Title
The economics of real estate brokerage and contracts

Permalink
https://escholarship.org/uc/item/9fc0608j

Author
Wignall, Christopher David

Publication Date
2009

Peer reviewed|Thesis/dissertation
The Economics of Real Estate Brokerage and Contracts

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Economics

by

Christopher David Wignall

Committee in charge:

Professor Joel Watson, Chair
Professor Vincent Crawford
Professor Gordon Dahl
Professor Branislav Slantchev
Professor Chris Woodruff

2009
The dissertation of Christopher David Wignall is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2009
DEDICATION

To E, A, and M.
# TABLE OF CONTENTS

Signature Page ........................................ iii
Dedication ........................................ iv
Table of Contents ....................................... v
List of Figures ......................................... vi
List of Tables ......................................... vii
Acknowledgements ..................................... viii
Vita and Publications ................................... ix
Abstract of the Dissertation ........................... x

Chapter 1  Word-of-Mouth Learning in Social Networks .... 1
  1.1 Introduction .................................. 1
  1.2 Conceptual Framework ...................... 3
  1.3 Background and Data ..................... 4
  1.4 Estimation and Identification .......... 8
  1.5 Evidence of Social Learning in Wards .... 10
  1.6 Social Learning Via Direct Communication 12
  1.7 Conclusion ................................ 14

Chapter 2  Hold-up and Durable Trading Opportunities .... 23
  2.1 Introduction ................................ 23
  2.2 The Benchmark Model of a Nondurable Trading Opportunity .... 26
  2.3 The Model of a Durable Trading Opportunity .... 29
  2.4 Clarification of the Outside Option Principle .... 32
  2.5 Strong Implementation with Non-Stationary Contracts .... 37
  2.6 Conclusion ................................ 38
  2.7 Appendix - Proof of Outside Option Results .... 40

Chapter 3  Efficiency and Entry in the Real Estate Brokerage Industry .... 43
  3.1 Introduction ................................ 43
  3.2 Real Estate Brokerage .................... 45
  3.3 Theoretical Framework ................... 50
  3.4 Empirical Results ......................... 52
  3.5 Conclusion ................................ 56
LIST OF FIGURES

Figure 1.1: Ward assignment of homes in a Utah County neighborhood. . . 16
Figure 1.2: Histogram of Distances between neighbors . . . . . . . . . . . 17

Figure 2.1: Regions of the space of outside-option payoff vectors. . . . . 35

Figure 3.1: Histogram - Commission . . . . . . . . . . . . . . . . . . . . . . 64
Figure 3.2: Scatter Plot - Commission vs. Sales Price . . . . . . . . . . . 64
Figure 3.3: Histogram - Listing Commission . . . . . . . . . . . . . . . . . . 65
Figure 3.4: Scatter Plot - Listing Commission vs. Sales Commission . . . 65
Figure 3.5: Model Equilibrium . . . . . . . . . . . . . . . . . . . . . . . . . 66
Figure 3.6: Time Series - Price and Quantity . . . . . . . . . . . . . . . . 67
Figure 3.7: Time Series - Price and Number of Agents . . . . . . . . . . 68
Figure 3.8: Time Series - Number of Agents and Quantity . . . . . . . . 68
Figure 3.9: Time Series - Price and Wages . . . . . . . . . . . . . . . . . . 69
Figure 3.10: Time Series - Wages . . . . . . . . . . . . . . . . . . . . . . . . 70
Figure 3.11: Time Series - Productivity . . . . . . . . . . . . . . . . . . . . 70
Figure 3.12: Time Series - Price and Productivity . . . . . . . . . . . . . . 71
Figure 3.13: Time Series - Number of Agents and Productivity . . . . . 71
Figure 3.14: Time Series - Quantity and Productivity . . . . . . . . . . . 72
# LIST OF TABLES

| Table 1.1: | Data Summary Statistics - Listings | 15 |
| Table 1.2: | Data Summary Statistics - Neighbors | 18 |
| Table 1.3: | Data Summary Statistics - Wards and Agents | 19 |
| Table 1.4: | Differences in Characteristics Across Neighbors | 20 |
| Table 1.5: | Linear Probability Model - Social Effects | 21 |
| Table 1.6: | Linear Probability Model - Information | 22 |
| Table 3.1: | Summary Statistics - Market Variables | 58 |
| Table 3.2: | Summary Statistics - Brokers | 58 |
| Table 3.3: | Summary Statistics - Commission Rates | 59 |
| Table 3.4: | Reproduction of results from Hsieh and Moretti (2003) | 59 |
| Table 3.5: | Regression: Productivity on Price | 60 |
| Table 3.6: | Regression: Number of Agents on Price and Quantity | 60 |
| Table 3.7: | Regression: Wage on Price, Quantity, and Number of Agents | 61 |
| Table 3.8: | Fixed Effects | 62 |
| Table 3.9: | Fixed Effects - Incumbents Only | 63 |
ACKNOWLEDGEMENTS

Many colleagues and friends have improved the quality of this research. Joel Watson has been a dedicated teacher, advisor, and example. Gray Calhoun, Gordon Dahl, Sam Dastrup, Silke Forbes, Graton Gathright, and Seth Pruitt have offered substantial advice and support.

I especially thank Graton Gathright, coauthor of Chapter 1, and Joel Watson, coauthor of Chapter 2. It is with their permission that I include that research in this dissertation.
VITA

2003  B. A. in Economics, Brigham Young University

2009  Ph. D. in Economics, University of California, San Diego
ABSTRACT OF THE DISSERTATION

The Economics of Real Estate Brokerage and Contracts

by

Christopher David Wignall

Doctor of Philosophy in Economics

University of California San Diego, 2009

Professor Joel Watson, Chair

This dissertation collects research on real estate brokerage and contract theory that was completed while the author was a student at the University of California, San Diego.

Chapter 1 presents an empirical analysis of the role of social networks in economic decision-making. Specifically, it is demonstrated that individuals are likely to choose the same real estate agent as a peer who is in the same social network. Evidence is presented that the social influence comes through peer-to-peer communication.

Chapter 2 offers a simple model of hold-up in a contracting environment where the contracted trade retains its value indefinitely. A simple option contract has both an efficient and an inefficient hold-up equilibrium. More complicated contracts can uniquely implement any efficient outcome.

Chapter 3 investigates whether entry into the real estate brokerage industry is inefficiently high. Previous research on this topic suffers from restrictive data. Use of data from a multiple listing service allows for heterogeneity between agents and addresses the omitted variable problem.
Chapter 1

Word-of-Mouth Learning in Social Networks

1.1 Introduction

When an individual chooses among options with unknown payoffs, she can often achieve a better expected payoff by first gathering information from peers who have chosen from the same set of options. A significant identification problem is endemic to studying such peer effects: unobserved characteristics that influence behavior may also influence which relationships form (see Manski, 1995). If some omitted variable leads two people to make similar decisions and also increases the probability that they become peers, then estimates of the peer effect will be biased upward. For example, if people with tastes for risky behavior are more likely to smoke and also tend to be friends with other risk lovers, then estimates of peer influence on smoking that fail to account for risk preferences will exhibit a positive omitted variable bias.

In this paper, we investigate social learning by home owners about the quality of real estate agents as the home owners choose agents to list their homes for sale. The social networks that we investigate are congregations of The Church of Jesus Christ of Latter-day Saints (Mormon). These congregations, called wards, are defined geographically in a manner such that, conditional on geography, the
assignment of homes to wards does not suffer from an unobserved selection process. As a result, once we control for the geographic selection of homes into wards, we can treat the assignment of residents to wards as essentially random, and we can identify the effect of social learning on the choice of real estate agents by home sellers.

Researchers have employed a variety of approaches to estimating social effects in the presence of omitted variables. For example, Duflo and Saez (2003) treat employees of academic departments at a large university as peers and use a randomization experiment to evaluate peer effects on attendance at a retirement benefits information fair. Sorensen (2006) looks at health plan choice by employees within academic departments of the University of California system. He uses the panel structure of his data to account for the unobserved heterogeneity between departments. Bayer et al (2005) and Hellerstein et al (2008) treat census blocks and tracts, respectively, as social networks where peers may learn about job opportunities.

Each Mormon ward is defined by a set of geographic boundaries, and each church member is assigned to the ward in which he resides. The exogeneity of ward assignment to real estate agent choice arises from the process by which geographic ward boundaries are specified. In localities where the concentration of church members is high, the process of ward boundary specification produces wards that typically enclose a small geographic area and whose boundaries are not coincident with significant neighborhood boundaries (such as major roads or subdivision boundaries). We focus on Utah County, Utah were the concentration of Mormons is approximately 89%. Consequently, in our sample, a typical home owner will have a set of geographically close neighbors who are fairly homogeneous, some in her ward and some in other wards.

We measure the influence of a home owner’s peers on his choice of real estate agent. We find that the average home seller is almost twice as likely to

---

1From the Religious Congregations and Membership in the United States, 2000, collected by Association of Statisticians of American Religious Bodies.

2The boundary discontinuity approach here is similar to the use of school district boundaries by Black (1999) to identify the value of public elementary schools to home owners.
choose the same real estate agent as a neighbor when they are both assigned to the same ward. The importance of the ward social network for real estate agent choice can be described in terms of the change in geography that will offset a ward relationship. For example, to be as influential as a ward neighbor that is 400 feet away, a neighbor assigned to a different ward must be 30% closer.

We also present evidence that home sellers respond to peers' private information about the quality of real estate agents, suggesting that at least some of the social learning that we find arises from word-of-mouth communication rather than from simply observing peers' behavior. This distinction has important welfare implications since pure observational learning faces a higher probability of an information cascade and inefficient herding\(^3\). Furthermore, direct communication between consumers concerning personal experience with real estate agents can provide reputational incentives to agents to please each client. These incentives may mitigate possible agency problems in real estate brokerage (see Levitt and Syverson, 2005).

In the next section, we outline our conceptual framework and predictions. In Section 3, we present background detail on both real estate brokerage and Mormon wards, and we describe the data that we employ. In Section 4, we discuss our approach to estimation and identification. We present evidence of social learning in Section 5. In Section 6 we present evidence of social learning in wards through direct communication. Section 7 concludes.

1.2 Conceptual Framework

Our objective is to evaluate how individuals are influenced by members of their social network in selecting a real estate agent to help sell a home. If a home seller learns about real estate agent quality through her social network, then her choice of real estate agent is more likely to be influenced by the choices of neighbors who belong to her social network.

This prediction can arise from both observational learning and direct communication between consumers concerning personal experience with real estate agents. These incentives may mitigate possible agency problems in real estate brokerage (see Levitt and Syverson, 2005).

\(^3\)For an overview of the literature on information cascades, see Bikhchandani et al (1998)
munication. In the case of purely observational learning, a home seller may notice the real estate agent choice of a neighbor and infer that the peer has private information that the chosen real estate agent is a high quality agent. If the home seller can more easily observe the choice of neighbors who belong to her social network, then she is more likely to choose the same agent as a neighbor if that neighbor belongs to her social network.

Social learning about the quality of real estate agents may also arise through direct communication between peers about personal experience with real estate agents. Depending on the content of the reports from neighbors about real estate agent quality, direct communication may increase or decrease the likelihood that a home seller chooses the same agent as a peer. If such reports tend to be positive, then, on average, home sellers will be more likely to choose the same real estate agent as neighbors who belong to the same social network.

The direct communication hypothesis provides a second prediction. If social learning occurs through direct communication about personal experience with agents, then a home seller is more likely to choose the same agent as a neighbor when the neighbor’s experience with the agent was positive. The home seller is less likely to choose the same agent as a neighbor whose experience with his agent was negative. If direct communication is more likely to occur between neighbors who belong to the same social network, then the effect of a neighbor’s experience on the home seller’s choice of agent will be stronger if they belong to the same social network.

1.3 Background and Data

Real Estate Brokerage

Nationwide, most home sellers employ a real estate agent to list their home. The contract between a home seller and her real estate agent is called a listing agreement. These contracts typically stipulate that the real estate agent will market the home in exchange for a payment, due at closing, that is expressed as a percentage of the sales price.
Virtually all real estate agents who list homes in Utah County belong to the only multiple listings service operating in the county, the Wasatch Front Regional Multiple Listings Service (WFRMLS). WFRMLS requires that its member agents add their new listings to the WFRMLS database within 72 hours of signing a listing agreement. We use data on all listings in the WFRMLS database of single family residences in Utah County from 1997-2006.

The data from WFRMLS for each listing include home characteristics (square footage, number of bedrooms, street address, etc.) and identifying information for the agents involved in the transaction. Each record also includes the date the property was listed and the asking price. For properties that resulted in a sale, we also have the sales price.

Because of the large number of listings in our sample, it is not computationally feasible to evaluate the relationship between every pair of listings. We limit our attention to pairs of listings that are located within one quarter mile and listed within five years of each other, and we call such pairs neighbors.

