.I.P.

ACKNOWLEDGEMENTS

First of all, I would like to express my gratitude to my advisor Professors Li Gan for guiding me during the entire PhD research. I was fortunate to meet him in the first semester of my PhD study at Texas A&M University and later became his student. He encouraged me to the empirical study of real estate markets and inspired me with his wisdom, knowledge and enthusiasm in our discussions. I appreciate all his encouragement and protection through these valuable years. I would also like to thank my co-chair Professor Qi Li and members of my doctoral committee, Professors Steven Puller and Haipeng Chen for their services and their helpful career advice and suggestions in general.

I am also grateful to my former research advisor in China, Professor Ran Tao. Ran is my best role model for an economist and a truly lifetime friend. His strong sense of responsibility and insights on the Chinese economic development and transition stimulated my enthusiasm in investigating economic and social issues and conducting scientific research and inspired me to the pursuant of a PhD in economics.

Finally, this dissertation is impossible without the love and endless support from my parents Xianfeng Zhao and Xiaoqing Ming in China, my grandma Suqing Yan (1932-2013) and my wife Yue Cui. To them, I am most indebted.

NOMENCLATURE

MLS	Multiple Listing Service
FTC	Federal Trade Commission
NAR	National Association of Realtors
FSBO	For-Sale-By-Owner
TOM	Time on the Market

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

1. INTRODUCTION: THE EFFICIENCY OF REAL ESTATE MARKETS

1.1 Introduction

The U.S. real estate markets have undergone substantial fluctuations over the past decades and especially in the recent years. Scholars have long attempted to understand and explain these dynamics and the difference across markets. This research work aims to investigate the micro level economic model and studies how market fundamentals affect residential real estate market outcomes and efficiency. I provide some theoretical and empirical evidences from analyzing home transactions in three aspects that may partially explain the variations of the outcomes in real estate markets. This dissertation can be separated into three chapters and each chapter is self-contained.

In the second chapter, I develop a directed search model and empirically tests the thick market effects on trading efficiency in housing market using home transaction data. Theoretical papers in search and matching literature suggest a positive return of market size on matching quality when the market exhibits frictions. The real estate market is a typical frictional market in which buyers and sellers actively search and match for trading partners. A number of previous papers have identified the positive effects of market size on matching quality from macro level analysis. But there is still lack of micro level evidence. In this chapter, I first build a novel directed search model that enables me to connect house list price, sale price and sale speed and control for unobserved seller's reservation price. This model is also consistent with the fact that house buyers often bid for houses and the resulting sale prices are higher than list prices. I then estimate the effects of market size on house list price, sale price and sale speed in Dallas housing market between 2006 and 2008 using the structural model. The size of the market is measured by defining a metric space of house characteristics and aggregating number of houses weighted by the distance measure. The estimation results show significant market size effects. I find that houses on thicker market are listed and sold at higher price. On average, one standard deviation increase in market size leads to about 7400 dollars increase in list price and 8000 dollars increase in sale price, or 4.3% and 5% respectively. The sale price to list price ratio also increases by 2%. In the meantime, when market size grows by one standard deviation, the sale time reduces by 1.6 weeks or 11 days, reflecting a faster matching between buyers and sellers.

The third chapter of this research adds real estate agents into the analysis and studies principal-agent problems in real estate markets. In fact, real estate agents are frequently involved in home transactions in residential real estate markets. According to National Association of Realtors (NAR), the real estate agent-assisted home sales account for 91% of total home sales in 2013. In a multi-year study of the residential real estate brokerage industry, the Federal Trade Commission (FTC) suggests that "*the market for real estate brokerage service does not accord with the customary model of competitively functioning markets*", indicating that the brokerage service may suffer from inefficiency. In the second chapter, I investigate the principal-agent problems arisen from dual agency. For a typical home sale, a seller real estate agent gets percentage commissions from the home owner and splits with the buyer agent in reward for producing the home buyer. Dual agency happens when the seller agent is directly contacted by a home buyer and in this situation, the agent gets all commissions. This dual agency provides incentive for the seller's agent to push the seller to accept offers from direct buyers rather than from buyers who have buyer agent, even if the offers are lower. In this chapter, I will introduce a theoretic model and present empirical evidence from Dallas metropolitan housing market on this principal-agent problem.

The fourth chapter of the dissertation deviates from the rational agent assumption and investigates the behavioral impacts of price endings on home sales. Recent literature in behavioral economics and marketing science suggests that price endings may have some psychological impacts on buyer's purchasing decision. In real estate markets, some home sellers list the home at some round prices (i.e, $200,000) while some sellers tend to use precise price strategy or general nine ending price strategy (i.e. $199,910). From the panel data model and regression discontinuity analysis of Dallas housing market transactions, I find homes listed with nine ending strategy are on average sold at 4.6% higher price than homes listed at round price, consistent with literature findings in retail markets. But precise prices do not always help the sale. When homes are listed at some precise prices exceeding nearby round prices, I do not find any behavioral influence buyer's negotiation and purchasing and final sale prices.

2. A DIRECTED SEARCH MODEL AND MARKET SIZE EFFECTS IN REAL ESTATE MARKETS

2.1 Introduction

How does market size affect house seller's strategic behaviors and how does it affect real estate market outcomes? The thick market effects have been investigated in numbers of theoretical and empirical work on some frictional markets. Early papers focus on theoretical modeling the thick market effects in labor market by introducing increasing return to scale matching technology in a search and matching framework, i.e. see Diamond [4], Pissarides [23], Mortensen [22] and their early research work. The market thickness is showed to encourage workers' and firms' search intensity and therefore reduce unemployment and vacancy spells and improve matching efficiency. On the empirical side, a few papers also attempt to discover some evidences of market size effects. For example, Gavazza [8] finds the existence of thick market effects in real assets market by developing a search and matching model for airplane traders. He finds that market size positively affects trading speed and airplane utilization rate. Gan and Zhang [6] investigate the thick market effects in labor market and find it reduces unemployment fluctuations over time. Gan and Li [5] developed and applied matching model to the U.S. academic market for new PhD economists and find that a field of specialization with more job openings and more candidates has a higher probability of matching.

In this chapter, I study the thick market effects on house seller's strategic pricing and search behavior and on trading efficiency. I first introduce the house sale game that that captures the institutional arrangements and conventions in U.S. housing market and build an equilibrium directed search model. In the model, house sellers post asking prices of houses and attract house visitors. In each selling period, randomly matched house buyers respond to the list price and compete for the house by making strategic counteroffers (bids) according to their valuations of the house. To be more specific, house visitors first decide whether to accept the list price. If multiple visiting buyers accept the list price, they bid for the property. If no buyers accept the offer, they may also bid for the house. The bidding rule is assumed to follow first-price sealed bid auction. The property therefore is entitled to the winning buyer with final transaction price either less, equal or more than the list price. The equilibrium condition shows that under the existence of competition and bidding among buyers, the list price serves to segment the high valuation buyers from low valuation buyers. The market size effect enters in my model through affecting the matching rate or offer arrival rate to the seller. Numerical simulations suggest that house sellers benefit from thick market effects by setting a higher list price to segment buyers and expected higher sale price and shorter time to sale. In addition, houses in thicker market are more likely to sell at or above list price even if they are listed at higher prices. And the ratio of expected sale price to list price also increase as market size increases. I then test the theoretic implications using a unique house transaction data set from Dallas Multiple Listing Service. I find empirical evidence of the prevalence of thick market effects. Specifically speaking, one standard deviation increase in market size or nine more listing houses in the sub-market leads to about 0.2% increase in sale price to list price ratio on average, equivalent to more than 4000 dollars increase in sale price. The probability to sell above list price also increases by 0.6%. In addition, one standard deviation increase of market size also shortens the marketing time by 1.6 weeks or 11 days.

This research makes several contributions to the literature. First, it adds to the search literature by studying strategic bidding among house buyers and the role of list price in segmenting buyers with heterogeneous valuations. Some early papers that attempt to study the effects of list price on house sale outcomes either fail to explain its role or lead to implications inconsistent with real house transaction results. For example, Horowitz [12] assumes a reduced form distribution of buyer's offer conditional on list price but he does not explain why and how offers are affected by list price. Knight [16] and Haurin et al. [10] posit that the list price truncates the buyer's counteroffer. But this assumption violates the fact that a substantial amount of houses are sold above list price. In this research instead, I explicitly introduce interactions between sellers and buyers in the house sale game and allow for potential competition among realized buyers. Therefore buyer's bidding behaviors are naturally introduced in the directed search model and the role of list price becomes clear under equilibrium. It also explains why sometimes the sale price is higher than house list price. Similar to our game specification, Albrecht et al. [1] also develop a directed search model in housing market that allows for bidding, but they introduce bidding rule to be ascending auction instead of first-price sealed bid auction so that pricing does not affect seller's revenue. They instead regard the role of list price as signal of seller's heterogeneous type. Second, to the best of our knowledge, this is the first project that studies how market size affect seller's pricing strategy in directed search model framework. Prior to my study, Gan and Zhang [7] study the impact of the unemployment rate on the housing market in the presence of the thick market effect. They calibrate a matching model using Texas city-level data and find that an increase in the unemployment generates poorer matching quality and as a consequence, prices and the transaction volume both decline more than in the absence of the thick market effect. this research connects market size with house list price, sale price, price reduction and selling speed and empirically investigate the thick market effects using micro level transaction data.

The remainder of this chapter is organized as follows. In the next section, I introduce the house sale game and seller's directed search model that captures the institutional arrangements in housing market and derive equilibrium conditions as well as some market outcomes. Since the system does not admit analytical expression, in section 3 I numerically solve the model and study the effect of market size on market outcomes. In section 4, I introduce the Dallas Multiple Listing Service (MLS) data set that are used for empirical analysis and discuss the constructions of some key variables. In section 5, I estimate the reduced form model and structural model. The last section concludes the paper.

2.2 The Directed Search Model

2.2.1 The House Sale Game and Equilibrium

The house sale game plays in the following steps.

Stage 1: A seller sets a list price a of the house.

Stage 2: In each selling period, a buyer can visit exact one house in each submarket and generate valuation x of the house independently. The matching process is random. Valuations are private information to the buyers. After the visit, each buyer decides whether to accept the list price a or not and the seller gets to know whether buyer's valuation is no less than a or below a.

Stage 3: There are three situations. If only one buyer accepts the list price a, the house is transferred to that buyer. If two or more buyers accept the list price in a selling period, they compete for the house by making counteroffers. The environment resembles first-price sealed bid auction. If none of visiting buyers accept the list price, buyers may also make counteroffers.

Stage 4: The seller can either accept the best counteroffer or deny all counteroffers. In the latter case, the sale goes on to the next period. In addition, if the house *attracts no visitors in a period, the seller retains the house for the next period.*

In this game, it is important to clarify the information availability to sellers and buyers. First, I assume that house characteristics are not fully observed by buyers before visit. Without visiting and valuating the house, buyers can not compare two houses with different list price. Instead, they are often focusing on some sub-markets for houses in their preferred location range and price range (say houses with value between 150,000 dollars to 200,000 dollars in some subdivision). The before-visiting incomplete information assumption suggests that buyer's visit to houses in one sub-market is random and the offer arrival rate to a single house would not be affected by list price. Second, I do not consider brokerage service in the housing market. Some papers find the existence of principal-agency problem and attempt to model it. For example, Hsieh and Moretti [13] find that real estate agents sell their own houses faster than their clints' houses. Yavas and Yang [32] instead modeled the agent's work effort into seller's pricing model. In our paper, since I focus on the directed search and market size, we do not include real estate agents in the model.

There are a few advantages of specifying the game rule explicitly in our analysis. First, the procedures of house sale game generally reflect the institutional arrangements and conventions in U.S. housing market. Second, it provides a structural approach to study the role of list price. The buyer's offer distribution conditional on list price can be derived from equilibrium conditions. In comparison to Horowitz [12] and Haurin et al. [10], the role of list price on buyer's counteroffer strategy tractable in our model. Second, since the game plays in discrete time version, it naturally allows for multiple visiting buyers and therefore introduced potential competitions among those buyers. Bidding among buyers automatically solves the Rothchild's paradox that the equilibrium offer distribution is non-degenerated. In addition, the determination of final price is consistent with actual house transaction data that houses may be sold less, equal or more than list price.

To derive the equilibrium conditions, I first study the buyer's optimal counteroffer strategy (or response function) to seller's list price a. We assume that buyer's valuation of the house after visit is generated independently from distribution with cdf $F(x)$ and pdf $f(x)$. If there are multiple visiting buyers, we denote $G^{(k-1)}(x)$ and $g^{(k-1)}(x)$ the cdf and pdf of the highest value among $k-1$ buyers. For buyer with valuation x less than list price, the symmetric equilibrium counteroffer (bidding) strategy $\beta(x)$ and expected payment $m(x)_{x \geq a}$ given k competing buyers are[1]

$$\beta(x) = a\frac{G^{(k-1)}(a)}{G^{(k-1)}(x)} + \frac{1}{G^{(k-1)}(x)}\int_a^x yg^{(k-1)}(y)dy; \tag{2.1}$$

and

$$m(x)_{x \geq a} = aG^{(k-1)}(a) + \int_a^x yg^{(k-1)}(y)dy. \tag{2.2}$$

For buyers with valuation x equal to or greater than list price, the counteroffer strategy and expected payment $m(x)_{x<a}$ are

$$\beta(x) = b\frac{G^{(k-1)}(b)}{G^{(k-1)}(b)} + \frac{1}{G^{(k-1)}(x)}\int_b^x yg^{(k-1)}(y)dy; \tag{2.3}$$

and

$$m(x)_{x<a} = bG^{(k-1)}(b) + \int_b^x yg^{(k-1)}(y)dy. \tag{2.4}$$

Summing up these two parts, the ex ante expected revenue to the seller given number

[1]See Krishna [17] for derivation details.

of bidders k is

$$
\begin{aligned}
E(\pi_k) =& k\{\int_a^\omega [aG^{(k-1)}(a) + \int_a^x yg^{(k-1)}(y)dy]f(x)dx \\
& + \int_b^a [bG^{(k-1)}(b) + \int_b^x yg^{(k-1)}(y)dy]f(x)dx\} \\
=& kaG^{(k-1)}(a)[1 - F(a)] + k\int_a^\omega y(1 - F(y))g^{(k-1)}(y)dy \\
& + kbG^{(k-1)}(b)[F(a) - F(b)] + k\int_b^a y(1 - F(y))g^{(k-1)}(y)dy \\
=& kaG^{(k-1)}(a)[1 - F(a)] + kbG^{(k-1)}(b)[F(a) - F(b)] \\
& + k\int_b^\omega y(1 - F(y))g^{(k-1)}(y)dy.
\end{aligned}
\tag{2.5}
$$

After deriving the buyer's counteroffer strategy given house list price a and fixed number of buyers, I now study the seller's search strategy. Since the number of visiting buyers k that a house seller encounters in a selling period is random at a given sub-market, we assume it follows Poisson distribution with arrival rate λ, i.e, the probability that k buyers visit the house in a single period is

$$p_k = Pr(n = k, \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}. \tag{2.6}$$

The arrival rate λ as we discussed, does not depend on variables other than the sub-market conditions. The seller's dynamic problem is therefore to set optimal list price and reservation price $\{a, b\}$ that maximize the following value function

$$
\begin{aligned}
V(a, b) = \max_{(a,b)}\{-c+\beta\{&\sum_{k=0}^{\infty} p_k F(b)^k V + \sum_{k=1}^{\infty} p_k kaG^{(k-1)}(a)[1 - F(a)] \\
&+ \sum_{k=1}^{\infty} p_k kbG^{(k-1)}(b)[F(a) - F(b)] \\
&+ \sum_{k=1}^{\infty} p_k k \int_b^\omega y(1 - F(y))g^{(k-1)}(y)dy\}\}.
\end{aligned}
\tag{2.7}
$$

After some algebra, the Bellman equation (2.7) can be expressed as

$$(b+c)\beta^{-1} = be^{-\lambda}e^{\lambda F(b)} + \lambda a e^{-\lambda}e^{\lambda F(a)}[1-F(a)] + \lambda b e^{-\lambda}e^{\lambda F(b)}[F(a)-F(b)] + \lambda^2 e^{-\lambda}\int_b^{\omega} y[(1-F(y)]e^{\lambda F(y)}f(y)dy. \tag{2.8}$$

The first order condition of the Bellman equation with respect to a gives us the optimal list price

$$bf(a) + \{\lambda a f(a)[1-F(a)] + 1 - F(a) - af(a)\}\, e^{\lambda[F(a)-F(b)]} = 0. \tag{2.9}$$

The optimal list price and reservation price $\{a, b\}$ are jointly determined by optimality conditions (2.8) and (2.9). It is readily to see that seller's optimal list price and reservation price depend on parameters including seller's waiting or search cost, offer arrival rate λ and the distribution of buyer's valuation F.

