The Dissertation Committee for Seyed Parviz Gheblealivand
certifies that this is the approved version of the following dissertation:

Essays in Entry and Exit, Social Inefficiency and Commission Rates in Housing Market

Committee:

Kenneth Hendricks, Supervisor

David S. Sibley

Eugenio J. Miravete

Thomas E. Wiseman

Randal B. Watson

Romana Khan

Essays in Entry and Exit, Social Inefficiency and Commission Rates in Housing Market

by

Seyed Parviz Gheblealivand, B.S., M.S., MBA

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2010

To:

MIR MASOUD, HAVA,

FIROOZ AND MARZIEH.

Acknowledgments

For all three papers, I am heavily indebted to Kenneth Hendricks for his advice and supports. For the first and second papers I would like to specially thank Eugenio Miravete for fruitful discussions. I am also very grateful to Mostafa Beshkar, Dean Corbae, Jim Gaines from Texas A&M Real Estate Center, David Sibley, Randal Watson, Tom Watson from TREC, Tom Wiseman and seminar participants in University of Texas at Austin, Bank of Canada, University of Wisconsin- Milwaukee, University of Wisconsin- Parkside and The Brattle Group. For the third paper, I am grateful to Tom Wiseman. Special thanks goes to Javad Yassari.

SEYED PARVIZ GHEBLEALIVAND

University of Texas at Austin

May 2010

Essays in Entry and Exit, Social Inefficiency and Commission Rates in Housing Market

Publication No. ____________________

Seyed Parviz Gheblealivand, Ph.D.
The University of Texas at Austin, 2010

Supervisor: Kenneth Hendricks

In the first paper, using a dataset of the records of Texas Real Estate Agents, I reexamine the findings of Hsieh and Moretti (2003) regarding the inefficiency of free entry in real estate industry: first, I point out one important source of misidentification in that paper's analysis of the relationship between home prices and the number of real estate agents in a city. This misidentification stems from not including the ratio of houses sold in a city to its labor force size as an explanatory variable. Failure to account for this variable will result in inflated coefficient for the effect of home prices on the percentage of real estate agents in a city's labor force. Second, I analyze the effect of home prices on productivity of real estate agents. Empirical evidence supports theory prediction of inverse relationship between home prices and productivity of its real estate agents (measured as the number of houses sold per agent) and the empirical results in Hsieh and Moretti (2003). Third, I investigate the relationship between the extra wages of real estate agents (defined as average

earning net of agents' outside option) and home prices in a city. In support for free entry, I find no evidence of any such relationship.

In theory, free entry potentially leads to social inefficiency. This paper finds strong empirical evidence consistent with excess entry into Texas Residential Real Estate Brokerage Industry and studies the effects of heterogeneity and future uncertainty on such inefficiencies. I develop a dynamic model of entry and exit with heterogeneous agents and modify the predictions of the earlier literature. I show that the heterogeneity among (real estate) agents results in a weaker relationship between the real estate commission fees and the number of real estate agents. I also show that the models developed for static cases in the previous papers are special cases of the more general model in this paper.

The model allows us to explain the lower business stealing effect compared to static and homogeneous models that is observed in the data. To address the issue of excess entry, I separate the business stealing effect from demand driven entry and find that on average 75 percent of entry is due to business stealing. To evaluate free entry, I control for agents' outside options and find that the extra wages of the real estate agents do not vary with housing prices.

The objective of the third paper is to study the determinants of commission rates in the two-sided market of real estate brokerage industry and explain the emergence of the MLS and its impact on commission rates. In addition to their commission rates, real estate agencies decide on their MLS policies as well: they can either list the property with the MLS and share information about it, or not list the property with the MLS. If a property is listed with the MLS, all MLS subscribers can see the listing and send their potential buyers to see that property. Potential buyers can go to any agency to purchase such a property. If the property is *not* listed with the MLS, to buy a house, a buyer must go to the same agency that the seller has signed up with.

Since sellers pay the commission fees, and buyers no longer have to go to the same agency, with MLS listing, buyers choose the closest agency regardless of the commission rates charged by the agencies. Therefore, changes in the commission rates only change the affiliation of the sellers and not that of the buyers. This leads to a softer competition under MLS listing as agencies compete only in the seller side of the market. The softer competition and resulting higher commission rates are desirable to the agencies. They prefer the MLS listing outcome and given the optimal strategies after observing each other's listing decisions, agencies weakly prefer listing to no listing. I show that the one period game has two Nash Equilibria in which either both real estate agencies choose to list their houses with the MLS, or both decide not to list their houses with the MLS. The no listing equilibrium forces buyers to work through that agency's agents and effectively ties the both sides of the market. The higher commission rate equilibrium of the game allows buyers to choose either agency and reduces the competition to the sellers side. Softer competition in turn, results in higher equilibrium commission rates and higher profits along the equilibrium path.

Table of Contents

List of Tables

List of Figures

Chapter 1

Can Free-Entry be Inefficient: Revisiting Hsieh and Moretti

1.1 Introduction

I use a new dataset from Texas Real Estate Market to replicate the analysis of Hseih and Moretti (JPE, 2003) in examining the efficiency of free entry by real estate agents.

Commission fees from the houses sold are the sources of income for the real estate agents. Higher prices along with a fixed commission rate create more commission fees and make the option of becoming a real estate agent more attractive. This would lead to a higher number of real estate agents as well as higher percentage of these agents among a city's labor force.

With fixed commission rate and free entry, even with the same number of houses sold, higher home prices lead to a more people to become real estate agents. In comparison between two cities that are identical except for their home prices, the one with higher housing price will attract more real estate agents as the total commission to be divided is larger, and therefore, will have a higher percentage of its workforce as real estate agents. An immediate corollary is that in high home price cities, each agent sells a lower number of houses, in other words, they are less ”productive”. Hsieh and Moretti (2003) find strong empirical support for these two claims in a dataset consisting of 282 Metropolitan Statistical Areas in 1980 and 1990 censuses. I adopt the same theoretical framework and use the dataset of Texas Real Estate Agents Record to evaluate their claims empirically.

I point out one important source of misidentification in their empirical analysis as they do not include the (normalized) number of houses sold in a city with its housing prices as an explanatory variable. Failure to account for this variable creates inflated coefficients for the effect of home price on the number of real estate agents and fails to meet the prediction in the theory. I show that once this issue is addressed, the coefficients get closer to theory prediction.

I find a very similar result as in Hsieh and Moretti (2003) for the relationship between agents’ productivity and housing prices in city. With higher prices, since there are more agents, fewer houses are needed to cover agents’ outside options. Therefore, in equilibrium, productivity of agents (defined as the number of houses sold per agent in a city) is lower in higher housing price cities. Empirical support is obtained for this result in Hsieh and Moretti (2003) and is confirmed in this paper. Similar to Hseih and Moretti (2003) I find that one percent increase in home prices in a city would lead to two thirds of percent drop in agents’ productivity, a result which is almost identical to that in Hsieh and Moretti paper.

The third claim in Hsieh and Moretti (2003) is to evaluate the free entry condition. In particular, a necessary condition for free entry is to have no extra wages for being a real estate agent. Hsieh and Moretti (2003) and this paper find strong evidence in support of this argument. Extra wages are defined as average agent's earning net of her outside option.

Assuming a fixed commission rate, changes in the amount of total commissions in a city can come from two factors: 1. It can either be due to a change in the number of houses sold, or 2. Because of a price change. Although both factors result in higher total commissions which in turn would be followed by higher number of real estate agents; they have totally different welfare implications. If total commissions increase only because of a price hike, the new (larger) population of real estate agents still provides brokerage service for the same number of houses. In other words, with the same number of houses sold in a city, the new real estate agents would have to steal business from the existing ones. This would result in lower number of houses sold per agent (which is my measure of agents' productivity). On the other hand, if total commissions increase because of higher number of houses sold in a market, this demand driven increase in the number of agents reflects no welfare loss to the society as the ratio of number of houses sold per agent does not change. Using the number of houses sold per agent as the measure of productivity, I differentiate between these two explanations of higher number of agents. I estimate that on average, business stealing effect accounts for about 65% of new entry into the real estate brokerage industry.

To evaluate the free entry condition, I construct the variable of "extra wage" as the average earning of real estate agents in a city net of their outside options. If entry to the real estate industry is free, then two conclusions can be drawn. First, the earning of the marginal type in the real estate industry should be equal to the marginal agents' outside options. Second, Controlling for outside options, agents' earning should not differ across cities; in particular, their extra wage should not depend on housing price. Data limitations only allow us to test the second prediction. I find strong support for this hypothesis over the cross sections in our data.

1.1.1 Empirical Literature on Social Wastes of Free Entry

A known property of perfect competition is the efficiency of its equilibrium outcome. The implication of this property is that if other conditions of perfect competition are met, any barrier to entry is harmful from the welfare standpoint.

However, if in a market firms do not behave ”competitively” (for example when they are not price takers anymore), total surplus could be improved by imposing some entry restriction into that market.[1] Social inefficiency of free entry can come from two sources:

1. The business stealing effect. In a market where new and incumbent firms' products are close substitutes, in addition to causing a potentially lower price, a new entrant can steal business from the incumbents. Bank regulations in Portugal in 1990s to limit the opening of new branches were partly aimed to prevent business stealing among rival banks.

2. When average costs are decreasing in output, a new firm will further move away the existing ones from their efficient output level. This is one of the differences between monopolistic competition and perfect competition mar-

[1] See Cabral (2000)

kets.[2]

Although potential inefficiency of free entry is well known in the theory; there are relatively few empirical studies to document this issue. Berry and Waldfogel (RAND, 1999) offers the first insight into estimating the social waste of free entry in radio broadcasting industry. Firms in this model are symmetric and play a static game. In this paper they estimate the fixed cost and demand function parameters and concludes that the welfare loss is 45% of revenue. Even with solid evidence of social inefficiency of free entry, the actual entry and exit rarely happens in their data due to FCC regulations.

Hsieh and Moretti (JPE, 2003) provide an analysis of social waste of free entry in real estate brokerage industry.

There are two interesting properties that make the study of real estate brokerage industry desirable. First, if acquiring a real estate license is perceived as entry, then the barriers to enter the real estate brokerage industry are relatively low. One only needs to attend a handful of classes and take an exam to be eligible for a real estate license. Notice that unlike physicians or attorneys board exam, real estate agency exams are not given by incumbent agencies and therefore, cannot be used to limit entry.

The second property is the seemingly unchanging commission rate for the real estate brokerage services. This rate is generally around 5 to 7 percent of the

[2]The extreme case in which free entry leads to excess number of firms is the standard circular city with zero marginal and positive fixed costs.

value of the property. I discuss this property in more detail in the coming section. Hsieh and Moretti (JPE, 2003) and Han and Hong (2008) also present historical evidences supporting persistency of the so called 6% commission rate over time, property values and in different cities. Fix commission rate implies that if there are two cities with the same number of houses for sale, the one with higher housing prices will attract more agents. Therefore, houses sold per agent will be lower in number in the high housing price city.

Hsieh and Moretti (2003) predicts that with fixed commission rates: 1) cities with higher housing prices would have more real estate agents; 2) agents in cities with higher home prices are less productive, in the sense that they sell fewer houses per; and 3) conditional on agents' outside options, due to free entry, agents' real wage should be independent of home prices and the same across cities. Their model studies the business stealing effect in real estate brokerage industry and abstracts away from the cost inefficiency.

Han and Hong (2008) highlight that Hsieh and Moretti (2003) suffers from the lack of cost inefficiency analysis. They show that the results of Hsieh and Moretti (2003) can be obtained using a simple model with no cost inefficiency. Focusing on this aspect of inefficiency, Han and Hong (2008) impose some structure on cost function and try to estimate the related parameters. They point out three sources of cost inefficiency as wasteful non-price competition, loss of economies of scale, and high fixed costs. They find direct evidence for all three. In particular, they "find that a one-standard-deviation increase in entry rate would increase the average variable costs by 28.96%, resulting from wasteful non-price competition."

Hsieh and Moretti (2003) starts with a static free entry condition with homogeneous agents. I show that the estimate of the coefficient of the effect of housing price on proportion of real estate agents in a city's workforce bears significant endogeneity. Failure to include the number of houses sold in a city and normalizing just one side in the free entry condition are the main causes of endogeneity. Using our dataset, we specify the correct relationship between the (normalized) number of real estate agents and (normalized) total available commission in cross sections of cities and find the coefficient close to theory prediction.

Another source of possible endogeneity is the relationship between the agents' outside options and average wage in a city. Both in Hsieh and Moretti (2003) and in my model, I assume that the two variables are exogenous. However, it is not hard to think of plausible scenarios in which the higher average wages in a city (the variable that Hsieh and Moretti (2003) and I use as the agents' outside options) affects the average housing prices in a city.

1.2 The Real Estate Brokerage Industry

1.2.1 Free Entry

As mentioned in the previous section, an interesting characteristic of the real estate brokerage industry is free entry. If entry is conceived as obtaining a real estate license, there are limited barriers to becoming an agent. To acquire a real estate salesperson license, one needs to pass a handful of real estate related courses and take the state license exam which is given multiple times a year.

The requirements are similar in other states. In addition, there is a large bulk of ”inactive” real estate agents who are ready to enter the market at any time.

1.2.2 Heterogeneity Among Real Estate Agents

Another property of the real estate brokerage industry is the great deal of heterogeneity among agents. Table (1.1) shows the distribution of income among real estate salespersons and brokers. Both 90/10 ratios[3] are significant and suggest that the homogeneity of real estate agents is not a very plausible assumption. The ratio however, is larger for salespersons than in brokers. Hsieh and Moretti (2003)

	Hourly Wage					
	Mean	10th	25th	Median	75th	90th
	\$/hr	\$/hr	\$/hr	\$/hr	\$/hr	\$/hr
Real Estate Brokers	44.71	28.69	30.57	33.72	48.31	>70
Real Estate Salespersons	30.91	6.73	10.47	17.71	44.43	>70

Table 1.1: Hourly wage of Real Estate Brokers and Salesperson in Austin, TX, May 2006. **Source:** US Bureau of Labor Statistics (BLS).

assume a continuum of homogeneous agents. In chapter 2 of this dissertation, I develop a heterogeneous model to evaluate the effect of changes in housing prices on entry of real estate agents. I point out the effect of heterogeneity on the results in next chapter and how heterogeneity modifies the results of Hsieh and Moretti (2003).

[3] The ratio of 90th percentile of earning to 10th percentile

1.2.3 Fixed Commission Rate

The other interesting characteristic of the real estate market is that the commission rate charged by the real estate agents is relatively fixed at around 6% of the value of the property. Figure (1.1) shows that the national median real estate commission rate is relatively constant over the period of 1991 to 2005. The median commission fees received by the realtors increases in the same period. For the purpose of this paper, I assume that the commission rate for the real estate brokerage service is constant.

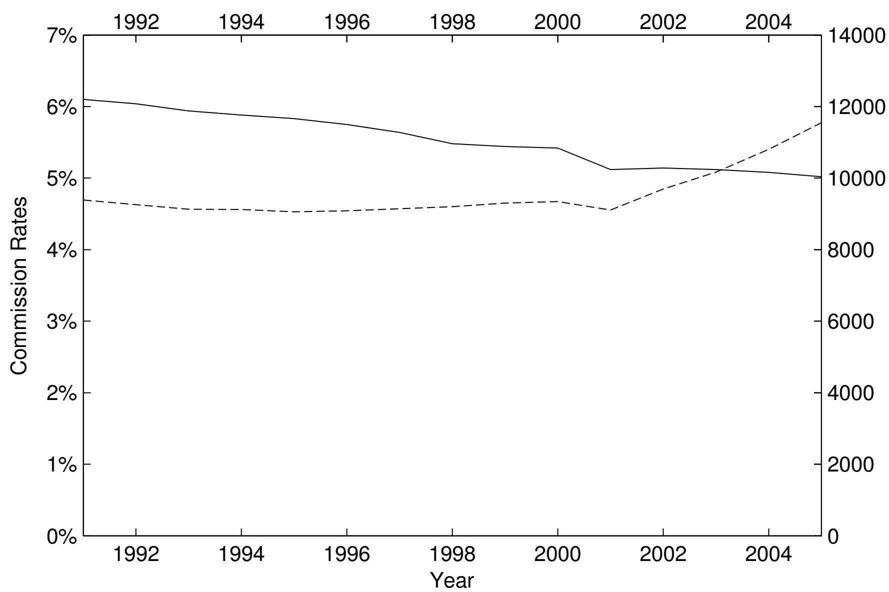

Figure 1.1: Median Commission Rates and median per house Commission Fees between 1991-2005. (**Sources:** Commission rates are from REAL Trends 500 ©; median home prices are from U.S. Department of Housing and Urban Development, U.S. Housing Market Conditions, 4th Quarter 2006, Tables 6-9 (Feb. 2007), and are a weighted average of new and existing home prices, based on annual sales; median home prices are converted into 2006 dollar; commission fees are calculated by multiplying commission rates by median home prices.