We employ several measures of geographic location of listings to account for the spatial relationship between properties. Based on street address, we place each listing on a map and calculate the distance between each pair of neighbors. Second, using geographic data from the Utah County Department of Information Systems, we determine whether each pair of homes is assigned to the same county-defined neighborhood. The county’s neighborhood definitions correspond to contiguous parcels of land that were developed contemporaneously. Finally, we use data from the U.S. Census Bureau (TigerLine) to determine whether neighbors belong to the same census block.

In Table 1, we present summary statistics on the characteristics of the homes in our sample of listings. The mean list price in our sample is $216,065. Fifty-nine percent of the listings result in a sale, and the mean sales price is $192,833.

Mormon Wards as Social Networks

Mormon wards are well-suited as a setting for investigating peer effects. Wards are important social networks to those who belong to them, and the assign-
ment of neighboring church members to wards is essentially random, conditional on the spatial relationship between homes.

Regular participation in one’s assigned ward involves frequent personal interaction with co-congregants. Since the Mormon Church has no paid clergy at the ward level, the wide array of leadership, teaching and other position are performed by individual lay members. For example, all adults are assigned a list of families that they are expected to visit on at least a monthly basis. According to a 2008 Survey by the Pew Research Center\textsuperscript{4}, 75\% of (self-reported) Mormons attend religious services at least once a week and 92\% of them are formal members of their congregations (wards). In addition, 77\% participate at least monthly in non-worship activities at church, including 63\% participating in social activities at church at least monthly.

There are at least two important reasons that, virtually without exception, practicing Mormons participate in the ward to which they are assigned. First, as mentioned above, the vast majority of church responsibilities are fulfilled by individual members. A church member is not typically eligible to perform any of these duties in a ward to which she is not assigned, and holding such a position is a hallmark of full fellowship. Another reason for participation in the assigned ward stems from the two levels of church worship in Mormon Theology. The first and most basic form of church worship is the weekly Sunday service, held in local chapels and open to the public. Each ward has its own set of meetings that are managed by the ward members and leaders. The second type of worship occurs in Mormon temples. Participation in temple worship is limited to members that are in good standing and approved by their ward leaders. One of the requirements for good standing is regular participation in the Sunday services of the ward to which they are assigned. It would be difficult to overstate the importance of temple worship in Mormon theology.

The geographic boundaries of wards are designed to include 300-500 members. In this paper, we focus our attention on Utah County, where approximately 89\% of the population is Mormon. This concentration of church members leads

\textsuperscript{4}US Religious Landscape Survey, Pew Forum on Religion and Public Life
to ward boundaries that enclose very small geographic areas, smaller than most subdivisions. For this reason, ward boundaries are not typically coincident with subdivision boundaries and major thoroughfares. Figure 1 illustrates the assignment of parcels to wards for a small region in Utah County.

A second reason ward boundaries are unlikely to coincide with important discontinuities in the spatial distribution of homes is that homogeneity across neighboring wards is an explicit objective of church leaders involved in the specification of ward boundaries\(^5\). In practice, this means that ward boundaries are likely to cut across neighborhoods. As an illustration, note that the clusters of very small parcels in Figure 1 are townhouses. The townhouses complexes are split and combined with neighboring detached residences to form wards similar in mix of property type.

We are able to determine the ward assignment of each property in our sample using the Church’s online ward assignment lookup tool\(^6\). In Table 2, we present summary statistics on neighbors of a typical listing. The average listing has 36 neighbors in the same ward and 42 neighbors assigned to a different ward. Neighbors in different wards are, on average 40% farther away than neighbors assigned to the same ward. Without conditioning on spatial relationship, a pair of neighbors assigned to the same ward is almost three times more likely to choose the same real estate agent as two neighbors in different wards.

In Table 3, we present summary statistics on wards and real estate agents. The homes in each ward are listed by a variety of real estate agents, suggesting that real estate agents do not specialize in particular wards.

The distribution of agent activity is highly skewed. More than half the agents in our sample list only one or two homes. Agents who listed more than two homes listed on average 21 homes.

---

\(^5\)Based on the authors’ private conversations with church leaders.

1.4 Estimation and Identification

Empirical Approach

To evaluate the influence of ward peers on real estate agent choice, we estimate the probability that a home seller chooses the same real estate agent as her neighbor. If ward members learn from each other, then the probability that neighbors choose the same real estate agent should differ depending on whether they are assigned to the same ward. We begin by estimating the following linear probability model:

\[ P(i \text{ and } j \text{ choose the same agent}) = \alpha + \beta W_{ij} + \gamma_1 D_{ij} + \gamma_2 D_{ij}^2 + \gamma_3 N_{ij} + \epsilon_{ij} \]  

where \( W_{ij} \) is equal to 1 if homes \( i \) and \( j \) are in the same ward, \( D_{ij} \) is the geographic distance between homes \( i \) and \( j \), and \( N_{ij} \) equals 1 if homes \( i \) and \( j \) are assigned to the same county-defined neighborhood. We measure distance in quarter miles so that \( D_{ij} \) lies in the interval from zero to one (since we only consider pairs of homes less than a quarter mile apart).

We calculate multi-way cluster-robust standard errors using the method developed by Cameron et al (2007) for this and all specifications. Observations are on pairs of listings, and each listing in the pair belongs to many different pairs. We estimate standard errors that are robust to clustering on both listings in the pair.

The estimation results for this specification are reported and discussed in Section 5 below.

To test whether the social learning in this setting arises from word-of-mouth information transmission we estimate the following variation on the regression in equation (1):

\[ P_{ij} = \alpha + \beta W_{ij} + \gamma_1 D_{ij} + \gamma_2 D_{ij}^2 + \gamma_3 N_{ij} + \delta_1 G_j + \delta_2 B_j + \delta_3 W_{ij} \ast G_j + \delta_4 W_{ij} \ast B_j + \epsilon_{ij} \]

where \( G_j = 1 \) if neighbor \( j \) had a good outcome (and therefore has positive information to report) and \( B_j = 1 \) if neighbor \( j \) had a bad outcome.

The estimation results for this specification are presented in Section 6.
Identification

We want to identify the impact of a neighbor’s choice of real estate agent on a home seller’s choice of real estate agent. The principal threat to identification in our setting is what Manski (1995) calls correlated effects. If a home owner’s neighbors who are assigned to his ward have homes that are systematically different than the homes of his neighbors who are assigned to different wards, then the effect that we estimate may represent correlations in behaviour due to correlations in unobserved home characteristics. An example of such an effect is small scale geographic specialization and marketing by agents to particular neighborhoods or types of homes.

Our identifying assumption is that, after we have conditioned on geography, a home owner’s intra-ward neighbors are not systematically different from his extra-ward neighbors.

One concern is that ward boundaries may coincide with unobserved discontinuities in the spatial distribution of house characteristics. Observed discontinuities in house characteristics include abrupt changes in house (and resident) characteristics at subdivision boundaries and geographic features like rivers, parks, and major roads.

Ward designers try to ensure homogeneity across ward boundaries, so adjacent neighborhoods are likely to be split into wards in a way that assigns some homes from each neighborhood to each ward. Practically, this means that ward boundaries are likely to cut across neighborhoods.

Our identification fails if, despite the planners’ objectives, there are unobserved neighborhood boundaries that are correlated with ward boundaries and that affect real estate agent selection. We investigate the extent to which this may occur by calculating the absolute difference in observed characteristics for each pair of neighbors and regress out the portion of those differences that are explained by their geographic relationship (distance, distance squared, and whether they are in the same county-defined neighborhood). We then calculate the means of these orthogonalized differences for neighbors in the same ward and for neighbors assigned to different wards and perform a t-test for the equality of those means (See Table
We reject the null hypothesis that the mean differences in observed home characteristics for neighbors assigned to the same ward are equal to the mean differences for neighbors assigned to different wards. Our large number of observations (over three million pairs of neighbors) means that we only fail to reject the null hypothesis for extremely small differences in the means.

In order to illustrate the magnitude of the difference between the means, we calculate the standard deviation of the distribution of differences of each characteristic for the set of neighbors of each individual listing. So, for example, for each listing $i$ we calculate:

$$\sigma_i(\text{Bedrooms}) = \sqrt{\frac{1}{J-1} \sum_{j \in J} [(\text{Bedrooms}_i - \text{Bedrooms}_j) - \text{Average Deviation}_i]^2}$$

where $J$ is the set of all neighbors of $i$ and also the number of elements in the set. We present the median individual standard deviation for each characteristic in Table 4. The largest difference between the means is less than 5% of the mean difference between homes and is less than 5% of the median individual standard deviation. The differences of means for the other characteristics are even smaller proportions.

### 1.5 Evidence of Social Learning in Wards

The results from estimating equation 1, reported in Table 5, suggest that a home seller is substantially more likely to choose the same real estate agent as her neighbor if they are assigned to the same ward. For an intermediate distance (one eighth of a mile, distance = .5), the probability that the home seller chooses the same agent as her neighbor is 1.4% if they are assigned to different wards and 2.4% if they are assigned to the same ward.

The distance between two neighbors is an important determinant of the probability that they choose the same real estate agent. As mentioned in the introduction, the importance of the ward social network can be described in terms of the change in geography that will offset a ward relationship. For example, to
be as influential as a ward neighbor that is 400 feet away, a non-ward neighbor must be 30% closer. Similarly, a ward neighbor 900 feet away is as influential as a non-ward neighbor that is 35% closer.

As additional evidence that our specification does not suffer from an omitted variable bias, we present the results from a long regression that includes differences in observed characteristics. If we have omitted a relevant neighborhood definition that is correlated with the ward definitions, then our estimate of $\beta$ will be biased. A neighborhood boundary is a discontinuous change in home characteristics, so the differences in home characteristics between two neighbors should be correlated with any omitted neighborhood definition (the differences should be larger for neighbors on opposite sides of the boundary). Including these differences in our regression, then, should attenuate any omitted variable bias. As column 4 of Table 5 shows, however, our estimate of the ward effect in this long regression is the same as the estimate from the short regression in column 3.

Census blocks are small neighborhoods bounded by geographic features (like roads, streams, and railroad tracks) and political boundaries (like city limits and property lines). In urban areas, the census block is often the same as the city block. The way the census blocks are defined means that homes in the same block are likely to be very similar in unobserved characteristics. In column 7 of Table 5, we present estimates of equation 1 on a subsample restricted to neighbors in the same census block. The persistence of the ward effect is additional evidence that our results are not due to bias from some omitted neighborhood definition.

Figure 2 illustrates that the geographic distribution of neighbors in the same ward differs significantly from that of neighbors in different wards. Neighbors in the same ward tend to be nearer to each other and the nearest neighbors are very likely to be assigned to the same ward. We are careful about how we control for the geographic relationship to ensure that the estimated ward effect is not an artifact of the spatial relationships between neighbors. We demonstrate in columns 5 and 6 of Table 5 that our estimation of $\beta$ is not sensitive to the specification.

---

7See Table 4 for the list of characteristics that are included.
8This description of the census block definition is based largely on information available from the US Census Bureau, www.census.gov
of the distance effect: it does not change when we include either a sixth degree polynomial (column 5) or a set of dummy variables representing a fine partition of the distances (0 to 33 feet, 34 to 66 feet, etc.).

We also estimate equation 1 on a subsample of our data that excludes the nearest neighbors (where the vast majority are in the same ward) and the most distant neighbors (where the majority are assigned to different wards). Column 8 presents estimates when we restrict attention to neighbors that are no less than 400 feet apart and no more than 900 feet apart (distance ∈ (.3, .7)). Our estimate of $\beta$ on this subsample is slightly larger than the estimate from the full sample, offering additional evidence that the geographic relationships do not drive our estimates of the ward effect.

1.6 Social Learning Via Direct Communication

We have offered evidence of social influence on real estate agent choice within wards. We now address the source of this influence. If home sellers are learning from peers only by observing choices and making inference about private information, then the choices of peers with identical characteristics will have identical influence. If, however, peers are communicating directly, then a home seller’s choices may respond to a peer’s private information, including information about outcomes.

Home sellers prefer a higher sales price, all else equal. If they are learning from their peers, they are more likely to choose the same real estate agent as a neighbor if that agent sold the neighbor’s house for a high price relative to the seller’s ex ante expectations and less likely to choose him if he sold the house for a low price relative to expectations. For each transaction, we calculate the percent difference between sales price and list price. We categorize transactions in the top decile of the distribution of percent difference as very high price and transactions in the bottom decile as very low price. We then analyze how these outcomes affect peer influence. The results are in column 1 of table 6.