Let $q = \sum_{k=0}^{\infty} Pr(n=k)F(b)^k$ denote the probability that the house is not sold in one period, we can also investigate how expected time to sale and the expected sale price are affected by these factors. For example, the expected sale price conditional on a sale can be expressed as

$$\begin{aligned} E(p|p>b) &= \frac{1}{1-q}\{\sum_{k=1}^{\infty} p_k k a G^{(k-1)}(a)[1-F(a)] + \sum_{k=1}^{\infty} p_k k b G^{(k-1)}(b)[F(a)-F(b)] \\ &\quad + \sum_{k=1}^{\infty} p_k k \int_b^{\omega} y(1-F(y))g^{(k-1)}(y)dy\} \\ &= \frac{(b+c)\beta^{-1} - be^{-\lambda(1-F(b))}}{1-e^{-\lambda(1-F(b))}}. \end{aligned} \tag{2.10}$$

And the expected time to sale is

$$E(TOM) = \sum_{t=1}^{\infty} tq^{t-1}(1-q) = \frac{1}{1-q} = \frac{1}{1-e^{-\lambda(1-F(b))}}. \tag{2.11}$$

2.2.2 Simulated Solution and Model Implications

From the optimality conditions we know that house list price and reservation price are jointly determined and market outcomes are affected by market size that affects offer arrival rate, seller's waiting or search cost and distribution of buyer's valuation. To better illustrate how these variables affects seller's list price, reservation price and market outcomes including final sale price and selling speed, we numerically solve the model. We first assume that the buyer's valuation x is uniformly distributed on $[\mu - \sqrt{3}\sigma, \mu + \sqrt{3}\sigma]$ with mean and variance are μ and σ^2. The optimality conditions can be simplified to

$$\begin{aligned}&[\mu + \sqrt{3}\sigma - b + \frac{4\sqrt{3}\sigma}{\lambda} - \frac{\lambda b(\mu + \sqrt{3}\sigma - a)}{2\sqrt{3}\sigma}]\exp(-\frac{\lambda(\mu + \sqrt{3}\sigma - b)}{2\sqrt{3}\sigma})\\ &\quad + \frac{\lambda a(\mu + \sqrt{3}\sigma - a)}{2\sqrt{3}\sigma}\exp(-\frac{\lambda(\mu + \sqrt{3}\sigma - a)}{2\sqrt{3}\sigma}) + \mu + \sqrt{3}\sigma - \frac{4\sqrt{3}\sigma}{\lambda} = \beta^{-1}(b+s)\end{aligned} \tag{2.12}$$

and

$$b + \left(\lambda a\frac{\mu + \sqrt{3}\sigma - a}{2\sqrt{3}\sigma} + \mu + \sqrt{3}\sigma - 2a\right)\exp\left(\frac{\lambda(a-b)}{2\sqrt{3}\sigma}\right) = 0. \tag{2.13}$$

The solutions to above system of equations do not admit any closed form express and therefore we find numerical solution using Matlab. Following convention, the discount rate β is set at 0.997. The initial offer arrival rate λ is set at 0.5 and initial search cost or waiting cost is assumed to be $c = 800$. The mean and standard deviation of buyer's valuation distribution are \$150,000 and \$50,000. We report the simulated solution in the first line of table 2.1. The house list price, buyer's reservation price and final sale price are \$219,300 ,\$201,200 and \$215,600 respectively. Roughly 34% of the houses are sold at or above the list price and the expected marketing time is 10.3 weeks or 72 days.

To understand how market size affects seller's pricing behavior and market outcomes, I simulate the effects by changing offer arrival rate parameter. The results are reported in the second panel of table 2.1. Holding the distribution of buyer's valuation unchanged, we increase the arrival rate from 0.1 to 0.7, the house list price increases from \$192,600 to \$222,900 and the expected sale price increases even more (from \$172,500 to \$221,800). The percentage of houses sell at or above list price sharply increases from 7% to 45% and the ratio of expected sale price to list price increases from 0.895 to 0.995, indicating that houses in thicker market enjoy more benefits from buyer's bidding. In addition, the expected time to sale also decreased sharply when offer arrival rate increases, even when the sellers strategically increase their reservation price of the house.

Some other factors can also affect market outcomes, including seller's search cost or waiting cost and the distribution of buyer's valuation. In the second panel of table 2.1, when seller's search cost increases from 100 dollars per week to 1000 dollars per week, seller tends to lower the list price and reservation price to sell the house faster, and the expected sale price also decreased. The heterogeneity of buyer's valuation also affects house sale. For example, Haurin et al. [10] find empirical evidence from Columbus, Ohio housing data set that houses that attract larger variance of buyer's valuation tends to have higher ratio of the list price to the expected sales and longer marketing time. In the second panel of table 2.1, I also simulate the market outcomes by increasing buyer's standard deviation from 30,000 dollars through 80,000 dollars and the ratio of expected sale price to list price and time to sale increase substantially.

Table 2.1: Simulations of the Effects of Offer Arrival Rate and Buyer Heterogeneity on Market Outcomes

	List price	**Res. price**	**E(Sale price)**	$Perc(s.price \geq l.price)$	$\frac{E(s.price)}{l.price}$	**Time on Mkt**
$\lambda = 0.5$, $\sigma = 50k$						
$c = 800$	221169	205151	215645	0.34	0.984	10.3
$\sigma = 50k$, $c = 800$						
$\lambda = 0.1$	192649	147597	172484	0.07	0.895	20.0
$\lambda = 0.2$	206180	174744	194011	0.21	0.940	14.5
$\lambda = 0.5$	221169	205151	215645	0.34	0.984	10.3
$\lambda = 0.7$	222858	208494	221808	0.45	0.995	9.3
$\lambda = 0.5$, $\sigma = 50k$						
$c = 100$	224123	211266	221702	0.36	0.990	14.2
$c = 500$	221169	205151	218042	0.35	0.985	13.2
$c = 800$	221169	205151	215645	0.34	0.984	10.3
$c = 1000$	218083	198711	214163	0.34	0.982	9.7
$\lambda = 0.5$, $c = 800$						
$\sigma = 30k$	189822	177107	188954	0.33	0.995	8.87
$\sigma = 50k$	221169	205151	215645	0.34	0.984	10.3
$\sigma = 80k$	264621	239762	257842	0.35	0.974	11.8

In the simulation, the discount rate β is set to be 0.997.

The buyer's valuation follows uniform distribution on $[\mu + \sqrt{3}\sigma, \mu - \sqrt{3}\sigma]$. μ is set to be 150k dollars.

2.3 The Data

2.3.1 The Dallas Multiple Listing Service

The data set I use comes from Dallas multiple listing service(MLS) database. The multiple listing service is an actively managed system that enables real estate brokers share information on properties they have listed and invite other brokers to cooperate in their sale in exchange for compensation if they produce the buyer. The multiple listing service disseminates listing information and sellers benefit by increased exposure to their property and buyers benefit because they can obtain information about all MLS-listed properties while working with only one broker.[2][31]. Current, about 90 percent of residential properties are listed and sold through over 800 MLSs nationwide.

The Dallas multiple listing service database focuses on the Dallas metropolitan area and its vicinity and records the listing information and transaction details of every residential property that is listed on Dallas MLS. A typical house listing record in Dallas MLS contains information on the type of house, house address, list date, off-market date, listing price, final sale price and a selection of house characteristics including home size, lot size, number of bedrooms, number of bathroom, etc.. It also includes some agents and broker information. We were able to get the MLS data on property transactions between the third quarter of 2006 to the second quarter of 2008.

Several points about the data set are worth mention. First, my data set includes only residential properties. I focus on residential single-family houses only and condos and other types of properties are excluded from our analysis. Second, there is no explicit geographic boarder that defines Dallas MLS market and our data set con-

[2]see http://en.wikipedia.org/wiki/Multiple_listing_service for detail functions of multiple listing service.

tains transactions in Dallas-Fort Worth-Arlington metropolitan area and vicinities. Instead, we have specific physical address for each house. Third, since our data set records all realized transactions, houses that are listed before the sample period and sold within such period are included. On the contrary, houses that are listed within the sample period but not sold are not captured by the data set.

Besides the key variables, I also have information on house characteristics, including number of bedrooms and number of bathrooms. The house square footage is available for about one third of the houses. For the rest of the houses we use regression-based method to impute approximate square footage.[3] The size of housing lot is recorded either in square footage or side lengths for some houses. We carefully identify the lot size and derive approximated lot size using imputation method[4]. Furthermore, we also generate an indicator that captures irregular lot[5]. The summary statistics of list price, sale price, time to sale and house characteristic information are provided in table 2.2.

2.3.2 Measure of Market Size

The key variable in my analysis is the size of the market. The variable of main interest is the the market size. The definition of a sub housing market, however, is not trivial since there is neither a well-defined geographic boundary of sub housing market nor an explicit set of house characteristics that segment sub-markets. In our paper, we define a sub-market as houses in a subdivision within certain price range. The measure of market size is also nontrivial. Even under the ideal measure of market size

[3] We first regress house square footage on number of bedrooms, number of bathrooms, house lot size (if available) and 5-digit zip code dummies. Then we use the estimated coefficients to predict the square footage for missing values.

[4] The size of housing lot is recorded in side lengths for some houses. If the lot is in rectangular shape, the lot size is directed calculated as the product of two adjacent sides. If the lot is non-rectangle, the lot size cannot be exactly computed without at least one known angle. We therefore approximate the lot size by taking average of two opposite sides and multiply the two averages.

[5] The irregular lot is defined as non-rectangular lot.

Table 2.2: Summary Statistics

Variable	Mean	Std. Dev.	N
List rice	214481	349253	191552
Sale price	201042	230265	191552
Price reduction	-13439	249875	191552
Sale price to list price ratio	0.94	0.10	191513
Percentage of houses (=LP)	8.9%	0.285	191552
Percentage of houses (¿LP)	14.6%	0.353	191552
Time on market (Weeks)	12.03	11.87	191552
Market size 1 (subdiv, 30k price range)	8.9	17.9	190243
Market size 2 (subdiv, 50k price range)	12.6	26.1	190243
Market size 3 (subdiv, 100k price range)	18.1	37.9	190243
Market size 4 (subdiv only)	28.9	55.3	190243
Number of bedrooms	3.4	0.73	191442
Number of bathrooms	2.5	0.93	191331
House square footage	2217	1239	51409
Imputed House square footage	2228	967	189433
House lot size (sqft)	14081	112989	75727
Percentage of irregular lots	6.5%	0.25	75727

as total number of market participants, to identify who are the market participants is not an easy question. In practice, economists adopt different measures of market size. In **?**]'s paper, the market size is measured by city population. Gavazza [8] instead measure the market size by airplane inventory. The former measure is suitable for cross-city level data but may not be applicable to the Dallas area data. Therefore in our paper, we measure the market size as number of house listing in the sub-market during the whole sample period. For example, one sub-market may be defined as the market of houses in subdivision "Village of Woodland Spring" with price range from $200,000$ to $300,000$ dollars. To test the robustness of the market size effects. we also vary the price bandwidths from 30,000 dollars, 50,000 dollars, 100,000 dollars to the whole price range and generate a set of market size measures. The average market sizes are 9, 13, 18 and 29 lists correspondingly. The standard deviations are 18,26,38 and 55, reflection a substantial variation in market size.

2.4 Estimation and Empirical Results

The theoretic model suggests that house list price, final sale price and selling speed are all jointly affected by offer arrival rate and distribution of buyer's valuation. In the mean time, it provides us three empirical implications on the thick market effects in housing market. First, the percentage of houses sold at or above list price increase as market size increases. Second, the ratio of sale price to list price also increases as the market size increases. Third, houses in thicker market are sold faster. Next, we first attempt to investigate these thick market effects using Dallas MLS data.

2.4.1 Market Size, List Price and Sale Price

As I discussed earlier, the theoretical model suggests that house sellers enjoy benefits from thick market effects by setting a higher list price and expected higher sale price. More interestingly, houses in thicker market are more likely to attract bids and sold above list price even their list price is already higher. In our data set, about 9% of the houses are sold at the list price and 15% are sold higher than list price, reflecting substantial amount of bidding behaviors. The average sale price to list price ratio is 0.94 and the sale price is about 6% less than the list price. The standard deviation is 10%, equivalent to 21,000 dollars.

I first regress the sale price to list price ratio on market size and house characteristics. Table 2.3 reports the estimated thick market effect on sale price to list price ratio after controlling time fixed effect and location dummies.. As discussed in the data section, we defines a sub housing market as houses in certain subdivision with certain price range. In the first column, the price bandwidth that defines a market is 30,000 dollars and market size is computed as number of listed houses in subdivision within such price bandwidths. The coefficient on market size is significantly positive. The magnitude shows that one standard deviation increase in market size (or equivalently nine more listing houses) of sub-market leads to about 0.2% increase in sale price to list price ratio on average, equivalent to more than 4000 dollars increase in final sale price. In column two through four of table 2.3, we gradually increase the price bandwidths that used to define the sub-market. The result is robust to the measures of market size.

Second, I estimate a probit model to test how market size affect the likelihood that houses are sold above list price. The estimation results are presented in table 2.4. The market size are defined in the same way as that in table 3. From the table we find that the coefficient on the market size in the probit model is 0.00147, corresponds to 0.6% increase in the likelihood of selling high when increasing the market size by one standard deviation.

2.4.2 Market Size and Selling Speed

Besides the effects on house list price and sale price, the market size also affect the selling speed. The increase of market size has two impacts that may affect house expected marketing time or selling speed. First, it increases the offer arrival rate and reduces the expected time to sale. Second, it also reservation price actually prolongs the house sale. The theoretic prediction that the former effect dominate the latter so that in overall, houses are sold faster in thick market than in thin market even when sellers strategically raise their reservation price.

To test this prediction, I again regress the time that houses stay on market (TOM) on market size and house characteristics. Table 2.5 reports the estimation results after controlling time fixed effect and location dummies. The effects are all negatively under all four measures of market size and the and significant in the first three columns. The result shows that on average, one standard deviation increase

Table 2.3: Estimation of Thick Market Effect on Sale Price to List Price Ratio

	(1)	(2)	(3)	(4)
VARIABLES	ratio	ratio	ratio	ratio
Market size 1	0.013%***			
	(2.71E-05)			
Market size 2		0.010%***		
		(1.79E-05)		
Market size 3			0.002%*	
			(1.21E-05)	
Market size 4				0.001%*
				(6.29E-06)
Number of bedrm	-0.00190***	-0.00192***	-0.00184**	-0.00178**
	(0.000736)	(0.000736)	(0.000736)	(0.000732)
Number of bathrm	-0.000914	-0.000892	-0.000921	-0.00091
	(0.00111)	(0.00111)	(0.00111)	(0.00111)
House lot size (sqft)	-3.58E-09	-3.57E-09	-3.61E-09	-3.55E-09
	(2.49E-09)	(2.49E-09)	(2.50E-09)	(2.50E-09)
Irregular lot dummy	0.00323***	0.00321***	0.00320***	0.00319***
	(0.00109)	(0.00109)	(0.00109)	(0.00109)
House size (sqft)	4.43e-06***	4.42e-06***	4.27e-06***	4.12e-06***
	(1.20E-06)	(1.20E-06)	(1.19E-06)	(1.19E-06)
Zipcode Year*Qtr	Yes	Yes	Yes	Yes
Constant	0.945***	0.945***	0.946***	0.946***
	(0.00212)	(0.00212)	(0.00211)	(0.0021)
Observations	74,459	74,459	74,459	74,459
R-squared	0.011	0.011	0.011	0.011

Robust standard errors in parentheses
*** p¡0.01, ** p¡0.05, * p¡0.1

Table 2.4: Probit Estimation of Thick Market Effect on Likelihood to Sell Above List Price

	(1)	(2)	(3)	(4)
VARIABLES	Probit	Probit	Probit	Probit
Market size 1	0.00147***			
	(0.000468)			
Market size 2		0.00114***		
		(0.000315)		
Market size 3			0.000399*	
			(0.00021)	
Market size 4				-1.80E-05
				-0.000125)
Number of bedrm	0.105***	0.105***	0.106***	0.106***
	(0.0126)	(0.0126)	-0.0126)	(0.0125)
Number of bathrm	0.0552***	0.0556***	0.0555***	0.0546***
	(0.0191)	(0.0191)	(0.0191)	(0.0191)
House lot size (sqft)	2.30e-07***	2.30e-07***	2.30e-07***	2.29e-07***
	(4.95E-08)	-4.95E-08)	-4.96E-08)	-4.98E-08)
Irregular lot dummy	-0.164***	-0.165***	-0.164***	-0.164***
	(0.0268)	(0.0268)	(0.0268)	-0.0268)
House size (sqft)	-0.000306***	-0.000306***	-0.000308***	-0.000308***
	(2.40E-05)	(2.40E-05)	(2.40E-05)	-2.40E-05)
Zipcode Year*Qtr	Yes	Yes	Yes	Yes
Constant	-0.900***	-0.900***	-0.894***	-0.890***
	(0.0335)	(0.0335)	(0.0334)	(0.0334)
Observations	74,448	74,448	74,448	74,448

Robust standard errors in parentheses
*** p¡0.01, ** p¡0.05, * p¡0.1

Table 2.5: Estimation of Thick Market Effect on House Selling Speed

	(1)	(2)	(3)	(4)
VARIABLES	TOM	TOM	TOM	TOM
Market size 1	-0.00815***			
	(0.00312)			
Market size 2		-0.00477**		
		(0.00207)		
Market size 3			-0.00300**	
			(0.00137)	
Market size 4				-0.000476
				(0.000763)
Number of bedrm	0.151*	0.152*	0.151*	0.159**
	(0.079)	(0.079)	(0.079)	(0.079)
Number of bathrm	0.399***	0.400***	0.402***	0.399***
	(0.121)	(0.121)	(0.121)	(0.121)
House lot size (sqft)	-2.06E-07	-2.07E-07	-2.05E-07	-2.06E-07
	(3.04E-07)	(3.04E-07)	(3.04E-07)	(3.04E-07)
Irregular lot dummy	-0.549***	-0.550***	-0.550***	-0.551***
	(0.155)	(0.155)	(0.155)	(0.155)
House size (sqft)	0.000523***	0.000519***	0.000514***	0.000505***
	(0.000146)	(0.000146)	(0.000146)	(0.000146)
Zipcode Year*Qtr	Yes	Yes	Yes	Yes
Constant	8.167***	8.180***	8.190***	8.215***
	(0.213)	(0.213)	(0.213)	(0.212)
Observations	73,730	73,730	73,730	73,730
R-squared	0.017	0.017	0.017	0.017

Robust standard errors in parentheses

*** p¡0.01, ** p¡0.05, * p¡0.1

in market size may shorten the selling time by 1.6 weeks or 11 days. Comparing the average marketing time of 80 days, market size improves selling speed and trading efficiency substantially.

2.4.3 Structural Estimation

Since the house list price and the seller's reservation price are jointly affected by market size and offer arrival rate and all market outcomes are influenced simultaneously, estimating the thick market effects on the sale price and sale speed separately may fail to control for unobservables and suffer from endogeneity issues. In this section, I specify a structural model that naturally derived from the theoretical directed search model to estimate the effects of market size.