The persistence of the 6% commission rate and its widely perceived high level (specially when compared with commission rates in other industrialized countries) is a huge puzzle on its own. One possible explanation for the question of why the commission rates show this level of rigidity is the collusion between the real estate agencies. Real estate brokerage industry is a two sided market in which firms provide a platform that matches buyers with sellers. Sellers pay the commission fees. Buyers go to the closest agency when firms share their properties on the Multiple Listing Services (MLS) and the changes in the commission rates can only affect the sellers side of the market. This softens the competition in the commission rates as the buyers side is split between the agencies evenly regardless of the commission rates charged. This softens the competition between real estate agencies and allows them to capture a higher level of profits. In fact agencies use MLS and information sharing as commitment device to limit competition to only one side of the market.

The second explanation for the 6% commission rule becomes clear by looking into the roles of different types of real estate agents: salespersons and brokers. Salespersons are known as the ”foot soldiers” in the real estate brokerage industry. They do not have the authority to post in the MLS, and therefore set any commission rate. Brokers on the other hand, list the properties, oversee the activities of salespersons and set the commission rates for all transactions. Figure (1.2) shows the number of real estate agents in the two metropolitan areas in our sample. Although the number of real estate agents have experienced a very sharp increase during the period studied in this figure, the (absolute) number of real estate **brokers** remain constant. This trend is similar in other cities in our dataset and the data I collected

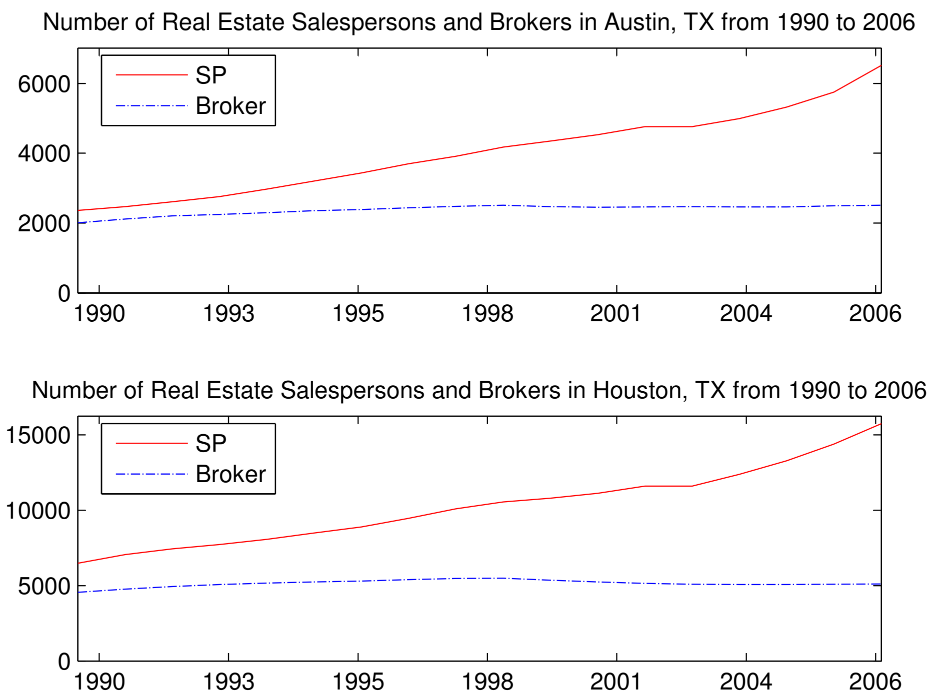

Figure 1.2: Number of real estate salespersons and brokers in Austin,TX and Houston, TX

from other states. Since the number of brokers determines the number of competitors in the real estate brokerage industry and this number is unchanged, there is no reason to change the commission rate because the real concentration of the market remained unchanged.

The third possibility is that the fixed commission rate simply reflects different elasticities of demand for real estate brokerage services. After all, it is not unrealistic to assume that the owner of a $500,000 house is willing to pay more for the brokerage service than someone who has a house valued at $100,000. However, in order for this to result in a fixed commission rate, the differences in elasticities have to be unrealistically enormous.

In any case, although the fix commission rate presents an interesting question, I abstract from the causes of this phenomenon and focus on the implications of the fixed commission rate. Chapter 2 of this dissertation develops a model in which information sharing is used as a collusive device between real estate agencies.

1.3 Data

The dataset used in this paper is the result of merger of three datasets: 1. The data on real estate agents; 2. Data on home prices; and 3. Data on wages.

1.3.1 Data on Real Estate Agents

The source of real estate agents data in this paper is the Texas Real Estate Commission's (TREC) records on real estate licensees. This dataset includes information on licenses of all real estate agents who were active at any point between 1980 and October 2007 as salespersons and brokers.

In this paper, I focus on real estate salespersons rather than brokers or real estate agencies. Real estate brokers need a different and longer formal training as well as minimum of few years of experience as salesperson to be eligible to receive broker license. I therefore, assume that entry into real estate brokerage as a real estate broker is not free.[4] Items like age, sex, address, DOB, license issue and

[4]From this point on, I use ”real estate agent” and ”real estate salesperson” interchangeably in

expiration dates, status (expired or active) and certification types are reported for each records.

---

this paper

Name	Address	License Issue Date	License Exp. Date	Certification type	Status
Parviz Alivand	Austin, TX 78712	3/14/2001	4/9/2007	SP	Expired

Table 1.2: A typical record in the real estate agents' data

Table (1.2) shows the age distribution of the of new agents. Han and Hong (2008) report a mean age of 48 years for real estate agents in their dataset. The median age of new and experiences salespersons is 49 years in our data in 2006. The General tendency among real estate agents to be older and more educated than the general population is documented in the literature. [5] Also women make 53% of real estate salespersons.

Depending on different factors, salesperson licenses can be valid for either one or two year(s). In this paper, an agent is considered in the market in year t if her license is valid at any time in that year. License numbers are used to track agents over time and provide longitudinal linkage in the data.

1.3.2 Data on Housing Prices

The housing market data is collected by the Multiple Listing Services (MLS) across the state of Texas. MLS's are "... private databases which allow real estate brokers representing sellers under a listing contract to widely share information about properties with real estate brokers who may represent potential buyers or wish to cooperate with a seller's broker in finding a buyer for the property.[6]" Notice that For Sale By Owner (FSBO) houses and most of newly constructed condos are not included in MLS's.

A "markets" is defined as an MLS. We use agents' addresses to locate them into respective MLS's. The data is from 42 MLS's in Texas. Figure (1) shows the

[5] See Han and Hong (2008) and Hsieh and Moretti (2003) for further evidence
[6] Wikipedia page for MLS

location of MLS's in a Texas map.

Real estate markets are local. I try to define markets as self contained as possible. Most of the MLS's are not in close proximity of other MLS's. However, some neighbor other MLS's. The two extreme examples are the two MLS's that are inside the city limits of Dallas and Houston Metroplexes respectively.

The 42 MLS's used in this research contain between 60 to 65 percent of all the salespersons in Texas. Remainder of real estate salespersons are in markets that are not represented by any MLS.

1.3.3 Data on Wages

I use the average wages in private sector in a city as the measure for real estate agents' outside options. The data is provided by the Texas Workforce Commission and is normalized by the Urban CPI (CPI-U); a measure calculated and published by the BLS.

Summary Statistics- 2006						
Year: 2006	**SP**	**Sales**	**Sales Volume (m $)**	**Average Price**	**Average Wage**	**Labor Force**
Min	32	153	19	105,400	23,064	12,635
25th Percentile	221	1,161	176	123,925	30,072	42,442
Median	366	2,528	324	134,600	34,484	82,745
75th Percentile	1,266	6,404	952	158,550	40,540	137,317
Mean	1,395	7,855	1,469	148,621	36,087	221,555
Max	14,398	80,994	15,816	239,100	53,044	2,146,547
St Dev	2,686	15,879	3,241	36,579	8,480	438,995

Table 1.3: Summary Statistics of the data for 2006. All monetary values are 2006 dollar amounts

Table (1.3.3) provides the summary statistics of the key variables. The dataset includes small (and usually isolated) towns with as few as 153 houses sold in 2006; up to the large metroplexes in Texas with over 80,000 houses sold during the same year. The distribution of total volume of house sales is right-skewed as can be seen from the table. On average, total sales will add up to almost $1.5b. This figure, which is above the 75th percentile of the data, is nearly 5 times the median of total dollar value of sold houses. This skewness is almost entirely due to a handful of very populated metropolitans in Texas; cities like Houston, Dallas, San Antonio and Austin.

In 2006, in an average city, we have close to 1,400 salespersons and the median population among the cities that are studied here is slightly over 200,000[7]. Table (1.3.3) shows that MLS's vary significantly in terms of their sizes; however, the housing prices demonstrate less dispersion.

Table (1.3.3) suggests an interesting trend. From the median point forward, the proportion of the total value of the sold houses to the number of real estate agents remains almost constant; suggesting that within larger cities earnings of salespersons do not differ significantly from one another. As will become apparent later in the paper, this can be explained by a simple static free entry condition.

Table (1.3.3) shows the changes in key variables in two periods, panel A for the period of 1990 to 1998; and panel B is the change from 1998 to 2006.

For the most part, this table documents significant growth in real estate sec-

[7] Longview-Marshall

tor in the time period of 1990 to 2006. This should come as no surprise as the state of Texas has been consistently among the states with above average and sometimes, hight growth rates both in 1990's and 2000's.[8]

[8]Source: BEA http://www.bea.gov/newsreleases/regional/gdp_state/gsp_newsrelease.htm and
http://www.bea.gov/newsreleases/regional/gdp_state/2005/gsp1005.htm

Panel A: Chages in Key Variables between 1990 and 1998						
1990-1998	**SP**	**Sales**	**Sales Volume**	**Average Price**	**Real Wage**	**Workforce**
Min	35.92%	-10.54%	-8.92%	-3.97%	-2.80%	7.43%
25th Percentile	43.37%	39.26%	54.32%	8.43%	4.88%	18.58%
Median	54.34%	60.29%	80.08%	13.44%	6.80%	21.74%
75th Percentile	57.44%	103.29%	132.79%	21.16%	10.14%	36.30%
Mean	54.80%	70.44%	95.28%	14.57%	7.94%	32.38%
Max	107.87%	226.56%	234.36%	37.64%	38.38%	90.89%
St Dev	14.85%	55.51%	65.43%	11.13%	7.48%	23.31%

Panel B: Chages in Key Variables between 1998 and 2006						
1998-2006	**SP**	**Sales**	**Sales Volume**	**Average Price**	**Real Wage**	**Workforce**
Min	-31.82%	-51.89%	-27.05%	7.26%	-2.44%	-3.54%
25th Percentile	2.55%	43.47%	68.01%	13.55%	4.99%	4.23%
Median	13.91%	58.15%	102.65%	18.62%	8.66%	9.48%
75th Percentile	26.91%	94.63%	133.41%	25.60%	10.78%	18.57%
Mean	15.82%	68.96%	104.14%	21.49%	8.06%	17.72%
Max	77.87%	180.46%	253.85%	58.55%	18.50%	90.81%
St Dev	22.14%	41.84%	51.37%	10.91%	5.04%	22.04%

Table 1.4: Percentage changes in key variables between 1990 and 1998 (Panel A); 1998 and 2006 (Panel B)

Between 1990 and 1998, the value of housing transactions conducted through MLS's in our dataset grew by at least 80% for over half of the MLS's. Meanwhile, the number of salespersons shows a sharp jump as well. This growth did not slow down in the subsequent 8 years and the value of housing sales at least doubles for most of the MLS's between 1998 and 2006.

Table (1.5) shows the distribution of age in new salesperson. New salespersons age distribution shows little change over time.

Age Distribution in New Salespersons						
Percentile						
Min	**10**	**25**	**50**	**75**	**90**	**Max**
21	26	31	39	48	55	68

Table 1.5: Age Distribution of New Salespersons

Table (1.5) shows the distribution of age in new salesperson. New salespersons age distribution shows little change over time.

Entry can be observable through agents' original issue dates. Salesperson licenses could be valid for one or two years. Their exit is observable through their license expiration date. Table (1.6) shows the entry and exit of agents into the real estate industry. Notice that some of the exit observed here is due to salespersons upgrade to broker status. For brokers, median tenure of salespersonship before turning into broker is 4 years. (Table (1.9)). Both entry and exit of agents demonstrate a sharply increasing trend. In absolute terms entry surpasses the entry generally but exit shows a steeper change. Median experience among exiting varies between 3 to 6 years (at different times)

Year	upgraded to broker	Total Exit from SP	SP Exit from Market	Number of new agents
1990	987	1148	161	3719
1994	576	856	280	4289
1998	625	3243	2618	6482
2002	787	6084	5297	4905
2003	748	6565	5817	8170
2004	813	7239	6426	9258
2005	1047	8120	7073	10506
2006	1287	10863	9576	13699

Table 1.6: Number of new and Exiting agents

Year	Min	10	25	Median	75	90	Max
1990	1	2	2	3	3	4	34
1994	1	2	5	6	7	7	8
1998	1	1	2	4	10	11	12
2002	1	1	2	4	14	15	16
2006	1	1	2	6	16	19	20

Table 1.7: Distribution of experience (years in the market) for exiting salespersons

Year	Min	10	25	Median	75	90	Max
1990	1	2	3	3	4	4	4
1994	1	3	6	7	7	8	8
1998	1	3	7	11	11	12	12
2002	1	2	6	14	15	16	16
2006	1	2	5	13	19	20	20

Table 1.8: Current agents' experience

Experience before turning into Broker						
Min	**10**	**25**	**Median**	**75**	**90**	**Max**
1	2	3	4	8	12	20

Table 1.9: Distribution of experience before turning into Broker

In our dataset, we can compute probability of exit at time $t + 1$ *conditional* on being in the market at t. Table (1.10) shows this conditional probability. We cannot observe the probability of entry into the market.

Women are very active in real estate industry. In fact they make more than half of real estate salespersons and make about one third of brokers.

Brokers' share in real estate agents has been declining persistently. This is not due to decline in the number of brokers but is rather due to explosive increase in the number of real estate salespersons. (Table 1.11)

Year	Probability
1997	0.04667
1998	0.10692
1999	0.14388
2000	0.13520
2001	0.12413
2002	0.11939
2003	0.11412
2004	0.11445
2005	0.12240

Table 1.10: Conditional Probability that a salesperon at time t exits market at $t + 1$

Year	percent of brokers in realtors
1990	0.4584
1994	0.4271
1998	0.3798
2002	0.3458
2005	0.3223

Table 1.11: Percent of brokers in realtors

Our focus in this paper is on licensees who have had valid licenses at any time between 1990 and 2006 and potential agents' decision to enter the real estate market as salesperson. Longitudinal linkage is provided through the real estate license number. In real world, agents can enter and exit multiple times, however, this is not observable in our data; therefore, I further assume that an agent is in the real estate sector the entire time between her license issue and expiration dates. In addition, we cannot distinguish between commercial real estate agents and those who are involved in residential real estate.

1.4 Theoretical Framework

Similar to Hsieh and Moretti (2003) I assume that there is a continuum of homogeneous agents in the market. Decision of a potential agent to enter the market depends on the revenues that she will receive after entry. In particular, if c is the (fixed) commission rate, S_j is the number of houses sold in city j, P_j is the average house prices in this city and w_j is agents' outside option, potential agents enter the market as long as they can earn positive rents. Therefore, in equilibrium, an agent's outside option is equal to her earning as a real estate agent. If b_j denotes the equilibrium number of agents in city j, total commission of selling S_j houses is cS_jP_j. Total commissions are divided equally between the existing b_j agents in the market. So each agent's share is cS_jP_j/b_j. In equilibrium this is equal to agent's outside option, w_j and we have:

$$w_j = \frac{cS_jP_j}{b_j}$$

After rearranging, the equilibrium number of agents, b_j, is given by:

$$b_j = \frac{cS_jP_j}{w_j} \tag{1.1}$$

Due to homogeneity assumption, there is no distinction between marginal and average agent. The case in which agents are heterogeneous in terms of their ability to sell properties are discusses in another chapter in this dissertation.