We find effects for peers with intermediate outcomes that are similar to the
effects in the baseline specification. In addition, for neighbors in the same ward we find a large premium associated with a very high price and a large penalty for a very low price. The additional influence of ward members with very high prices is double that of those with intermediate outcomes and the penalty from a very low price almost cancels the ward effect completely.

We modify our categorization of outcomes by considering different cutoffs for the definitions of very high price and very low price. In column 2 of Table 6 we define these outcomes using the top and bottom quartiles of the distribution of deviations. In column 3 we use the ninety-fifth and fifth percentiles. We find that the impact of good information is smaller with the lower threshold and larger with the higher threshold. While the impact of bad information increases slightly with the higher threshold definition, it does not change with the bottom quartile definition of column 2.

The markup over list price might reflect market conditions that may influence whether two neighbors select the same real estate agent. The specification in column 4 of Table 6 includes dummy variables for the year that each of the neighbors listed her home. We use the same definitions of very high price and very low price as in column 1 (the top and bottom deciles). Our estimate of the impact of good information outside the ward decreases substantially (0.9 to 0.3), but the impact of good information in the ward changes very little (1.4 to 1.3) and the impact of bad information in the ward doesn’t change at all.

We also consider an alternative definition of outcome. We estimate a linear hedonic model of the natural log of sales price using the observed characteristics of homes and times and ward fixed effects. We then use the deviations from the predicted values (the residuals from the hedonic regression) to categorize outcomes. In column 5 of Table 6, we say that a home sells for a very high price if the difference between log sales price from the predicted log sales price is in the top decile of its distribution. We say that it sells for a very low price if it is in the bottom decile of its distribution. Column 7 of Table 6 uses the ninety-fifth and fifth percentiles. The qualitative results are not sensitive to the benchmark used - the estimates from column 5 are very similar to those in column 1.
1.7 Conclusion

We have presented evidence of word-of-mouth learning in a social network. The principal challenges to identification of this type of social effect are distinguishing which individuals are peers and separating social effects from correlations in behavior of peers that arise from unobserved similarities\(^9\). We address both of these issues by taking Mormon wards as our setting. The ward constitutes a social network for which the group composition is known and for which we can construct a control group of individuals who differ essentially from ward members only in their ward assignment.

The results that we present suggest that social learning plays a role in a home seller’s selection of an agent to assist in an important transaction. We have also presented evidence that personal referrals are part of this social learning. It seems unlikely that this phenomenon is particular to Mormon wards. Congregations in other denominations and other types of social networks (school, athletic, service, etc) may function in similar ways.

We have exploited the geographic assignment of Mormons to congregations to identify social effects in real estate agent choice. There are many other decisions of economic importance that may be subject to social effects. The natural experiment investigated here holds promise for identifying social effects in many such decisions.

I thank Graton Gathright, coauthor of this research, for permission to include it in this dissertation.

\(^9\)See Manski (2000).
Table 1.1: Data Summary Statistics - Listings
Summary Statistics of Homes - Data are from the Wasatch Front Regional
Multiple Listing Service and include single family residences listed in Utah
County between 1997 and 2007.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>List Price</td>
<td>216065</td>
<td>125632</td>
<td>174900</td>
</tr>
<tr>
<td>Square Feet</td>
<td>2675</td>
<td>1255</td>
<td>2416</td>
</tr>
<tr>
<td>Acres</td>
<td>0.35</td>
<td>1.0</td>
<td>0.22</td>
</tr>
<tr>
<td>Bedrooms</td>
<td>3.9</td>
<td>1.2</td>
<td>4</td>
</tr>
<tr>
<td>Bathrooms</td>
<td>2.5</td>
<td>1.0</td>
<td>2</td>
</tr>
<tr>
<td>Garage Capacity</td>
<td>1.5</td>
<td>0.80</td>
<td>2</td>
</tr>
<tr>
<td>Patios</td>
<td>0.46</td>
<td>0.50</td>
<td>0</td>
</tr>
<tr>
<td>Decks</td>
<td>0.28</td>
<td>0.45</td>
<td>0</td>
</tr>
<tr>
<td>Wet Bars</td>
<td>0.40</td>
<td>0.52</td>
<td>0</td>
</tr>
<tr>
<td>Fire Places</td>
<td>0.64</td>
<td>0.76</td>
<td>0</td>
</tr>
<tr>
<td>Year Built</td>
<td>1984</td>
<td>23</td>
<td>1994</td>
</tr>
<tr>
<td>Sold Indicator</td>
<td>0.59</td>
<td>0.49</td>
<td>1</td>
</tr>
<tr>
<td>Sold Price</td>
<td>192833</td>
<td>99903</td>
<td>163000</td>
</tr>
</tbody>
</table>
Figure 1.1: Ward assignment of homes in a Utah County neighborhood.
Figure 1.2: Histogram of Distances between neighbors

Histogram of Distances between neighbors in the same and in different wards.
Table 1.2: Data Summary Statistics - Neighbors

Summary Statistics of Neighbors - Geographic location of properties from Google Maps. Ward affiliation from www.lds.org. Neighborhood definition from the Utah County Assessor’s Office. Census Block data from the US Census Bureau (TigerLine). Two Homes are “neighbors” if they are located within a quarter mile and sold within five years of each other. Distance is measured in quarter miles.

<table>
<thead>
<tr>
<th>Mean (Standard Deviation)</th>
<th>In Ward</th>
<th>Out of Ward</th>
<th>In Census Block</th>
<th>In Neighborhood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>36</td>
<td>42</td>
<td>20</td>
<td>52</td>
</tr>
<tr>
<td>Number of Neighbors</td>
<td>(29)</td>
<td>(34)</td>
<td>(37)</td>
<td>(44)</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>35</td>
<td>8</td>
<td>43</td>
</tr>
<tr>
<td>Percent of Neighbors</td>
<td>(23)</td>
<td>(23)</td>
<td>(25)</td>
<td>(30)</td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>55</td>
<td>13</td>
<td>73</td>
</tr>
<tr>
<td>Distance Between Neighbors</td>
<td>(0.25)</td>
<td>(0.21)</td>
<td>(0.26)</td>
<td>(0.26)</td>
</tr>
<tr>
<td></td>
<td>0.47</td>
<td>0.75</td>
<td>0.46</td>
<td>0.59</td>
</tr>
<tr>
<td>Percent that Chose the Same Agent</td>
<td>(9.9)</td>
<td>(4.6)</td>
<td>(11)</td>
<td>(9.8)</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Variable</td>
<td>Mean</td>
<td>Standard Deviation</td>
<td>Median</td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>------</td>
<td>--------------------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td><strong>Ward</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(N = 832)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Listings</td>
<td>72</td>
<td>42</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>Number of Agents</td>
<td>51</td>
<td>26</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>Listings per Agent</td>
<td>1.3</td>
<td>1.0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>Agent</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(N = 5904)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Listings</td>
<td>10</td>
<td>31</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Number of Wards</td>
<td>7</td>
<td>19</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Listings per Ward</td>
<td>1.3</td>
<td>1.0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>Agent with at least three listings</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(N = 2649)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Listings</td>
<td>21</td>
<td>44</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Number of Wards</td>
<td>14</td>
<td>27</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Listings per Ward</td>
<td>1.5</td>
<td>1.5</td>
<td>1.2</td>
<td></td>
</tr>
</tbody>
</table>
Table 1.4: Differences in Characteristics Across Neighbors

Differences in Characteristics Across Neighbors - We calculate the absolute difference in observed characteristics for each pair of neighbors. We regress out the portion of this difference that is explained by their geographic relationship (distance, distance squared, and whether they are in the same county-defined neighborhood). We present the mean of these orthogonalized differences in characteristics for neighbors in different wards and neighbors in the same ward. We present the t-statistic for a test that the means are equal. We also calculate the standard deviation of the differences for each individual property. For example, \( \sigma_i(\text{Beds}) = \sqrt{\left(\sum_{j \in J} [(\text{Beds}_i - \text{Beds}_j) - (\sum_{j \in J} [\text{Beds}_i - \text{Beds}_j])/\#J]^2\right)} \) where \( J = \{\text{All neighbors if } i\} \). We present the median standard deviation for each characteristic to illustrate the variation between a typical listing and its neighbors.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Out</th>
<th>In</th>
<th>Difference</th>
<th>t-stat</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Square Feet</td>
<td>0.261</td>
<td>0.255</td>
<td>0.006</td>
<td>17.5</td>
<td>0.201</td>
</tr>
<tr>
<td>Bedrooms</td>
<td>1.005</td>
<td>0.989</td>
<td>0.016</td>
<td>13.9</td>
<td>0.796</td>
</tr>
<tr>
<td>Bathrooms</td>
<td>0.707</td>
<td>0.700</td>
<td>0.008</td>
<td>8.6</td>
<td>0.601</td>
</tr>
<tr>
<td>Year Built</td>
<td>8.373</td>
<td>8.025</td>
<td>0.348</td>
<td>20.9</td>
<td>7.538</td>
</tr>
<tr>
<td>Log List Price</td>
<td>0.258</td>
<td>0.261</td>
<td>-0.003</td>
<td>-10.2</td>
<td>0.163</td>
</tr>
<tr>
<td>Log Acres</td>
<td>0.562</td>
<td>0.566</td>
<td>-0.004</td>
<td>-6.5</td>
<td>0.292</td>
</tr>
<tr>
<td>Garage Capacity</td>
<td>0.490</td>
<td>0.471</td>
<td>0.019</td>
<td>21.8</td>
<td>0.532</td>
</tr>
<tr>
<td>Fireplaces</td>
<td>0.404</td>
<td>0.398</td>
<td>0.006</td>
<td>8.8</td>
<td>0.507</td>
</tr>
<tr>
<td>Wet Bars</td>
<td>0.420</td>
<td>0.418</td>
<td>0.002</td>
<td>3.3</td>
<td>0.489</td>
</tr>
<tr>
<td>Deck</td>
<td>0.299</td>
<td>0.296</td>
<td>0.003</td>
<td>5.5</td>
<td>0.430</td>
</tr>
<tr>
<td>Patio</td>
<td>0.450</td>
<td>0.441</td>
<td>0.009</td>
<td>15.7</td>
<td>0.494</td>
</tr>
</tbody>
</table>
Table 1.5: Linear Probability Model - Social Effects

Linear Probability Model - Social Effects Parameter estimates from $y_{ij} = X_{ij}\beta + \varepsilon_{ij}$ where $y_{ij} = 1$ if neighbors $i$ and $j$ choose the same real estate agent to list their homes. All standard errors are robust to clustering on the identity of each listing in the pair of neighbors.

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>In the Same Ward</td>
<td>1.2**</td>
<td>1.9**</td>
<td>1.0**</td>
<td>1.0**</td>
<td>1.0**</td>
<td>1.5**</td>
<td>1.1**</td>
<td>0.8**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.12)</td>
<td>(0.06)</td>
<td>(0.04)</td>
<td></td>
</tr>
<tr>
<td>In the Same Neighborhood</td>
<td>0.6**</td>
<td>0.3**</td>
<td>0.03</td>
<td>0.3**</td>
<td>0.3**</td>
<td>0.9**</td>
<td>0.4**</td>
<td>0.2**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.14)</td>
<td>(0.07)</td>
<td>(0.04)</td>
<td></td>
</tr>
<tr>
<td>Distance (in 1/4 Miles)</td>
<td>-17.9**</td>
<td>-15.9**</td>
<td>-13.7**</td>
<td></td>
<td>-15.9**</td>
<td>-7.5**</td>
<td>-14.3**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.52)</td>
<td>(0.50)</td>
<td>(0.46)</td>
<td></td>
<td>(0.74)</td>
<td>(1.40)</td>
<td>(0.49)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance²</td>
<td>11.6**</td>
<td>10.5**</td>
<td>9.2**</td>
<td></td>
<td>10.1**</td>
<td>4.3**</td>
<td>9.5**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.38)</td>
<td>(0.36)</td>
<td>(0.34)</td>
<td></td>
<td>(0.62)</td>
<td>(1.33)</td>
<td>(0.36)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>In the Same Census Block</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.3**</td>
<td></td>
</tr>
<tr>
<td>Difference in Observed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.07)</td>
<td></td>
</tr>
<tr>
<td>Characteristics</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance Polynomial</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Distance Dummies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>7.5**</td>
<td>0.6**</td>
<td>6.4**</td>
<td>8.2**</td>
<td>13.2**</td>
<td>11.2**</td>
<td>6.4**</td>
<td>3.7**</td>
<td>5.7**</td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(0.04)</td>
<td>(0.17)</td>
<td>(0.20)</td>
<td>(0.57)</td>
<td>(0.54)</td>
<td>(0.26)</td>
<td>(0.36)</td>
<td>(0.17)</td>
</tr>
<tr>
<td>N</td>
<td>3,604,384</td>
<td>3,029,737</td>
<td>3,029,737</td>
<td>2,861,739</td>
<td>3,029,737</td>
<td>3,029,737</td>
<td>871,981</td>
<td>1,322,795</td>
<td>3,029,737</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.014</td>
<td>0.005</td>
<td>0.012</td>
<td>0.017</td>
<td>0.012</td>
<td>0.012</td>
<td>0.012</td>
<td>0.003</td>
<td>0.013</td>
</tr>
</tbody>
</table>
Table 1.6: Linear Probability Model - Information

Linear Probability Model - Information

Results from linear probability model: \( y_{ij} = X_{ij}\beta + \varepsilon_{ij} \) where \( y_{ij} = 1 \) if neighbors \( i \) and \( j \) chose the same real estate agent. We consider various definitions of "very high price" and "very low price." See text for details.