There are two main issues needed to be address in the structural estimation. First, similar to Horowitz [12] and labor market data, I only observe the buyer's counteroffers (final transaction prices) that are accepted by sellers and do not observe rejected prices. In another words, I do not observe buyers' counteroffers that are less than seller's reservation price. And the seller's reservation price is the not directly observed either. Therefore the reservation price as truncation point of the buyer's counteroffers needs to be estimated in the structural model. Second, the theoretical model assumes homes are identical and ignores the heterogeneity in home characteristics. In empirical structural analysis, I need to introduce house characteristics into the econometric model. Specifically, I assume that

$$y_i = X_i\beta + \varepsilon_i; \tag{2.14}$$

where y_i is the house final sale price and X_i is a vector of house characteristics. ε_i reflects the heterogeneity in buyer's valuation of the house. Noticing that since the observed sale price are truncated at seller's reservation price, I can specify the log-likelihood function as follow

$$
\begin{aligned}
LL = \sum_{i=1}^{N} & - \lambda(TOM_i - 1)[1 - F(b_i|a_i;\theta)] \\
& - 1[y_i = a_i] \times \lambda[1 - F(y_i|X;\theta)]log[\lambda[1 - F(y_i|X;\theta)]] \\
& - 1[y_i \neq a_i] \times [\lambda[1 - F(y_i|X;\theta)]log[\lambda^2(1 - F(y_i|X;\theta))f(y_i|X;\theta)]];
\end{aligned}
\tag{2.15}
$$

where a_i, b_i and TOM_i refer to observed house list price, reservation price and weeks on the market respectively. θ is a set of parameters. I further assume that the market size M_i enters into the model by affecting the market arrival rate, i.e., $\lambda = \gamma M_i$. The heterogeneity of buyer's valuation on each house ε_i follows normal distribution $N(0, \sigma^2)$, so that $F(y_i|X;\theta) = \Phi((y_i - X_i\beta)/\sigma)$ and $f(y_i|X) = (1/\sigma)\phi((y_i - X_i\beta)/\sigma)$.

As I discussed earlier, the seller's reservation price is not observed and therefore I consider it as a parameter to be estimated. From the first order conditions in previous section, I find that the reservation price b_i and list price a_i are jointly determined. Therefore the availability of list price information could also help the identification of reservation price. Specifically, I follow the literature to assume that $b_i = \gamma a_i$ and substitute this relation into the likelihood function and the parameters γ in the function (and so forth the reservation price) can be estimated together with the rest parameters in the model. The structural estimation results are reported in table 2.6 and the standard deviations are derived from bootstrapping.

From table 2.6, I find that the market size positively affect the sale price and sale speed, consistent with the reduced form results. It suggests that homes on the largest 5% market enjoy a market size premium of about 5000 dollars than homes on the smallest 5% market and are sold at about 12 days faster on average. These micro level findings together with macro level findings in previous literature show that both the matching quality and matching speed are positively affected by the

Table 2.6: Structural Estimation of Thick Market Effects

	(1)	(2)
VARIABLES	SP/LP	TOM
Market size 1 (subdiv, 30k)	0.0158%***	-0.0202***
	(3.41E-04)	(0.00237)
Home characteristics	Yes	Yes
Zipcode and Year*Qtr Dummies	Yes	Yes
Constant	0.928***	8.767***
	(0.00212)	(0.413)
Observations	73,730	73,730

Robust standard errors in parentheses
*** p¡0.01, ** p¡0.05, * p¡0.1

market size.

2.5 Conclusion

This chapter proposes a house sale game that that captures the institutional arrangements and conventions in U.S. housing market and build an equilibrium directed search model to study the thick market effects on house seller's strategic pricing and search behavior and trading efficiency. The bidding occurs when multiple buyers compete for property and buyer's strategic counteroffers are affected by seller's choice of house list price. In equilibrium, it allows sellers to adopt list price to segment buyers by their valuations. And the final sale price, consistent with fact of the real transaction data, could be either lower, equal or higher than list price depending on realized number of buyers and their valuations in a selling period.

By numerically simulating the equilibrium directed search model, I find thick market effects prevail in housing market. First, house sellers benefit from thick market effects by setting a higher list price and expected higher sale price and shorter time to sale. Second, houses in thicker market are also more likely to sell at or above list price even if they are listed at higher prices. And the ratio of expected sale price to list price also increase as market size increases.

I then empirically investigate the thick market effects in housing market using unique data set from Dallas Multiple Listing Service. In reduced form analysis, I find that one standard deviation increase in market size increase sale price to list price ratio by 0.2% on average, which is equivalent to more than 4000 dollars increase in sale price. And the probability to sell above list price also increases by 0.6%. After controlling the unobservable effects, the results from structural model are also consistent with the reduced form findings. In the meantime, I find that while the thick market effects raise seller's reservation price and increases offer rejection probability, one standard deviation increase in market size still shortens the marketing time by 1.6 weeks or 11 days. These micro level findings together suggest that the existence of thick market effects in real estate markets.

3. DUAL AGENCY AND PRINCIPAL-AGENT PROBLEMS IN REAL ESTATE MARKETS

3.1 Introduction

In housing market, real estate agents are usually hired to facilitate home sales. According to National Association of Realtors (NAR), the real estate agent-assisted home sales account for 91% of total home sales in 2013.[1] Real estate agents usually have superior network in home sales, more knowledge on the real estate market, better marketing strategy and superior negotiation skills than typical home owners. When representing home owners, the agents are delegated exclusive authority to list the homes on the multiple listing service, a home listing platform and invite other agents to cooperate in their sales. The seller agents also provide a bundle of brokerage services including determining appropriate listing price, advertising and marketing and property maintenance. In addition, they also assist with legal documentation of home sales. In most cases, the seller agents are paid a fraction of final sale price from the sellers as commissions.

Not only home owners hire real estate agents to assist the home sales, the home buyers also frequently consult agents in their home searching process. Sometimes the real estate agents will serve as the buyer agents and help them on home evaluation and price negotiation. The buyer agents are not paid by buyers for their assistance, but split the commissions with the seller agents as reward for producing the matchings between home sellers and home buyers. In some other circumstances, if the agents recommend their own listings to the buyers or if the buyers contact

[1] Agent-assisted home sales accounts for 91% and FSBOs accounts for 9% of total home sales. The typical FSBO home sold for $\$184,000$ compared to $\$230,000$ for agent-assisted home sales. See more details at http://www.realtor.org/field-guides/field-guide-to-quick-real-estate-statistics.

them directly for their listings, they play role as dual agency that represents both the sellers and the buyers. And in this scenario, the agents get all commissions.

In US, the conventional commission rate is 6 percent of the final sale price. In recent years, the competition among real estate agents has driven the average commission rate down to 5.1%.[2]. From the seller agent's point of view, the commissions from home sales could differ substantially depending on whether they serve as dual agent or they sold with help of other buyer agents. The difference in commissions resulting from two types of agency structures provide the seller agents a strong incentive to sell the home to direct buyers than to buyers that are assisted by buyer agents, even at lower prices.

The principal-agent problem may arise in two folds. First, the information asymmetry between the home owners and the seller agents allows the seller agents to persuade the owners to accept lower offers from direct buyers and they enjoy higher commissions with the sacrifice of the owners' interest. Second, as a consequence and from the home buyer's point of view, the home buyers may not benefit from the assistance of buyer agents. Although the buyer agents do not necessarily collude with the seller agents, they do not bring any extra discount to the buyer. Instead, by bringing in more agent in the sale, the home buyers are paying more on the sale price.

In this chapter, I empirically test the effects of agency structure on home sales, including home sale price and speed of sale. To illustrate the idea, I first build a simple search model to show that the agency structure could create principal-agent problems through affecting the commission structure. In each period, the seller

[2] In some large cities such as Los Angeles and New York City, the real estate commission rate is further dropped to 4.5% to 5% for some listings. In most circumstances, the agents split the commission when two agents are involved in the transaction. In some other cases, the buyer's agents are often guaranteed to get 2.5% to 3% commission rate when they produce the buyer

agent will set two reservation prices contingent on the buyer type, a direct buyer or a buyer who has a buyer agent. An commission maximized seller agent will set a lower reservation price for direct buyers and higher reservation price for buyers with agent, and enjoy more commissions when the home is sold to a direct buyer than to buyer with agent. The testable implication is that the average sale price of home sales in the dual agency case will be lower than the price of home sales that involve two different agents. In addition, although the agency structure distorts the seller agent's incentive on sale price, it has null effect on sale speed, since at each period, the agency structure in the next period is not deterministic but random.

The data set I use for empirical study is from Dallas multiple listing service database in 2007. The data set offers several advantages in studying the agency structure and home sales. First, I have the information on the ID's of the real estate agents that are involved in each home sale. The agent's ID is unique for each agent and it helps me to separate the 16% of the home sales that are assisted by only seller's agent from the rest 84% home sales that are assisted by two agents. Moreover, the brokerage office IDs are also recorded in each sale and I can test if there is any collusion between agents. Second, the home listing information and sale outcomes including home list price, sale price, list date and off-market date are all available. I am able to construct the sale discount and speed of sale from these variables. Third, each home listing also contains the information on house characteristics including home size, location, number of bedrooms, number of bathrooms, etc.. I am able to test if there is any connection between the agency structure and home characteristics. I can also control for these variables in the analysis.

My identification strategy is to compare the outcomes of home sales in the dual agency situation with that of home sales that involve two different agents. If the principal-agent problem exists, I would expect the dual-agent-assisted home sales have lower sale price or deeper discounts than do the two-agent-assisted home sales. The identification strategy relies on two assumptions. First, there is no sorting on home characteristics. If homes with lower quality attract more direct buyers than other homes do, the difference in sale price may be explained by the difference in home value rather than agency structure. I test the validity of this assumption and find no correlation between agency structure and home observed characteristics. Second, the agency structure is not predictable and it is a stochastic consequence. If the agency structure is predictable at the time of home listing, the agents may strategically set the home list price. The relation between sale price and agency structure may be attributed to the relation between list price and sale price. I also do not find any evidence connects the list price with the agency structure. In addition, the agent's ability also matters in home sales. If some agents are more knowledgable and more famous, they may generate more home sales without cooperation of other agents. Failing in capturing the agent heterogeneity could result in a spurious relation between agency structure and sale price. In my regressions, I control the unobserved agent level information through the agent fixed effects. If all these assumptions hold, the assignment of the agency structure is as good as random conditional on the agent fixed effects.

Overall, the empirical study supports the existence of the principal-agent problem on the home sale price that is caused by the agency structure. After controlling the agent fixed effects, I find the home sales in the dual agency scenario have 1.7% to 2.6% or equivalently 3,400 to 5,000 dollars deeper discount on the sale price than that of the home sales which involve two different agents, adding to the average discounts of 6%. In addition, the seller agent's ability to push the home owners for more discounts also depends on the competition among buyers. Focusing on the home sales which attract fewer buyers by excluding home sales of which the sale price is higher than the list price incurred by bidding among multiple buyers, I find that the discounts are even deeper in the dual agency sales. Also, consistent with the theoretical prediction, I do not find any correlation between the time on the market and agency structure after controlling the agent fixed effects. These results are robust to different model specifications.

Another plausible explanation to the difference in the sale price between dual-agent-assisted home sales and two-agent-assisted sales is the price collusion the buyer agents and the seller agents. They have incentives to collude since they would both enjoy greater commissions through a higher sale price. The collusion story indicates another possible source of principal-agent problem and I am not able to fully separate this effects with other effects. Instead, I test this alternative explanations by comparing the sales that involve agents from same office with the sales that involves agents from different offices. If the collusion story were true, I would expect the agents from the same brokerage office collude more often on sale price. The sale price would be higher than the sale price of home sales that are assisted by agents from two offices. On the contrary, I find the home sales that involve two agents from the same office are sold 0.3% or 600 dollars less, which does not favor the collusion story.

My empirical findings provide some evidence to a number of theoretical papers that studies how the commission structure creates misaligned incentives for the home owners and the agents and how it induces principal-agent problems in housing market. For example, Zorn and Larsen [34] build a standard agency model and show that both flat-fee and percentage commission systems do lead to an alignment of the sellers and brokers interests and the broker's amount of search and listing price might also be below the sellers optimal price. Arnold [2] further compares the fixed-percentage commission, flat-fee, and consignment systems and concludes that fixed-percentage commission system is the only one of the three systems considered that can induce a first-best, incentive-compatible contract, giving its distortion on pricing. The discussion of the costs of principal-agent problems is extended to a general framework in Lazear [18]'s recent paper. In the paper, he investigates the informational benefits from hiring an agent and the costs of incentive misalignment resulting from commission structures. He finds that when agents have superior information or cheaper time than owners, the percentage commissions may induce a larger costs from price distortion than the flat fee commissions. Although the latter system may sacrifices other motivating aspects of the performance pay. My research adds to this discussion that the agency structure, when it's combined with the percentage commissions, may cause more severe principal-agent problems in real estate market.

My findings also adds to the growing empirical studies on the impact of information distortion. The information asymmetry between agents and the home owners leads to a distortive sale price because it enables the agents to push the home owners to accept a lower price from direct buyers. This information distortion is also find in some other markets. For example, excess capacity in hospital leads doctors to induce demand for more expensive services from their clients [9]. In vehicle inspection market, the inspectors tend to let vehicles pass the inspections to earn repeat business [14]. Particular in real estate markets, a number of recent papers also find evidence of the information distortion. Rutherford et al. [25] compares the agent-owned home sales and the client-owned home sales and find that the agents on MLS of one metropolitan area in Texas do not sell their homes faster than their client-owned homes, but they do use their information advantage to sell at a price premium of approximately 4.5%. Levitt and Syverson [19] also find evidence consistent to Rutherford et al. [25]'s findings. They examine nearly 100,000 home sales on MLS Illinois and find that the agents are more impatient in home sales and is willing to accept lower prices for a shorter sale time. Their work shows that the homes owned by real estate agents sell for 3.7% more than other homes and stay on the market 9.5 days longer. Two recent papers that study the dual agency structure are more related to my study. Prince et al. [24] examine the effects of the regulation of dual agency in residential real estate transactions in Long Island, New York between 2004 and 2007 but do not find any effect on sale price. Johnson et al. [15] find that the impact of dual agency depends on property ownership. They conclude that the dual agent is associated with a 7.19 percent price premium on agent owned properties, a 13.06 percent price discount on government-owned properties and a 7.04 percent discount on bank owned properties. While the price differences are quite economic significant among three types of homes, the results should be interpreted with cautious. Although the neighborhood fix effects are controlled in their analysis, if the ownership sorts on unobserved home characteristics such as home quality, the difference in sale price may not reflect the differential effects of dual agency. In contrast to their findings, I find the dual agency structure do have a negative effect on the sale price, although it does not affect sale speed. It supports that the information distortion not only presents in the way that agents spend less efforts on selling their clints homes than their own homes, but also in the circumstance when they sell the houses to direct buyers at lower price.

My findings together with these studies in information distortion also fit into the context of studying the efficiency (or inefficiency) of the brokerage service. Hsieh and Moretti [13]'s paper argues that the real estate brokerage is generally inefficient in terms of sale price and sale speed. They present across-city evidence that entry of new real estate brokers could further drive down the agent's returns and incentives and aggravate the inefficiency. Hendel et al. [11] instead compares the home sales on MLS and home sales on a For-Sale-By-Owner platform and find that although

MLS provide full brokerage services and homes are sold faster, the MLS listings do not generate a higher sale price. But their findings are less reliable to draw conclusions on inefficiency of the MLS service if the sellers on two platforms are different. Bernheim and Meer [3] also compare for-sale-by-owner (FSBO) homes with agent-assisted homes on the Stanford University campus. They find that the add value of brokerage service bundle in this non-MLS market do not justify the commissions. Their findings should also be interpreted and generalized to other real estate market with cautious if the Stanford housing market differs substantially from traditional MLS markets.

The rest of this chapter is organized as follows. In section 2, I provide a simple model to illustrate the misalignment between home owner's incentive and agent's incentive. I show that agents are willing to accept offers from direct buyers than from buyer's agents, even if the prices are lower. But the speed of sale does not depend on agency structure. In section 3, I introduce the sample set from Dallas Multiple Listing Service (MLS) database and provide summary statistics. In section 4, I present empirical evidence on the principal-agent problem and check the sensitivity of our results. The last section is the conclusion.

3.2 A Simple Model

In this section, we present a simple theoretical model to illustrate the mechanism of agency structure (commission structure) on home sales. The basic idea is that a profit maximization seller's agent will set lower reservation price to a buyer and higher reservation price to a buyer's agent. Therefore, the seller agent is more willing to accept offers from direct buyers than from other agents that delegate buyers. But since the probabilities of getting a direct buyer and getting a buyer's agent is determined by market conditions rather than home listing information, the speed of sale should not be depending on the agency structure.

Consider the decision making process of a home seller agent. The objective function for the agent is to maximize the commissions, which comes as a percentage of the home sale price. In each period, the agent draws a pair (θ, p) of visitor type and a home buying offer. The type of visitor θ can be either a buyer's agent or a buyer along. In the former case, if the agent accepts the offer, he gets a percentage α of sale price. In the latter case, the agent gets 2α. I assume the probabilities are q and $1 - q$ respectively. The house buying offer p is drawing from some distribution with c.d.f. $F(p)$. The agent can either accept an offer or decline and offer. Furthermore, the draws of θ and p are simultaneous and independent. The agent can draw a new offer if and only if he also draws a new visitor type.

A seller's agent chooses visitor type-offer $v(\theta, p)$ pairs subject to the following conditions. First, if the home is sold to visitor with type θ and offer p, the agent gets a total amount θp of commission and the sale ends. Therefore the worker maximizes $E \sum_{t=0}^{\infty} 1[s_t = sold] * \theta_t p_t$. Second, the visitor type takes binary outcomes. If the visitor is a buyer's agent, $\theta = \alpha$ and $\theta = 2\alpha$ if the visitor is a buyer. Third, an offer is a draw of p from c.d.f. F. Successive draws are independent with $F(0) = 0$ and $F(B) = 1$, where B is some upper bound. The agent decides at each period whether to accept a type-offer pair or decline the offer. If the offer is declined, the agent draws a new pair at the beginning of next period.