Equation (1.1) also assumes that the commission rate charged by the real estate agents as the percentage of the value of the properties is fixed. Anecdotal evidence in support of a fix commission rate is presented earlier in this paper. From computational standpoint a fix commission rate reduces one dimension in which an equilibrium is sought.

In addition, I must assume that aside from agents' outside option, there is no significant cost associated with entry or exit. I find this assumption plausible because the (nominal) costs of obtaining and renewing real estate licenses are very small relative to the foregone wages that a potential agent has to incur if she decides to become a real estate agent.[9] I also assume that potential agents can enter this market instantaneously.

1.5 Estimation Results

1.5.1 Number of Real Estate Agents

1.5.1.1 Levels

The first hypothesis of Hsieh and Moretti (2003) is that the cities with high housing price attract more real estate agents. Hsieh and Moretti (2003) runs a regression of percentage of real estate agents from a city's workforce on average home prices. Higher home prices make real estate brokerage industry more attractive for potential agents and increases the likelihood of pursuing this career versus other options. If LF denotes the size of workforce, from the theoretical framework presented in the previous section, we have the following relationship between the

[9] In October 2009, the cost of obtaining a real estate salesperson for the first time (an "original application") in Texas, was $97 and its renewal was either $57 or $107. During the same period, the average application fee of an original for real estate broker license was $300 and the cost of renewal was $500.

percentage of real estate agents and home prices:

$$\frac{b_j}{LF} = \frac{cS_jP_j}{LF \times w_j} \Rightarrow \ \log\left(\frac{b_j}{LF}\right) = \log(c) - \log(w_j) + \log(P_j) + \log\left(\frac{S_j}{LF}\right)$$

In other words, even after including agents' outside options and a constant term in the explaining variables, theory predicts a coefficient of 1 in the regression of $\log(b_j/LF)$ on $\log(P_j)$, **only** if the ratio of $\frac{S_j}{LF}$ is constant. Therefore, unless the ratio of number of houses sold in a city to the size of labor force is the same (across the cities, for cross-sectional estimates), based on the correlation of $\frac{S_j}{LF}$ with P_j, the coefficient of $\log(P_j)$ in the aforementioned regression differs from 1.

In the dataset used in this paper, the correlation between $\log(S_j/LF)$ and $\log(P_j)$ is 0.31. So, by omitting $\log(S_j/LF)$, the explanatory power of variation in this variable will be "picked up" by inflated or deflated coefficients of other explaining variables. Because of the positive correlation, the theory predicts the coefficient of $\log(P_j)$ to be larger that 1. The correlation between $\log(S_j/LF)$ and $\log(w_j)$ in my data is very close to zero (-0.02) and the omission of $\log(S_j/LF)$ does not affect coefficient of $\log(w_j)$.

Notice that if S_j and P_j are independent, the theory predicts a coefficient of one in the regression of $\log(b_j)$ on $\log(P_j)$. However, if in fact S_j and P_j are correlated, then that coefficient will be above (below) unity if $corr(S_j, P_j) > 0$ ($corr(S_j, P_j) < 0$). Data shows a correlation of 0.7 between the (log) of number of houses and (log) of the housing price, so an above 1 coefficient for $\log(P_j)$ is predictable. I run (1.1) for each cross-section between 1990 and 2006. A typical table of results (for year 2000) is shown in table (1.12):

Dependent Variable: $\log(b_j/LF)$ in year 2000			
	coef	stdev	z-val
$\log(p_j)$	2.553	0.682	3.743
$log(w_j)$	-1.752	0.750	-2.333
const	-17.090	6.954	-2.458

Table 1.12: The estimates for coefficients of equation (1.1) in year 2000.

The coefficient for $\log(p_j)$ is 2.553 (0.682), much larger than the prediction in theory when $\log(S_j/LF)$ is included.

The following histogram (figure (1.4)) shows the coefficient of $log(p)$ in cross-sectional regression of $\log(b_j/LF)$ on $log(p)$, $\log(w)$ and a constant term. The hypothesis that the coefficient of $\log(p)$ is equal to 1 in those regression is rejected for all the years in our data.

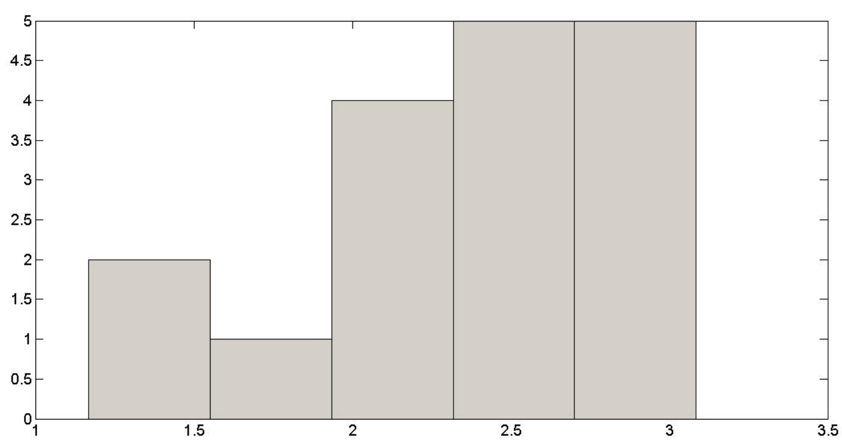

Figure 1.3: Histogram of estimates of cross-sectional regression of $\log(b_j)$ on $\log(p_j)$

Hsieh and Moretti (2003) reports coefficients of 0.623 (0.058), 1.142 (0.097) and 0.917 (0.078) for OLS regressions of (logs) percentage of real estate agents in a city on housing prices in 1980, 1990 and the change between 1980 to 1990 respectively. That paper does not fully specify the independent variables and whether $\log(w)$ is included in explanatory variables or not.

This is well below our estimates for the similar coefficients on **level** regressions. Hsieh and Moretti do not comment on the control variables that they use in the regression of $\log(b)$ on $\log(p)$. It is helpful to point out one of the distinction between the dataset used in Hseih and Moretti (2003) and the dataset used in this paper. The dataset used in Hseih and Moretti (2003) is rich in both cross-sectional and over-time variations. Since my dataset shows little over time variation, I cannot get any significant coefficients for regression on differences.

The right structural regression in this setting is the regression of $\log(b_j)$ on log of total volume of home sales at city j. We denote this total volume by TV_j and will look for the coefficient of $\log(b_j)$ on $\log(TV_j)$. Notice that in equation (1.1), $TV_j = S_j \times p_j$ and we can predicted coefficient for $\log(TV_j)$ from this equation is 1. Following table provides the results of this regression for the previous cross-section: With this specification, I get the following histogram for the coefficient

Dependent Variable: $\log(b_{2000})$			
	coef	stdev	z-val
$\log(TV_{2000})$	0.843	0.075	11.246
$\log(w_{2000})$	-0.031	0.453	-0.069
const	-9.678	3.862	-2.506

Table 1.13: The estimates for coefficients of $\log(b)$ on $\log(TV)$ for year 2000.

of regression of the (log of) number of real estate agents on total commissions available and agents' outside options:

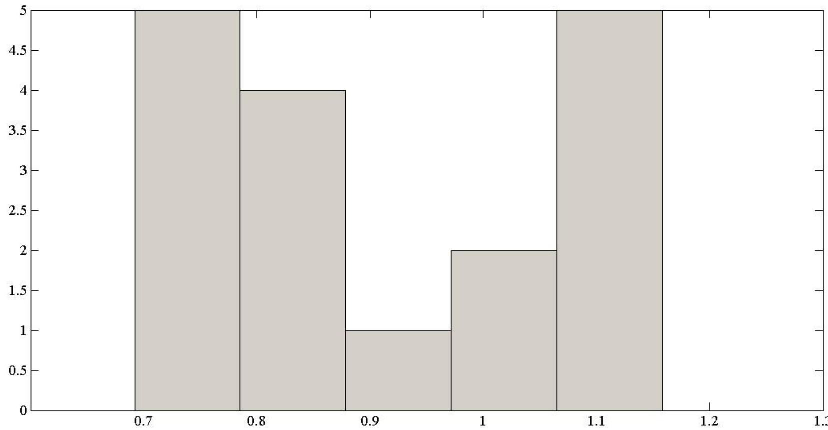

Figure 1.4: Histogram of estimates of cross-sectional regression of $\log(b_j)$ on $\log(TV_j)$

1.5.1.2 Differences

Hsieh and Moretti (2003) report a coefficient of 0.917 (0.078) for the regression of changes in the share of real estate agents in a city over changes in housing prices. Most of the variation my dataset comes from the cross-sectional changes whereas the dataset in Hseih and Moretti (2003) is rich in terms of over-time variation.

1.5.2 Productivity of Real Estate agents

Increase in the number of real estate agents can come from two sources; price increases and increases in the number of houses sold in an MLS. So far I did not differentiate between these two sources. However, they have quite different implications in terms of social welfare. If housing prices are fixed and increases in total commission solely come from the higher number of home sales; then each agents will sell as many houses as she would prior to the hike in the sales number. This in turn is an indication of absence of any business stealing effect among the agents. In the alternative scenario, where price jump is the only cause of increase in total commissions, the same number of houses are sold by a higher number of real estate agents, new agents root for business in the pool of houses that could have been sold by a lower number of agents and therefore business stealing effect is going on.

In the absence of any business stealing effect, and when the only form of entry in demand driven, the ratio of $\frac{S}{b}$ must remain constant; in particular, it must be independent of housing prices. On the other hand, if the number of real estate agents increase purely due to the business stealing effect, then the model predicts the coefficient of regression of $\log(\frac{S}{b})$ on $\log(p)$ to be equal to one. In between cases will be characterized by other coefficients. The closer this coefficient to one means that a larger portion of entry is due to business stealing effect.

Hsieh and Moretti (2003) considers two measures for productivity. First, is the number of houses sold per agent in a city and the second is the number of houses sold per each hour of labor by real estate agents. The second measure has the advantage of taking the variation of work schedules, particularly part-time real estate agents into account.

In the one-shot, homogeneous agent game, with fixed commission rate of 6%, controlling for agents' outside options, agents productivity should demonstrate a one-to-one negative correlation with average housing price in a city. In other words, one percent increase in housing price in a city reduces the productivity of agents in that city by one percent. Cross-sectional estimates in Hsieh and Moretti (2003) support this theoretical result. The coefficients of regressing log(productivity) on log(price) for 1980 and 1990 respectively are -0.929 (0.059) and -1.098 (0.049).

Dependent Variable: $\log(S/b)$ **in Year 2000**					
	coef	stdev	z-val	95% Confidence Interval	
$\log(p)$	-0.649	0.409	-1.586	[-1.451	0.153]
$\log(w)$	0.422	0.450	0.938	[-0.460	1.305]
cons	4.624	4.174	1.108	[-3.556	12.804]

Table 1.14: Estimates of coefficients of $\log(S/b)$ on $\log(p)$. Theory predict this coefficient to be 1 for pure business stealing effect.

Table (1.14) presents the estimate results of regression of productivity of real estate agents on average housing price in the year 2000. The estimated business stealing effect is -0.649. In other words, almost two-thirds of the entry into real estate brokerage industry is due to business stealing effect. If housing prices go up by one percent, then the average number of houses sold per agent will decrease by two-thirds of a percent. The basic reason is that since the commission rate is almost unchanging, any increase in the housing price will generate more revenues to be divided among real estate agents. Notice that the revenue will increase even with the same number of houses for sale. Higher revenues will attract new agents into the market and therefore, decrease the number of houses sold per agent.

1.5.3 Extra Wages of the Real Estate Agents

The assumption of free entry into the real estate brokerage industry has two implications: first, within a city, marginal real estate agents' earnings must be equal to that of marginal non-agents. Since agents are assumed to be homogeneous, in equilibrium, each agent must be indifferent between staying in the market and leaving. Second, once we control for outside options, real estate agents' earnings should not differ across cities; in particular, it should not depend on variation in housing prices.

Since my dataset is limited to agents and average wage in the private sector, I cannot test the first hypothesis. The second implication, however, is testable using my data.

It must be noted that since the second implication of free entry is only a necessary condition, its empirical validation (tested here) does not guarantee that free entry exists in the market. It merely provides evidence **consistent** with free entry.

I define the "extra wage" as the average wage earned by the real estate agents in a city minus their outside option ($\frac{cSP}{b} - w$). From equation (1.1), extra wage must be zero:

$$\frac{cSP}{b} - w = 0. \tag{1.2}$$

Table (1.15) shows the results of the cross sectional regression of extra wages on the housing prices for the year 2000: the evidence is strongly in favor of no relationship between extra wage and home price variations across cities in my dataset.

Dependent Variable: Extra Wage 2000					
	coef	Std Err	z-val	95% Conf. Interval	
log(price)	-0.033	0.329	-0.101	[-0.678	0.612]
cons	0.445	3.818	0.116	[-7.039	7.929]

Table 1.15: Cross-sectional estimation of extra wage on housing price for the year 2000. Theory predicts this coefficient not to be significantly different from zero

The results of table (1.15) are reiterated throughout all the cross sections in our data.

In figure (1.5), I present the distribution of the coefficients of the (log) total commissions in the regression of the (log) extra wages on (log) total commissions. Due to free entry, the coefficient for $\log(TV)$ in this regression must be insignificant too, which is confirmed in our regressions.

Failure to account for lower ”time in the market” with more agents

1.6 Conclusion

With fixed commission rate and free entry, even with the same number of houses sold, higher home prices lead to a more people to become real estate agents. In comparison between two cities that are identical except for their home prices, the one with higher housing price will attract more real estate agents as the total commission to be divided is larger, and therefore, will have a higher percentage of its workforce as real estate agents. An immediate corollary is that in high home

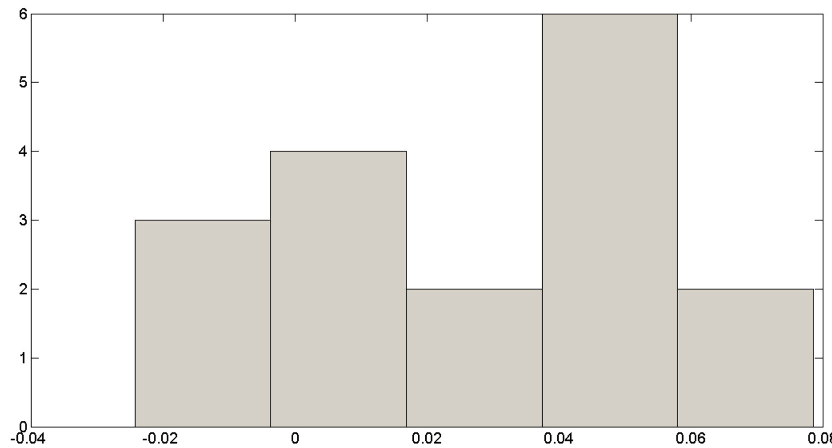

Figure 1.5: Histogram of estimates of cross-sectional regression of extra wage on total commissions

price cities, each agent sells a lower number of houses, in other words, they are less ”productive”. Hsieh and Moretti (2003) find strong empirical support for these two claims in a dataset consisting of 282 Metropolitan Statistical Areas in 1980 and 1990 censuses. I adopt the same theoretical framework and use the dataset of Texas Real Estate Agents Record to evaluate their claims empirically.

I point out one important source of misidentification in their empirical analysis as they do not include the (normalized) number of houses sold in a city with its housing prices as an explanatory variable. Failure to account for this variable creates inflated coefficients for the effect of home price on the number of real estate agents that fail meet the prediction in the theory. I show that once this issue is addressed, the coefficients get closer to theory prediction.