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>In the Same Ward</td>
<td>1.2**</td>
<td>1.3**</td>
<td>1.2**</td>
<td>1.2**</td>
<td>1.1**</td>
<td>1.074**</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.06)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.051)</td>
</tr>
<tr>
<td>Distance (in 1/4 Miles)</td>
<td>-17.8**</td>
<td>-17.8**</td>
<td>-17.8**</td>
<td>-17.2**</td>
<td>-17.9**</td>
<td>-17.9**</td>
</tr>
<tr>
<td></td>
<td>(0.52)</td>
<td>(0.52)</td>
<td>(0.52)</td>
<td>(0.49)</td>
<td>(0.52)</td>
<td>(0.52)</td>
</tr>
<tr>
<td>Distance^2</td>
<td>11.6**</td>
<td>11.6**</td>
<td>11.6**</td>
<td>11.2**</td>
<td>11.6**</td>
<td>11.6**</td>
</tr>
<tr>
<td></td>
<td>(0.37)</td>
<td>(0.37)</td>
<td>(0.37)</td>
<td>(0.36)</td>
<td>(0.37)</td>
<td>(0.37)</td>
</tr>
<tr>
<td>Very High Price</td>
<td>0.9**</td>
<td>0.2**</td>
<td>0.3*</td>
<td>0.3**</td>
<td>0.6**</td>
<td>0.7**</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.06)</td>
<td>(0.13)</td>
<td>(0.09)</td>
<td>(0.08)</td>
<td>(0.10)</td>
</tr>
<tr>
<td>Very Low Price</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>-0.03</td>
<td>-0.2*</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.05)</td>
<td>(0.15)</td>
<td>(0.09)</td>
<td>(0.07)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>Ward * Very High Price</td>
<td>1.4**</td>
<td>0.5**</td>
<td>2.5**</td>
<td>1.3**</td>
<td>1.3**</td>
<td>1.6**</td>
</tr>
<tr>
<td></td>
<td>(0.21)</td>
<td>(0.14)</td>
<td>(0.33)</td>
<td>(0.21)</td>
<td>(0.17)</td>
<td>(0.18)</td>
</tr>
<tr>
<td>Ward * Very Low Price</td>
<td>-1.0**</td>
<td>-1.0**</td>
<td>-1.1**</td>
<td>-1.0**</td>
<td>-0.8**</td>
<td>-1.0**</td>
</tr>
<tr>
<td></td>
<td>(0.14)</td>
<td>(0.10)</td>
<td>(0.18)</td>
<td>(0.13)</td>
<td>(0.16)</td>
<td>(0.18)</td>
</tr>
<tr>
<td>Time Dummies</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>7.4**</td>
<td>7.4**</td>
<td>7.4**</td>
<td>14.9**</td>
<td>7.4**</td>
<td>7.4**</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.18)</td>
<td>(0.17)</td>
<td>(0.98)</td>
<td>(0.17)</td>
<td>(0.17)</td>
</tr>
<tr>
<td>R^2</td>
<td>0.014</td>
<td>0.014</td>
<td>0.014</td>
<td>0.031</td>
<td>0.015</td>
<td>0.015</td>
</tr>
</tbody>
</table>
Chapter 2

Hold-up and Durable Trading Opportunities

2.1 Introduction

The hold-up problem arises in situations in which contacting parties can renegotiate their contract between the time they make relationship-specific investments and the time at which they can trade. To illustrate, consider a simple model of trade between a buyer and a seller. First, the parties write a contract specifying the terms of trade. Second, the seller chooses a level of investment, which is observable to the buyer. Next the parties can renegotiate their contract. Finally, the parties have the opportunity to trade and an external enforcer (such as a court) enforces the contract. If the parties cannot write an enforceable contract that is contingent on the investment decision, the buyer can extract concessions from the seller under the threat of blocking trade. Anticipating this hold-up, the seller may not invest optimally ex ante. Hold-up is particularly serious in settings with “cross-investment,” when the seller’s investment directly affects the buyer’s valuation of trade.

Many authors have analyzed the severity of hold-up in various contractual settings. Prominent models yield different predictions, but their disparate assumptions make comparisons difficult. Joel Watson (2007) shows that essen-
tial differences between various models in the literature lie in the modeling of the actions that consummate trade. In particular, “public action” models may incorporate artificial constraints on contracting by not recognizing the opportunity of the contracting parties to use individual trade actions (such as the buyer’s action of whether to physically accept delivery) as options.\(^1\) Watson (2007) thus identifies the technology of trade—the exact nature of the parties’ individual trade actions—as a key determinant of whether hold-up interferes with investment incentives.\(^2\)

In this paper, we examine a particular dimension of the trade technology: the extent to which the trading opportunity is durable. In relationships with a durable trading opportunity, if the parties do not complete the trade in a given period then they will have another opportunity to do so in the future. Most of the literature on investment and hold-up focuses on models with a nondurable trading opportunity (in which trade can only take place at a single point in time). However, the motivations for these analyses often involve stories of a durable trading opportunity. For example, Georg Nöldeke and Klaus Schmidt (1998) argue that suitably defined option contracts function to induce efficient investment when parties can consummate trade at any time. Edlin and Hermalin (2000), on the other hand, argue that a party can effectively let an option expire and then renegotiate from scratch. Che and Joszef Sákovics (2007), in summarizing the literature, point to durability as a major contributor to the hold-up problem. Thus, many researchers view durability as playing a key role in the hold-up problem, although there has been little formal analysis of settings with durable trading opportunities.\(^3\)

Our objective here is to provide a precise analysis of durability and its effect on the hold-up problem. We construct and analyze a model that explicitly accounts for the technology of trade, including the durability of the trading opportunity. Trade occurs when the buyer installs the good. The parties write contracts that condition transfers on the buyer’s verifiable installation action. By comparing the

---

\(^1\)Prominent public-action models include Yeon-Koo Che and Donald B. Hausch (1999), Aaron Edlin and Benjamin Hermalin (2000), Edlin and Stefan Reichelstein (1996), and Ilya Segal (1999)

\(^2\)Thomas Lyon and Eric Rasmusen (2004) also argue that models in the previous literature make unrealistic assumptions about how option contracts function. They note that options can be exercised even following impasse in renegotiation.

\(^3\)Abraham Wickelgren (2007) is an exception that we discuss later in this paper.
durable trade environment with a benchmark model that has only one trading opportunity, we develop our main theme: that durability of a trading opportunity does not per se affect the hold-up problem. Specifically, if an outcome is supported in the nondurability environment then it is also supported in the environment in which the trading opportunity is durable. In general, multiple equilibria exist for simple stationary contracts, but we show that efficient outcomes can be uniquely implemented by using some nonstationary contracts.

We begin the analysis by describing the benchmark model with a nondurable trading opportunity. In Section 3 we describe our model of a durable trading opportunity and compare it to the benchmark model. Analysis of simple open-ended option contracts demonstrates that both the outcome described by Nöldeke and Schmidt (1998) and the outcome described by Edlin and Hermalin (2000) can arise in equilibrium. The significant result of this section is that there is a contract and an equilibrium that leads to efficient investment and trade. Therefore, hold-up is not necessarily a problem in environments with durable trading opportunities, even with very simple contracts.

Some recent papers, including Edlin and Hermalin (2000) and Wickelgren (2007), have motivated the link between durability and hold-up by appealing to the “outside option principle” from bargaining theory. The outside option principle basically states that a bargainer’s ability to opt out of negotiation does not affect the outcome of bargaining if the option gives a sufficiently low payoff. In Section 4, we provide some general analysis of bargaining with outside options that serves to clarify the outside option principle and its implications for hold-up. We show that, for cases relevant in contractual settings, there exist interesting multiple equilibria in the bargaining game.

Section 5 presents our analysis of nonstationary option contracts and unique implementation. We describe a specific nonstationary contract that induces a unique equilibrium outcome and gives efficient investment incentives. We conclude that durability of a trading opportunity generally does not complicate the prospect of inducing efficient investment and that option contracts can avoid the hold-up problem in a wide range of settings.
In this simple model, trade is irreversible; once the buyer has installed, the interaction ends. More generally, there may be reversible trade actions. In Section 6, we offer concluding remarks and a brief discussion of differential reversibility and hold-up.

2.2 The Benchmark Model of a Nondurable Trading Opportunity

We start with an individual-action model with unverifiable investment and a verifiable individual trade action as in Watson (2007). Suppose a buyer and a seller interact as described in the introduction. The order of actions is:

- **Period 0**
  - The parties write a contract $C$ specifying an externally enforced monetary transfer $m$ from the buyer to the seller to be compelled if the buyer installs the good in period 1. We normalize to zero the payment specified for the contingency in which the buyer does not install the good. The contract may also specify an up-front transfer, which will not affect the subsequent analysis.$^4$
  - The seller makes an investment choice, selecting between H at personal cost $c$ and L at zero cost. Let $\theta$ denote the seller’s choice, which we call the *state*. This investment action is observed by the buyer but is unverifiable to the external enforcer.

- **Period 1**
  - The parties have an opportunity to renegotiate the contract, altering the specification of $C$ and potentially including an up-front transfer.

$^4$The assumption of zero payment conditional on no trade is therefore without loss of generality.
- The buyer chooses whether to install the good. This choice is verifiable.
- The external enforcer compels the contractually specified monetary transfer, which is \(m\) if the good was installed and zero otherwise.

We suppose that if the seller invests in period 0 (so that \(\theta = H\)), then the installed good gives benefit \(x > c\) to the buyer. If the seller does not invest (\(\theta = L\)), then the good is worthless. The joint value of high investment and trade is \(x - c\). The specification of the buyer’s installation choice as an individual action is the key to having a precise account of the technology of trade. We have not included in the specification of the game a phase in which the parties can make announcements because messages will not be required for our results.

The contract \(C\) is called a forcing contract if it specifies either \(m < 0\) (giving the buyer the incentive to install regardless of the seller’s investment level) or \(m > x\) (never giving the buyer the incentive to install). We call \(C\) an option contract if it specifies \(m \in [0, x]\), since the buyer has the incentive to install in state \(H\) and not in state \(L\).

The seller’s investment action and the buyer’s choice of whether to install are assumed to be consistent with sequential rationality: each player selects his individual action to maximize his expected payoff and anticipates rational behavior in the future.

If the contract would induce an inefficient installation action, the parties will renegotiate. The outcome of renegotiation is assumed to be consistent with a “black-box” cooperative bargaining solution in which the players divide surplus according to fixed bargaining weights \(\pi_B\) and \(\pi_S\) for the buyer and seller. Surplus is defined relative to the continuation value of proceeding under the original contract. This characterization of renegotiation along with the sequential rationality conditions identify a contractual equilibrium (see Watson 2006)\(^5\).

\(^5\)Because the players are risk neutral in money, the cooperative solution yields the same expected payoffs as does the following non-cooperative specification of negotiation: Nature selects one of the players to make an ultimatum offer to the other player, who either accepts or rejects it; Nature selects player \(i\) with probability \(\pi_i\). We assume that (i) on the self-enforced component of contract, players behave as agreed whenever this is consistent with individual rationality; and (ii) if an offer is rejected, then the equilibrium in the continuation of the game does not depend on the identity of the offerer or on the nature of the offer. These are the Agreement and Disagreement
A value function gives the continuation payoffs from the start of period 1 as a function of the state. Since we are not interested in contracts that force installation, we assume \( m \geq 0 \). In state L, therefore, the buyer does not have the incentive to install under the initial contract (he gets zero benefit in state L) and no surplus can be obtained from renegotiating. In equilibrium, then, the state L continuation values are zero and we can restrict attention to the continuation payoff vector from period 1 in state H, which we denote \( \mathbf{v} = (v_B, v_S) \in \mathbb{R}^2 \). In addition, let \( a \) be the buyer’s action in state H, with \( a = 1 \) denoting installation and \( a = 0 \) denoting non-installation.