Let $v(\theta, p)$ be the optimal value of the problem at the beginning of a period for a agent with visitor type-offer pair (θ, p) who is about to decide whether to draw a new visitor type and buying offer. The Bellman equation is

$$v(\theta, p) = \max \left\{ \theta p; -c + \beta \int \int \theta' p' dF(p') G(\theta') \right\} \tag{3.1}$$

where $G(\theta = \alpha) = p$ and $G(\theta = 2\alpha) = 1 - p$. β is the time discount factor and c is the waiting cost. The maximization is over the above two choices, i.e., either accepting θp or draw a new pair (θ', p') in next period. The agent's optimal strategy is to set a reservation price $\bar{p}$ depending on the type of visitor, such that

$$v(\theta, p) = \begin{cases} \theta\bar{p} = -c + \beta \int\int \theta' p' dF(p')G(\theta') & : p \leq \bar{p} \\ \theta p & : p > \bar{p} \end{cases} \tag{3.2}$$

and accept offer only when it exceeds the reservation price. I denote $\bar{p}(\theta = \alpha) = \bar{p_1}$ be the reservation price conditioned on receiving offer from a buyer's agent and $\bar{p}(\theta = 2\alpha) = \bar{p_2}$ be the reservation price conditioned on receiving offer directly from a buyer. The reservation prices are determined by the following equations,

$$\alpha\bar{p_1} = -c + (2 - q)\alpha\beta \int_0^B pdF(p) \tag{3.3}$$

$$2\alpha\bar{p_2} = -c + (2 - q)\alpha\beta \int_0^B pdF(p) \tag{3.4}$$

After some algebra, we find that $\bar{p_1}$ and $\bar{p_2}$ satisfy conditions

$$\bar{p_1} = -c/\alpha\gamma_1 + (2 - q)\beta/\gamma_1 \int_{\bar{p_1}}^B (p - \bar{p_1})dF(p) \tag{3.5}$$

and

$$\bar{p_2} = -c/\alpha\gamma_2 + (2 - q)\beta/\gamma_2 \int_{\bar{p_2}}^B (p - \bar{p_2})dF(p), \tag{3.6}$$

where $\gamma_1 = 1 - (2 - q)\beta$ and $\gamma_2 = 2 - (2 - q)\beta$ or $\gamma_2 = \gamma_1 + 1$. Therefore, the agent's optimal choice is to accept offer from a buyer's agent only if the offer exceeds p_1 and accept offer from buyer if buyer's offer exceeds p_2. Noticce that α is usually equal or less than 3%. We have the following propositions.

Proposition 3.2.1 *Given c.d.f F, $q \in [0,1]$ and $\alpha \in (0, 0.1)$, we have $\bar{p_1} > \bar{p_2}$ and $E(p|buyer'sagent) > E(p|buyer)$. Furthermore, as q decreases, both $\bar{p_1}$ and $\bar{p_2}$ decrease and $\bar{p_2}$ decreases more.*

This proposition says that the seller agent will set lower reservation price to a buyer than the reservation price to a buyer's agent to maximize his profit. It predicts that holding other things constant, the sale prices of one-agent-assisted homes will be on average smaller than the sales that involve both a seller's agent and a buyer agent. It also suggests that the arrival rate of house visitors can affect the reservation prices and in hotter market, the difference in sale price between one-agent-assisted homes and two-agent-assisted homes is smaller.

Since in each period, the seller agent receives an independent draw of the agency structure and offer. I denote the probability of sale in each period as $Prob(sale)$ and

$$Prob(sale) = q(1 - F(\bar{p_1})) + (1 - q)(1 - F(\bar{p_2})). \tag{3.7}$$

This leads to the following proposition.

Proposition 3.2.2 *Define T as time on the market, we have $E(T) = 1/Prob(sale)$ and $E(T|\theta_T = 2\alpha) = E(T|\theta_T = \alpha) = E(T)$.*

The speed of sale, or the time on the market is the reciprocal of the probability of sale. Since the probabilities of getting a direct buyer and getting a buyer agent is determined by market conditions rather than home listing or seller's agent's reservation price, the speed of sale would not be different, whether the home is sold directly to a buyer, or sold through a buyer agent.

3.3 The Data and Variables

3.3.1 The Dallas Multiple Listing Service

The data set I use comes from Dallas multiple listing service(MLS) database. The multiple listing service is an actively managed home listing platform that enables real estate brokers share information on properties they have listed and invite other brokers to cooperate in their sale in exchange for compensation if they produce the buyer. MLS is the primary source of information of homes currently for sale. According to NAR's 2006 survey of home buyers and sellers, 88 percent of sellers reported that their home was listed in the MLS. The MLS disseminates listing information and sellers benefit by increased exposure to their property and buyers benefit because they can obtain information about all MLS-listed properties while working with only one broker.[3][31]. In current, about 90 percent of residential properties are listed and sold through over 800 MLSs nationwide.

The Dallas multiple listing service database focuses on the Dallas-Fort Worth-Arlington metropolitan area and its vicinity and records the listing information and transaction details of every residential property that is listed on Dallas MLS. A typical house listing record in Dallas MLS contains information on the type of house, house address, list date, off-market date, listing price, final sale price and a selection of house characteristics including home size, lot size, number of bedrooms, number of bathroom, etc.. I get the MLS data on 88,788 home sales in 2007. The S&P/Case-Shiller home price index shows that the residential real estate in Dallas market between this period is relatively stable[4].[30]

[3]see http://en.wikipedia.org/wiki/Multiple_listing_service for detail functions of multiple listing service.

[4]The S&P/Case-Shiller Dallas home price index measures the average change in value of existing single-family housing stock in Dallas given a constant level of quality. During 2007, the index changes moderately from 122.64 in January to 120.78 in December. The peak value is 126.47 in June and trough is 120.78 in December.

The novelty of this data set is that it includes the agents information that can help me identify the one-agent-assisted home sales and two-agent-assisted home sales. Foe each home sale, both the seller agent ID and buyer agent ID are attached. This ID is unique for real estate agent. If the ID on the seller agent is the same as the buyer agent, the home is sold to a direct buyer. If the IDs are different, the home sale would have involved two agents. Moreover, both the seller's brokerage office ID and buyer's brokerage office ID are also recorded in each sale and I can test the collusion between agents.

3.3.2 The Data and Variables Summary

In the sample set, about 96% of the properties are single-family residential houses and the rest 4% are condominiums. We focus on residential single-family home sales in our analysis. The condominiums are excluded from our analysis because the condominium market may behave quite different from the major residential market. The outcome variables in our analysis are the home list price, sale price and time on the market (TOM). The house list price and sale price are directly observed from the data, and the time on the market is measured number of the weeks between the house off-market date and the list date. We make the following restrictions to the estimation sample. First, we drop sales with list prices and sale price that are below the 1st and above 99th percentiles respectively. Second, we dropped sales with sale price to list price ratio that is less than 0.5 or greater than 2 to exclude suspicious recordings. Third, we drop sales where the time on the market is greater than 2 years. Those sales account less than 0.1% of the total sample size. In total, we reduce the sample size to 84,348 home sales. On average, the home list price is around 200,000 dollars and the final sale price has 6% discount. A typical home stays on the market for 79 days and 99% of the homes are sold within a year. The summary statistics of house list price, sale price and time on the market (TOM) is presented in table 3.1.

Table 3.1: Summary Statistics: Home Sales and Home Characteristics

Variable	**Mean**	**Std. Dev.**	**N**
List price (dollars)	199826	140503	84348
Final sale price (dollars)	188826	133247	84348
Price change (dollars)	11000	17652	84348
Sale to List price ratio	94%	8%	84348
Time on the market (days)	79	76	84348
Number of bedrooms	3.4	0.71	84313
Number of bathrooms	2.5	0.87	84279
Home size (sqft)	2285	960	8799
Imputed Home size (sqft)	2205	810	84266
Lot size (sqft)	10266	7690	31946

The variable of main interest is the agency structures of the home sales. During the year of 2007, in total 12,721 seller agent actively practiced in home sales. Among the homes, 85% are bought and sold with the assistance of buyer agents and seller agents, while the rest 15% are assisted by dual agents. From the agent's perspective, on average the seller agent assisted 118 of home sales. But this number is driven up by some star agents and the median agents sold 16 homes per year. Among the agents who sold more than one homes in 2007, 82% of them have both dual-agent-assisted sale and two-agent-assisted sale records. The median agent has 11% of the former sales and 89% of the latter. There is a substantial variation across agents and the standard deviation is 17%. The summary statistics of the agents performance can be found in table 3.2.

In analysis, we use a binary variable to indicate whether a home sale is assisted with only a seller agent or both seller agent and buyer agent. To be specific, we

Table 3.2: Summary Statistics: Agent Sales

Variable	**Mean**	**Std. Dev.**	**N**
Number of home sales	118	287	84348
Homes sold by dual agents(%)	15%	18%	84348
Agent average dual agencies	18	54	84348
Agents involved in both agency structure (%)	82%	39%	84348
Homes sold by agents from the same office (%)	5%	23%	84348

define

$$\textit{dual agency dummy}_{ij} = \begin{cases} 1 & : \text{if home sale is assited by dual agent} \\ 0 & : \text{if home sale is assisted by two different agents} \end{cases} \tag{3.8}$$

where ij denotes agent i and j denotes house listing j. We also define another a dummy variable to indicate whether a home sale is assisted by two agents from the same brokerage office or from different brokers as follows.

$$\textit{same office dummy}_{ij} = \begin{cases} 1 & : \text{if home sale assisted by two agents in same office} \\ 0 & : \text{if home sale assisted by agents in different brokers} \end{cases} \tag{3.9}$$

Of the 84,348 total home sales, 4,963 sales are assisted by two agents from the same brokerage office and 68,044 sales are assisted by by agents from different brokers.

Besides the key variables, the data set also provides us information on the house characteristics, including the number of bedrooms and number of bathrooms. However, the home size is available for about one third of the houses. We use regression-based method to approximate home square footage for houses that have missing data.[5] The house lot size is sometimes directly recorded in square footage and some-

[5] We first regress house square footage on number of bedrooms, number of bathrooms, house lot size (if available) and 5-digit zip code dummies. Then we use the estimated coefficients to predict the square footage for missing values.

times include the lengths of all sidelines of the lot. We carefully derive approximated lot size using imputation method.[6] The average home size in our sample set is about 2200 square feet with 3.4 bedrooms and 2.5 bathrooms, and the lot size is about 10000 square feet. The summary statistics of house characteristics are also presented in table 3.1.

3.4 Empirical Results

3.4.1 Identification Strategy

My identification strategy is to compare the outcomes of home sales in the dual agency situation with that of home sales that involve two different agents. If the principal-agent problem exists, I would expect the dual-agent-assisted home sales have lower sale price or deeper discounts than do the two-agent-assisted home sales. A general econometric model that describes the relation between dual agency and sale price is as follows.

$$\mathit{ln(sale\ price)}_{ij} = f(\mathit{dual\ agency\ dummy}_{ij}; X_j, z_j, c_i) + u_{ij} \tag{3.10}$$

In above model, the subscript i refers to agent i and j refers to house listing j. The dependent variable is the log of sale price. The *dual agency dummy* is the dummy variable defined earlier that equals to one if only the seller's agent assisted the home sale and 0 otherwise. X_j is a set of observed house characteristics and z_j is unobserved house information. The information on seller's agent is also unobserved and denotes as c_i. u_{ij} is the stochastic term.

Our identifying assumption is that conditional on the agent fixed effects, the

[6]The size of housing lot is recorded in side lengths for some houses. If the lot is in rectangular shape, the lot size is directed calculated as the product of two adjacent sides. If the lot is non-rectangle, the lot size cannot be exactly computed without at least one known angle. We therefore approximate the lot size by taking average of two opposite sides and multiply the two averages.

agency structure for each home sale is assigned randomly. This assumption relies on three facts. First, the agent's ability is heterogenous and it affects home sales. Some agents are more knowledgable and more famous while some are less experienced. Star agents may be able to generate more home sales without cooperation of other agents and at the same time, they also sell homes more effectively. Failing in capturing the agent heterogeneity could result in a spurious relation between agency structure and sale price. Therefore in the main regressions, I control the unobserved agent level information through the agent fixed effects.

The second fact is that at the time when the each home were listed, neither the home owner nor the seller agent could perfectly predict whether the home is going to be sold to a direct buyer or to a buyer with an agent. Although some seller agents may have better chance to be served as dual agent than other agents, the possibility of dual agency for each agent is still random. Therefore, the seller agents would apply similar pricing and marketing strategy to all the listing homes. A number of papers have found that seller agents strategic set the list price to attract buyers, for example, see Yavas and Yang [32], Haurin et al. [10] and Zhao [33]. And if the agency structure is random, it should not affect the list price determination.

In table 3.3, I specify a log-linear regression model and test the relation between list price and agency structure. The first column is a pooled regression that regress the list price on dual agency dummy and observed home characteristics. The coefficient on the dual agency dummy is 0.06%, indicating that it has a negligible impact on list price, and it is insignificant. The home size and number of bathrooms are positively correlated to the list price. The number of bedrooms are negatively correlated with list price because more luxury homes tend to have fewer bedrooms given the home size. In the second column and third column, we add month fixed effects and agent fixed effects to control time-varying market conditions and agent characteristics that may affect the agency structure and list price. In the last column, we also include the home lot size and a dummy for irregular lot. The sample size is reduced to 32,143. In all these model specifications, the coefficients on dual agency dummy are insignificant.

Table 3.3: Log List Price and Agency Structure

	(1) ln(LP)	(2) ln(LP)	(3) ln(LP)	(4) ln(LP)
Dual agency dummy	0.060%	0.11%	-0.46%	-1.51%
	(0.06)	(0.10)	(-0.95)	(-1.92)
Number of bedrooms	-0.0189**	-0.0186**	0.0267***	0.0254***
	(-3.12)	(-3.06)	(5.36)	(3.66)
Number of bathrooms	0.355***	0.356***	0.254***	0.268***
	(39.44)	(39.27)	(35.87)	(25.05)
Log home size	0.354***	0.353***	0.319***	0.219***
	(17.53)	(17.43)	(18.89)	(8.95)
Log lot size				0.132***
				(18.34)
Irregular lot dummy				0.0828***
				(7.44)
Month fixed effects	No	Yes	Yes	Yes
Agent fixed effects	No	No	Yes	Yes
		(-0.32)	(-0.26)	(-0.58)
Constant	8.499***	8.480***	7.980***	8.612***
	(67.58)	(67.48)	(20.90)	(14.03)
N	84244	84244	84244	32143

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

t Note: LP refers to list price

The third fact that justifies our identification assumption is that I do not find any sorting issue on home observed characteristics. In my regression model, if the unobserved home characteristics is correlated with the agency structure, my identification strategy could be threatened and the results are less valid. For example, if homes with lower quality that are often on discounts attract more direct buyers than other homes do, the difference in sale price may be explained by the difference in home value rather than the agency structure. Given the data limitation, I would not be able to completely ruled out the sorting possibility un home unobservables. Instead, I test the correlation between agency structure and home observed characteristics.

Table 3.4: Probit Model of Agency Structure and Home Characteristics

	(1) Dual agent	(2) Dual agent	(3) Dual agent	(4) Dual agent
Number of bedrooms	-0.0155	-0.0159	0.0310	0.0307
	(-1.22)	(-1.25)	(1.53)	(1.52)
Number of bathrooms	0.187*	0.186*	0.214	0.211
	(1.71)	(1.70)	(0.78)	(0.77)
Log home size (log sqft)	-0.614	-0.610	-0.676	-0.667
	(-1.35)	(-1.33)	(-0.96)	(-0.94)
Log lot size			0.161**	0.161**
			(2.84)	(2.83)
Irregular lot dummy			-0.0133	-0.0151
			(-0.36)	(-0.41)
Month fixed effects	No	Yes	No	Yes
Constant	3.159***	3.173***	1.929***	1.950***
	(10.95)	(11.01)	(4.08)	(4.13)
N	84244	84244	32143	32143

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.4 presents the estimation results from a probit model. In the first column, only the number of bedrooms, number of bathrooms and log home size are included in the regression and only the number of bathrooms has marginally significant positive correlation on the agency structure. If the number of bathroom is positively correlated with home luxury, it may indicate a marginally sorting issue. In the second column, we add month fixed effects and the results are similar to the first column. In the third and fourth column, we add home lot information, the coefficient on number of bathroom becomes insignificant. Instead, the coefficient on lot size becomes significant. It still provides a scenario that there might be a mild sorting on home characteristics. Although the agency structure does not sort on a number of other home characteristics, the results should be interested with cautious.

3.4.2 The Home Sale Price

The economic model predicts that the sale agents are more likely to accept offers from direct buyers even if the offers are lower because they enjoy higher commissions in dual agency case. It sheds light on the relation between home sale price and the commission rate. Ideally, if I were able to see the commission rate in each transaction, I would directly test how the relation between changes in commission rate and the home sale price. Instead, I focus on the effect of agency structure on home sale price and specify a log-linear regression model as follows.

$$\mathit{ln(sale\ price)}_{ij} = \beta_0 + \mathit{dual\ agency\ dummy}_{ij}\beta_1 + X_j\gamma + c_i + u_{ij} \tag{3.11}$$

As explained before, we control the agent fixed effects c_i and home observed characteristics in our analysis. The estimation results are reported in table 3.5. All standard deviations are clustered at the agent level. Overall, I find a strong evidence that support the existence of the principal-agent problem in the dual agency home sales. In the first column, I regress the log sale price against dual agency dummy without controlling other variables. The coefficient on the dual agency dummy is -6.3% and it is significant at 1% level. It suggests that the homes are sold at -6.3%

lower in dual agency case, a strong evidence that the seller agents leverage the information advantage to distort the sale price and earn more commissions. In column 2, home characteristics and agent fixed effects are added into the regression model. The coefficient on the dual agency dummy dropped to about -1.7%. The large drop in the magnitude may be mainly attributed to the inclusion of agent fixed effects, since the marketing or discount strategy may vary substantially across the seller agents. In column 3 of table 5, the month fixed effects are also controlled and the dual-agent-assisted home sales are still associated with about 1.6% more discounts than the two-agent-assisted home sales. With the inclusion of log home lot size and dummy for irregular lot, the sample size is reduced more than a half, but the effect of agency structure on sale price becomes -2.6%. The results also imply that with the assistance of an additional buyer agent, the home buyers do not benefit from buyer agent services. Instead, they pay about 1.7% to 2.6% more on the final price.

Some other home attributes are all positively significant in the regressions, including the number of bedrooms, the number of bathrooms, the log of home size, the log of lot size and irregular lot dummy. Larger homes are unsurprisingly sold at higher price. On average, a 10% increase in home size leads to roughly 2.6-3.5% increase the home sale price. Homes that have larger yard also tends to be more popular. A 10% increase in home lot size leads to about 1.3% increase in the home sale price. Besides home size and lot size, the bedrooms and bathrooms also add value to the home sales. An additional bedroom adds about 2.5% value to the home sales. The coefficient on number of bathroom is surprisingly 0.24, about ten times the size of coefficient on bedroom and suggests that each additional bathroom will bring 24% increase in home sale price. One possible explanation is that luxury homes usually have more bathrooms and when it is unobserved, the coefficient on the number of bathrooms picks up the unobserved effects.