I find a very similar result as in Hsieh and Moretti (2003) for the relationship between agents' productivity and housing prices in city. With higher prices, since there are more agents, fewer houses are needed to cover agents' outside options. Therefore, in equilibrium, productivity of agents (defined as the number of houses sold per agent in a city) is lower in higher housing price cities. Empirical support is obtained for this result in Hsieh and Moretti (2003) and is confirmed in this paper. Similar to Hseih and Moretti (2003) I find that one percent increase in home prices in a city would lead to two thirds of percent drop in agents' productivity, a result which is almost identical to that in Hsieh and Moretti paper.

The third claim in Hsieh and Moretti (2003) is to evaluate the free entry condition. In particular, a necessary condition for free entry is to have no extra wages for being a real estate agent. Hsieh and Moretti (2003) and this paper find strong evidence in support of this argument. Extra wages are defined as average agent's earning net of her outside option.

Chapter 2

Housing market and Social Inefficiencies of Free Entry

In theory, free entry potentially leads to social inefficiency. This paper finds strong empirical evidence consistent with excess entry into Texas Residential Real Estate Brokerage Industry and studies the effects of heterogeneity and future uncertainty on such inefficiencies. I develop a dynamic model of entry and exit with heterogeneous agents and modify the predictions of the earlier literature. I show that the heterogeneity among (real estate) agents results in a weaker relationship between the real estate commission fees and the number of real estate agents. I also show that the models developed for static cases in the previous papers are special cases of the more general model in this paper.

The model allows us to explain the lower business stealing effect compared to static and homogeneous models that is observed in the data. To address the issue of excess entry, I separate the business stealing effect from demand driven entry and find that on average 75 percent of entry is due to business stealing. To evaluate free entry, I control for agents' outside options and find that the extra wages of the real estate agents do not vary with housing prices.

2.1 Introduction

I develop a simple dynamic entry and exit model with heterogeneous agents to evaluate the effects of heterogeneity and dynamics on social wastes of free entry in the real estate brokerage industry.

In this model agents are subject to different productivity shocks. Commission fees- which are assumed to be the only source of income for real estate agents- are divided according to the number of agents and their productivity shocks. The model generates entry and exit even in the steady state and predicts a proportional relationship between the number of agents and the value of home sales. Assuming ”sticky” types, I show that the proportionality is always less in the dynamic model and is decreasing in heterogeneity. The types are sticky when the probabilities of receiving the same productivity shocks for incumbents are higher than the probability of drawing a new productivity shock of that type for potential new agents. In Markov Perfect Equilibrium (MPE), low type agents exit and high type agents stay in the market. The incentives for potential agents to enter the market are lower with heterogeneity because high type incumbents are more likely to stay high type than a potential agent becoming one. Therefore, incumbent high type agents are more likely to capture a larger share of any new rents. This in turn reduces the size of the (potential) pie for those contemplating to enter the market.

Empirical results provide strong support for the model. Estimates of the proportionality factor are around 0.8 to 0.9, consistent with calibration and predictions of the model. I also use the model to explain the effects of heterogeneity on business stealing effect and show that heterogeneity leads to lower business stealing. I define the measure of non-price competition as the ratio of the number of agents to the number of houses sold in a city and show that any given increase in average price will lead to a lower increase in competition when agents are heterogeneous. Models with no heterogeneity predict that since agents do not compete in commission rates, any increase in home prices will go to social waste. Inclusion of heterogeneity and dynamics in the model allows us to explain why we do not observe such behavior in the data.

One conclusion of free entry assumption is that the average earning of agents should not differ across cities. I evaluate this necessary condition and show that free entry is consistent with the observations from our data. In particular, the extra wage of being real estate (the difference between average earning of real estate agents and their outside options) does not depend on housing prices or total available commissions.

2.1.1 Empirical Literature on Social Wastes of Free Entry

A known property of perfect competition is the efficiency of its equilibrium outcome. The implication of this property is that if other conditions of perfect competition are met, any barrier to entry is harmful from the welfare standpoint.

However, if in a market firms do not behave "competitively" (for example when they are not price takers anymore), total surplus could be improved by imposing some entry restriction into that market.[1] Social inefficiency of free entry can

[1] See Cabral (2000)

come from two sources. First, is the business stealing phenomenon. In a market where new and incumbent firms' products are close substitutes, in addition to (potentially) lower price, a new entrant can steal business from the incumbents. Bank regulations in Portugal in 1990s to limit the opening of new branches were partly aimed to prevent business stealing among rival banks. Second, when average costs are decreasing in output, a new firm will further move away the existing ones from their efficient output level.[2]

Although potential inefficiency of free entry is well known in the theory; there are relatively few empirical studies to document this issue. Berry and Waldfogel (RAND, 1999) offers the first insight into estimating the social waste of free entry in radio broadcasting industry. Firms in this model are symmetric and play a static game. In this paper they estimate the fixed cost and demand function parameters and concludes that the welfare loss is 45% of revenue. Even with solid evidence of social inefficiency of free entry, the actual entry and exit rarely happens in their data due to FCC regulations.

Hsieh and Moretti (JPE, 2003) provide an analysis of social waste of free entry in the real estate brokerage industry.

There are two interesting properties that make the study of the real estate brokerage industry desirable. First, if acquiring a real estate license is perceived as entry, then the barriers to entering the real estate brokerage industry are relatively low. One only needs to attend a handful of classes and take an exam to be eligible

[2]The extreme case in which free entry leads to excess number of firms is the standard circular city with zero marginal and positive fixed costs.

for a real estate license. The second property is the seemingly unchanging commission rate for the real estate brokerage services. This rate is generally around 5 to 6 percent of the property value. Hsieh and Moretti (2003) and Han and Hong (2008) present historical evidence supporting persistency of the so called 6% commission rate over time, regardless of property values, and in different cities. A fixed commission rate implies that in two cities with the same number of houses for sale, the one with higher housing prices will attract more agents. Therefore, the number of houses sold per agent is lower in the high housing price city.

Hsieh and Moretti (2003) predict that with free entry and fixed commission rates: 1) cities with higher housing prices would have more real estate agents; 2) agents in cities with higher home prices are less productive, in the sense that they sell fewer houses per agent, and 3) conditional on agents' outside options, agents' real wages should be independent of home prices and be the same across cities. Their model studies the business stealing effect in the real estate brokerage industry in a static environment with homogeneous agents. They also abstract from cost inefficiency.

Han and Hong (2008) highlight that Hsieh and Moretti (2003) suffer from a lack of cost inefficiency analysis. They show that the results of Hsieh and Moretti (2003) can be obtained using a simple model with no cost inefficiency. Focusing on this aspect of inefficiency, Han and Hong (2008) impose some structure on cost function and try to estimate the related parameters. They point out three sources of cost inefficiency as wasteful non-price competition, loss of economies of scale, and high fixed costs and find direct evidence for all three. In particular, they "find that a one-standard-deviation increase in entry rate would increase the average variable costs by 28.96%, resulting from wasteful non-price competition."

2.2 The Real Estate Brokerage Industry

The real estate brokerage industry has three main properties that make it very useful in our analysis:

1. Free Entry
2. Heterogeneity Among Real Estate Agents
3. Fixed Commission Rate

All these properties are discussed in detail in chapter 1 of this dissertation.

2.3 Model

The model is based on Hopenhayn (1992). Agents supply the real estate brokerage service and households demand it. Commission fees are the sources of income for the real estate agents. R_t is the total commission fees at time t. Agents receive productivity shocks ϕ in each period.[3] The productivity shocks can be either high or low, $\phi \in \{\phi^H, \phi^L\}$ and follow a Markov process independent across agents. I normalize $\phi^L = 1$. For the incumbents, the probability of staying in the same type is denoted by θ:

$$\Pr(\phi_{t+1} = \phi | \phi_t = \phi) = \theta$$

[3]I refer them as "types" as well

Agents are identical within each type; those with higher productivity shocks can sell more houses. Specifically, for each house sold by a low type agent, the number of houses sold by a high productivity shock agent is $\phi^H > 1$. Therefore, if each low type agent sells s houses, and if N^H and N^L respectively denote the number of high type and low type agents, the total number of houses sold will be:

$$s(\phi^H N^H + N^L)$$

Total commission fees, R, is divided according to the number of houses sold. Hence, each low type agent receives:

$$R^{'} \equiv \frac{R}{\phi^H N^H + N^L}$$

and each high type agent receives:

$$\frac{R}{\phi^H N^H + N^L} = \phi^H R^{'}$$

Each agent pays the continuation cost of c_f in each period. We need to have $c_f > 0$ in order to have any exit taking place.

Potential agents can enter the market at any time by paying the entry cost of c_e. The probability of receiving a high productivity shock is η for the new entrants. The discount rate is $0 < \beta < 1$.

Timing: Incumbents

1. Enter t with ϕ_{t-1}
2. decide whether to exit

3. receive the same productivity shock with probability θ; or receive the other productivity shock with probability $1-\theta$.

4. receive payment and pay c_f for the period.

5. enter $t+1$ with ϕ_t

Timing: Potential Entrant

1. Decide whether to pay c_e and receive this period's shock ϕ_t with density η.

2. same as 4-5 above.

Incumbents need to decide whether to stay in the market or leave before realizing their current shocks. If they decide to stay in the market, they would have to pay c_f regardless of their type. If V_L and V_H stand for the value functions of high and low types, we would have:

$$V_H = \frac{\phi^H R}{\phi^H N^H + N^L} - c_f + \max\{0, \beta(\theta V_H + (1-\theta)V_L)\} \tag{2.1}$$

$$V_H = \phi^H R' - c_f + \max\{0, \beta(\theta V_H + (1-\theta)V_L)\} \tag{2.2}$$

and for low type: $$V_L = \frac{R}{\phi^H N^H + N^L} - c_f + \max\{0, \beta(\theta V_L + (1-\theta)V_H)\} \tag{2.3}$$

$$V_L = R' - c_f + \max\{0, \beta(\theta V_L + (1-\theta)V_H)\} \tag{2.4}$$

Potential entrants enter as long as their expected income is non-negative:

$$\eta V_H + (1-\eta)V_L - c_e \geq 0$$

Notice that the decision on whether to stay in the market for the next period is conditional on current period's productivity shock. An agent with type ϕ^i will stay in the market iff

$$\theta V^i + (1-\theta)V^j \geq 0, \;\; i \in \{L, H\}, \;\; j \neq i.$$

and exit iff:

$$\theta V^i + (1-\theta)V^j \leq 0, \;\; i \in \{L, H\}, \;\; j \neq i.$$

It is important to notice the importance fixed commission rate here. If the commission rate charged by the real estate agents was not fixed, then the agents would have to include the choice of commission rate, c, in their maximization problem and the commission fees, R, will be a function of c. Computationally, this will eliminate one state variable.

Equilibrium

I construct a Markov Perfect Equilibrium (MPE) in which low type agents exit and only high type agents stay:

$$\theta V_H + (1-\theta)V_L \geq 0 \tag{2.5}$$

$$\theta V_L + (1-\theta)V_H \leq 0 \tag{2.6}$$

Later, I will characterize the necessary and sufficient conditions that guarantee the existence (and uniqueness) of such equilibrium.

In this equilibrium, the number of agents of each type will follow the following equations:

$$N^H_{t+1} = \theta N^H_t + \eta k_{t+1} \tag{2.7}$$

$$N^L_{t+1} = (1-\theta)N^H_t + (1-\eta)k_{t+1} \tag{2.8}$$

Here, k_t denotes the number of new entrants in period t. N_t^H and N_t^L are the number of high and low type agents in period t respectively. Equations (2.7) and (2.8) reiterates the basic dynamics of the equilibrium: the population of next period's high type agents consists of two parts: first, the incumbent high type agents in the each period, who, with probability θ, will remain high type, and second, the new agents, who will be of high type with probability η. On the other hand, the low type agents at time $t+1$, are either the incumbent high type agents, who stay and with probability $1-\theta$ receive a low shock the following period, or the new coming agents who draw a low productivity shock with probability $1-\eta$. Notice that the exit of the low type agents is key for equations (2.7) and (2.8) to hold. Otherwise, we would need to include the incumbent low type agents who would stay and receive their productivity shocks in period $t+1$.

In steady state, $k_{t+1} = k_t = k$ and we must have $N^L = k$. In other words, the number of new comers k, is equal to the number of low type agents (who leave the market). For N^H we must have:

$$N^H = \frac{\eta k}{1-\theta} \tag{2.9}$$

Equations (2.9) is the rearrangement of:

$$(1-\theta)N^H = \eta k$$

which is the necessary condition for the number of high type agents to stay the same in the steady state.

Total number of agents, N is:

$$N \equiv N^H + N^L = k\left(\frac{1-\theta+\eta}{1-\theta}\right) \tag{2.10}$$

Given the discussion so far, the equilibrium in this game is characterized by the following equations:

$$
\begin{aligned}
N &\equiv N^H + N^L && (2.11)\\
N^L &= k && (2.12)\\
N^H &= \frac{\eta k}{1-\theta} && (2.13)\\
\phi^L &= 1, \phi^H > 1 && (2.14)\\
\frac{\phi^H N^H}{\phi^H N^H + N^L} + \frac{N^L}{\phi^H N^H + N^L} &\equiv 1 && (2.15)\\
\theta V_H + (1-\theta) V_L &> 0 && (2.16)\\
\theta V_L + (1-\theta) V_H &< 0 && (2.17)\\
V_E &= \eta V_H + (1-\eta) V_L - c_e = 0 && (2.18)\\
V_H &= \frac{\phi^H R}{\phi^H N^H + N^L} - c_f + \beta(\theta V_H + (1-\theta) V_L) && (2.19)\\
V_L &= \frac{R}{\phi^H N^H + N^L} - c_f && (2.20)
\end{aligned}
$$

I set $c_e = 0$ in the rest of this paper. It can be shown that if $c_e \neq 0$, the results will only get stronger.[4].

As discussed earlier, for an agent with type ϕ, the decision to stay in the market during the next period depends on her expectations from the evolution of her type. An incumbent high type agent will stay in the market if the probability of receiving another high productivity shock in the next period is higher than that of a new potential agent. Otherwise, high type agents will exit and the market will be filled with new agents. In other words, $\theta > \eta$ is necessary condition for

[4]Especially the results in proposition (2.23)

high type agents to stay in the market. On the other hand, if an agent with low productivity shock exits the market, she will be replaced by a new agent. For this to happen, a new agent must have a higher probability of becoming a high type than an incumbent low type agent. The following lemma proves that the intuitions that I just explained do indeed guarantee the entrance and exit of agents with different types as expected and therefore, are the necessary and sufficient conditions for the existence and uniqueness of the equilibrium.

Lemma 2.3.1. *The equilibrium in which low type agents exit and high type agents stay in the market exists if and only if:*

$$\theta > \eta \tag{2.21}$$

$$\textit{and } \theta > 1-\eta \tag{2.22}$$

This equilibrium is unique.

Proof. See the appendix. □

With conditions (2.21) and (2.22) being met, the payoff for the potential agents, $\eta V_H + (1-\eta)V_L$ lies between those of low and high types agents:

$$(1-\theta)V_H + \theta V_L < \eta V_H + (1-\eta)V_L < \theta V_H + (1-\theta)V_L$$

It is also easy to see that the inequalities (2.21) and (2.22) cannot be satisfied simultaneously unless $\theta > \frac{1}{2}$.

The following proposition determines the equilibrium number of agents, N:

Proposition 2.3.2. *If $c_e = 0$, the equilibrium number of agents satisfying equations (2.11) through (2.20) is given by:*

$$N = B\frac{R}{c_f} \tag{2.23}$$

where N is the number of agent in equilibrium, R is the total commission fees and c_f is the continuation cost. $B = A(\phi^H) \times L$ where $A(\phi^H)$ and L are defined as follows:

$$A(\phi^H) = \frac{1 - \beta\theta + \beta\eta + \eta\phi^H - \eta}{1 + \eta\phi^H - \theta} \tag{2.24}$$

$$L = \frac{1 + \eta - \theta}{1 - \beta\theta + \beta\eta} \tag{2.25}$$

Proof. See the appendix. □

The proportionality factor of B, summarizes the relationship between available commission fees and continuation costs, and the number of real estate agents, N. Among other things, this relationship depends on the level of heterogeneity represented by ϕ^H in this model. The proportionality factor shows how many new agents would enter into the market if the commission fees R, goes up by c_f. The following lemma shows that higher heterogeneity will translate into weaker relationship (or lower B):

Lemma 2.3.3. *In an equilibrium in which low type agents exit and high type agents stay in the market, the proportionality factor (B), between the number of agents on one hand, and available commission fees and continuation costs on the other hand* *is decreasing in the heterogeneity among agents. It reaches its maximum when agents are homogeneous and its maximum is 1. In other words:*

$$B \leq 1 \tag{2.26}$$

$$\textit{and } \phi^H = 1 \Rightarrow B = 1 \tag{2.27}$$

Proof. See the appendix. □

When agents are homogeneous, the proportionality factor will be equal to unity. When $(B = 1)$, equation (2.23) implies that an increase in R by the continuation cost c_f, will result in one more agent being in the market in the stationary equilibrium. However, when agents are heterogeneous the proportionality factor will be less than unity $(B < 1)$ meaning that an increase by the amount of c_f in R *is not enough* to cause more agents to opt into becoming real estate agents.