The value \( \mathbf{v} \) is said to be implemented by contract \( C \) if there is an equilibrium (combining sequential rationality and the bargaining solution) of the game from period 1 that achieves this payoff vector. Our interest is in finding a contract that gives the seller the incentive to invest efficiently, meaning that it implements \( \mathbf{v} \) satisfying \( v_S \geq c \). From the time of his investment action in period 0, the seller will obtain \(-c + v_S\) by investing at level H and will obtain zero (no cost and zero value from period 1) by investing at level L.

For our simple model, it is not difficult to verify the existence of a contract that induces efficient investment and trade.

**Proposition 1:** In the model with nondurable trade, any \( \mathbf{v} \) satisfying \( v_S \in [0, x] \) and \( v_B = x - v_S \) can be implemented. Specifically, there is an option contract that induces efficient investment and trade.

**Proof:** Take any vector \( \mathbf{v} \) that satisfies the conditions of the proposition. Let the contract \( C \) specify \( m = v_S \). Then in state H, the buyer has the incentive to install in period 1 if the contract stands, so let us specify \( a = 1 \). No renegotiation will take place in this state (because the buyer’s action to install yields an efficient outcome), so the buyer’s payoff from the beginning of period 1 is \( x - m \), the seller’s payoff is \( m \), and therefore \( \mathbf{v} \) is implemented. Note that specification of \( v_S \in [c, x] \) induces efficient investment and trade. \( \square \)

---

6Since the value function is defined from the beginning of period 1, it does not incorporate any cost of investment that the seller may incur in period 0.
2.3 The Model of a Durable Trading Opportunity

We next consider a durable trading opportunity in which the opportunity to trade persists indefinitely. Time is discrete and the parties discount the future using discount factor \( \delta \). Interaction in each period \( t = 1, 2, \ldots \) is identical to period 1 of the model without durable trade except that, from a given period \( t \), the game continues into period \( t + 1 \) if and only if the buyer does not install the good in period \( t \). The game ends at the end of the period in which the good is installed:

- **Period 0**
  - The parties write a contract \( C \) which specifies an externally enforced monetary transfer \( m \) to be compelled if the buyer installs the good.
  - The seller makes an investment choice, selecting between H at personal cost \( c \) and L at zero cost. Let \( \theta \) denote the seller's choice, which we call the state. This investment action is observed by the buyer but is unverifiable to the external enforcer.

- **Period \( t = 1, 2, \ldots \)**
  - The parties have an opportunity to renegotiate the contract.
  - The buyer chooses whether to install the good. This choice is verifiable.
  - The external enforcer compels the contractually specified monetary transfer, which is \( m \) if the good is installed and zero otherwise. If the buyer installs, then he obtains the benefit of the good and the relationship ends. Otherwise, the relationship continues into the next period.

In this description of the setting with a durable trading opportunity, we have limited attention to simple open-ended contracts. These contracts specify a single price \( m \) that the buyer must pay when he installs the good; the price \( m \) is constant across periods and the contract does not expire. We use this model to
evaluate how the durability of the trading opportunity affects the hold-up problem relative to the model without durable trade, and to demonstrate that there is a contract that gives the seller incentive to make the efficient investment.

As with the model with a nondurable trade opportunity, our analysis will be in terms of equilibrium continuation values. Because we now have an infinite number of periods, there will be a sequence of state-contingent continuation values. As before, the continuation values in state L will always be zero, so we can express the relevant values in terms of a sequence \( \{v_t^L\}_{t=1}^\infty \) of continuation values in state H. For each \( t \), \( \mathbf{v}^t = (v_B^t, v_S^t) \in \mathbb{R}^2 \) is the continuation payoff vector from period \( t \) in state H. We can also represent the buyer’s individual behavior in state H as a sequence \( \{a_t^H\}_{t=1}^\infty \), with the interpretation that if the game reaches period \( t \) then the buyer will take trade action \( a_t \) in this period.

We model the players’ behavior as in Section 2. In each period, the buyer’s choice of whether to install is made to maximize his payoffs. Furthermore, the resolution of renegotiation in each period is consistent with the bargaining solution that divides surplus according to the fixed bargaining weights. In state H it is rational for the buyer to install in period \( t \) if and only if \( x - m \geq \delta v_{t+1}^H \), because when the buyer does not install he gets zero in the current period and then waits for the continuation value from the start of the next period. Thus, in equilibrium, the buyer’s individual behavior and continuation payoffs must satisfy

\[
a_t^H = 1 \quad \text{only if} \quad x - m \geq \delta v_{t+1}^H,
\]

\[
a_t^H = 0 \quad \text{only if} \quad x - m \leq \delta v_{t+1}^H.
\] (2.1)

The opportunity for renegotiation in period \( t \) implies that, in state H,

\[
\mathbf{v}^t = \mathbf{w}^t + \pi[x - w_B^t - w_S^t],
\] (2.2)

where

\[
\mathbf{w}^t = \begin{cases} 
(x - m, m) & \text{if } a_t^H = 1 \\
\delta v_{t+1}^H & \text{if } a_t^H = 0.
\end{cases}
\]

In these expressions, \( \mathbf{w}^t \) denotes the disagreement point of negotiation, which is the value of continuing in the current period under the original contract. The parties can achieve the maximum joint surplus by renegotiating to a contract that forces trade, so the surplus of renegotiation in state H is \( x - w_B^t - w_S^t \).
In summary, an equilibrium is characterized by a sequence \( \{a^t, v^t\} \) such that conditions 2.1 and 2.2 hold for all \( t \). The H-state value \( v \) is said to be implemented by contract \( C \) if there is an equilibrium (combining sequential rationality and the bargaining solution) of the game from period 1 such that \( v = v^1 \).

Now that we have described the trade environment and precisely accounted for the trade technology, we can compare (i) the implementable set in the environment with a durable trading opportunity and (ii) the implementable set in the benchmark model with a nondurable trading opportunity.

**Proposition 2:** Any vector \( v \) that is implementable in the benchmark (nondurability) model is also implementable in the environment with a durable trading opportunity.

**Proof:** Consider a contract that specifies \( m \in [0, x] \). Note that this implements \( v = (x - m, m) \) in the model without durable trade. We shall demonstrate that the open-ended version of the contract implements the same H-state payoff vector in the durable trade setting. Specify \( \{a^t, v^t\} \) such that \( (a^t, v^t) = (1, (x - m, m)) \) for all \( t \). We will show that \( \{a^t, v^t\} \) is an equilibrium by checking that Conditions 2.1 and 2.2 hold for all \( t \). The first condition reduces to \( x - m \geq \delta (x - m) \) and so clearly holds. Note that \( w^t = (x - m, m) \) so the surplus of renegotiation is zero and the second condition holds. \( \| \)

Thus, there is an equilibrium of the game with a durable trading opportunity in which hold-up is not a problem and the parties invest and trade efficiently. This verifies Nöldeke and Schmidt’s (1998) intuition regarding open-ended option contracts.

Interestingly, in addition to demonstrating that durability does not necessarily exacerbate the hold-up problem, our model also gives some support to Edlin and Hermalin’s (2000) intuition about the buyer being able to credibly commit to refrain from installing the good until the contract is renegotiated. More precisely, a simple open-ended option contract may give rise to multiple equilibria, including one that is consistent with Edlin and Hermalin’s story.
Proposition 3: An open-ended option contract specifying a price of \( m \geq (1 - \delta \pi_B)x \) implements \( \hat{v} = (\pi_Bx, \pi_Sx) \).

Proof: Specify \( a^t = 0 \) and \( v^t = (\pi_Bx, \pi_Sx) \) for all \( t \). Under the assumption \( m \geq (1 - \delta \pi_B)x \), we have \( x - m \leq \delta \pi_Bx \) and so Condition 2.1 holds for all \( t \). Since \( a^t = 0 \) for all \( t \), \( w^t = \delta v^{t+1} \) and so condition 2.2 clearly holds as well. ||

If the players are patient (\( \delta \) is close to one) and the buyer has substantial bargaining power relative to the seller’s cost of high investment, then an open ended option with \( m \) near \( x \) supports an equilibrium in which the players divide the trade value according to their bargaining weights. In this equilibrium, unless \( c < \pi_sx \), the seller will not have the incentive to invest efficiently.

In summary, option contracts support equilibria with efficient outcomes as well as equilibria with hold-up. It is important to note, however, that the parties have an ex ante interest in selecting an equilibrium that induces an efficient outcome. If hold-up persists in environments with durable trading opportunities, it is because of adverse equilibrium selection.

2.4 Clarification of the Outside Option Principle

Many authors have noted the relationship between contracting models with renegotiation and models of bargaining with outside options. For example, Edlin and Hermalin (2000) appeal to the outside option principle, which holds that, with some bargaining protocols, only sufficiently large outside options affect the outcome relative to bargaining without outside options. In effect, the option serves as a constraint on the players’ payoffs, in that the player with the option must get at least his outside option payoff. If this payoff is less than the subgame perfect equilibrium payoff when there is no outside option, then any threat to opt out is non-credible and it does not affect the bargaining.

Though the outside option principle is motivated by a non-cooperative bargaining game, the set of equilibrium outcomes is very sensitive to assumptions about the timing of the option.\(^\text{7}\) This game also assumes that a single player has

\(^\text{7}\)For an analysis of the different cases, see Martin Osborne and Ariel Rubinstein (1990).
an outside option which, if taken, gives a positive payoff only to this player. In contrast, in the contract renegotiation under investigation here, opting out means exercising the contractual option (by installing the good) which generally gives positive payoffs to both players. In fact, the outside option yields a payoff vector that is close to the efficient frontier. We now introduce a simple model of bargaining that demonstrates how the outside option affects the outcome of bargaining.

Suppose player 1 and player 2 are bargaining over a surplus of size one. In each period one of the players is randomly selected to propose a division of the surplus. If the other party accepts the proposal, the game ends and the parties get the proposed payoffs. If the proposal is rejected, then player 1 has an opportunity to take an outside option and end the game. If the proposal is rejected and the outside option is not taken, the interaction repeats in the next period. Players discount future periods using discount factor $\delta$. The probability that player 1 gets to make the offer in any given period is $\pi_1$, whereas $\pi_2$ is the probability that player 2 gets to make the offer. If player 1 takes the outside option, he gets a payoff of $w_1$ and player 2 gets a payoff of $w_2$.

The bargaining model described here assumes that the offers made by the parties are efficient (the proposed payoffs sum to one). In addition, we focus on equilibria in which the outcome is efficient from the start of any given period. More generally, there may exist equilibria with delay (and therefore equilibrium payoffs that sum to less than one). The following two results characterize the equilibrium payoffs; the first result, which illustrates the basic idea of the outside option principle, is a special case of the second. See Figure 2.1 for a graphical depiction. Define $p^* \equiv \delta \pi_1/(1 - \delta \pi_2)$.

---

8In particular, equilibria with delay will arise in the cases with multiple equilibria described below.
Result 1: Suppose $w_2 = 0$. Then:

1. For $w_1 < \delta \pi_1$, the unique equilibrium payoff vector is $\pi = (\pi_1, \pi_2)$.

2. If $\delta \pi_1 \leq w_1 \leq p^*$ then, for all
   
   \[ z \in \left[ w_1 + (1 - \delta)\pi_1, \frac{p^*}{\delta} \right], \]
   
   there exists an equilibrium with payoffs $(z, 1 - z)$.

3. For $p^* < w_1$, the unique equilibrium payoff vector is $\left( \pi_1 + \pi_2 w_1, \pi_2(1 - w_1) \right)$.

Result 1 demonstrates the basic idea behind the outside option principle: the ability to opt out only affects the outcome of the bargaining when the value of the option is greater than the bargaining outcome in the game without an outside option (that is, when $w_1 > \delta \pi_1$). Settings of contract renegotiation, however, generally do not have the property that $w_2 = 0$. The following result describes the equilibria with general outside option payoff vectors.

Result 2: Consider any $w = (w_1, w_2)$.

1. (Region A) If $w_1 \leq \delta \pi_1$ and $w_1 + p^* w_2 \geq p^*$, then for all
   
   \[ z \in \left[ \pi_1(1 - w_2) + \pi_2 w_1, \pi_1 \right], \]
   
   there exists an equilibrium with payoff vector $(z, 1 - z)$.

2. (Region B) If $w_1 > \delta \pi_1$ and $w_1 + p^* w_2 > p^*$, then the unique equilibrium payoff vector is $\left( \pi_1(1 - w_2) + \pi_2 w_1, \pi_1 w_2 + \pi_2(1 - w_1) \right)$.

3. (Region C) If $w_1 < \delta \pi_1$ and $w_1 + p^* w_2 < p^*$, then the unique equilibrium payoff vector is $\pi$.