Table 3.5: Log Sale Price and Agency Structure

	(1)	(2)	(3)	(4)
	ln(SP)	ln(SP)	ln(SP)	ln(SP)
Dual agency dummy	-6.28%**	-1.72%***	-1.63%***	-2.62%**
	(-3.11)	(-3.51)	(-3.33)	(-3.25)
Number of bedrooms		0.0245***	0.0248***	0.0209**
		(4.99)	(5.05)	(3.01)
Number of bathrooms		0.240***	0.241***	0.255***
		(33.93)	(33.90)	(22.62)
Log home size (log sqft)		0.352***	0.350***	0.262***
		(20.81)	(20.62)	(10.07)
Log lot size				0.125***
				(18.18)
Irregular lot dummy				0.0813***
				(7.10)
Month fixed effects	No	No	Yes	Yes
Agent fixed effects	No	Yes	Yes	Yes
Constant	11.97***	7.468***	7.438***	8.043***
	(684.76)	(18.88)	(18.91)	(12.24)
N	84348	84244	84244	32143

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

t Note: SP refers to sale price

Although the information asymmetry allows the seller agents to persuade the home owners to lower the sell price when they serve as dual agents, their capability of leveraging the information advantage may be affected by competition among buyers. For example, more competition may reveal the market information to the home owners and reduce the seller agent's power to push the home owners accepting low price. I separate the home sales into two groups. In one group, the sale price of homes is bid above the list price, which implies those homes attract more competition among home buyers. In another group, the home sales are associated with less competition among buyers and the sale price is below the list price. I test the effect of the agency structure for each group in table 3.6.

The first two columns of table 3.6 report the estimates of homes that are sold at sale price less than the list price. In the first column, the coefficient on dual agency dummy is 2.6% and it is significant at 1% level, suggesting that the seller agents can better exert the information advantage and give 2.6% more discounts to direct buyers and enjoy double commission benefits from dual agency. In the second column when the home lot size information is added into the regression, the discounts become even deeper. In the column 3 and column 4, the same regression models are applied to the homes that attract more competition among buyers and are sold above list price. The coefficient on dual agency structure becomes positive. It may suggest that that when the information advantage is reduced by competition among buyers, the seller agents in dual agency case tend to persuade the buyer to offer a higher price. But the effect is statistically insignificant.

An alternative explanation to our results that the two-agent-assisted homes are sold at a higher price than dual-agent-assisted homes is that the real estate agents collude on home sale price. In fact, both agents have incentives to agree on a higher sale price since they both benefit from such sale. This explanation provide another

Table 3.6: Buyer's Competition, Log Sale price and Agency Structure

	$SP < LP$		$SP >= LP$	
	(1)	(2)	(3)	(4)
Dual agency dummy	-2.60%***	-3.33%***	1.81%	0.09%
	(-4.55)	(-3.48)	(1.60)	(0.05)
Number of bedrooms	0.0150**	0.0104	0.0599***	0.0593***
	(2.78)	(1.32)	(7.02)	(4.92)
Number of bathrooms	0.225***	0.247***	0.295***	0.284***
	(30.32)	(19.93)	(19.48)	(13.44)
Log home size (log sqft)	0.417***	0.308***	0.126***	0.101
	(21.77)	(10.07)	(3.46)	(1.90)
Log lot size		0.126***		0.120***
		(17.76)		(7.30)
Irregular lot dummy		0.0773***		0.0948***
		(5.81)		(3.87)
Month fixed effects	Yes	Yes	Yes	Yes
Agent fixed effects	Yes	Yes	Yes	Yes
Constant	6.647***	7.395***	10.11***	10.38***
	(15.43)	(10.13)	(19.58)	(12.84)
N	64882	24741	19362	7402

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

t Note: SP refers to sale price

mechanism that may create the principal-agent problem from the buyer's point of view. Unfortunately, I am unable to directly identify the collusion sale and non-collusion sale to test this explanation. However, an possible way to test is to compare the home sales in which the seller agents and the buyer agents are more likely to collude with the home sales in which they are less likely to collude. Agents from same brokerage office are more likely to collude since they are known to each other and more often cooperate in sales. Therefore I test the collusion story by comparing the home sales that are assisted by agents from the same brokerage office to the home sales that are assisted by agents from different brokerage offices. As described in previous section, a dummy variable *same office* $dummy_{ij}$ is used to capture the variation of brokerage office and it equals one if agents are from the same office.

The estimation results are summarized in table 3.7. Column 1 and column 3 are the estimation results from the OLS regression. The coefficients on the same office dummy ranges from -0.2% to 0.5% and both are economically and statistically insignificant. After including the agent fixed effects, the coefficients on the same office dummy reported in column 2 and column 4 remain insignificant. It suggests that the agents from same office may not collude on sale price, at least they do not behave differently on sale price from the agents from different offices. A less intuitive results in the regression is that the coefficient on the number of bedrooms is significantly negative in column 1 and 3 without controlling the agent fixed effects. In fact, this is caused by the correlation between agent unobserved characteristics and the home characteristics. In column 2 and 4, after controlling the agent fixed effects, the effect of the number of bedrooms become positively significant on the home sale price.

In addition to the above explanation, I also checked the possible relations between the buyer's heterogeneity in patience, the agency structure and the sale price, since

Table 3.7: Home Sales and Collusion in Office

	(1)	(2)	(3)	(4)
	ln(SP)	ln(SP)	ln(SP)	ln(SP)
Same office dummy	-0.21%	-0.66%	0.56%	-0.17%
	(-0.17)	(-1.05)	(0.32)	(-0.16)
Number of bedrooms	-0.0219***	0.0261***	-0.0622***	0.0224***
	(-3.58)	(5.12)	(-6.68)	(3.31)
Number of bathrooms	0.336***	0.233***	0.391***	0.257***
	(34.92)	(30.93)	(30.16)	(20.82)
Log home size (log sqft)	0.405***	0.364***	0.271***	0.254***
	(19.25)	(20.06)	(8.35)	(8.75)
Log lot size			0.112***	0.126***
			(8.60)	(16.23)
Irregular lot dummy			0.137***	0.0725***
			(7.87)	(6.43)
Month fixed effects	Yes	Yes	Yes	Yes
Agent fixed effects	No	Yes	No	Yes
Constant	8.083***	7.149***	8.089***	7.655***
	(59.54)	(15.90)	(29.80)	(10.68)
N	72923	72923	27802	27802

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

the positive correlation between dual agency and sale price discount may also be attribute to the buyer's patience. For example, if the buyers of more expensive homes are busier and less patient, they are more likely to delegate the search to the buyer agents. At the same time, they may accept a higher price and therefore the price of home sales with the assistance of two agents appear to be higher than home sales that have dual agents. From the buyer's point of view, it provides another channel through which the buyer agents can take more advantage from information asymmetry and buyer's impatience and it may lead to more severe principal-agent problem. In figure 2, use home list price as a proxy for home value and plot the relation between home value and dual agency percentage. It shows that when homes are less than 100,000, the dual-agent-assisted home sales take up about 20% of the total home sales. It suggests that buyers for cheaper homes are more likely to be . But when the home value is above 100,000, there is no significant relation between the home value and adopting of dual agency.

3.4.3 The Speed of Home Sale

Since at the time of listing, neither the home owners nor the seller agents know the agency structure of the home sales and the buyer type only realized when the homes are sold, the speed of sale should not be correlated with agency structure. I also test the effect of agency structure on the time on the market in this section. If there is no correlation between the agency structure and the speed of sale, it adds more support to the randomness of the agency structure in home sale. I use time one the market to measure the speed of sale and duration model may particular fit the data. I specify the Cox proportional hazards model as follows.

$$\lambda(t; W_{ij}) = \kappa(W_{ij})\lambda_0(t) \tag{3.12}$$

where $W = (dual\ agency\ dummy_{ij}, X_j, c_i, u_{ij})$.

The estimation results are summarized in table 3.8. All the coefficients on the dependent variables have been translated from hazard ratio to percentage effects. In the first column, the dual agency dummy together with the home size, the number of bedrooms, the number of bathrooms and the month fixed effects are included in the regression model. Initially the coefficient on dual agency dummy is significant without controlling the agent fixed effects and the magnitude reduces mildly to about -9% when I add the log of home lot size and the lot shape in the third column. Without considering the seller agent information, it suggests that on average, the dual-agent-assisted homes stay on the market for -11.7% shorter time than two-agent-assisted homes. Although the coefficient is significant, it may reflect spurious relation between the agency structure and the time on the market. To test this, I further control the agent fixed effects and re-estimate the model in column 2. The result shows that after controlling the agent fixed effects, the coefficient drops to -0.3% and the effect of dual agency on the sale speed becomes both economically and statistically insignificant. When I add the home lot size and lot shape dummy into the regression in column 4, the coefficient on dual agency dummy remains insignificant and the sign even flips. This exercise implies that the home sale speed varies across agents, but it does not depend on the agency structure. This finding in general support our assumption that the agency structure is randomly determined and it should not affect the time on the market theoretical model.

Besides the agency structure, most of the home characteristics in the sale speed regressions are significantly correlated with the home time on the market. First, larger homes spend more time to sell. On average, 10% increase in the home square footage will prolong the marketing time by about 10 to 19 percent, or 8 to 16 days. Second, fixing the home size, homes with more bedrooms and more bathrooms are

Table 3.8: Proportional Hazard Model Estimation of TOM and Agency Structure

	(1) TOM	(2) TOM	(3) TOM	(4) TOM
Dual agency dummy	-11.7%***	-0.40%	-9.08%***	2.43%
	(-11.46)	(-0.32)	(-5.50)	(1.22)
Number of bedrooms	-0.0218**	-0.0140	-0.0658***	-0.0562***
	(-3.01)	(-1.73)	(-5.59)	(-4.25)
Number of bathrooms	-0.104***	-0.118***	-0.118***	-0.131***
	(-10.83)	(-10.76)	(-7.40)	(-7.26)
Log home size (log sqft)	0.0985***	0.0427	0.190***	0.139**
	(3.88)	(1.49)	(4.64)	(3.04)
Log lot size			-0.0705***	-0.0274
			(-6.42)	(-1.87)
Irregular lot dummy			0.0503*	0.0116
			(2.13)	(0.33)
Month fixed effects	Yes	Yes	Yes	Yes
Agent fixed effects	No	Yes	No	Yes
				(1.55)
N	83618	83618	31908	31908

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: the coefficients have been translated from hazard ratio to percentage.

more favored in the market and are sold faster. The effect of each additional bedroom on the time on the market ranges from 2 to 6 percent or 1 to 4 days depending on model specifications. And each additional bedroom reduces the time on the market by about 10% to 13% or 8 to 10 days.

3.5 Conclusion

In this chapter, I test the effects of agency structure on home sales. I first build a search model for the seller agent and shows that when agents are delegated pricing authority, they tends to accept lower offers from direct buyers than offers from buyers who have other assisting agents and they enjoy all commissions from dual agency home sales. From the buyer's point of view, they do not benefit from assistance of buyer agents neither.

I have employed a unique data set that contains the home information, sale information and agents information from Dallas-Fort Worth Metropolitan MLS listings to empirically examine the effect of agency structure on home sales. I find supporting evidence from the null effect of agency structure on list price that at the time when homes were listed, the agency structure was not known to the agents. Conditioning on the agent fixed effects, whether an agent serves as dual agent in a home sale do not depend on home observed characteristics. By comparing the homes sales that are assisted by dual agents to the home sales that are assisted by two different agents, I find that the dual-agent-assisted home sales are associated with 1.7% to 2.6% or equivalently 3,400 to 5,000 dollars deeper discounts on the sale price than that of the home sales which involve two agents. This finding supports the existence of a severe principal-agent problem.

I also find empirical evidence that the agent's ability to leverage the information advantage depends on the competition among buyers. When there are fewer competitions and when the homes are sold below the list price, the seller agents can give 2.6% to 3.3% more discounts to direct buyers. In contrast, when there are more bidding among buyers when the homes are bid above the list price, the agents can not exert their information advantage on setting different sale price for different types of buyer. I have also tested the collusion between buyer agents and seller agents that could possibly drive the sale price high in the home sales that are associated with two agents and I find no evidence that favors this story.

The effect of agency structure on the time on the market is also tested in this paper using a Cox-proportional hazard model and consistent with the theoretical implication, I do not find any correlation between the time on the market and agency structure after controlling the agent fixed effects. It further supports my identification assumption that the agency structure is randomly decided. In addition, home characteristics and seller agents matter to the sale speed.

4. PRICE ENDINGS AND HOME SALES

4.1 Introduction

Recent literature in behavioral economics and marketing science suggests that when sellers decide the list price of some goods, the price endings may not be trivially decided since it may have some psychological impacts on buyer's purchasing decision. In this chapter, I attempt to study how price endings lead to difference in sale outcomes in real estate markets.

The literature in studying the ending digit effects on sales has suggested several explanations. First, the price ending affects the buyer's cognitive ability and it affects the magnitude judgement. For example, it may take longer time for buyers to process the magnitude information when the price is precise than when it is round. For example, buyers may spend longer time to compare 241,873 with 241,349 than to compare 240,000 with 250,000. Second, some precise prices may be perceived cheaper than the actual price, especially when prices end with 9. For example, the perceived difference between \$2.99 and \$3.00 may be more than one cent. A number of empirical papers have found that sellers in the retail market strategically choose the price endings and take advantages of the phycological impacts on buyer's decisions. Some of the representative discussions can be found at Schindler and Kirby [26], Stiving and Winer [27], Mazumdar and Papatla [21], Thomas and Morwitz [28], Manning and Sprott [20], Thomas et al. [29], etc.

The behavioral effects of price endings on home sales in real estate markets for several reasons. First, the homes are large items and the buyers may spend more effort on interpreting the price information and comparing other properties. It is not clear whether the price ending effect in small item retail markets would exist in the housing market. Second, unlike in the retail market that the list price is usually a take-it-or-leave-it offer, the home list price may be

A number of papers attempt to explore how the home list price is determined and how it affects home sales. Knight [16] and Haurin et al. [10] consider a sequential search problem for a typical home seller. In each period, the home seller attracts a visiting buyer with some probability. The lower the list price, the higher the probability that it attracts a buyer. On the other hand, the list price is the ceiling price of buyer's counteroffer and lower list price reduces the seller's expected revenue. The optimal list price then balances these two effects. In the second chapter, I also provided an explanation to the list price in the directed search model framework. This chapter deviates from the rational agent models and investigate how some prices may potentially generate psychological effects and how sellers may take advantages of these list pricing strategies in home sales. It also help us to interpret the heaping phenomenon of the list prices (say, 200,000 and 199,000) in real estate markets.

The regression discontinuity analysis and fixed-effect panel data model analysis of the Dallas multi-year housing market transactions present two main findings. First, consistent with the precise-price effects in retail markets, when the homes that are priced as precise price lower than the round price, or in particularly 9-ending prices, they tend to have fewer discounts and are sold at higher prices than homes that are priced in thousand-ending. It suggests that the precise price may help the seller with the negotiation process and the home sale. Second, if the sellers set the price higher than round prices (thousand-endings), the precise price strategy does not provide additional benefit in sale prices.

The rest of this chapter is organized as follows. In the next section, I introduce the sample set from Dallas Multiple Listing Service (MLS) database and provide summary statistics. In section 3, I present empirical evidences on of the price ending effects on sale prices from the regression discontinuity results and panel data model estimations. The last section is the conclusion.

4.2 The Data and Variables

The data set I use comes from Dallas multiple listing service(MLS) database. The multiple listing service is an actively managed home listing platform that enables real estate brokers share information on properties they have listed and invite other brokers to cooperate in their sale in exchange for compensation if they produce the buyer. MLS is the primary source of information of homes currently for sale. In current, about 90 percent of residential properties are listed and sold through over 800 MLSs nationwide and this data set represents the majority of home sales in the Dallas metropolitan area. A typical listing record in my data set contains information on the type of house, house address, list date, off-market date, listing price, final sale price and a selection of house characteristics including home size, lot size, number of bedrooms, number of bathroom, etc.. A majority of the homes are either listed with round price, or listed with 900 as ending. There is substantial amount of homes that are listed using other ending digits than these two (figure C.1). I separate the data into two groups, the homes that are listed using round price and the homes that are listed using precise price and provide the summary statistics in table 4.1.

Table 4.1: Summary Statistics by Price Ending

Price ending	**List price**	**Sale price**	**TOM**	**Bedrm**	**Bathrm**	**SQFT**
Round price	248246	234683	71.69	3.42	2.28	2305
	(312834)	(288287)	(67.88)	(.72)	(.80)	(1164)
Precise price	180080	172412	73.47	3.45	2.19	2168
	(108819)	(103234)	(66.18)	(.67)	(.62)	(809)

From above table, it is quite obvious that the homes that are listed using precise price are generally cheaper than homes that are listed with round price, indicating that the sellers may strategically set the price ending digit depending on home characteristics. Therefore a direct comparison of the final sale prices may not reflect a causal relation between price ending and sale price. I will discuss more about the identification strategies in the next section.

Sale frequency	**Count**	**Percentage**
1	263796	85.87
2	40930	13.32
3	2406	0.78
4	76	0.02
Total	307208	100.00

Table 4.2 reports the home sale frequencies in the data set. This data set spans from 2002 to 2008 and in total contains over 30 thousands home transaction records. About 87% of the homes are sold once during this period. For the rest 13% of the homes, we observe multiple transactions. These repeated sales provide a unique advantage that it enables us to specify a panel data model and control the unobserved home characteristics in analysis. I also presents the summary of home prices and characteristics in table **??**. It shows that the homes that are sold multiple times have similar list price, sale price and characteristics to the homes that are sold once during this period. The results from later panel data analysis may be valid for the whole sample.

Sale freq.	List price	Sale price	TOM	Bedrm	Bathrm	SQFT
1	200255	190761	74	3.4	2.21	2209
	(196666)	(182407)	(67)	(.69)	(.68)	(945)
2	213946	203785	67	3.45	2.24	2238
	(226567)	(208199)	(62)	(.69)	(.70)	(914)
3	236732	225637	64	3.47	2.28	2278
	(337933)	(322257)	(59)	(.71)	(.77)	(995)
4	184176	174651	83	3.41	1.98	2073
	(144377)	(139339)	(71)	(.62)	(.64)	(868)
Total	202360	192766)	72	3.44	2.22	2213
	(202438)	(187594)	(67)	(.69)	(.69)	(942)

4.3 Estimation Results

4.3.1 Identification Strategies

Conceptually, my identification strategy is to compare the sale outcomes of homes that are listed using precise price with the sale outcomes of homes that are listed using round price. But the direct comparison may suffer from the fact that cheaper homes tend to be listed at precise prices and I would mistakenly conclude that precise prices have detrimental effects on home sales.