Notice that possible risk aversion plays no role in this conclusion as agents are considered risk neutral in the model. Any possible risk aversion effect will be in addition to what this equilibrium has predicted and in that sense, B provides the most conservative estimates for the heterogeneity effect.

For an equilibrium to exist, we need $\theta > \eta$. When θ is almost equal to η and $c_e \approx 0$, the expected payoff for a new entrant, $\eta V_H + (1 - \eta)V_H$, is almost equal to $\theta V_H + (1 - \theta)V_L$, which is the expected payoff for a current high type agent. As the difference between θ and η grows, the gap between the expected payoff for a current high type agent and that of a potential new entrant will increase. So, in equilibrium, we should expect to see higher proportion of current relative to the new agents. Higher proportion of high type agents means that if commission fees increase, the high productivity agents capture most of the newly created fees. This means that the relationship between the commission fees and the number of real estate agents is weaker for larger differences between θ and η. In other words, if we define $x = \theta - \eta$, then B should be decreasing with respect to x. This intuition is supported in the model as:

$$\frac{\partial B}{\partial x} = \frac{-\eta(1-\beta)(\phi^H - 1)(1 + \eta\phi^H - \eta - \beta(\theta - \eta)^2)}{(1 - \beta\theta + \beta\eta)^2(1 - \eta\phi^H - \theta)^2} \leq 0 \tag{2.28}$$

Notice that with homogeneity, ($\phi^H = 1$), $\frac{\partial B}{\partial x} = 0$. This should come as no surprise since when agents are homogeneous, both states of the world are the same and therefore, the probability of reaching one state should not play any role in determining the relationship of the commission fees and the number of real estate agents.

2.3.1 Hazard rate implications of the model

In this section I show that the implicit assumption of instantaneous adjustment to the market factors results in a constant hazard rate.

Assume that k_s is the number of new entrants in period s. The number of survivors of this cohort at time $s + t - 1$ is $\eta\theta^{t-2}k_s$ for $t > 2$; therefore, a total of $\eta\theta^{t-2}(1 - \theta)k_s$ will exit the market at period $s + t$, given the survival up to $t - 1$. Now, if we consider all the possible values for s, the hazard rate for the population that has survived $t-1$ periods from entry but exits in period t after entry is:

$$h(t) = \frac{\sum_s \eta\theta^{t-2}(1-\theta)k_s}{\sum_s \eta\theta^{t-2}k_s} \tag{2.29}$$

$$h(t) = \frac{\eta\theta^{t-2}(1-\theta)\sum_s k_s}{\eta\theta^{t-2}\sum_s k_s} \tag{2.30}$$

$$\text{or:} \quad h(t) = 1-\theta \tag{2.31}$$

2.3.2 Evaluating the effect of heterogeneity and future uncertainty

Equation (2.23) predicts a less than one-to-one relationship between total commission fees and the number of real estate agents. This relationship is such because the equation includes heterogeneity and future uncertainty in agents' optimal behavior. However, the effect of each factor is not identifiable.

Using equation (2.31) along with the average hazard rate of 0.1, I calibrate the type persistence probability, θ to be around 0.9.[5] I calibrate ϕ^H to the ratio of 75th percentile to 25th percentile earning of real estate agents. From table (1.1), this ratio is $\phi^H = 4.5$. The discount rate is $\beta = 0.85$, and the probability of high state draw for new salespersons, $\eta = 0.4$. With these specifications, I can calibrate the overall effect of heterogeneity to $B = 0.9$.

2.4 Data

The dataset used in this paper is the result of merger of three datasets: 1. The data on real estate agents; 2. Data on home prices; and 3. Data on wages. These three datasets are discussed in detail in chapter 1 of this dissertation.

[5] See the appendix on data for the average hazard rate

2.5 Estimation Results

2.5.1 Number of real estate agents

I begin by estimating the relationship between the number of real estate agents and the total commissions available in each MLS. Equation (2.23) predicts that these two variables are linearly and positively correlated to each other. Once agent's outside option is fixed, this equation predicts that the effect of total commissions on the number of real estate agents cannot be more than one-to-one. ($B \leq 1$). This is the direct result of both heterogeneity and future uncertainty faced by each current and potential agent.

Our fundamental value function equations (2.1) and (2.3) show that the agents with higher productivity shocks can seize a higher share of the total commissions. This, along with the assumption of "sticky" types, implies that any increase in total commissions, whether it is due to a price hike, or rise in the number of houses being sold in an MLS, will mostly be captured by the agents with high productivity shocks.

Dependent Variable: N					
	coef	stdev	z-val	95% conf. interval	
R/c_f	0.83	0.03	26.77	[0.87	0.79]

Table 2.1: Cross-sectional regression of the number of real estate agents on total commissions in a city/ outside option for the year 2000.

Table (2.1) presents the estimation results of the number of real estate agents on the total available commissions for the year 2000. The goal here is to estimate the proportionality factor, B, in the following equation between number of real estate agents, N and total commissions available to be divided among real estate agents, R:

$$N = B\frac{R}{c_f} \tag{2.32}$$

The basic result of this table, that the coefficient of total commission in the regression of N on R/c_f is generally less than 1, is confirmed by Figure (2.1) which shows the estimates in other cross sections.

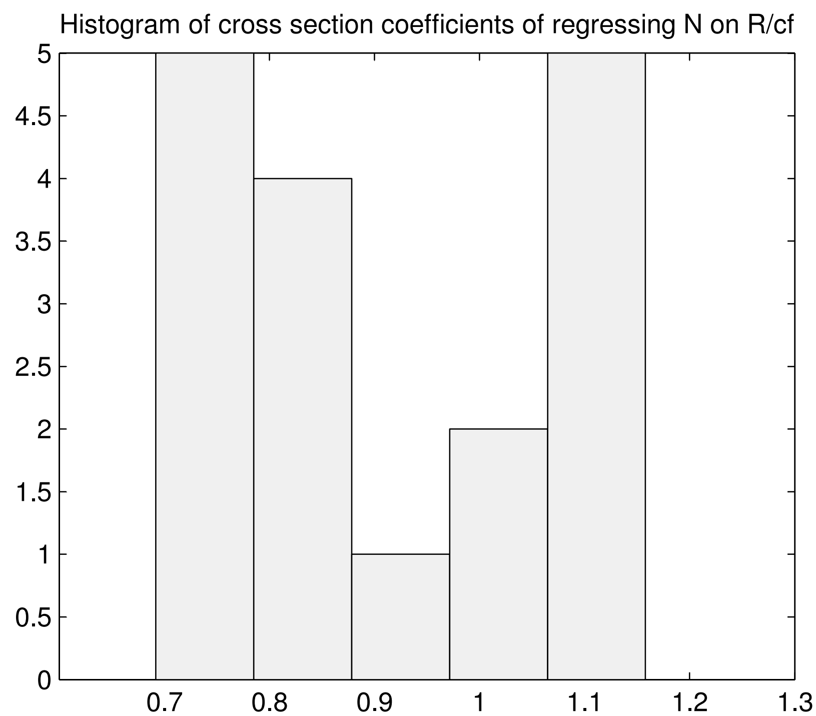

Figure 2.1: Histogram of estimates of cross-sectional regressions of N on R/c_f.

The estimates of agents' city-specific outside options, used throughout this paper, are the weighted average of wages of all private sector workers in the same city.

Results in figure (2.1) show little to no sensitivity to different methods of counting agents, weighting and normalization. In order to reduce the effects heteroskedasticity, I normalized both sides of equation (2.23) by the workforce size of the cities.

Notice that failure to account correctly for the effects of normalization could potentially lead to significant endogeneity errors. Hsieh and Moretti (2003) estimate the regression of the percentage of workforce employed as real estate agents on housing price in each city. From equation (2.23):

$$\log(\frac{N}{WF}) = \log(cB) + \log(\frac{cp}{c_f}) + \log(\frac{S}{WF}) \tag{2.33}$$

Unless $\frac{S}{WF}$ is constant, regressing $\log(\frac{N}{WF})$ on $\log(\frac{cp}{c_f})$ will suffer from endogeneity. The coefficient of $\log(\frac{cp}{c_f})$ will be larger than unity if $corr(\frac{S}{WF}, p) > corr(\frac{S}{WF}, c_f)$, which is exactly the case in our data.

2.5.2 Business Stealing Effects

2.5.2.1 Demand Driven vs. Business Stealing Effects

As discussed earlier, agents enter the market if more commission fees become available. With a constant commission rate, the only way for commission fees to increase is either through an increase in the number of houses for sale or a hike in housing prices. Although the number of agents in a city depends on the total available commissions and not the number of house sales or house prices separately, the welfare implications are quite different for these two paths of increasing the commission fees.

For a fixed number of houses in a market, a price hike will cause an increase in total available commission fees, causing more agents to enter the market. new agent entry however, is socially inefficient; a higher number of agents will compete for the same houses (that required a lower number of agents for brokerage services). Here, the social inefficiency can be observed by the increased competition among agents in cities with higher prices. The new agents would have to *steal* business from the existing ones.

On the other hand, if housing prices are fixed, an increase in the number of real estate agents, due to more houses being available for sale, does not constitute any social inefficiencies. In this scenario, new agents enter the market in response to higher demand for real estate brokerage services.

I use the ratio of number of agents in a city divided by the number of houses sold, $\frac{N}{S}$ to differentiate between business stealing and demand driven entry. This ratio is the number of agents who compete to sell each house on the market and is my measure of competition. In the case of pure business stealing, this fraction goes up as the price of housing increases. However, from equation (2.23), we can see that the fraction will be less sensitive with higher heterogeneity among agents:

$$\frac{N}{S} = B \times \frac{cp}{c_f} \tag{2.34}$$

With higher heterogeneity, B is lower and the effects of change in price will be reflected by lower changes in competition, $\frac{N}{S}$. In other words, heterogeneity blunts the edge of competition by reducing potential entrants incentive to enter the market.

This is so because ex ante, a potential agent is less productive than an incumbent high type agent. Since low type agents exit in equilibrium, any new comer must compete with people who are more likely to ”succeed”.

2.5.2.2 Estimates of Business Stealing Effect

I estimate B from equation (2.34) to measure the business stealing effect. This equation has the number of competitors for an average house is in the left hand side. On the right hand side, we have the normalized commission per house. It predicts that when agents are heterogeneous, even with free entry and no competition in commission rate, not all price increases result in social waste. This is different from the homogeneous case where in absence of price (commission rate here) competition, any price increase will lead to social waste.

To elaborate, consider a market in which agents are homogeneous and the price changes such to lead the commission fees earned on each house to support a new agent (i.e. the new price $p^{'}$ is such that $\frac{cp^{'}}{c_f} = \frac{cp}{c_f} + 1$. Then, since $B = 1$, equation (2.34) predicts that for each house, in a steady state, a new agent will enter the market. However, when agents are heterogeneous and in equilibrium only low type agents exit the market, the same increase in price will not increase the competition at the same level as it did with homogeneous agents. The reason again is potential agents lower expectations for receiving a high productivity shock compared to incumbent (and currently high type) agents.

From the theoretical standpoint equations (2.23) and (2.34) are equivalent. However, they are quite different from empirical point of view. In equation (2.23), the number of houses sold, S explains some of the variation in N. However, S is absent from the explanatory variables in equation (2.34). This guarantees higher standard errors for the estimates of B which can show itself in less frequency of observing significant coefficient in our cross-sections. The following table presents statistically significant[6] coefficients of regressing N/S on cp/c_f:

Dependent variable: $\frac{N}{S}$, Independent variable: $\frac{cp}{c_f}$				
Year	coef.	st. error	z-value	p-value
1990	0.846	0.589	1.438	0.925
1993	0.719	0.472	1.525	0.936
1994	0.612	0.444	1.378	0.916
1995	0.920	0.436	2.108	0.982
1996	0.795	0.515	1.544	0.939
2003	0.557	0.431	1.291	0.902

Table 2.2: Estimates of the effect of increase per house commissions on the level of competition in the market.

The median estimations for B is 0.75, suggesting that around 75% of the increase in commission fees earned per house relative to agents' outside options will go toward increasing the (non-price) competition among agents which is socially wasteful. To put this figure in perspective, I measure the number of real estate agents who enter the market to steal business from existing ones with a \$10,000 increase in average prices in Austin, Texas. With 30,000 houses sold in 2006 in Austin, a \$10,000 increase in average housing price creates \$18 million new commission fees that needs to be divided among real estate agents. With pure business

[6] At 90% significance level

stealing effect ($B = 1$), this will induce 382 potential agents to become Realtors. However, the heterogeneity among agents decreases the incentive for potential agents and only 287 enter the market (with $B = 0.75$). In other words, instead of an increase of 6.5% in the number of agents, we will observe less than 5% increase in this figure.

This measure of competition relies on the assumption of homogeneity in the efforts needed to sell houses across cities. Heterogeneity among cities in that respect blunts the precision of such measure. If houses in high housing cities require more effort to sell, then the difference in competition measure between such cities and those with low housing cost is partly because houses with higher cost "naturally" require more agents, not because high prices can accommodate more agents. This would lead to an overestimation of the business stealing effect.

2.5.3 Extra Wages of the Real Estate Agents

The assumption of free entry into the real estate brokerage industry has two implications: first, within a city, marginal real estate agents' earnings must be equal to that of marginal non-agents. Second, once we control for outside options, real estate agents' earnings should not differ across cities; in particular, it should not depend on variation in housing prices.

Data limitations do not allow for the evaluation of the first implication. But the second implication is testable using our data.

I define the "extra wage" as the average wage earned by the real estate agents in a city minus their outside option ($\frac{R}{N} - c_f$). Assuming no entry cost ($c_e = 0$), we

Dependent Variable: Extra Wage 2000					
	coef	Std Err	z-val	95% Conf. Interval	
log(price)	-0.033	0.329	-0.101	[-0.678	0.612]
cons	0.445	3.818	0.116	[-7.039	7.929]

Table 2.3: Cross-sectional estimation of extra wage on housing price for the year 2000. Theory predicts this coefficient not to be significantly different from zero

can show from equation (2.23) that:

$$\frac{R}{N} - c_f = c_f \left(\frac{1}{B} - 1 \right) \tag{2.35}$$

which requires that the extra wage is positive (because $B \leq 1$), and is independent of the total commissions.

Table (2.3) shows the results of the cross sectional regression of extra wages on the housing prices for the year 2000. The results of this table hold throughout all the cross sections in our data. In the following figure, I present the distribution of the coefficients of the total commissions in the regression of the extra wages on total commissions.

2.5.4 Estimation errors and source of endogeneity

The model implicitly assumes instantaneous adjustment to the market variables of price, outside option of the agents and number of houses for sale. This section addresses the potential problems that can arise in the estimation.

Current and potential real estate agents adjust instantaneously to the housing market variables. This implicit assumption is central to the results in this paper. If there is a delay between the actual exit of agents and observing this in the data,

Figure 2.2: Histogram of estimates of cross-sectional regression of extra wage on total commissions

the relationship will be weaker in reality than predicted and estimated by the model. For instance, consider an agent who exits the market at time t because of unfavorable market conditions, but her license expires at $t + 1$. Although total available commissions are lower, pointing for a lower number of real estate agents; failure to account for her exits results in higher number of agents that would have been predicted if the exit was correctly at t. Similar reasoning can be made for the case of entry and higher available commissions. It is difficult to come up with a measure to observe agents' exit more accurately than setting license expiration as their exit date. In addition, since the salesperson licenses are valid for either one or two years, this effect should not be very significant.