4. (Region D) If $w_1 \geq \delta \pi_1$ and $w_1 + p^* w_2 \leq p^*$, then for all
   
   \[ z \in \left[ w_1 + (1 - \delta)\pi_1, \frac{p^*(1 - w_2)}{\delta} \right], \]
   
   there exists an equilibrium with payoff vector $(z, 1 - z)$.
Figure 2.1: Regions of the space of outside-option payoff vectors.
To understand the implications of this result, consider a setting in which an efficient outcome occurs if player 1 takes the outside option; that is, \( w_1 + w_2 = 1 \). Recognize that this is essentially the contractual situation evaluated in the Section 2.3 with the contract specifying \( m \in [0, x] \). Refining further, consider the case in which \( w_1 < \delta \pi_1 \) (which corresponds to having \( m \) close to \( x \) in the contracting model). Then Region A is the relevant region of Result 2 and we have a continuum of equilibrium payoff vectors. There is an equilibrium in which player 1’s payoff \( z \) is equal to \( w_1 \), another equilibrium in which player 1’s payoff is \( z = \pi_1 \), and a continuum of equilibria between (that is, each \( z \in [w_1, \pi_1] \) can be supported in an equilibrium).

This analysis demonstrates that arguments about the outside option principle, for example by Edlin and Hermalin (2000), are implicitly arguments about equilibrium selection. Since player 2 gets a positive payoff if player 1 takes the option, and because the sum of the players’ outside option payoffs equals the whole surplus, Result 2 applies and there exists an equilibrium of the bargaining game (specifically, with \( z = w_1 \)) that gives the seller sufficient ex ante incentive to invest. Edlin and Hermalin’s story possibly relates to another equilibrium of the bargaining game, in which \( z = \pi_1 \). But note that, during the initial contracting stage the parties prefer to coordinate on first equilibrium because it gives optimal investment incentives.

Wickelgren (2007) also appeals to the outside option principle in his analysis of buyer option contracts. He mistakenly asserts that the seller’s valuation of the outside option does not affect the set of equilibrium payoffs. In addition, his discounting rule is equivalent to an assumption that the outside option yields a payoff vector on the line \( w_1 + w_2 = \delta \), meaning that by renegotiating the players somehow can trade sooner. Following Watson (2007), we view it more realistic to consider that the trade action takes place at a fixed point in time, whether or not renegotiation occurs before. More generally, we would be hard pressed to motivate how renegotiation activity could decrease the time to the next trading opportunity rather than increase it.
2.5 Strong Implementation with Non-Stationary Contracts

Returning to the contracting model, we next examine the issue of unique implementation. As we demonstrated earlier, stationary (constant price) contracts induce multiple equilibria for relatively high prices. Ex ante, the buyer and the seller may wish to ensure selection of an equilibrium that gives the seller the incentive to invest efficiently. We shall demonstrate that a nonstationary option contract can be structured so that, in each state, there is a unique equilibrium from period 1 and, furthermore, the seller invests efficiently.

We first expand the model from Section 3 by introducing more general contracts. Let $C = \{m^t\}_{t=1}^{\infty}$, so that the price specified by the contract depends on the period in which the buyer installs the good; that is, $m_t$ is the price paid if the buyer installs the good in period $t$. Equilibrium conditions 2.1 and 2.2 now become:

$$a^t = 1 \text{ only if } x - m^t \geq \delta v^{t+1}_B, \text{ and}$$

$$a^t = 0 \text{ only if } x - m^t \leq \delta v^{t+1}_B. \quad (2.3)$$

and

$$v^t = w^t + \pi [x - w^t_B - w^t_S], \quad (2.4)$$

where

$$w^t = \begin{cases} 
(x - m^t, m^t) & \text{if } a^t = 1 \\
\delta v^{t+1} & \text{if } a^t = 0.
\end{cases}$$

We say a contract $C$ uniquely implements a value function $v$ if there is one and only one sequence $\{a^t, v^t\}_{t=1}^{\infty}$ that satisfies Conditions 2.3 and 2.4 for all $t$, and such that $v^1 = v$.\footnote{This is a stronger notion of unique implementation than is typically studied, because we are insisting that the entire sequence $\{a^t, v^t\}_{t=1}^{\infty}$ be uniquely determined, rather than just $v^1$.}

**Proposition 4:** In the setting with a durable trading opportunity, any $v = (v_B, v_S)$ satisfying $v_S \in [0, x)$ and $v_B + v_S = x$ can be uniquely implemented (by a nonstationary contract).
Proof: Consider any $v = (v_B, v_S)$ such that $v_S \in [0, x)$ and $v_B + v_S = x$. Select an integer $T$. For all $t \geq T$, let $C$ specify $m^t = 0$. In state H, therefore, the buyer will exercise the option in period $T$ and will also do so in every subsequent period (Condition 2.3 is met). Clearly, continuation values in all periods $t \geq T$ are uniquely determined.

Next, define $\{m^t\}_{t=1}^{T-1}$ inductively so that, for each $t < T$,

$$m^t = \min \left\{ \frac{m^{t+1}}{2} + \frac{\delta m^{t+1} + x(1 - \delta)}{2}, v_S \right\}.$$  

By construction, we have

$$x - m^t > \delta(x - m^{t+1})$$

for every $t$, which means that the buyer strictly prefers to install in period $t$ rather than wait to do so in period $t + 1$. Because continuation values from period $T$ are uniquely determined, this implies (via backward induction) that continuation values in all previous period are also uniquely determined. In each period, contingent on trade not occurring earlier, the buyer will install. Letting $a^t \equiv 1$ for all $t$, we therefore know that $\{a^t, v^t\}_{t=1}^{\infty}$ uniquely satisfies Conditions 2.3 and 2.4.

By choosing $T$ sufficiently large, we have $m^1 = v_S$ and the desired value function $v^1$ is uniquely implemented.

Note that, since it would leave the buyer with a payoff of zero (and be indifferent between actions), a value function with $v_S = x$ cannot be uniquely implemented. ||

2.6 Conclusion

We have analyzed a model of unverifiable investment to demonstrate that durability of the trading opportunity is not necessarily a factor in the hold-up problem. In fact, simple open-ended option contracts can implement both the efficient (high investment) outcome and the problematic hold-up outcome. Furthermore, nonstationary option contracts can ensure that the seller has optimal incentives and that the efficient outcome is attained.

It is important to emphasize that these results are contingent on the specified trade technology: verifiable trade consummated by the buyer. As Watson
(2007) discusses, other trade technologies may lead to a more severe hold-up problem. Our point here is that durability of the trading opportunity does not on its own complicate the hold-up problem; rather, other components of the trade technology (in particular, who has the actions that consummate trade) are the key determinants of the severity of the hold-up problem. In other words, durability in itself is not a cause of trouble for contracting parties.

In our model of trade, installation by the buyer ends the interaction. This action, therefore, is not reversible (once trade has occurred, the parties cannot revert to the no-trade outcome). Watson and Wignall (2007) analyze a more general model of the durable trade interaction with reversible trade actions and analyze how differential reversibility affects the scope of contracting and the hold-up problem.

I thank Joel Watson, coauthor of this research, for permission to include it in this dissertation.
2.7 Appendix - Proof of Outside Option Results

Result 1 follows immediately from Result 2 (setting $w_2 = 0$), so we proceed to prove Result 2. The proof uses the standard technique of constructing relations between bounds on equilibrium payoffs in various subgames. We focus on equilibria in which the outcome is efficient in the continuation from the start of any given period.

Let $\eta$ and $\underline{\eta}$ be the maximum and minimum of player 1’s payoff from the beginning of a period, over all efficient subgame perfect equilibria in the game from this point. (The construction will establish that the bounds are met, so the maximum and minimum exist.) Let $\overline{\gamma}$ and $\underline{\gamma}$ be the maximum and minimum of player 1’s continuation values from the stage at which he has the opportunity to take the outside option. Similarly, let $\overline{\mu}$ and $\underline{\mu}$ be the maximum and minimum of player 2’s continuation values from the option stage. Then we have

$$\eta = \pi_1 (1 - \mu) + \pi_2 \overline{\gamma} \quad \text{and} \quad \underline{\eta} = \pi_1 (1 - \underline{\mu}) + \pi_2 \underline{\gamma}.$$  \hspace{1cm} (2.5)

For intuition, consider the first of these and note that when player 1 is selected to make the offer, the best equilibrium outcome for him is to hold player 2 down to her worst equilibrium outcome, which gives player 1 the payoff $1 - \mu$.

Our analysis continues by separately examining three cases having to do with whether player 1 has the incentive to take the outside option in a given period:

Case 1: $w_1 > \delta \overline{\eta}$, so player 1 always takes the outside option.

Case 2: $w_1 < \delta \underline{\eta}$, so player 1 never takes the outside option.

Case 3: $w_1 \in [\delta \underline{\eta}, \delta \overline{\eta}]$, so player 1 may take the outside option following some histories and forego the outside option following others.

In Case 1, the equilibrium continuation values are nailed down uniquely:

$\overline{\gamma} = \underline{\gamma} = w_1, \overline{\mu} = \underline{\mu} = w_2,$ and

$$\overline{\eta} = \underline{\eta} = \pi_1 (1 - w_2) + \pi_2 w_1.$$  

Define $p^* = \delta \pi_1 / (1 - \delta \pi_2)$. Note that, by substituting for $\overline{\eta}$, the presumption for this case ($w_1 > \delta \overline{\eta}$) becomes the necessary condition $w_1 + p^* w_2 > p^*$. 

The continuation values are also uniquely determined in Case 2. We obtain 
\( \bar{\gamma} = \delta \bar{\eta}, \gamma = \delta \eta, \bar{\mu} = \delta (1 - \eta), \) and \( \underline{\mu} = \delta (1 - \bar{\eta}). \) The presumption for this case becomes the necessary condition \( w_1 < \delta \pi_1. \)

Case 3 is more complicated. Considering player 1’s incentive in the option stage of a period, we see that the best equilibrium continuation for player 1 from the option stage is to forego the outside option and get \( \eta \) from the start of the next period. Thus, we have \( \bar{\gamma} = \delta \bar{\eta}. \) Likewise, the worst continuation has player 1 receiving \( \eta \) from the next period, which would motivate player 1 to take the outside option, implying \( \gamma = w_1. \) Player 2’s worst and best continuation values from the option stage will depend on the relative magnitude of \( w_2: \)

\[
\bar{\mu} = \max \{ w_2, \delta \left( 1 - \frac{w_1}{\delta} \right) \}
\]
\[
\underline{\mu} = \min \{ w_2, \delta (1 - \bar{\eta}) \}
\]

Here, attaining the bounds requires a selection over equilibria in the continuation from the next period.

We continue the analysis of Case 3 by exploring the possibilities for \( w_2. \) The inequality \( w_1 \leq \delta \bar{\eta} \) implies that \( \delta (1 - \bar{\eta}) \leq \delta (1 - w_1/\delta), \) and therefore there are three subcases to consider:

Subcase 3a: \( w_2 \leq \delta (1 - \bar{\eta}). \)

Subcase 3b: \( \delta (1 - \bar{\eta}) < w_2 < \delta \left( 1 - \frac{w_1}{\delta} \right). \)

Subcase 3c: \( \delta \left( 1 - \frac{w_1}{\delta} \right) \leq w_2. \)

Working through the straightforward implications of these inequalities, we obtain the following. In Subcase 3a, we have \( \bar{\mu} = \delta - w_1 \) and \( \underline{\mu} = w_2. \) Substituting for \( \bar{\mu}, \underline{\mu}, \bar{\gamma}, \) and \( \gamma \) in equations 2.5 and solving for \( \bar{\eta} \) and \( \underline{\eta} \) then yields

\[
\bar{\eta} = \frac{p^* (1 - w_2)}{\delta} \quad \text{and} \quad \underline{\eta} = w_1 + (1 - \delta) \pi_1.
\]

The presumptions for Case 3 and Subcase 3a translate into the necessary conditions \( w_1 \geq \delta \pi_1 \) and \( w_1 + p^* w_2 \leq p^* \). (The inequality defining Subcase 3a is implied by these two inequalities.)

Subcases 3b and 3c are handled the same way. In Subcase 3c, we obtain

\[
\bar{\eta} = \pi_1 \quad \text{and} \quad \underline{\eta} = \pi_1 (1 - w_2) + \pi_2 w_1 \text{ with necessary conditions } w_1 \leq \delta \pi_1 \text{ and}
\]
\( w_1 + p^* w_2 \geq p^* \). (The inequality defining Subcase 3c is implied by these two inequalities.) Subcase 3b turns out to be vacuous because solving for \( \bar{\eta} \) and \( \underline{\eta} \) and substituting in the presumptions for this subcase yields a contradiction.