The relation between precise price strategy and home value is more obvious in figure C.2 where the percentage of homes that uses precise price is negatively correlated with list price and in figure C.3 where the percentage of homes that uses round price is negatively correlated with home size (square footage).

I avoid this issue and adopt two alternative approaches. The first approach is the regression discontinuity method. Instead of comparing the average sale prices of all precise-priced homes and round-priced homes, I focus on narrower list price ranges and studies whether the precise-price home sales are different from round-priced home sales. Through this approach, I mitigated the confounding effects of unobserved home value on sale prices. The estimation details and results are discussed in the next subsection.

The second approach is to take advantage of the multiple sales in my data set and adopt panel data models in analysis. As I discussed in the data section, about 13% of the homes are sold multiple times in eight years and I am able to control for home unobserved factors in the panel data models. The source of identification would then come from the homes that are sold using different pricing strategies over this sample period.

4.3.2 Regression Discontinuity Results

In figure C.4, I plot the average price discount (defined as list price minus sale price) along the price ending line. For example, if the list price is 199,600 and the price discount is 8000 dollars, it will fall into the left part of the graph. If the list price is 200,100, it will fall into the right part of the graph. And if it is a round price (end in thousands), i.e., 201,000, it will fall exactly into the middle. The dots represent the average price discount within each price ending cell and the curves are quadratic fitted curves. Figure C.5 also presents similar analysis but instead use price discount percentage.

These two graphs present two findings. First, when the homes that are priced as precise price lower than the round price, or in particularly 9-ending prices, they tend to have fewer discounts and are sold at higher prices than homes that are priced in thousand-ending. It suggests that the precise price may help the seller with the negotiation process and the home sale. Second, if the sellers set the price higher than round prices (thousand-endings), the precise price strategy does not provide additional benefit in sale prices.

In table 4.3.2, I statistically estimated the discontinuity gap of the sale price

Table 4.2: Log Sale Price and Price Discount

	(1) log sale price	(2) Price discount in percentage	Observations
Bandwidth=1000	.01067***	-.58%***	193060
	(5.53)	(-10.19)	
Bandwidth=800	0.0271***	-0.82%***	151557
	(11.87)	(-12.49)	
Bandwidth=600	0.0273***	-0.91%***	146001
	(10.20)	(-11.75)	
Bandwidth=400	0.0255***	-0.95%***	139329
	(7.62)	(-9.88)	
Bandwidth=200	0.0173***	-1.10%***	132868
	(3.90)	(-7.76)	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Average treatment effect, regression adjustment for home characteristics

and percentage price discount using different bandwidths. The results show that the precise-price homes are on average sold at about 0.6% to 1.1% higher than the round-priced or above round-priced homes.

In figure C.5 to figure C.10, I estimate the precise-price effects on the price discount percentage at different price levels and in figure C.11 to figure C.15, I estimated the precise-price effects on the log sale price. The results are consistent to the main findings and are robust.

4.3.3 Panel Data Model Results

Although the discontinuity approach has provided a direct way of testing the precise-price effects on home sales, it may still suffer from the unobservable issues. Even within each price level, homes can still be substantially different in many aspects and analysis that fails to control for these covariates would be less sound.

As I discussed in the data section and in the identification stately section, I adopt the panel data feature to statistically estimate the precise price effects as robustness checks. I create two dummy variables. The first dummy variable is the precise price dummy that equals to one if the list price is a precise price and zero otherwise. The second dummy variable is the lower than round price dummy that equals to one if the precise price is lower than its corresponding thousand-ending price and zero otherwise. The regression results are reported in table 4.3.

Table 4.3: Log Sale Price, Price Discount and Price Ending

	(1) ln(SP)	(2) ln(SP)	(3) ln(SP)	(4) Price discount
Precise price	-0.0453***	0.0581***	0.00342	0.887%
	(-39.85)	(4.99)	(0.10)	(0.86)
Precise price×			0.0601***	-1.79%***
Lower than round price			(5.14)	(-5.01)
bedrooms	-0.121***			
	(-110.57)			
bathrooms	0.0990***			
	(76.71)			
Home size (sqft)	0.000518***			
	(475.30)			
Constant	11.07***	11.98***	11.98***	5.05%***
	(3719.86)	(1566.68)	(1567.19)	(21.67)
Observations	288096	42798	42798	42798

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The first column the table 4.3 is the results from pooled regression estimation of log sale price. I find that the precise price is negatively correlated with the sale price and it is because that the home value are negatively correlated with precise price strategy so that the regression picks up such negative relation. In the second and third column, I estimated a fixed effect panel data model. The coefficient on precise price become significant positive in the second column, suggesting that once I control the home value, the precise price have positive effect on the sale price. In the third column, when I include the interaction term of the precise price and lower than round price dummies into the regression, the interaction term became significant and the precise price dummy became insignificant. This result is consistent with my previous regression discontinuity analysis that the precise price strategy helps with the sale only when it is lower than the round price, or generally when it is a 9-ending price. The fourth column uses the price discount as dependent variable and it shows that the results are still robust.

4.4 Conclusion

In this chapter, I study how the price ending affect sales in housing market. Recent literature in behavioral economics and marketing science suggests that the price endings may have some phycological impacts on buyer's purchasing decision and sellers may strategically determine the price endings and help the sales.

I analyze a data set that contains multi-year home transaction in the Dallas Metropolitan area and find positively significant precise-price effects on home sales. The regression discontinuity analysis and fixed-effect panel data model analysis of the Dallas multi-year housing market transactions present two main findings. First, consistent with the precise-price effects in retail markets, when the homes that are priced as precise price lower than the round price, or in particularly 9-ending prices, they tend to have fewer discounts and are sold at higher prices than homes that are priced in thousand-ending. It suggests that the precise price may help the seller with the negotiation process and the home sale. Second, if the sellers set the price higher than round prices (thousand-endings), the precise price strategy does not provide additional benefit in sale prices.

Although the precise-price effects are significant and consistent with the findings in the retail markets in previous behavioral economics and marketing literature, they should be cautiously interpreted. The list prices in the real estate markets are different from the list prices in most retail markets that the former is a starting price for negotiation and the latter is a take-it-or-leave-it offer. We would require more theoretical and experimental analysis to understand how precise price affects the negotiation and the home sale.

REFERENCES

[1] James Albrecht, Pieter Gautier, and Susan Vroman. Directed search in the housing market. *Working Paper*, 2012.

[2] Michael Arnold. The principal-agent relationship in real estate brokerage services. *Real Estate Economics*, 20(1):89–106, 1992.

[3] Douglas Bernheim and Jonathan Meer. Do real estate brokers add value when listing services are unbundled? *Economic Inquiry*, 51(2):1166–1182, 2013.

[4] Peter Diamond. Unemployment, vacancies, wages. *American Economic Review*, 101(4):1045–1072, 2011.

[5] Li Gan and Qi Li. Efficiency of thin and thick markets. *Working paper*, 2004.

[6] Li Gan and Qinghua Zhang. The thick market effect on local unemployment rate flucturations. *Journal of Econometrics*, 133(127-152), 2006.

[7] Li Gan and Qinghua Zhang. The market thickness and the impact of unemployment on housing market outcomes. *NBER Working Paper*, (12134), 2013.

[8] Alessandro Gavazza. The role of trading frictions in real asset markets. *American Economic Review*, 101(4):1106–1143, 2011.

[9] Jonathan Gruber and Maria Owings. Physician financial incentives and cesarean section delivery. *Rand Journal of Economics*, 27(1):99–123, 1996.

[10] Donald Haurin, Jessica Haurin, Taylor Nadauld, and Anthony Sanders. List prices, sale prices and marketing time: An application to u.s. housing markets. *Real Estate Economics*, 38(4):659–685, 2010.

[11] Igal Hendel, Aviv Nevo, and Francois Ortalo-Magne. The relative performance of real estate marketing platforms: Mls versus fsbomadison.com. *American Economic Review*, 99(5):1878–1898, 2009.

[12] Joel Horowitz. The role of the list price in housing markets: Theory and econometric model. *Journal of Applied Econometrics*, 7(2):115–129, 1992.

[13] Chang-Tai Hsieh and Enrico Moretti. Can free entry be inefficient? fixed commissions and social waste in the real estate industry. *Journal of Political Economy*, 111(5):1076–1122, 2003.

[14] Thomas Hubbard. An empirical examination of moral hazard in the vehicle inspection market. *RAND Journal of Economics*, 29:406–426, 1998.

[15] Ken Johnson, Zhenguo Lin, and Jia Xie. Dual agent distortions in real estate transactions. *Real Estate Economics*, forthcoming, 2014.

[16] John Knight. Listing price, time on market, and ultimate selling price: Causes and effects of listing price changes. *Real Estate Economics*, 30(2):213–237, 2002.

[17] Vijay Krishna. *Auction Theory*. Academic Press, 2 edition, 2010.

[18] Edward Lazear. The impatient salesperson and the delegation of pricing authority. *NBER Working Paper No. 20529*, September 2014.

[19] Steven Levitt and Chad Syverson. Market distortions when agents are better informed: The value of information in real estate transactions. *The Review of Economics and Statistics*, 90(4):599–611, 2008.

[20] Kenneth C. Manning and David Sprott. Price endings, left-digit effects, and choice. *Journal of Consumer Research*, 36(2):328–335, 2009.

[21] Tridib Mazumdar and Purushottam Papatla. An investigation of reference price segments. *Journal of Consumer Research*, 37(6):246–258, 2000.

[22] Dale Mortensen. Markets with search friction and the dmp model. *American Economic Review*, 101(4):1073–1091, June 2011.

[23] Christopher Pissarides. Equilibrium in the labor market with search frictions. *American Economic Review*, 101(4):1092–1105, 2011.

[24] Jeff Prince, Vrinda Kadiyali, and Daniel Simon. Is dual agency in real estate a

cause for concern? *Real Estate Finance and Economics*, 48(1):164–195, 2014.

[25] R.C. Rutherford, T.M. Springer, and A. Yavas. Conflicts between Principals and Agents: Evidence from Residential Brokerage. *Journal of Financial Economics*, 76(3):627–665, June 2005.

[26] Robert Schindler and Patrick Kirby. Patterns of right most digits used in advertised prices. *Journal of Consumer Research*, 24(1):192–200, 1997.

[27] Mark Stiving and Russell Winer. An empirical analysis of price endings with scanner data. *Journal of Consumer Research*, 24(1):57–67, 1997.

[28] Manoj Thomas and Vicki Morwitz. Penny wise and pound foolish: The left-digit effect in price cognition. *Journal of Consumer Research*, 32(1):55–64, 2005.

[29] Manoj Thomas, Daniel Simon, and Vrinda Kadiyali. The price precision effect: Evidence from laboratory and market data. *Marketing Science*, 29(1):175–190, 2010.

[30] Wikipedia. S&p/case-shiller dallas home price index. *http://us.spindices.com/indices/real-estate/sp-case-shiller-tx-dallas-home-price-index*, October 2014.

[31] Wikipedia. Multiple listing service. *http://en.wikipedia.org/wiki/Multiple_listing_service*, October 2014.

[32] Abdullah Yavas and Shaiwee Yang. The strategic role of listing price in marketing real estate: Theory and evidence. *Real Estate Economics*, 23(3):347–268, 1995.

[33] Xi Zhao. The directed search model and market size effects – evidence from dallas housing market. *Working Paper*, 2014.

[34] Thomas Zorn and James Larsen. The incentive effects of flat-fee and percentage commissions for real estate brokers. *Real Estate Economics*, 14(1):24–47, 1986.

APPENDIX A

APPENDIX FOR CHAPTER 2

A.1 Proofs

In this appendix, I derive the optimality Conditions of the Directed Search Model. The seller's dynamic problem can be described in the following Bellman equation

$$\begin{aligned} V(a,b) = \max_{(a,b)}\{-c + \beta\{\sum_{k=0}^{\infty} p_k F(b)^k V + \sum_{k=1}^{\infty} p_k k a G^{(k-1)}(a)[1-F(a)] \\ + \sum_{k=1}^{\infty} p_k k b G^{(k-1)}(b)[F(a)-F(b)] + \sum_{k=1}^{\infty} p_k k \int_b^{\omega} y(1-F(y)) g^{(k-1)}(y) dy\}\}. \end{aligned} \tag{A.1}$$

In above equation, the seller sets list price a and reservation price b to maximize the value function. In each period, the house is either sold or unsold. The latter case happens if no buyer's visit the seller or one or more buyers visit the seller but make offers that do not meet seller's reservation price. If the house is unsold, it stays on the market for the next period. Under other circumstances, the house is sold to the highest bidder at a, b or the highest bid according to the game rule specified earlier. c in the Bellman equation denotes seller's waiting cost per search period and β is the time discount factor. p_k denote the probability that k buyers visit the seller in one period, and we assume that it follows poisson distribution, i.e., $p_k = Pr(n = k, \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$. Fix the number of bidders k, $G^{(k-1)}(x)$ and $g^{(k-1)}(x)$ denote the cdf and pdf of the highest value among $k-1$ draws. We assume that the buyers' valuations are independently generated from some distribution with cdf

$F(x)$ and pdf $f(x)$. From order statistics, we know that

$$G^{(k-1)}(x) = F(x)^{k-1}, \tag{A.2}$$

$$g^{(k-1)}(x) = (k-1)F(x)^{k-2}. \tag{A.3}$$

Therefore the Bellman equation can be simplified as follows.

$$\sum_{k=0}^{\infty} p_k G_1^{(k)}(b)V = Ve^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda F(b))^k}{k!} = Ve^{-\lambda} e^{\lambda F(b)}; \tag{A.4}$$

$$\begin{aligned}\sum_{k=1}^{\infty} p_k k a G^{(k-1)}(a)[1-F(a)] &= a(1-F(a)) \sum_{k=1}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} kF(a)^{k-1} \\ &= \lambda a(1-F(a))e^{-\lambda} \sum_{k=1}^{\infty} \frac{(\lambda F(a))^{k-1}}{(k-1)!} \\ &= \lambda a(1-F(a))e^{-\lambda} e^{\lambda F(a)};\end{aligned} \tag{A.5}$$

$$\begin{aligned}\sum_{k=1}^{\infty} p_k k b G^{(k-1)}(b)[F(a)-F(b)] &= b[F(b)-F(a)] \sum_{k=1}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} kF(b)^{k-1} \\ &= \lambda b[F(b)-F(a)]e^{-\lambda} \sum_{k=1}^{\infty} \frac{(\lambda F(b))^{k-1}}{(k-1)!} \\ &= \lambda b[F(b)-F(a)]e^{-\lambda} e^{\lambda F(b)};\end{aligned} \tag{A.6}$$

and

$$\begin{aligned}&\sum_{k=2}^{\infty} p_k k \int_b^{\omega} y(1-F(y))g^{(k-1)}(y)dy \\ &= e^{-\lambda} \int_b^{\omega} y(1-F(y)) \sum_{k=2}^{\infty} \frac{\lambda^k}{k!} k(k-1)F^{(k-2)}(y)f(y)dy \\ &= \lambda^2 e^{-\lambda} \int_b^{\omega} y(1-F(y)) \sum_{k=2}^{\infty} \frac{(\lambda F(y))^{k-2}}{(k-2)!} f(y)dy \\ &= \lambda^2 e^{-\lambda} \int_b^{\omega} y(1-F(y))e^{\lambda F(y)} f(y)dy.\end{aligned} \tag{A.7}$$

Replacing expressions (A.4) through (A.7) into the value function (A.1), we get

$$(b+c)\beta^{-1} = be^{-\lambda}e^{\lambda F(b)} + \lambda ae^{-\lambda}e^{\lambda F(a)}[1-F(a)] \\ + \lambda be^{-\lambda}e^{\lambda F(b)}[F(a)-F(b)] + \lambda^2 e^{-\lambda}\int_b^{\omega} y[(1-F(y)]e^{\lambda F(y)}f(y)dy. \tag{A.8}$$

Furthermore, the first order condition of above equation with respect to a gives us the optimal list price,

$$bf(a) + \{\lambda af(a)[1-F(a)] + 1 - F(a) - af(a)\}\, e^{\lambda[F(a)-F(b)]} = 0. \tag{A.9}$$

Finally, the pair $\{a, b\}$ that satisfies conditions (A.8) and (A.9) is seller's optimal list price and reservation price.