Another source of error in the estimations is the potential endogeneity between agents' outside options and housing prices. In each city, I set agents' outside options as the average annual wage in private sector of that city. It is important to notice that wages and housing prices are not independent. Higher wages usually lead to immigration and higher willingness to pay for housing, boosting its price, which in turn increases total commissions and attracts more agents.

On the other hand, a direct result of increase in agents' outside options is lower incentive to enter the market by the potential agents because of higher attaraction of other sectors. Higher outside options should lead to a lower number of real estate agents in this sense.

Although not very small, I assume the first effect to be negligible. I only consider the second (direct) effect of agents' outside options.

2.6 Conclusion

I develop a simple dynamic model of entry and exit to evaluate the effects of heterogeneity and future uncertainty on social wastes of free entry and business stealing effect and the response of current and potential real estate agents to changes in market variables (price, competitors, number of houses for sale and total home sale value)

In this model agents receive productivity shocks. Commission fees- which are assumed to be the only source of income for real estate agents- are divided according to the number of agents and their productivity shocks. The model predicts a proportional relationship between the number of agents and the value of home sales, while under certain assumptions[7] the proportionality factor is always lower than that in a static and homogeneous model and is decreasing in homogeneity. I construct a Markov Perfect Equilibrium (MPE) in which low type agents exit and high type agents stay in the market and show that this behavior is equivalent to sticky types assumption. The higher probability of an incumbent high type agents in receiving a high productivity shock reduces the size of (potential) pie for those contemplating to enter the market. Therefore, in the MPE in this paper the incentives for potential agents to enter the market are lower with higher heterogeneity.

I test the predictions of the model by applying it to real estate markets in Texas. The dataset is very extensive, it includes the license records of over 210,000 real estate agents. The records include entry and expiration dates, date of birth and address of each individual agent. In addition, I have collected data on average selling price, the number of houses sold and the value of home sales in 42 markets[8] in Texas. Since the licenses have to be renewed every one or two years, we can locate agents in markets and test the predictions of the model in terms of sensitivity of the number of real estate agents to the available commissions in each market.

The results provide strong support for the model. Estimates of the propor-

[7]Namely, the assumption of sticky types which holds when the probabilities of receiving the same productivity shocks for incumbents are higher than the probability of drawing a new productivity shock of that type for potential new agents.

[8]Each MLS is defined as a market

tionality factor are around 0.8 to 0.9. These estimates are consistent with calibration and predictions of the model. In addition, I use the model to explain the effects of heterogeneity on business stealing effect and show that business stealing is lower when heterogeneity is present. When agents are homogeneous, new entries that stem from price increases will be higher. I define the measure of non-price competition in market as the ratio of the number of agents to the number of houses sold in a city and show that with heterogeneity, we need a higher price increase to prompt a certain increase in competition for an average house. Models that assume homogeneity predict that since agents do not compete in commission rates, any increase in home prices will go to social waste. Inclusion of heterogeneity and dynamics in the model allows us to explain why we do not observe such behavior in the data.

A necessary condition for free entry that under free entry the average earnings of real estate agents are independent of housing price differences. Otherwise, I test this condition in the following way: I construct the variable of extra wage as the difference between average earning of real estate agents and their outside options and show that this variable is independent of price variation and value of home sales across cities.

Chapter 3

Information Sharing and Higher Commission Rates in the Two-Sided Real Estate Market

The objective of this paper is to study the determinants of commission rates in the two-sided market of real estate brokerage industry and explain the emergence of the MLS and its impact on commission rates. In addition to their commission rates, real estate agencies decide on their MLS policies as well: they can either list the property with the MLS and share information about it, or not list the property with the MLS. If a property is listed with the MLS, all MLS subscribers can see the listing and send their potential buyers to see that property. Potential buyers can go to any agency to purchase such a property. If the property is *not* listed with the MLS, to buy a house, a buyer must go to the same agency that the seller has signed up with.

Since sellers pay the commission fees, and buyers no longer have to go to the same agency, with MLS listing, buyers choose the closest agency regardless of the commission rates charged by the agencies. Therefore, changes in the commission rates only change the affiliation of the sellers and not that of the buyers. This leads to a softer competition under MLS listing as agencies compete only in the seller side of the market. The softer competition and resulting higher commission rates are desirable to the agencies. They prefer the MLS listing outcome and given the optimal strategies after observing each other's listing decisions, agencies weakly prefer listing to no listing. I show that the one period game has two Nash Equilibria in which either both real estate agencies choose to list their houses with the MLS, or both decide not to list their houses with the MLS. The no listing equilibrium forces buyers to work through that agency's agents and effectively ties the both sides of the market. The higher commission rate equilibrium of the game allows buyers to choose either agency and reduces the competition to the sellers side. Softer competition in turn, results in higher equilibrium commission rates and higher profits along the equilibrium path.

3.1 Introduction

The objective of this paper is to study the determinants of commission rates in the two-sided market of real estate brokerage industry and explain the emergence of the MLS and its impact on commission rates. In addition to their commission rates, real estate agencies decide on their MLS policies as well: they can either list the property with the MLS and share information about it, or not list the property with the MLS. If a property is listed with the MLS, all MLS subscribers can see the listing and send their potential buyers to see that property. Potential buyers can go to any agency to purchase such a property. If the property is *not* listed with the MLS, to buy a house, a buyer must go to the same agency that the seller has signed up with.

Since sellers pay the commission fees, and buyers no longer have to go to the same agency, with MLS listing, buyers choose the closest agency regardless of the commission rates charged by the agencies. Therefore, changes in the commission rates only change the affiliation of the sellers and not that of the buyers. This leads to a softer competition under MLS listing as agencies compete only in the seller side of the market. The softer competition and resulting higher commission rates are desirable to the agencies. They prefer the MLS listing outcome and given the optimal strategies after observing each other's listing decisions, agencies weakly prefer listing to no listing. I show that the one period game has two Nash Equilibria in which either both real estate agencies choose to list their houses with the MLS, or both decide not to list their houses with the MLS. The no listing equilibrium forces buyers to work through that agency's agents and effectively ties the both sides of the market. The higher commission rate equilibrium of the game allows buyers to choose either agency and reduces the competition to the sellers side. Softer competition in turn, results in higher equilibrium commission rates and higher profits along the equilibrium path.

Initially, real estate markets were one-sided markets: sellers contracted with a real estate agency to sell their homes and the agency charged the seller a commission rate for its services. Buyers searched the market for houses they liked. If they liked a house, they would contact the seller's agent to gain access to the house. If they wanted to buy the house, they would negotiate with the seller's agent.

Eventually, real estate agencies starting using Multiple Listing Services centers (MLS). The MLS provided comprehensive information about all of the houses sold by its subscribers. In listing a house on MLS, the seller's real estate agency agreed to share the commission with the agent of the buyer, if he has one. As a result, real estate agents had an incentive to sign up buyers and help them search the market. The MLS transformed the real estate market into a two-sided market in which the real estate agencies are platforms through which buyers and sellers interact. It reduced buyer search costs, increased the probability of a sale, and improved the quality of the matches between buyers and sellers.

Two-sided markets are roughly defined as markets in which one or several platforms enable interactions between end-users, and try to get the two (or multiple) sides ”on board” by appropriately charging each side. (Rochet and Tirole (2004)). Platforms compete and charge user in each sides appropriate fees to facilitate the interaction between the end users in the two sides. One can think about these platforms agents who match or facilitate the match of the end users in different sides. Therefore, consumption by agents in one side of the market has a great deal of complementarity with consumption on the other side. Notice that unlike consumers in a market in which firms produce multiple products, in two-sided markets these ”externalities are not internalized by the end user” (Rochet and Tirole (2004)).

Bernheim and Whinston (1990) analyze the effect of multimarket contact between firms in their ability to collude. Their famous irrelevance result is that ”when markets are identical, firms are identical and technology exhibits constant return to scale, then multimarket contact does not enhance firms' abilities to sustain collusive prices”. If a firm decides to cheat on collusive agreement in one market, it will cheat in all markets. Therefore, in order for multimarket contact to increase the ability of firms to sustain collusion, technology, markets or costs should be different.

This paper takes a different approach. Consumption of the service in one side of the market is a perfect complement of its consumption in the other side of the market. In regular markets, the only contact between firms is through price. In the two-sided markets discussed here however, in addition to the traditional tool of price (commission rate here) real estate agencies can also use sharing information about their properties as another tool to interact with the other agency. This will only lead to a softer competition in the game: firms use information sharing as a tool to reciprocate ”good” behavior of charging high price.

Real estate agencies have the ability to extend their market power on the sellers who sign up with them to the interested buyer by tying the two sides of the market together. I show that although tying two sides of the market is a profitable deviation in the stage game and the forms one Nash equilibrium, firms can achieve higher commission rates and higher profits along the equilibrium path if they untie the sides and reduce the competition to only one side of the market. This goal is achieved through sharing information through shared databases.[1]

3.2 The Model

I model the real estate brokerage industry as a two sided market in which real estate agencies match buyers and sellers of real estate properties (houses). I assume that real estate agencies use a common platform to facilitate transactions. The

[1] commonly known as the Multiple Listing Services or MLS

main question of interest here is the impact of the MLS on equilibrium commission rates.

3.2.1 Environment

Houses are differentiated by their location but are otherwise identical and sell for price p. This price is set by the sellers and is assumed to be exogenous. The locations of the houses are distributed uniformly on the unit interval. Seller x is the owner of a house located x miles from the left end point. Buyers are also distributed uniformly on the unit interval. Buyer x is someone who wants to buy the house at location x. There are two real estate agencies, *l* and *r,* that are located respectively at the left and right end of the unit interval. Agencies have the technology to match buyers and sellers. Sellers sign exclusive contracts with the agencies to sell their houses and buyers sign exclusive contracts with the agencies to help them buy a house. For a given transaction, an agency can represent a buyer, a seller, or both. I differentiate between two types of transactions: internal and external. Transactions are called ”internal” if the buyer and seller are represented by the same agency and they are called ”external” if the buyer and seller are represented by different agencies.

Agencies derive revenues from the commission fees that they charge sellers and buyers for their services. The commission fee is a percentage of value of the house. Each agency charges the same commission rate to its buyers as it does to its sellers. Let c_i denote the commission rate of agency i, $i = l, r$. The seller of the house pays her commission fee and that of the buyer. Thus, a seller who signs with agency i pays $2c_ip$ to agency i on internal transactions and pays $(c_i + c_j)p$ on external transactions. In the latter case, c_ip goes to agency i and c_jp goes to the buyer's agency. The agency's costs of service are normalized to zero.[2]

In addition to setting a commission rate, each agency also has to decide whether or not to list its houses on the MLS. Let $d_i = 1$ if agency i lists its houses with the MLS and $d_i = 0$ is agency i decides not to list its houses on the MLS. This decision is a "sweeping" one: an agency either lists all of its houses or none of them. By listing its houses on the MLS, agency i gives buyers who have signed with the other agency access its houses and forgoes that buyer's commission if the buyer is with the other agency.

Both sellers and buyers incur a linear travel cost in signing with a real estate agency. The travel cost per mile is denoted by t.

3.2.2 Timing

The timing of the game is as follows:

1. Agencies decide whether or not to list their houses on MLS.
2. After observing the listing decisions of the agencies, they then decide on their commission rates.
3. After observing the agencies' commission rates and their listing decisions, sellers choose an agency.

[2]Although one can reasonably argue that it is more costly to match externally than internally.

4. After observing the decisions of the agencies and sellers, buyers choose an agency.

3.2.3 Strategies, Payoffs and Equilibrium

We begin by characterizing the optimal decision of the buyers. Buyer x wants to be matched with seller x. If seller x signs with an agency that chooses to list their properties with the MLS, then buyer x can go to either agency. Since travel is costly, she signs with the agency that is closest to her location. Consequently, if both agencies list with the MLS, then the indifferent buyer is at the midpoint and each agency's share of the buyers market is half. If both agencies choose not to list, then buyer x signs with the agency that lists the house of seller x.

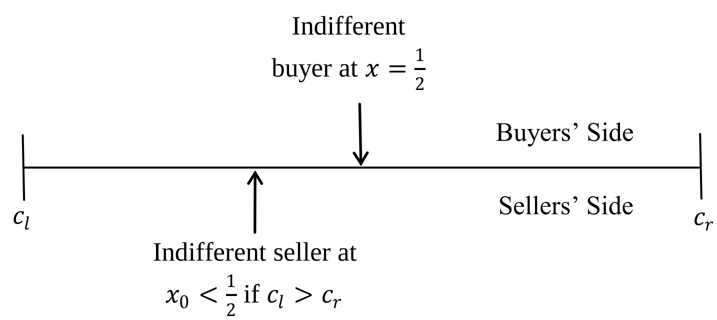

Figure 3.1: Indifferent buyer is always at the midpoint and the only side affected by the commission rates is the sellers side.

Consider next the decision of sellers. Given the decision rule of the buyers, sellers are free to sign with the agency that offers the lowest total costs: travel costs plus commission fees. Let x_0 denote the location of the seller who is indifferent between signing with agency l or r.

3.2.3.1 Case 1: $d_l = d_r = 1$.

Here both firms list properties with the MLS, agencies. Buyers sign with the closest agency independently of the commission rates. As a result, each agency gets half of the buyers. Assume without loss of generality that $c_l \geq c_r$ so the indifferent seller lies to the left of $1/2$. If she signs with agency l, then she pays a commission fee of $2c_l p$; if she signs with agency r, then she pays a commission fee of $(c_l + c_r)p$. Therefore, she is indifferent if

$$\begin{aligned} 2c_l p + tx &= (c_l + c_r)p + t(1-x) && (3.1) \\ \Leftrightarrow x_0^L &= \frac{(c_r - c_l)p}{2t} + \frac{1}{2} && (3.2) \end{aligned}$$

The payoff of agency l is

$$\pi_l(c_r, c_l; 1, 1) = pc_l x_0^L + \frac{1}{2}pc_l \qquad (3.3)$$

and the payoff to agency r is

$$\pi_r(c_r, c_l; 1, 1) = pc_r(1 - x_0^L) + \frac{1}{2}pc_r. \qquad (3.4)$$

3.2.3.2 Case 2: $d_l = d_r = 0$

Here both agencies do not use the MLS. In this case, buyer x has to sign with the agency that sells the house located at x. In other words, buyers will divide themselves between the two agencies in the same way that sellers do. Seller

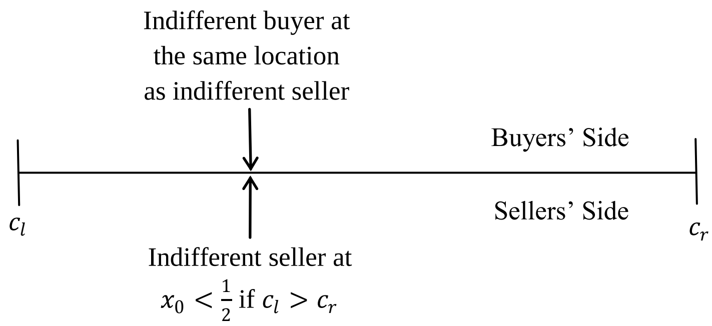

Figure 3.2: The location of indifferent seller and buyer are exactly the same when agencies do not list properties with the MLS.

x chooses the agency that minimizes its total cost: the commission fees and the traveling costs. If seller x signs with agency l, he pays a commission fee of

$$2c_l p + tx.$$

If seller x signs with agency r, he pays a commission fee of

$$2c_r p + t(1 - x).$$

Equating the two commission fees and solving for the location of the indifferent seller yields

$$x_0^N = \frac{2(c_r - c_l)p + t}{2t} \tag{3.5}$$

All sellers (and buyers) with $x \leq x_0^N$ choose agency l and all sellers (and buyers) with $x > x_0^N$ choose agency r. The payoff to agency l

$$\pi_l(c_l, c_r; 0, 0) = 2c_l p x_0^N \tag{3.6}$$

and the payoff of agency r is

$$\pi_r(c_l, c_r; 0, 0) = 2c_r p(1 - x_0^N) \tag{3.7}$$

3.2.3.3 Case 3: $d_l = 0,\ d_r = 1$

Here agency l does not list its houses with the MLS and agency r does. Seller x pays a commission fee of $2c_l p$ if she signs with agency l since she knows that the matching buyer will also sign with agency l. If seller x signs with agency r, then her commission fee depends upon her location. She pays the commission fee $2c_r p$ if her location is to the right of 1/2 and the commission fee $(c_l + c_r)p$ if her location is to the left of 1/2. Therefore, if she is to the left of 1/2, then her location is the same as x_0^L. But, if she is to the right of 1/2. then her location is x_0^N. The payoff of agency l is:

$$\pi_l(c_l, c_r; 0, 1) = \begin{cases} pc_l x_0^L + \frac{1}{2} pc_l \text{ if } c_l > c_r \\ 2pc_l x_0^N \text{ if } c_l < c_r \end{cases} \tag{3.8}$$

and the payoff to agency r

$$\pi_r(c_r, c_l; 0, 1) = \begin{cases} pc_r(1 - x_0^L) + \frac{1}{2} pc_r \text{ if } c_l > c_r \\ 2pc_r(1 - x_0^N) \text{ if } c_l < c_r \end{cases}$$

The derivation of payoff when agency l lists with the MLS and agency r does not list with the MLS is similar to Case 3.