Consider the four regions described in Result 2. Note that Region A satisfies the necessary conditions for Case 3c. In addition, the interior of Region A satisfies the necessary conditions for Cases 1 and 2 which identify equilibrium payoffs that are contained in the set identified by Case 3c. Thus, Case 3c gives the set of equilibrium payoffs for Region A. Region D is consistent with the necessary conditions of Case 3a and only this case. Region B is consistent with the necessary conditions of only Case 1, and Region C satisfies the necessary conditions of only Case 2. These facts imply the results.
Chapter 3

Efficiency and Entry in the Real Estate Brokerage Industry

3.1 Introduction

In this paper, I investigate the welfare effects of entry into the residential real estate brokerage market. My analysis of a unique data set finds no evidence that entry of new brokers decrease market productivity or efficiency. My findings contrast with the results of two other analyses of real estate brokerage. I argue that the discrepancy arises from data limitations and econometric specification errors of previous research.

In the neo-classical model of a market, free entry is one of the conditions of a market in long-run competitive equilibrium.\(^1\) There are, however, environments in which too many firms enter a market. Mankiw and Whinston (1986) show that in the presence of imperfect competition, entry can decrease social welfare. In some markets, new firms “steal business” from incumbents, decreasing those firms’ output and profits. Since this cost is not borne by the new firm, the private cost of entry is lower than the total social cost.

In addition, firms that are making positive profits may engage in costly

\(^1\)Mas-Colell, Whinston, and Greene define a market price, quantity, and set of firms as a long run competitive equilibrium if: 1. The firms are maximizing profits, 2. The market clears, and 3. There is free entry of firms
and unproductive rent-seeking behavior. Entry can lead to increased rent-seeking activity, further decreasing efficiency.

Since the theoretical prediction of the welfare effects of entry is ambiguous, analysis of these effects is necessarily empirical. To my knowledge, the first research of this type is Berry and Waldfogel’s (1999 and 2001) investigation of American radio stations. A more recent pair of papers, Hsieh and Moretti (2003) and Han and Hong (2008) evaluate the real estate brokerage market.

Following this latter pair of papers, I focus my attention on entry in residential real estate brokerage. There are two general characteristics of this market that, taken together, make inefficiently high entry seem like a possible outcome. First, real estate agents are typically paid a percentage of the sales price of any deals they broker. As I will discuss in the next section, this commission rate is almost always 6% of the total sales price (3% per broker if both sides are represented) and does not appear to vary within or between markets. Since the total number of transactions (home sales) and the house price are set in the housing market, the total revenue available to real estate agents is essentially exogenous to the brokerage market. At the same time, entry into the brokerage market is quite low cost. As Hsieh and Moretti argue, an increase in the price of housing leads to an increase in the agents’ revenues and induces entry. This entry is inefficient since it is pure “business stealing.”

Researchers have taken two different approaches in this analysis of real estate brokerage. Hsieh and Moretti take an indirect approach: they present a simple model of real estate brokerage and derive its testable implications. Their three main predictions, which I will discuss in depth in Section 4, are:

1. Markets with higher home prices have more real estate agents (per capita).
2. Increases in house price are not associated with increases in brokers’ profits.
3. Brokers in markets with higher home prices are less productive.

Their empirical results are consistent with their predictions. Using the estimated relationships between the variables, they offer estimates of the social cost of inefficient entry between $1.2 Billion and $8.2 Billion.
Han and Hong analyze entry by directly estimating brokers’ cost functions. If operating at a point on their cost curve where returns to scale are increasing, then “business stealing” is socially wasteful. In addition, they look for costly rent-seeking behavior. They find evidence of both of these inefficiencies and estimate that 47% of brokers’ average variable cost is wasteful.

In this paper, I follow an approach similar to Hsieh and Moretti: I investigate whether the brokerage industry has the observable characteristics of a market with too many firms. Unlike the above papers, however, I find little evidence of inefficient entry.

Data limitations and econometric specification issues in earlier studies can explain the difference in results. Both Hsieh and Moretti and Han and Hong rely on US Census data which, as I discuss in depth in Section 2, carries two substantial limitations. I discuss specific problems with the previous research in Section 4. My main points are that, first, Hsieh and Moretti’s econometric model suffers from severe specification problems, including omitted variable bias. Second, Han and Hong’s assumption that entrants’ expected revenue is a function of observable characteristics (which do not include incumbency) is implausible. In effect, they assume that entrants have the same expected revenue as incumbents with the same observable characteristics. This leads them to over estimate entrants’ expected revenue and, ultimately, costs.

In the next section, I discuss the real estate brokerage industry and present evidence for the fixed commission structure and free entry. I also describe my data and discuss its strengths and weaknesses relative to the census data used by Hsieh and Moretti and Han and Hong.

### 3.2 Real Estate Brokerage

A real estate agent’s tasks include marketing, negotiating, and helping with paperwork. One of the most valuable services offered is access to the multiple listings service (MLS), a collection of the details of the homes listed by a large group of local brokers. The MLS gives buyers a single source of available properties
and is, therefore, an important marketing outlet for sellers. Access to the MLS is typically restricted to people who are represented by a member of the MLS.

The contract between a home seller and his broker promises the broker a percentage of the sales price of the home in exchange for representation. If she is a member of the MLS, the agent is required to post the listing within a few days of signing the listing agreement. Perhaps the most important component of an MLS property listing is the commission offered to the buyer’s broker. In effect, the listing is an open contract between the listing agent and any member of the MLS that brings a buyer for the home. In this sense, the broker who represents the home buyer is an agent of the listing agent, who sets the terms of the contract and pays the commission. A typical listing agreement will promise 6% of the sales price to the listing agent. The listing agent then, in turn, will post a promised commission to the buyer’s broker of 3% - half of her own commission.

In this paper, I use data from the Wasatch Front Regional MLS (WFRMLS), the largest MLS in Utah and the only MLS in the Wasatch Front region (which includes about 80% of the state’s population). Each observation is a listing, and includes information on the listing agent, the location of the property, and the home characteristics. For listings that successfully sell, the data include the final sales date, price, and the identity of the buyer’s agent. My data include listings from June 1996 through the November 2008. Since I only have half of 1996 and since many of the listings from the end of 2008 were still active when the data were collected, listings from those years will be left out of much of the following analysis. Table 1 presents summary statistics of the listings. Table 2 summarizes information on the individual agents. The statistics represent the average annual activity and earnings of agents who have at least one transaction recorded in the MLS.

In the early 1980’s, the Federal Trade Commission conducted an investigation of the real estate brokerage industry for anti-trust violations. One of the FTC’s fundamental concerns was the uniformity of commission rates, which was documented in the final report. There have been anecdotal reports that some inno-

\[\text{If a buyer does not have his own agent, the listing agent acts as the buyer’s agent and the MLS is annotated to reflect this.}\]
vations in the brokerage industry, such as online MLS access and internet discount brokers, have put downward pressure on the commission rate and may have broken the fixed rate structure. The WFRMLS data do not corroborate these stories.

Figure 1 presents a histogram of the commission offered to the buyer’s broker in my data. The vast majority of offers are exactly 3% and over 90% are between 2.5% and 3.5%. Table 3 shows how the data vary over some specific subsamples. The mean commission offered is 2.9% and does not vary over time. Figure 2 shows a scatter plot of the commission against the sales price of a randomly selected subsample of 1,000 homes. Since the commissions are concentrated at 3%, the points in the scatter are jittered by adding white noise to each variable. The figure illustrates that the offers are fairly uniform and do not vary much with house price.

The total commission offered to the seller’s agent (which includes any commission she may offer to a buyer’s broker) is specified in the listing agreement. There is a field in the WFRMLS form that allows agents to enter the commission received by the seller’s broker, but the field is not required and it is available for fewer than 2% of the listings. These 10,000 observations are still informative and are, in fact, more information about the listing agent’s commission than is typically available to researchers. Figure 3 presents a histogram of the reported commissions. Again, the majority are exactly 3% and almost all are between 2.5% and 3.5%.

Figure 4, a scatter plot of the reported list agent’s commission against the commission offered to the buyer’s broker, tells an interesting story about the agents’ behavior. Obviously, most of the listings have 3% of the sales price going to each broker. Other than that, though, there are two patterns. First, there is a tendency to offer 3% to the buyer’s broker, so that if a list agent lowers her commission rate, she takes the total cut herself. See, for example, the large group of observations where the list agent gets 2.5% and still offers 3% to the buyer’s agent. This is consistent with the marketing and matching use of the MLS. List agents are trying to induce buyers’ brokers to bring their clients to see listings, so they must compete on price (the commission offered) with other listing agents.
The second pattern of behavior is the group of points along the 45 degree line. Some of this may represent data entry errors - agents or their administrative assistants accidentally entering the commission offered to the buyer’s broker into both fields. It may also, however, represent a tendency to split total commission evenly between the agents. This second interpretation is consistent with the location of the masses of observations. The largest group is at (3%, 3%), which is consistent with both stories. The other two large groups are at (3%, 2%) and (2.5%, 2.5%), which represent a total commission of 5% divided according to the two different methods.

The MLS data on commissions show quite convincingly that commissions are relatively fixed in the sample. The vast majority of offered commissions are at exactly 3% with some small variation above and below. The limited information on total commission shows that the majority of listing agreements are for a total of 6% which the listing agent splits evenly with the buyer’s broker. There is some variation in the total commission, but nearly all are either 5% or 6%.

There are many possible explanations of the fixed commission structure in real estate brokerage and how it is maintained in the presence of the low barriers of entry that characterize the industry. The important fact for this paper, however, is simply that the commission rate is steady and remains so despite changes to the market structure.

The Federal Trade Commission and the US Department of Justice published a joint report in 2007 that characterized residential real estate brokerage as an industry with free entry: “the nearly universal opinion is that there are no significant barriers to entry, if entry is construed as gaining a license in order to practice.” The state licensing requirements are typically very limited. In Utah, the requirements are basically a high school diploma, 90 hours of certified training, a state examination, and $147.3

Data collected from the Utah Division of Real Estate indicate that agents are entering the market frequently. The licensing data that is publicly available is of limited use in this analysis because, first, it is not matched to the MLS data,

---

3(from the Utah Division of Real Estate)
second, state law requires real estate licenses for professionals outside the scope of this study (such as commercial agents, leasing agents, and some administrative staff), and, finally, the data only report the date the license was granted and the date it expires. Individuals must pay a nominal annual fee to keep their licenses active, so many agents will suspend their license during periods of inactivity. This activity is not publicly available from the Division. The data indicate a significant increase in the number of licenses granted: 1,603 new sales agent licenses in 1996 and 3,309 2007.

The MLS data I use offer some substantial benefits relative to the census data used by Hsieh and Moretti and Han and Hong. Hsieh and Moretti use information from the 5% sample of the 1980 and 1990 Census of Population of Housing. Han and Hong use mainly the same sample of the 2000 Census. Real estate agents are identified by their occupational category (“Real Estate Sales occupation”) and their wages and hours are reported. The number of homes transacted is imputed by counting the number of home owners in the survey who moved within the last year. The average house price is calculated as the average home value reported.

The census data offer several benefits relative to the MLS data. First, the census has data for a large number of markets across the US while I only have information for the relatively small Utah market. The census data has information on hours worked which is not available through the MLS. Finally, the census has information on wages that I have to impute from the MLS data. This means that their earnings figures are net of costs while mine are gross.

Despite these limitations of the MLS data, it still has several characteristics that make it more appropriate for the investigation of the welfare effects of entry. In particular, the MLS data offer individual level observations of productivity. I can see how many homes each agent has listed and sold. This means that, unlike Hsieh and Moretti, I can allow for heterogeneous performance of agents. Since I observe each agent over time, I can allow for agent heterogeneity in estimation. In addition, I can use an agent’s first observed activity as a measure of their market entry. As I discuss in Section 4, the market performance of incumbents and entrants is substantially different and failure to account for this can cause
problems in testing and estimation. Finally, the MLS data are more informative about the actual set of agents in the market. In particular, the 1980 and 1990 “Real Estate Sales occupation” includes, along with sales agent, appraisers and leasing agents.

3.3 Theoretical Framework

While Han and Hong directly estimate brokers’ costs, Hsieh and Moretti present a simple model of inefficient entry in real estate brokerage and derive its testable predictions. They extend their simple model to include heterogeneity of agents, but they do not present a solution to the extended model or discuss its implications.

In this section I present both models. I describe the three testable implications of the simple model that are at the core of Hsieh and Moretti’s empirical approach. I also characterize the solution of the heterogeneous agents model and compare the empirical predictions of the two models.

Real Estate Brokerage with Homogeneous Agents

In a simple market, \( S \) identical homes transact at price \( P \). \( N \) identical real estate agents, with no costs of production or capacity constraints, split the market evenly, so that each has a share equal to \( Q/N \). Agents receive \( c \cdot P \) for every transaction. Individuals outside of the brokerage industry receive market wage \( w \).