A.2 Figures

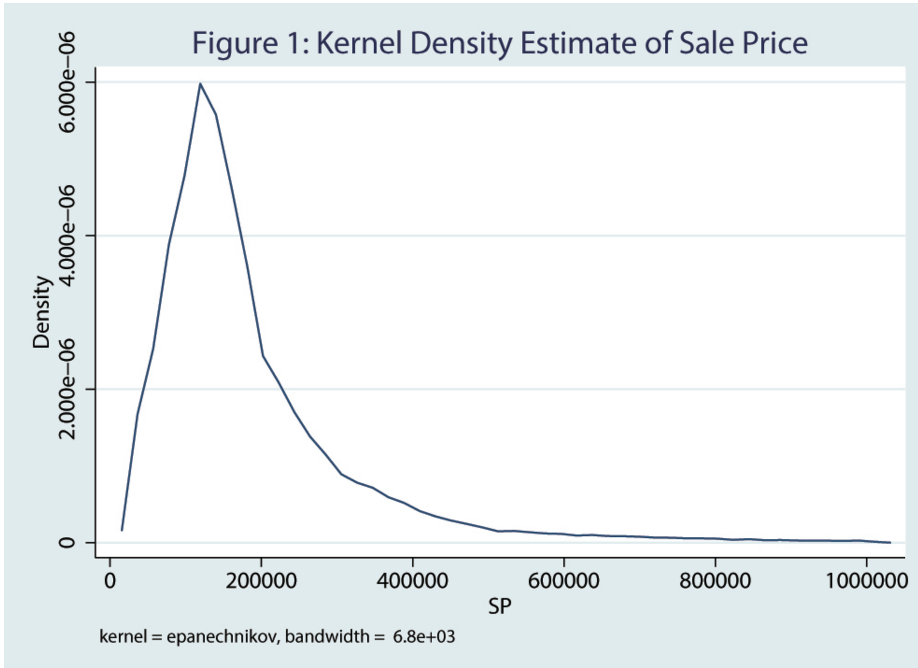

Figure A.1: Kernal Density Estimation of Sale Price

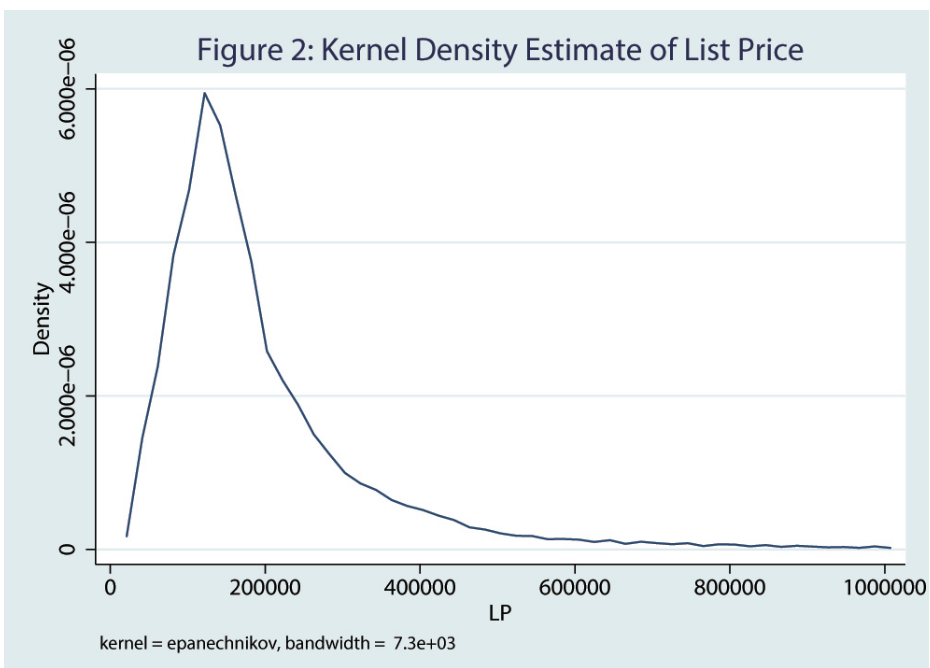

Figure A.2: Kernal Density Estimation of List Price

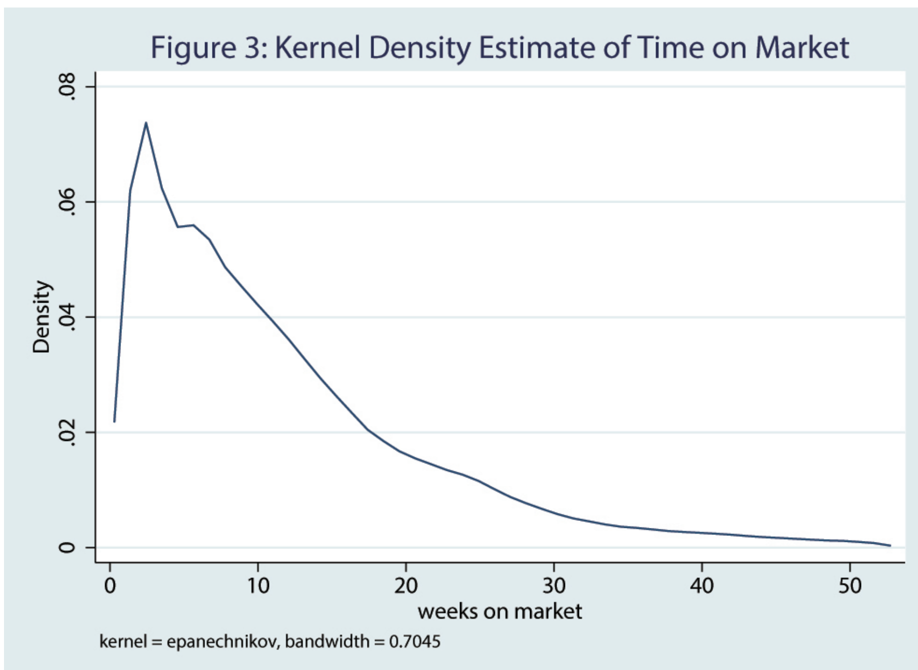

Figure A.3: Kernal Density Estimation of Time on the Market

APPENDIX B

APPENDIX FOR CHAPTER 3

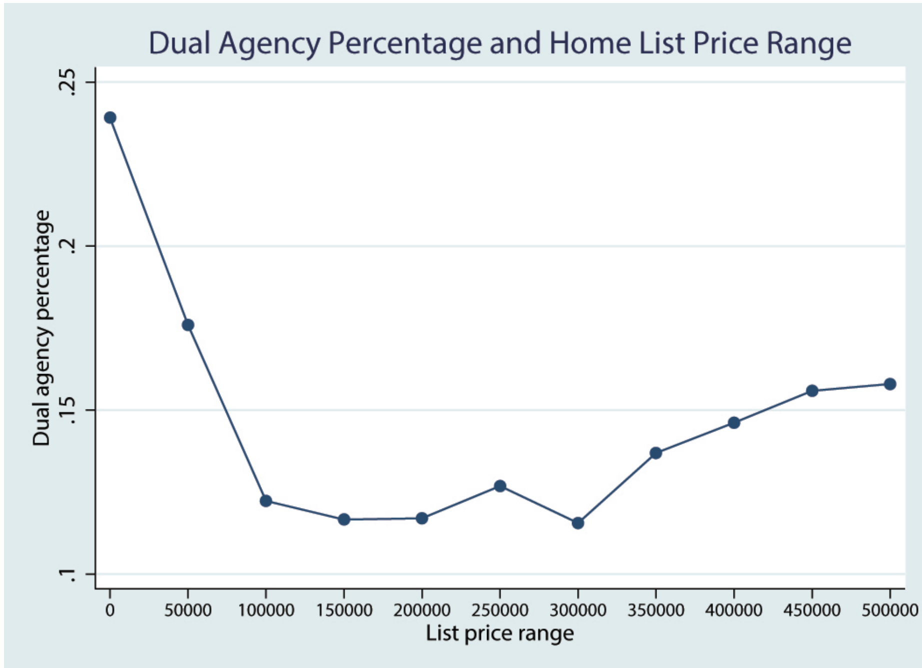

Figure B.1: Dual Agency and Home List Price

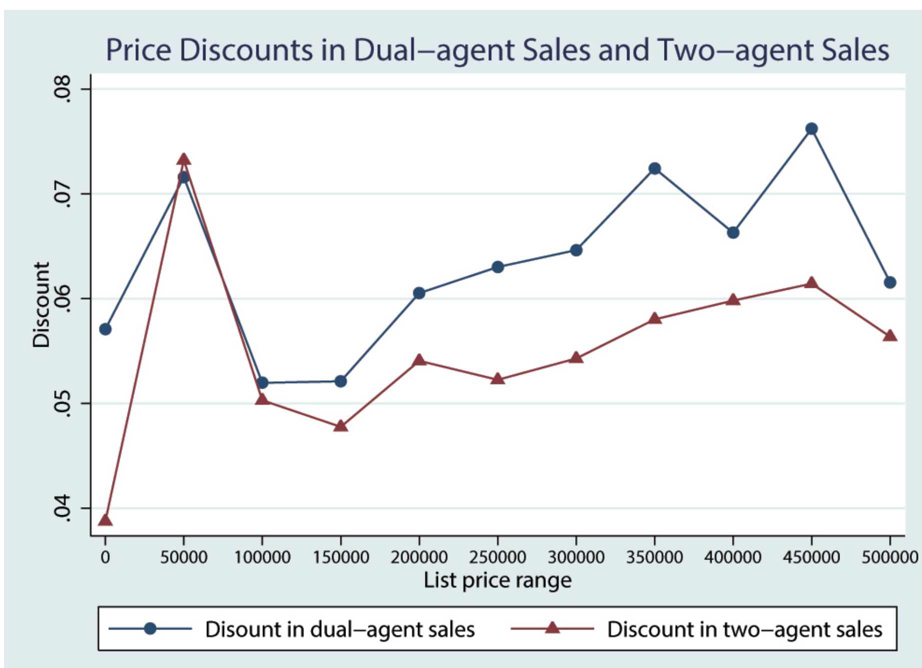

Figure B.2: Price Discounts and Agency Structure

APPENDIX C

APPENDIX FOR CHAPTER 4

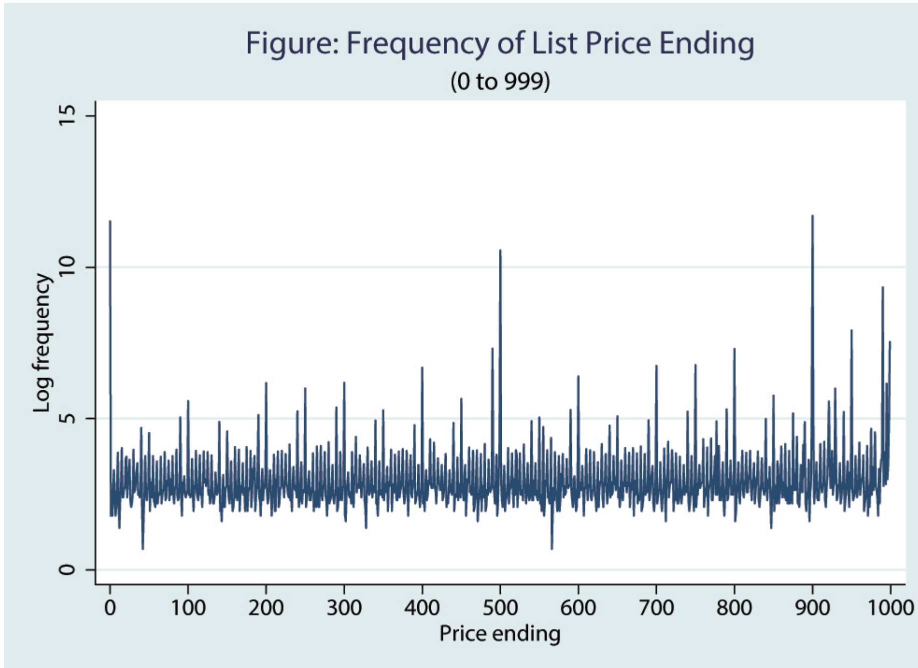

Figure C.1: Frequency of price endings

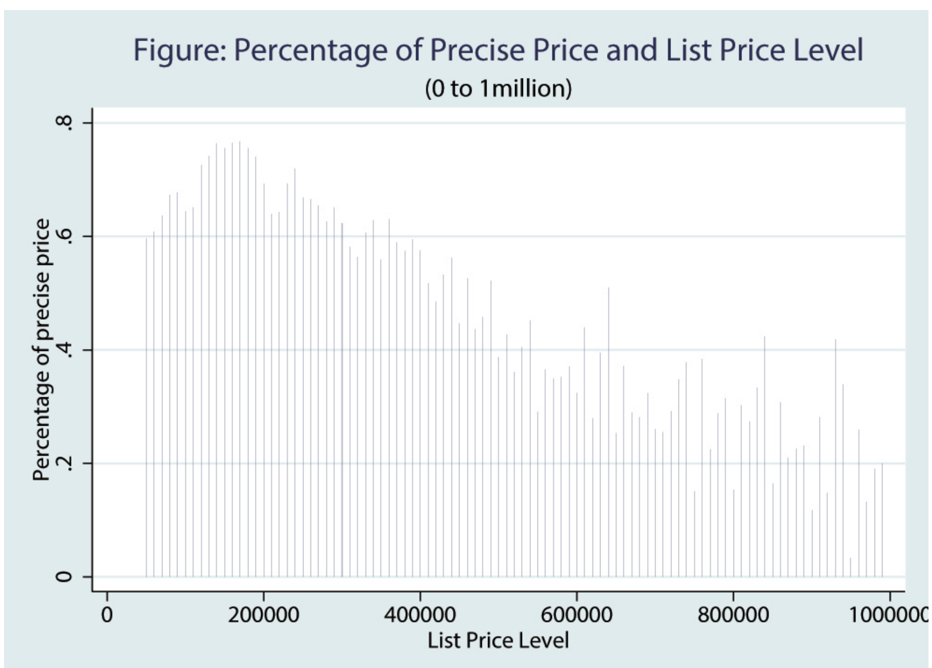

Figure C.2: Home list price and precise price

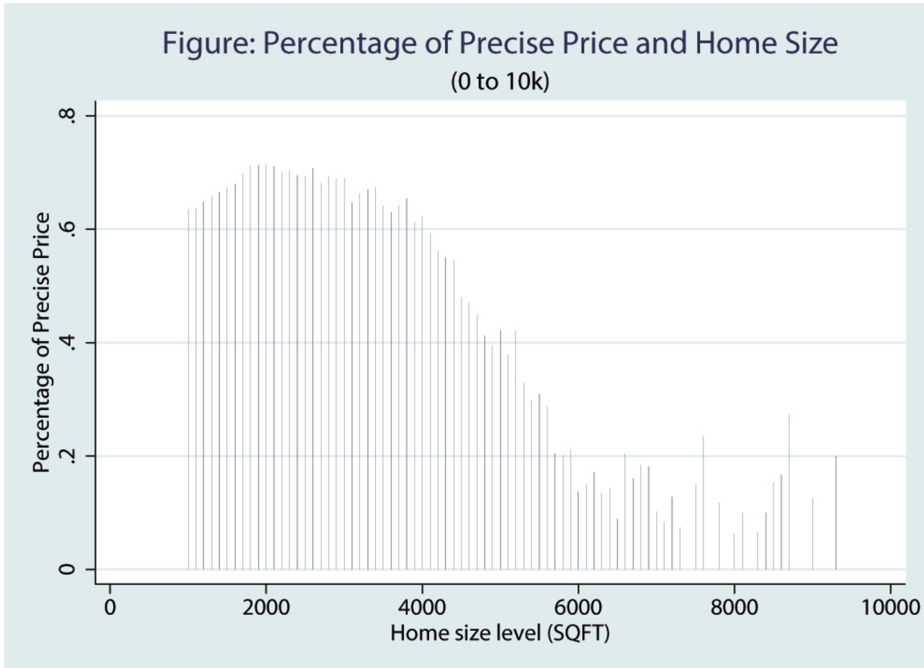

Figure C.3: Home size and precise price

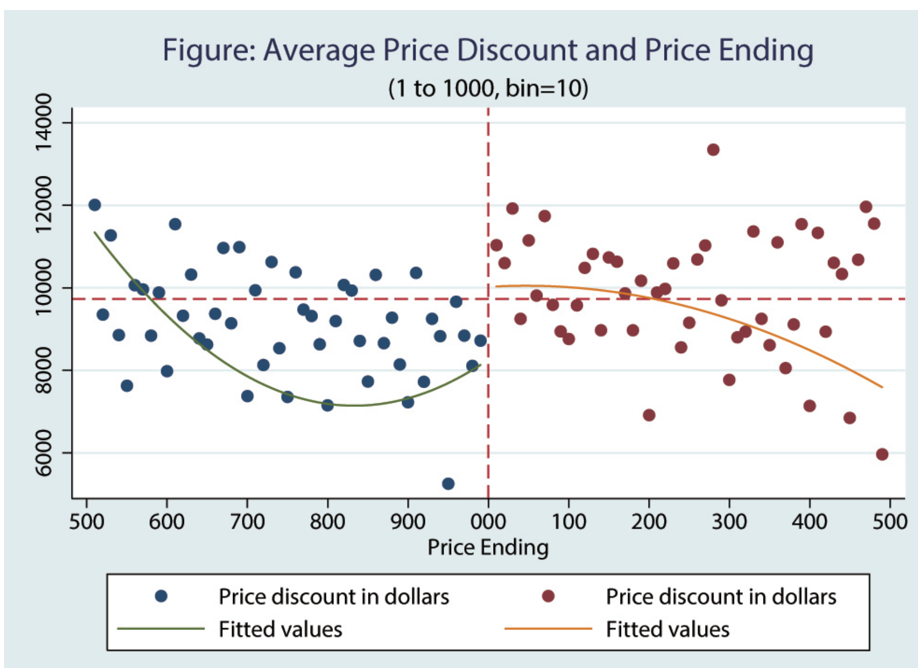

Figure C.4: Discontinuity in Price Decrease: Heterogeneous precise price effects

Figure C.5: Discontinuity in Sale-price-list-price ratio: Heterogeneous precise price effects