Given the optimal decision rules of the buyers and sellers, the subgame perfect equilibrium of the game consists of a profile of commission rates $(c_l^*(d_i, d_j), c_r^*(d_i, d_j))$

such that

$$\pi_i(c_i^*, c_i^*; d_i, d_j) \geq \pi_i(c_i, c_j^*; d_i, d_j), \tag{3.9}$$

and a pair of MLS listing decisions (d_l^*, d_r^*) such that:

$$\pi_i(c_i(d_i^*, d_j^*), c_j(d_i^*, d_j^*); d_i^*, d_j^*) \geq \pi_i(c_i(d_i, d_j^*), c_j(d_i, d_j^*); d_i, d_j^*)$$
$$\forall i, j \in \{l, r\}\ j \neq i,\ c_i \in \mathbb{R}^+, d_i \in \{0, 1\}.$$

The first condition states that the pair of commission rates in each of the four possible subgames is a Nash equilibrium. The second condition states that the MLS listing decisions are mutual best responses when agencies anticipate that the commission rates in the second stage are determined by Nash equilibrium play.

3.3 The Equilibrium

In this section, I solve for the subgame perfect equilibrium of the game. I first solve for the Nash equilibrium of each of the four possible subgames that can arise from the listing decisions of the agencies. Using the equilibrium payoffs for these subgames, I then solve for the equilibrium listing decisions.

3.3.1 The $(d_l = d_r = 0)$ Subgame

I consider first the subgame in which both agencies choose not to use the MLS. In this case, there are only internal transactions in the market. Agency i

chooses c_l to

$$\max_{c_l} \pi_l(c_l, c_r; 0, 0) = 2c_l p x_0^N = 2c_l p\left(\frac{2(c_r - c_l)p + t}{2t}\right) \tag{3.10}$$

$$\text{FOC} : \frac{p}{t}\left(2(c_r - c_l)p + t - 2pc_l\right) = 0 \tag{3.11}$$

Therefore, the best response of agency l is given by

$$c_l = \frac{2c_r p + t}{4p} \tag{3.12}$$

Substituting this equation into equation (3.12) determines the location of the indifferent seller:

$$x_0^N = \frac{2(c_r - c_l)p + t}{2t} = \frac{2pc_r + t}{4t}. \tag{3.13}$$

Imposing the boundary condition of $x_0^N \leq 1$, equation (3.12) is agency l's best response only if $c_r \leq \frac{3t}{4p}$. For higher values of c_r, chooses its commission rate such that the indifferent seller is at $x_0^N = 1$. It gets all the sellers (and buyers) and charges them the highest commission rate that it can. So the best response of agency l is:

$$c_l^N = BR^N(c_r) = \begin{cases} \frac{2c_r p + t}{4p} & \text{if} \quad 0 \leq c_r \leq \frac{3t}{2p} \\ c_r - \frac{t}{2p} & \text{if} \quad c_r > \frac{3t}{2p} \end{cases} \tag{3.14}$$

Notice that for positive values of c_r the other boundary condition for the indifferent seller ($x_0^N \geq 0$) is automatically satisfied.

Symmetry implies that agency r's best response to the commission rate of c_l: [3]

$$c_r^N = BR^N(c_l) = \begin{cases} \frac{2c_l p + t}{4p} & \text{if} \quad 0 \leq c_l \leq \frac{3t}{2p} \\ c_l - \frac{t}{2p} & \text{if} \quad c_l > \frac{3t}{2p} \end{cases} \tag{3.15}$$

[3] In this case, the location of the indifferent seller in terms of c_l is $x_0^{NL} = \frac{3t - 2pc_l}{4t}$

The symmetric equilibrium of this subgame is

$$c_l = c_r = \frac{t}{2p}. \tag{3.16}$$

Thus, each agency gets half of the sellers and buyers, and their equilibrium profits are

$$\pi_r^N = \pi_l^N = \frac{t}{2}. \tag{3.17}$$

3.3.2 The $(d_l = d_r = 1)$ Subgame

I consider next the subgame in which both agencies list their houses with the MLS. Recall that the indifferent buyer is at the midpoint and the location of the indifferent seller is:

$$x_0^L = \frac{p(c_r - c_l)}{2t} + \frac{1}{2} \tag{3.18}$$

Agency l chooses its commission rate to

$$\max_{c_l} \quad pc_l \left(\frac{(c_r - c_l)p + t}{2t} \right) + \frac{1}{2} pc_l \tag{3.19}$$

$$\text{FOC} \quad : \quad \frac{p}{2t} \left((c_r - c_l)p + 2t - pc_l \right) = 0 \tag{3.20}$$

Solving the first order condition yields

$$c_l = \frac{pc_r + 2t}{2p}. \tag{3.21}$$

Substituting this equation into equation (3.18) determines the location of indifferent seller:

$$x_0^L = \frac{p(c_r - c_l)}{2t} + \frac{1}{2} = \frac{pc_r}{4t}$$

As in the previous case, we can impose the boundary condition $x_0^L \leq 1$ and obtain $c_r \leq 4t/p$. For $c_r > 4t/p$, agency l sets its commission rate such that $x_0^L = 1$. Thus, the best response function of agency l is

$$c_l^L = BR^L(c_r) = \begin{cases} \frac{pc_r+2t}{2p} & \text{if} \quad 0 < c_r \leq \frac{4t}{p} \\ c_r - \frac{t}{p} & \text{if} \quad c_r > \frac{4t}{p} \end{cases} \tag{3.22}$$

Symmetry implies that agency r's best response function:

$$c_r^L = BR^L(c_l) = \begin{cases} \frac{pc_l+2t}{2p} & \text{if} \quad 0 < c_l \leq \frac{4t}{p} \\ c_l - \frac{t}{p} & \text{if} \quad c_l > \frac{4t}{p} \end{cases} \tag{3.23}$$

Thus, the symmetric equilibrium to the subgame in which both agencies use the MLS is

$$c_l = c_r = \frac{2t}{p}. \tag{3.24}$$

Each agency gets half of the buyers and sellers and earns

$$\pi_l = \pi_r = 2t. \tag{3.25}$$

3.3.3 The $(d_l = 0,\ d_r = 1)$ Subgame

When the left agency does not list with the MLS while the agency r does, the payoff to agency l is determined by the following set of equations:

$$\pi_l(c_l, c_r; 0, 1) = \begin{cases} pc_l x_0^L + \frac{1}{2}pc_l \text{ if } c_l \geq c_r \\ 2pc_l x_0^N \text{ if } c_l \leq c_r \end{cases} \tag{3.26}$$

and the payoff to agency r

$$\pi_r(c_r, c_l; 0, 1) = \begin{cases} pc_r(1 - x_0^L) + \frac{1}{2}pc_r \text{ if } c_l \geq c_r \\ 2pc_r(1 - x_0^N) \text{ if } c_l \leq c_r \end{cases}$$

If $c_l \geq c_r$ then the payoffs to the agencies are:

$$\pi_l(c_l, c_r; 0, 1) = pc_l x_0^L + \frac{1}{2}pc_l \tag{3.27}$$

$$\pi_r(c_r, c_l; 0, 1) = pc_r(1 - x_0^L) + \frac{1}{2}pc_r \tag{3.28}$$

From the previous section, we know that in this situation, agencies best respond to each other when $c_l = c_r = 2t/p$. Each firm receives the payoff of $2t$.

This outcome however, is not the equilibrium of this subgame. Consider the following perturbation by the left agency: instead of charging $c_l = 2t/p$, the left agency charges slight less commission rate of $c_l = (2 - \epsilon)t/p$. Keeping $c_r = 2t/p$, the payoff to the left agency is now determined by:

$$\begin{aligned}
\pi_l\left(\frac{(2-\epsilon)t}{p}, \frac{2t}{p}; 0, 1\right) &= 2pc_l x_0^N \\
&= 2 \times p \times \frac{(2-\epsilon)t}{p} \times \left(\frac{1}{2} + \frac{p}{t}(c_r - c_l)\right) \\
&= 2 \times (2-\epsilon)t \times \left(\frac{1}{2} + \frac{p}{t} \times \frac{\epsilon t}{p}\right) \\
&= (2 + 3\epsilon - 2\epsilon^2)t \\
&> 2t \text{ for } \epsilon \simeq 0
\end{aligned}$$

In other words, agency l can increase its payoff by undercutting agency r. Therefore, the agencies cannot reach an equilibrium by best responding to each other while splitting the buyers side and agency l charging $c_l \geq c_r$. So we must have $c_l \leq c_r$ and the payoffs to the agencies are determined by the following equations:

$$\pi_l(c_l, c_r; 0, 1) = 2pc_l x_0^N \tag{3.29}$$

$$\pi_r(c_r, c_l; 0, 1) = 2pc_r(1 - x_0^N) \tag{3.30}$$

The equilibrium of this subgame is similar to the equilibrium of the subgame of $d_l = d_r = 0$ and happens when both agencies charge $c_l = c_r = t/2p$ and each receive the payoff of $t/2$.

The results derived above indicate that the equilibrium commission rates when both agencies use the MLS are four times higher than they are when one or both agencies do not use the MLS. The intuition is as follows. When the agencies use the MLS, they compete for the marginal seller but not for buyers. Each agency gets one half of the buyers, and the commissions that go with those buyers, independently of the commission rates that they set. This is not the case when agencies do not use the MLS. In competing for sellers, they also compete for buyers since buyers have to sign with the agency that sells the houses that they want. In other words, the agencies only compete on the sellers side of the market when they use the MLS whereas they compete on both sides of the market when they do not use the MLS. The competition is ”stiffer” in the latter case, which leads to lower commission rates.

3.3.4 Equilibrium

The following proposition characterizes the set of (pure) subgame perfect equilibria.

Proposition 3.3.1. *When agencies decide on their listing policies before their commission rates, then the stage game has two Nash Equilibria in which both agencies list properties with the MLS or both do not list properties with the MLS and charge commission rates specified by equations (3.16) and (3.24) respectively.*

Proof. The payoff matrix for agencies when they decide on their commission rates after they know about each other's MLS policies is summarized in the following table:

Agency R \ **Agency L**	Listing	No Listing
Listing	$2t,\ 2t$	$\frac{t}{2},\ \frac{t}{2}$
No Listing	$\frac{t}{2},\ \frac{t}{2}$	$\frac{t}{2},\ \frac{t}{2}$

Table 3.1: Payoff matrix for Agencies in the commission rate setting stage.

As table (3.1) shows, listing is a best reply to listing and not listing is a best reply to not listing. Therefore, the game has two pure equilibria. Q.E.D. □

The propositions establishes that two outcomes are possible. One outcome is that both agencies list their houses with the MLS, charge commission rates of $c_l = c_r = 2t/p$, and receive the payoff of $2t$. The second outcome is that neither agency lists its houses with the MLS, charge commission rates $c_l = c_r = t/2p$, and receive the payoff of $t/2$. Both agencies earn higher payoffs when they use the MLS. In fact, listing is a weakly dominant strategy. Therefore, it is not too surprising that agencies in the real estate market quickly adopted the MLS when it became available.

3.4 The Assumptions

In this section I discuss how the results depend upon the assumptions of the model.

3.4.1 Timing

The preceding analysis assumes that the agencies can commit to using the MLS *before* they set the commission rates. An alternative approach is to assume that the agencies choose their listing policies and commission rates independently and simultaneously. Since the optimal decision rules of the buyers and sellers depend only upon the listing decisions and the commission rates, they are not affected by the change in the timing of the decisions by the agencies. Therefore, given these rules, the game is a simultaneous move game. A Nash equilibrium in this game is a profile $(c_l^*, d_l^*, c_r^*, d_r^*)$ such that

$$\pi_i(c_i^*, c_i^*; d_i^*, d_j^*) \geq \pi_i(c_i, c_j^*; d_i, d_j^*), \forall i, j \in \{l, r\}\ j \neq i,\ c_i \in \mathbb{R}^+, d_i \in \{0, 1\}. \tag{3.31}$$

The following proposition establishes that the unique Nash equilibrium consists of both agencies not listing their houses with the MLS and charging the "low" commission rate.

Proposition 3.4.1. *When agencies decide on the commission rates and MLS policies simultaneously, the unique Nash equilibrium consists of* $d_l^* = d_r^* = 0$ *and* $c_l^* = c_r^* = t/2p$

Proof. Fix c_r and suppose $d_r = 1$. The payoff to agency l if it lists with the MLS and sets $c_l = BR^L(c_r)$ is given by

$$\pi_l^L(c_r; 1, 1) = \frac{(pc_r + 2t)^2}{8t}.$$

Its payoff if it does not list with the MLS and sets $c_l = BR^N(c_r)$ is

$$\pi_l^N(c_r; 0, 1) = \frac{(2pc_r + t)^2}{8t}.$$

Therefore, agency l prefers not to list if and only if

$$\frac{(2pc_r + t)^2}{8t} > \frac{(pc_r + 2t)^2}{8t} \Leftrightarrow c_r > \frac{t}{p}. \tag{3.32}$$

Recall that the $c_l = c_r = 2t/p$ are the equilibrium commission rates when both agencies are committed to listing their houses on the MLS. Since this price exceeds $\frac{t}{p}$, there is no Nash equilibrium in which both firms list.

Suppose $d_r = d_l = 0$. Then the equilibrium commission rates are $c_l = c_r = t/2p$. Since this price is strictly less than t/p, the inequality derived above implies that each agency prefers not to list if the other agency does not list. Q.E.D. □

The fact that virtually all real estate agencies use the MLS supports the assumption that agencies can commit to using the MLS prior to setting commission rates.

3.4.2 ”Naive” buyers

The game discussed in this paper nests the case of ”naive” buyers: buyers who allocate themselves to the closest agency regardless of their commission rate and MLS listing policies. The outcome of such a game would be the same as the outcome of the game with listing: naive buyers simply do not respond to changes in the commission rates and therefore any change in commission rates would only affect the seller. As a result, the competition is only in the sellers side. The only Nash equilibrium is the one in which both firms list their properties in MLS.

3.4.3 Heterogeneous Homes

Except for their location, houses are assumed to be identical here. Buyers are also assumed to be identical except for their location. In addition, buyers are assumed to know exactly what they want and the probability of matching buyer x with the house located at x is one. These assumptions lead to the conclusion that in the equilibria of the game, all houses are sold and all buyers are matched to their desired house.

In reality, houses differ in characteristics other than location and the buyers need to search on more than one dimension to determine the best match. Uncertainty about which house to buy in turn means that the buyer does not know which real estate agency is listing her best match. Hence, under no MLS listing, agencies could end up with a bulk of unmatched buyers and unmatched houses. This issue is not a problem if real estate agencies use MLS as one could reasonably argue that the probability of finding a match is higher if buyers could search over the properties of two agencies rather than just one[4]. Therefore, to reduce the size of their unmatched buyers and properties, agencies' incentives to list their properties with MLS is higher when the matching is probabilistic. One needs to explicitly account for the search process and its cost to see if MLS listing can be supported as part of

[4]We could assume that internal search costs are zero but it is costly if buyers search other firms' listing with the MLS via a real estate agency.

a subgame perfect equilibrium path in the stage game.