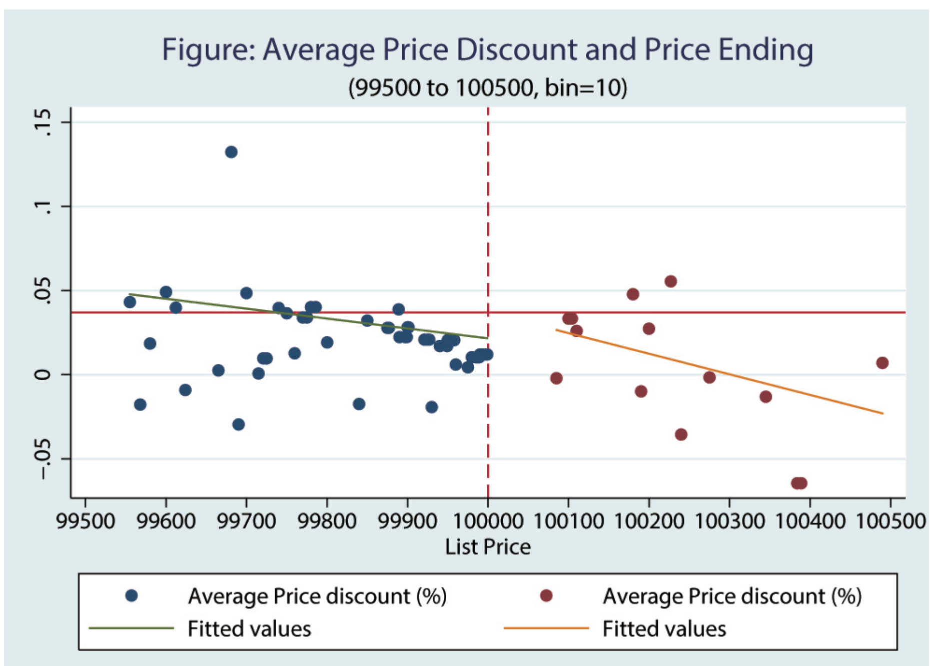

Figure C.6: Discontinuity: Heterogeneous precise price effects, $100k

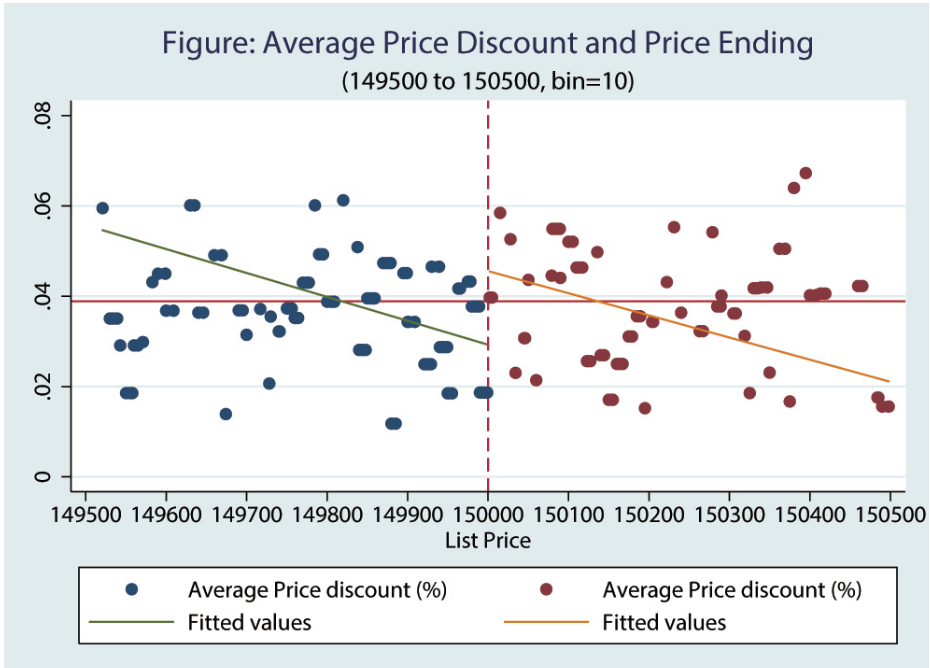

Figure C.7: Discontinuity: Heterogeneous precise price effects, $150k

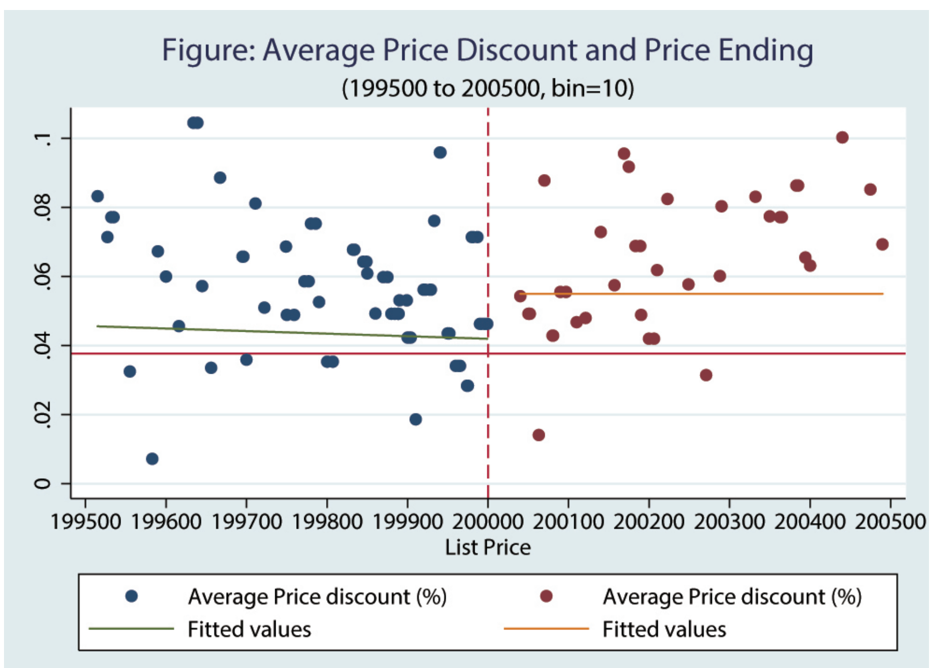

Figure C.8: Discontinuity: Heterogeneous precise price effects, $200k

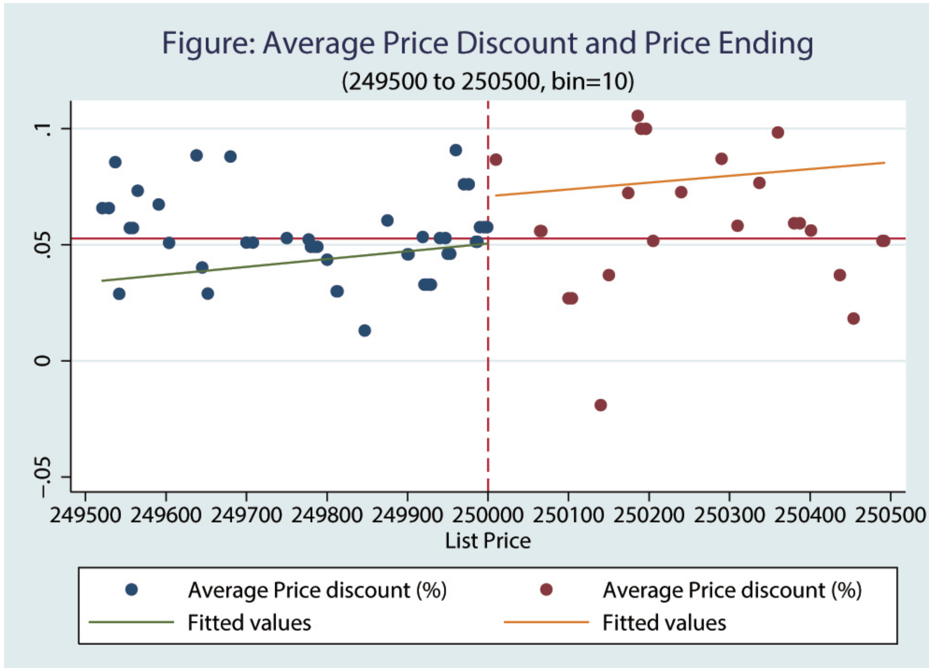

Figure C.9: Discontinuity: Heterogeneous precise price effects, $250k

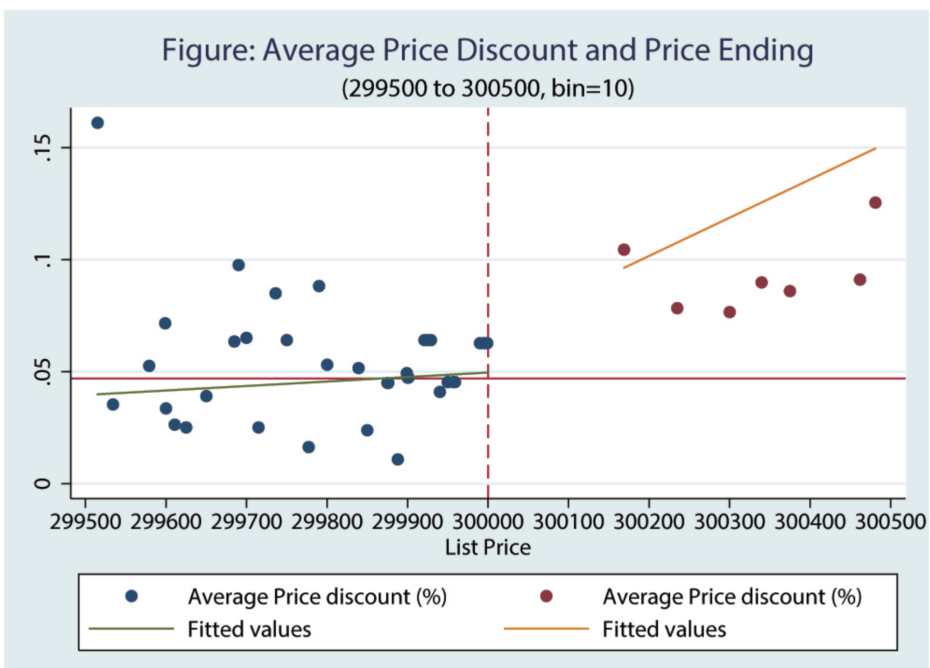

Figure C.10: Discontinuity: Heterogeneous precise price effects, $300k

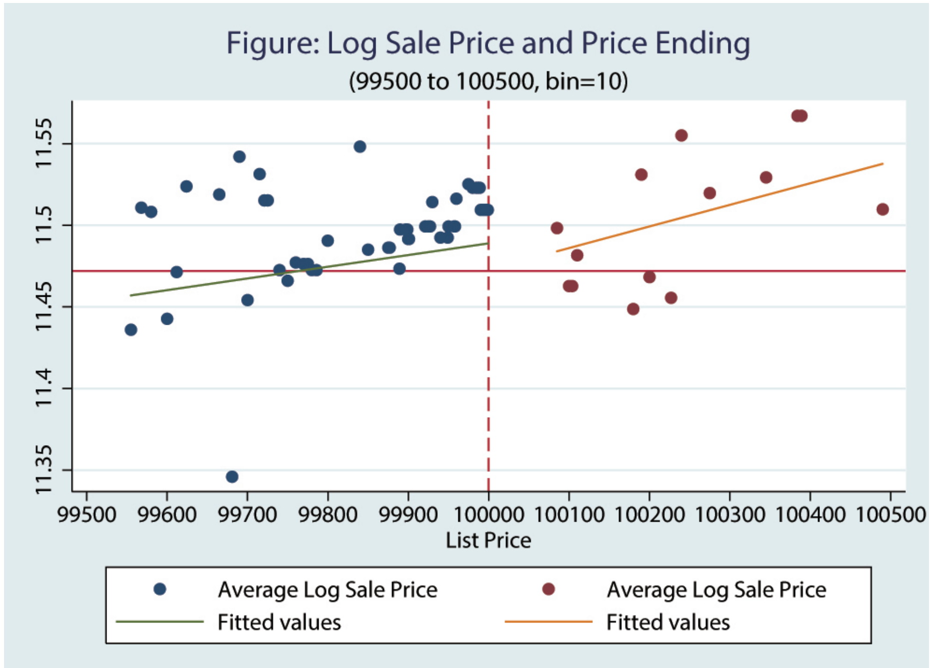

Figure C.11: Discontinuity: Heterogeneous precise price effects, ln(sale price) $100k

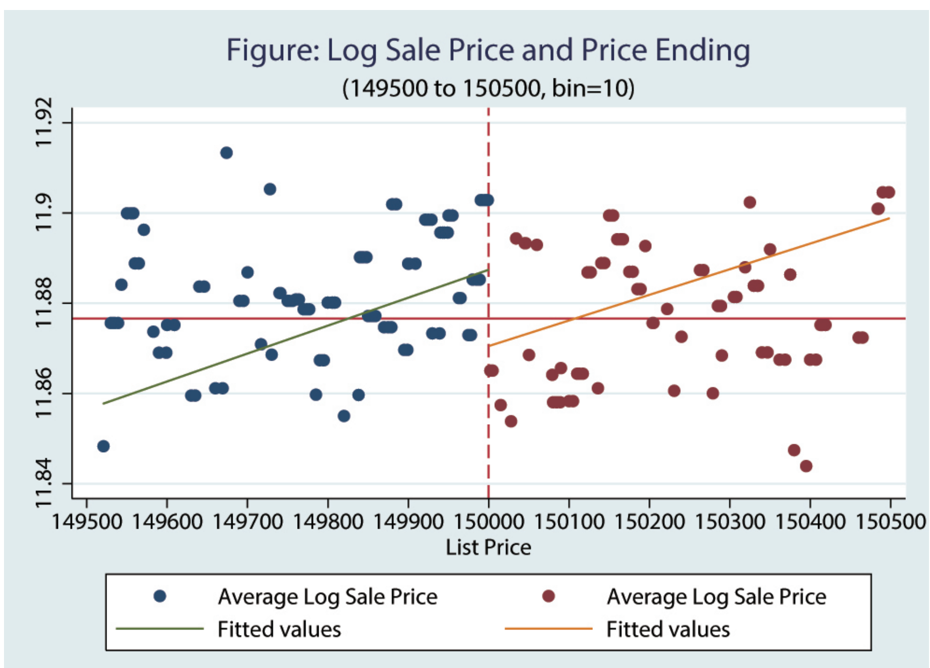

Figure C.12: Discontinuity: Heterogeneous precise price effects, ln(sale price) $150k

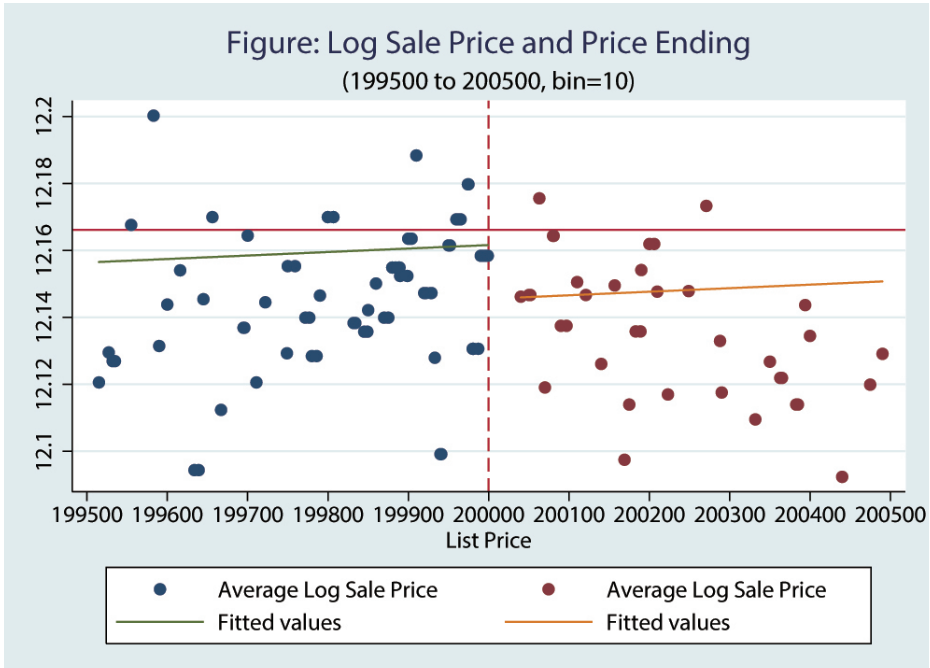

Figure C.13: Discontinuity: Heterogeneous precise price effects, ln(sale price) $200k

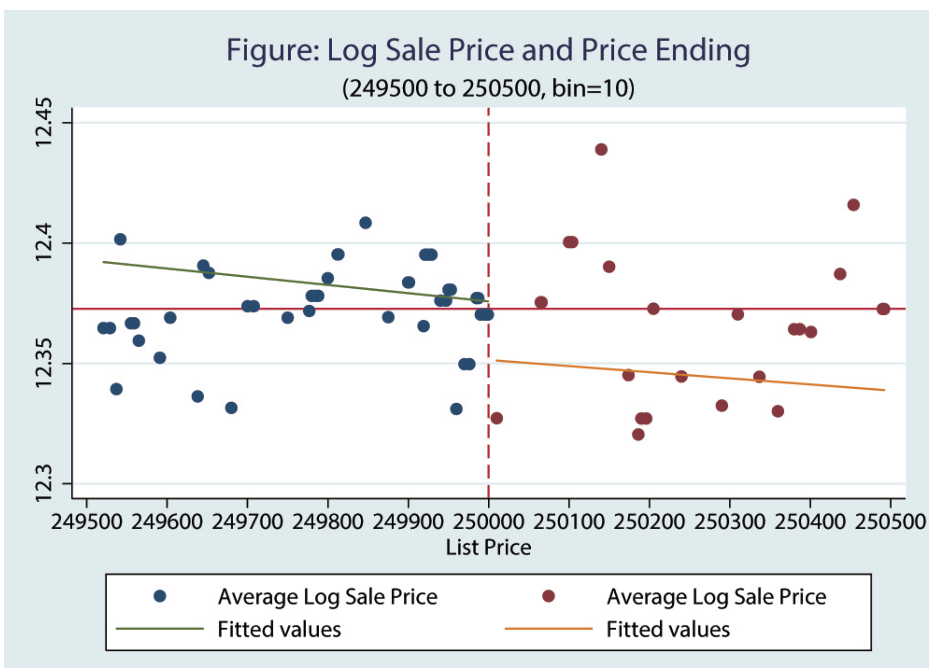

Figure C.14: Discontinuity: Heterogeneous precise price effects, ln(sale price) $250k

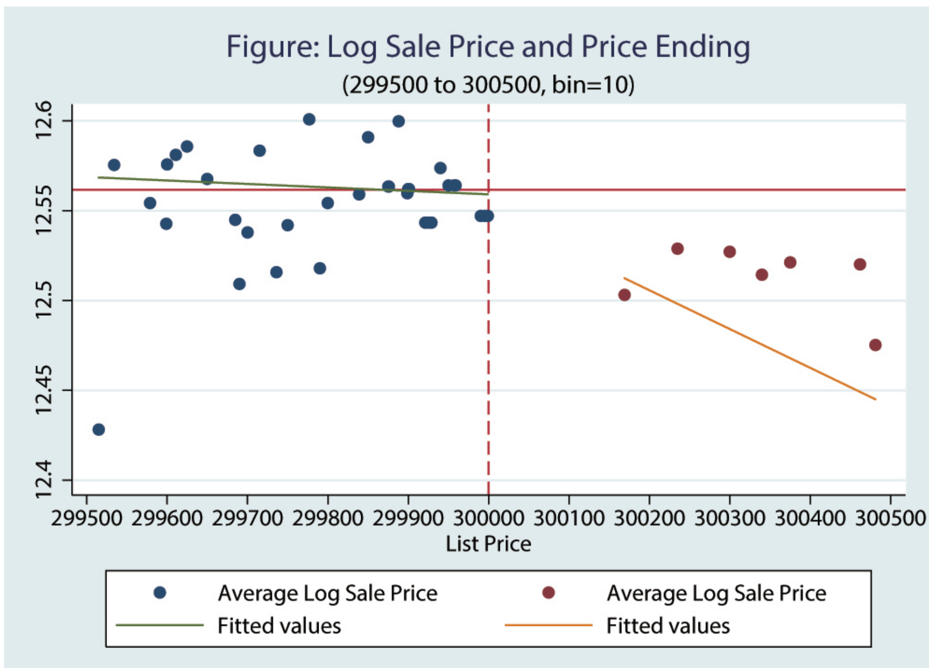

Figure C.15: Discontinuity: Heterogeneous precise price effects, ln(sale price) $300k