Under probabilistic matching, social welfare is not independent of MLS listing policies anymore. MLS listing increases the chances of sale for a given house and higher number of houses are expected to be sold in the market when agencies use MLS. There is a loss for each house that could have been sold under MLS listing (that is not sold because agencies decide to not list). On the other hand, the search, especially external search over the MLS, is costly. The net effect of MLS listing policies depends at least on these two parameters: the size of the search cost and its effect of search in increasing the probability of sale.

3.4.4 New Entrant

Sunk costs are not explicitly modeled in this paper. Ability of potential entrants to enter the market depends on the size of the sunk cost of entry as well as the post-entry game payoffs. One interesting question is whether the incumbent agencies can prevent entry by credibly threatening not to show homes listed with entrant to their buyers. In a model where houses are identical, I do not believe that this threat is credible: sellers will switch, knowing the buyers will switch with them. Except for their location, houses that are represented by the new entrant are exactly the same the ones represented by incumbents and as likely to sell.

However, if buyers are uncertain about their preferences, then they may not search houses listed by the entrant which in turn means that sellers will not list with the entrant. It becomes a self-fulfilling expectation. In this situation, the decision of incumbent agencies to exclude entrant's houses decrease the probability of finding a match for the houses that are listed with the entrants. The outcome of the new game depends on sunk costs as well as the search process and probability of finding a match, but a successful entry blocking is more likely when homes are more heterogeneous.

3.5 Conclusion and Extension

In the preceding sections we have analyzed competition between real estate agencies in a static model of the real estate market. We have shown that there are two equilibria: the high commission rate equilibrium in which the agencies use the MLS and the lower commission rate equilibrium in which the agencies do not use the MLS. When agencies use the MLS, buyers are free to sign with the agency that they prefer rather than the agency that sells the house that they want to buy. As a result, the agencies share the market for buyers and compete only in the market for sellers. By contrast, when agencies do not use the MLS, the two sides of the market are tied: a buyer has to sign with the agency that sells the house that she wants. This means that whenever an agency reduces its commission rate and converts a seller, it also attracts the buyer who wants to purchase the seller's house. The agencies compete on both sides of the market, which leads to stiffer competition and lower commission rates. From the agencies' perspective, the MLS allows them to soften competition and earn higher commission rates. By restricting firms to internal transactions, policymaker can force agencies to compete more aggressively and lower commission rates.

Appendices

Appendix A

Appendix to Chapter 2

A.1 Proofs of Lemmas and Proposition

A.1.1 Proof of Lemma 2.3.1

Proof. It is necessary and sufficient to show that equations (2.21) and (2.22) are equivalent to:

$$\theta V_H + (1-\theta)V_L \geq 0$$

$$\theta V_L + (1-\theta)V_H \leq 0$$

Starting with the first inequality:

$$\begin{aligned} \theta V_H + (1-\theta)V_L &= \left(\frac{\theta(\phi^H + \beta(1-\theta))}{1-\beta\theta} + (1-\theta)\right) R' \\ &- c_f\left(\frac{\theta(1+\beta(1-\theta))}{1-\beta\theta} + (1-\theta)\right) \\ &= \left(\frac{1+(\theta-1)\theta}{1-\beta\theta}\right) R' - \frac{c_f}{1-\beta\theta} \\ &= \frac{c_f\,(1-\beta\theta+\beta\eta)\,(\phi^H-1)(\theta-\eta)}{(1-\beta\theta)(1-\beta\theta+\beta\eta+\phi^H\eta-\eta)} > 0 \end{aligned}$$

To pass from second to third line, we need to use the equilibrium value of R' from equation (A.5). Calculations show that persistence in types is necessary and sufficient for the high type agents to stay for the next period.

For the second inequality:

$$\begin{aligned}\theta V_L + (1-\theta)V_H &= \left(\frac{(1-\theta)(\phi^H + \beta(1-\theta))}{1-\beta\theta} + \theta\right) R^{'} \\ &- c_f\left(\frac{(1-\theta)(1+\beta(1-\theta))}{1-\beta\theta} + \theta\right) \\ &\vdots \\ &= \frac{c_f(\phi^H - 1)(1-\theta-\eta)}{(1-\beta\theta+\beta\eta+\phi^H\eta-\eta)} \leq 0\end{aligned}$$

Therefore the conditions stated in the lemma are necessary and sufficient for the existence and uniqueness of the equilibrium. □

A.1.2 Proof of Proposition 2.3.2

Proof. We "construct" an equilibrium in which agents exit in L state but remain in the market in H state.

Notice that by defining $R^{'} \equiv \frac{R}{\phi^H N^H + N^L}$, value functions become:

$$V_H = \phi^H R^{'} - c_f + \beta(\theta V_H + (1-\theta)V_L) \tag{A.1}$$

$$V_L = R^{'} - c_f \tag{A.2}$$

We can solve for V_L and V_H from equations (A.1) and (A.2):

$$V_L = R^{'} - c_f \tag{A.3}$$

$$V_H = \frac{R^{'}(\phi^H + \beta(1-\theta)) - c_f(1+\beta(1-\theta))}{1-\beta\theta} \tag{A.4}$$

The solution for $R^{'}$ from the free entry condition (equation (2.18)) is:

$$R^{'} = \frac{(1-\beta\theta+\beta\eta)c_f + (1-\beta\theta)c_e}{1-\beta\theta+\beta\eta+\eta\phi^H-\eta} \tag{A.5}$$

For simplicity, I assume that $c_e = 0$.[1] With this assumption, we can solve for the equilibrium number of real estate agents as follows:

$$N = \left(\frac{1 - \beta\theta + \beta\eta + \eta\phi^H - \eta}{1 + \eta\phi^H - \theta}\right)\left(\frac{1 + \eta - \theta}{1 - \beta\theta + \beta\eta}\right)\frac{R}{c_f} \tag{A.6}$$

Set:

$$A(\phi^H) = \frac{1 - \beta\theta + \beta\eta + \eta\phi^H - \eta}{1 + \eta\phi^H - \theta} \tag{A.7}$$

$$L = \frac{1 + \eta - \theta}{1 - \beta\theta + \beta\eta} \tag{A.8}$$

and we will have:

$$N = A(\phi^H) \times L \times \frac{R}{c_f} \tag{A.9}$$

□

A.1.3 Proof of Lemma 2.3.3

Proof. Notice that with the assumption of $\theta > \eta$ (the new entrants are less likely to draw a high shock than current H types), $A(\phi^H)$ is decreasing in ϕ^H because:

$$\frac{dA}{d\phi^H} = \frac{(1 - \beta)\eta(\eta - \theta)}{(1 + \eta\phi^H - \theta)^2} < 0 \tag{A.10}$$

Since by definition, $\phi^H \geq 1$, therefore, the maximum value of $A(\phi^H)$ is obtained when $\phi^H = 1$:

$$A(1) = \frac{1 - \beta\theta + \beta\eta}{1 + \eta - \theta} = \frac{1}{L} \Rightarrow A(\phi^H)L \leq 1 \tag{A.11}$$

[1] Results regarding negative V_L and satisfying basic inequalities hold stronger with positive entry cost of c_e.

For $\phi^H = 1$:

$$N = A(1) \times L \times \frac{R}{c_f} = \frac{R}{c_f} \tag{A.12}$$

$\square$

A.2 More types and more states

In this subsection, I consider the effect of extending the model to include more types for the agents and more than one realization for R. I show that in the steady-state equilibrium, having more states (more than one R) does not change agents' decision on whether to stay in or leave the market. The main reason is the assumption of instantaneous adjustment to the market variables by the agents.

I assume that there are L states that could be realized and agents' types could be one of possible K types. Transition probability matrix between the realizations of states of R is given by $\Pi = [\pi_{lj}]_{L\times L}$ $\Pr(R' = R_j | R = R_l) = \pi_{lj}$. Agents' types evolve according to the following matrix: $\Theta = [\theta_{ki}]_{K\times K}$. A type i agent's problem in state j is to maximize its value function of:

$$V_{ij} = \frac{R_j}{N_j}\phi_i - c_f + \beta \max\left\{0, \sum_{k=1}^{K}\sum_{l=1}^{L} \theta_{ki}\pi_{lj}V_{kl}\right\} \tag{A.13}$$

Lemma A.2.1. *In stationary equilibrium,* $\frac{R_j}{N_j}$ *is constant and* V_{ij} *is independent of* j*; or;* $V_{ij} = V_{is}$ *where* $j \neq s$.

Proof. I show that $\frac{R_j}{N_j}$ being constant is consistent with the stationary equilibrium and results in the stationary equilibrium of the case where there is only one state for

R. Since there is only one solution to (A.13), we can conclude that the solution for the one state problem solves the multi state problem and $\frac{R_j}{N_j}$ has to be constant.

Assume $\frac{R_j}{N_j} = \frac{R}{N}$ is constant $\forall\, j \in \{1, \cdots, M\}$

$$\begin{aligned} V_{ij} = V_i &= \frac{R_j}{N_j}\phi_i - c_f + \beta \max\left\{0, \sum_{k=1}^{N}\sum_{l=1}^{M}\theta_{ki}\pi_{lj}V_{kl}\right\} && \text{(A.14)} \\ &= \frac{R}{N}\phi_i - c_f + \beta \max\left\{0, \sum_{k=1}^{N}\theta_{ki}V_k\sum_{l=1}^{M}\pi_{lj}\right\} && \text{(A.15)} \\ &= \frac{R}{N}\phi_i - c_f + \beta \max\left\{0, \sum_{k=1}^{N}\theta_{ki}V_k \times 1\right\} && \text{(A.16)} \end{aligned}$$

Therefore, a solution for

$$V_i = \frac{R}{N}\phi_i - c_f + \beta \max\left\{0, \sum_{k=1}^{N}\theta_{ki}V_k\right\} \text{ for } i \in \{1, \cdots, N\} \tag{A.17}$$

is also a solution to (A.13). □

With the assumption of instantaneous adjustment to the market variables, agent's current and future payoffs do not depend on the state of the market. The number of agents in the market N differs across different states because different states imply different total commissions. The number of real estate agents, N adjusts such that the ratio of $\frac{R}{N}$ remains the same across the possible states of the world.

Appendix B

Appendix to Chapter 3

B.1 Repeated Game Equilibrium

One of the implications of the static subgame perfect equilibrium when agencies set their MLS policies and commission rates simultaneously is that the MLS is never used. This section explains the emergence of MLS as part of a the subgame perfect equilibrium in the repeated game in such environment and outlines the conditions under which listing can be supported as part of a repeated game equilibrium. It is clear that the repetition of the stage game subgame prefect equilibrium is a subgame perfect equilibrium for the repeated game. I show that the agencies can use that equilibrium as the punishment for deviating from the listing equilibrium and support MLS listing in the repeated game's subgame perfect Nash Equilibrium.

Proposition B.1.1. *For the discount rate of higher than 2/5, the MLS policy of listing properties is a subgame perfect Nash Equilibrium of the repeated game.*

Proof. Agencies earn the profit of $\pi = 2t$ when charge the commission rate of $c_l = c_r = 2t/p$. This outcome is a Nash Equilibrium if agencies are required to list their properties with the MLS. In the static game however, firms have the incentive to not share their properties in the MLS and deviate from this constrained Nash

Equilibrium. Such deviation brings the profits of $pi^{cheating} = 3t$ to the deviating agency while the other agency earns zero.

If firm j implemented the following strategies in period $l \in \{0, 1, \cdots, n - 1\}$:

$$c_j = \frac{2t}{p},\ \ d_j = 1,\ \ \forall l \leq n - 1$$

then firm i will choose the following strategies in period n:

$$c_i = \frac{2t}{p},\ \ d_i = 1,\ \ \text{in period } n$$

otherwise:

$$c_i = \frac{t}{2p},\ \ d_i = 0,\ \ \text{in period } n$$

Firm i charges commission rate. consistent with listing equilibrium as long as the other firm does the same. If firms deviate from such strategy, they are back to no-MLS-listing equilibrium.

On the equilibrium path, as long as firms charge $c = 2t/p$, they earn the profit of $\pi_{nf} = 2t$. By deviating, firms charge the commission rate of $c = \frac{3t}{2p}$ and earn the profits of $\pi^{cheating} = 3t$ for one period, after which, firms revert back to no listing equilibrium and earn $\pi_f = t/2$ forever afterward. Firms must be "patient" enough to be able to avoid deviation. In fact, discount rate must satisfy the following inequality:

$$\frac{\pi^L}{1-\beta} > \pi^{cheating} + \frac{\beta}{1-\beta} \times \pi^{NL} \tag{B.1}$$

$$\frac{2t}{1-\beta} > 3t + \frac{\beta}{1-\beta} \times \frac{t}{2} \tag{B.2}$$

$$\Rightarrow \beta > \frac{2}{5} \tag{B.3}$$

□

Using grim-trigger strategy and reversion to the no-listing subgame perfect Nash Equilibrium (which has the lower profits) as punishment, firms can support listing equilibrium in the repeated game for a wide range of discount rate.

Bibliography

[1] Bernheim, B. Douglas, and Whinston, Michael D., 1990, *"Multimarket contact and collusive behavior"*, The RAND Journal of Economics, Vol. 21, pp. 1-26

[2] Berry, Steve and Waldfogel, Joel, 1999, *"Free Entry and Social Inefficiency in Radio Broadcasting"*, The RAND Journal of Economics, 30, pp. 397-420.

[3] Bresnahan, Timothy F., and Reiss, Peter C., 1991, *"Entry and Competition in Concentrated Markets."* Journal of Political Economy, Vol. 99, 9771009.

[4] Cabral, Luis M. B., 2000, *"Introduction to Industrial Organization"*, MIT Press.

[5] Collard-Wexler, Allan, 2006, *"Plant Turnover and Demand Fluctuations in the Ready-Mix Concrete Industry"*, Working paper, NYU.

[6] Dixit, Avinash K., and Stiglitz, Joseph E., 1997, *"Monopolistic Competition and Optimum Product Diversity.* American Economic Review, Vol. 67, pp. 297308.

[7] Dixit, Avinash,1989, *"Entry and Exit Decisions under Uncertainty"*, Journal of Political Economy, Vol. 97, No. 3. pp. 620-638.

[8] Han, Lu and Hong, Seung-Hyun, 2008, *"Entry and Inefficiency in the Real Estate Brokerage Industry"*, Working Paper.

[9] Hopenhayn, Hugo A., 1992, *"Entry, Exit, and firm Dynamics in Long Run Equilibrium"*, Econometrica, Vol. 60, No. 5, pp. 1127-1150.

[10] Hsieh, Chang-Tai and Moretti, Enrico, 2003, *"Can Free Entry Be Inefficient? Fixed Commissions and Social Waste in the Real Estate Industry"*, Journal of Political Economy, vol. 111, no. 5, pp. 1076-1122.

[11] Mankiw, N. Gregory, and Whinston, Michael D., 1986, *"Free Entry and Social Inefficiency."* Rand Journal of Economics, Vol. 17, pp. 4858.

[12] Rochet Jean-Charles, and Tirole, Jean, 2003, *"Platform Competition in Two-Sided Markets"*, Journal of the European Economic Association, Vol. 1, pp. 990-1029.

[13] Rochet Jean-Charles, and Tirole, Jean, 2004, *"Two-sided markets: an overview"*, mimeo.

[14] Spence, A. Michael, 1976, *"Product Differentiation and Welfare."* American Economic Review, Papers and Proceedings, Vol. 66, pp.40714.

[15] U.S. Department of Housing and Urban Development, 2007, *"U.S. Housing Market Conditions, 4th Quarter 2006"*.

[16] Whinston, Michael D., 1990, *"Tying, Foreclosure, and Exclusion"* The American Economic Review, Vol. 80, pp. 837-859

Vita

Seyed Parviz Gheblealivand was born in Tehran, Iran in February 1979. After attending Alborz High School from 1993 to 1997 in Tehran, Iran, he entered Sharif University of Technology and receive his Bachelor of Science in Civil Engineering in 2001, and his Master in Business Administration (MBA) in 2003 from that school. In August 2003 he enrolled into the University of Texas at Austin for the Graduate Program in Economics.

Permanent address: No. 52, Habibzadegan St.
Azadi Ave.
Tehran 14588
Iran

This dissertation was typeset with LaTeX[†] by the author.

[†]LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.