The housing market determines \( P \) and \( S \); the commission rate \( c \) is fixed. Total revenue in the brokerage industry, \( R \equiv c \cdot P \cdot S \), is therefore exogenous. Agents can enter freely, so equilibrium is characterized by \( N \) such that:

\[
w = \frac{c \cdot P \cdot \frac{S}{N}}{N}
\]

\[\Rightarrow N = c \left( \frac{P}{w} \right) S\]
Define average productivity as \( S/N \). Then rearranging (2) yields

\[
Productivity = \frac{1}{c} \frac{1}{P/w}
\] (3.3)

These equations together generate the three testable predictions that motivate Hsieh and Moretti’s empirical approach. These equations hold in each housing market. When a group of such markets is observed, the model predicts that, ceteris paribus,

1. Markets with higher home prices have more real estate agents (per capita).
2. Increases in house price are not associated with increases in brokers’ profits.
3. Brokers in markets with higher home prices are less productive.

**Real Estate Brokerage with Heterogeneous Agents**

Agents are not identical and do not have the same profits or number of transactions. Hsieh and Moretti present the following model of real estate brokerage with heterogeneous agents.

There is a unit mass of individuals, \( i \sim Uniform[0,1] \). The index \( i \) represents idiosyncratic ability that impacts market wages and brokerage earnings. The outside wage for individual \( i \) is:

\[
w_i = w + \theta \cdot i
\]

Let \( A \subset [0,1] \) be the set of brokers. Brokerage revenue for person \( i \) is:

\[
\frac{i}{\int_A i\,di} \cdot R
\]

See Figure 5 for a graphical representation. Since both of these revenue functions are linear in \( i \) and, for \( w > 0 \), the market wage for \( i = 0 \) is greater than brokerage earnings, there are two possible solutions:

1. No entry (the lines don’t cross)
2. Entry for all \( i > a \) for some \( a \in (0,1) \) (the lines cross at \( i = a \)) \(^4\)

\(^4\)Note that the solution described by Hsieh and Moretti does not exist.
Equilibria with entry are fully characterized by \( a \), where \( a \) solves:

\[
w + \theta \cdot a = \frac{2a}{1 - a^2} \cdot R
\]

An increase in \( R \) corresponds to a shift to the left and a flattening of the brokerage revenue line. This means that, as before, and increase in \( R = (c \cdot P \cdot S) \) leads to entry. In contrast with the prediction of the first model, however, the net effect on wages is positive: broker \( i \)'s revenue is increasing with \( R \). In this case, only the marginal broker’s wages are unresponsive to \( R \).

The functional form guarantees that market share has a constant elasticity with respect to \( R \). When new brokers enter in response to price changes, they “steal business” from incumbents at a rate proportional to the incumbents’ market share. This means that more successful brokers (higher \( i \)) are actually hurt more by entry than their competitors.

The model with heterogeneity adds substantial insight into the empirical implications of the homogeneous agent model. The predicted relationship between home price and the number of real estate agents is a implication of the free entry assumption: An increase in \( P \) leads to an increase in \( R \), which, under free entry, induces more people to become agents. The models make different predictions about the effect of price changes on profits, so brokers earnings can increase with house prices in a market with inefficient entry. Entry in response to changes in house price alone necessarily decreases productivity in both models.

### 3.4 Empirical Results

In this section I present my empirical analysis of entry in real estate brokerage. I argue that the data offer little evidence that new agents are causing inefficiency. I will also discuss in detail some problems with the analysis of Hsieh and Moretti and Han and Hong that may account for the difference in results.

Figures 6 through 14 preview the discussion in the remainder of this section. The don’t represent rigorous analysis of the empirical relationships, but the pictures clearly demonstrate two of the main characteristics of the data: first, many of the variables move together and, second, incumbents have higher wages and are
Figure 6 plots the average price and the number of homes sold. An important characteristic of the data is illustrated—while I assume that the housing market variables are exogenous to the real estate brokerage market, they are not independently distributed. They reflect the relationship between supply and demand of housing and are correlated (in this case, negatively correlated).

Figures 7 and 8 give a substantial insight into the impact of market variables on broker entry. While entry clearly traces the market quantity (Figure 8), the response to price is much less obvious. Figure 9 shows that, on the other hand, brokers’ commissions do appear to move with price.

Figures 10 and 11 illustrate the heterogeneity in brokers’ outcomes. Incumbent brokers (those who have been in the market for at least one year) have higher wages and are more productive than the average broker while entrants do much worse than average.

Finally, Figures 12 through 14 illustrate the relationships between individual productivity and the market variables. Specifically, the relationship between average price and average productivity appears much weaker than the relationships between productivity and quantity or the number of agents.

Consider the following model:

\[ y_{ijt} = \alpha_i + \alpha_j + \alpha_t + \beta_1 P_{jt} + \beta_2 Q_{jt} + \beta_3 N_{jt} \]

where \( i \) indexes the individual agent, \( j \) indexes the market/city, and \( t \) indexes the year. \( y_{ijt} \) is one of three possible measures of productivity for agent \( i \) in market \( j \) during year \( t \): the total number of transactions, the number of homes sold, or the number of times the agent represented a home buyer who successfully purchased...
a home. Results from fixed effects estimation are presented in Table 8 and show no negative effect of changes in price on productivity of agents.

While this estimation (particularly columns 1 through 3) shows that home price is not negatively correlated with productivity, the number of agents in the market does have a negative relationship with productivity. When new brokers enter the market in response to an increase in the housing market, do they drive down productivity? To answer this question, I estimate the portion of $N$, the number of agents, that can be explained by the house price. Table 10 shows that this type of entry (due to changes in price alone) is not associated with reduced productivity.

The observation that the variables from the housing market can be arbitrarily correlated in any sample is critical to my analysis of Hsieh and Moretti. They want to evaluate the relationships between the market house price and the number, wages, and productivity of real estate agents. By taking logs of the model equations, they propose the following econometric models of the relationships:

1. $\log(N_j) = \alpha_1 + \beta_1 \log \left( \frac{P_j}{w_j} \right) + \varepsilon_j$
2. $\log(\bar{\pi}_j) = \alpha_2 + \beta_2 \log(P_j) + \varepsilon_j$
3. $\log \left( \frac{Q_j}{N_j} \right) = \alpha_3 + \beta_3 \log \left( \frac{P_j}{w_j} \right) + \varepsilon_j$

where $N_j$ is the number of agents, $P_j$ is the average house price, $Q_j$ is the number of home transactions, $w_j$ is the outside wage, where $\bar{\pi}_j$ is the average profits in market $j$ (calculated as the difference between the average reported earnings and the average outside wage).

Hsieh and Moretti formulate their three predictions as tests of the estimated parameters from the above predictions, specifically that $\beta_1 > 0$, $\beta_2 = 0$, and $\beta_3 < 0$. Their main results are reproduced in Table 4. The results from the difference specification (using the differences in the variables from 1980 to 1990) are consistent with their predictions. These results, however, should not be interpreted as evidence for inefficiently high entry in real estate brokerage. First, each of these specifications suffers from at least one significant error that means that the estimated parameter does not actually represent the relationship of interest.
Second, they fail to test their model against the main alternative hypothesis: a standard neo-classical labor supply model.

Each of the three models above includes the natural log of the ratio of two variables. In the first and third regressions, for example, the regressor is \( \log \left( \frac{P_j}{w_j} \right) \), which is equivalent to \( \log(P_j) - \log(w_j) \). This restriction of the model means that the estimates of \( \beta \) reflect the relationships between the regressand and both regressors, and hypothesis tests are actually tests of this joint relationship. When Hsieh and Moretti reject \( \beta = 0 \), they are actually rejecting the restricted model.

The problem can be clearly illustrated with a simple thought experiment. Suppose first that the number of agents is negatively correlated with the outside wage (so that when the outside wage goes up, fewer people become agents) but not correlated with the house price at all. The estimate of \( \beta_1 \) will be positive, even though the underlying data are completely inconsistent with the entry model. Alternatively, suppose that the model is true. When house prices go up by 1%, the number of agents goes up by 1%. Similarly, when the outside market wage goes up by 1%, the number of agents goes down by 1%. In this case, the restricted model is true, so the estimated \( \beta \) is zero.

The problem is more complicated when the log ratio is on the left hand side of the equation. Consider the following three regressions:

\[
\begin{align*}
\log \left( \frac{Q_j}{N_j} \right) &= \alpha + \beta \log(P_j) \\
\log(Q_j) &= \gamma_1 + \delta_1 \log(P_j) \\
\log(N_j) &= \gamma_2 + \delta_2 \log(P_j)
\end{align*}
\]

An estimate of \( \beta \) will simply reflect the relationships in the second and third regressions: \( \beta = \delta_1 + \delta_2 \). Table 5 presents the results of a simple replication of this type of regression. In order to be consistent with Hsieh and Moretti’s approach, I use the differences in market variables from 1997 to 2007. Since I only have 65 cities in the sample, estimates are very imprecise but illustrative. As before, the model cannot make any refutable predictions about the parameters of the regression with the log ratio.

The natural log specification makes the characterization of the problem
straightforward, but it is not the problem. A regression of $Q_j/N_j$ on $P_j$ would face similar problems - the parameter estimate would partially reflect the individual relationship between $Q_j$ and $P_j$ as well as the relationship between $N_j$ and $P_j$. The limitations of the census data make it incapable of addressing this point since it has no alternative definition of productivity.

This criticism leads immediately to the next point: the market variables $P$ and $Q$ are correlated, and yet $Q$ is omitted from all three regressions. Tables 6 and 7 present my simple illustration of this point for both of the first two regressions: there is clearly an omitted variable bias. In the third regression, there are actually two omitted variables - $N$ and $Q$. Since these variables are correlated with each other and with $P$, estimates of the relationship between $Q/N$ and $P$ will include a bias for each of these omissions.

Again, in this third case the census data cannot overcome this problem. The only available definition of productivity involves the market variables, which are all related.

Hsieh and Moretti’s empirical approach is frustrated by these two problems. Even though respecification is possible for the first two predictions, their most important test - that productivity is negatively impacted by entry - cannot be evaluated with their data. The MLS data, however, do provide an alternative definition of productivity that allows investigation of this issue. By considering individual agents, I can evaluate how productivity is impacted by entry and the housing market.

3.5 Conclusion

I have argued that there is no evidence of inefficiently high entry in the real estate brokerage market. My analysis of data from the Wasatch Front Regional MLS finds that the relationship between house prices and agent productivity is weak. Since entry induced by changes in price alone does not negatively impact productivity but does impact wages, it may be the case that there are significant barriers to entry despite the minimal licensing requirements.
These results contrast with other empirical findings. I’ve argued, however, that this previous research faces severe data limitations that force researchers to make strong assumptions that are not consistent with the observable characteristics of the market.

Entrants, on average, engage in fewer transactions and generate much less revenue than incumbents. There may be other characteristics of the brokerage industry that serve as barriers to entry, since they preserve incumbent profits from the threat of new brokers. Possible explanations include reputation effects (see Gathright and Wignall(2009)).
Table 3.1: Summary Statistics - Market Variables

Summary Statistics - Market Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Entire Sample</th>
<th>Salt Lake City</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Listings</td>
<td>665208</td>
<td>118707</td>
<td>8109.4</td>
<td>14687.4</td>
</tr>
<tr>
<td>Total Sold</td>
<td>350520</td>
<td>66468</td>
<td>4306.8</td>
<td>8262.7</td>
</tr>
<tr>
<td>Average Nominal Price</td>
<td>183018.8</td>
<td>184675.3</td>
<td>193772.4</td>
<td>60460.8</td>
</tr>
<tr>
<td>Average Real Price</td>
<td>155833.6</td>
<td>159237.9</td>
<td>162941.7</td>
<td>50051.3</td>
</tr>
<tr>
<td>Total Agents</td>
<td>24450</td>
<td>11415</td>
<td>2019.4</td>
<td>1897.2</td>
</tr>
</tbody>
</table>
Table 3.2: Summary Statistics - Brokers

<table>
<thead>
<tr>
<th>Variable</th>
<th>All Agents</th>
<th>Incumbents</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>Homes Listed</td>
<td>6.901</td>
<td>15.79</td>
</tr>
<tr>
<td>Homes Sold</td>
<td>4.033</td>
<td>12.577</td>
</tr>
<tr>
<td>Homes Bought</td>
<td>3.481</td>
<td>4.942</td>
</tr>
<tr>
<td>Earnings</td>
<td>42110.307</td>
<td>64304.103</td>
</tr>
<tr>
<td>Years Active</td>
<td>4.091</td>
<td>3.46</td>
</tr>
</tbody>
</table>

Table 3.3: Summary Statistics - Commission Rates

Table of means for the commission rate offered in the Wasatch Front Regional Multiple Listing Service (WFRMLS). Subsamples include only listings that sold, only listings that didn’t sell, listings in Salt Lake City, and listings by year for each year in the sample.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsold Homes</td>
<td>.030</td>
<td>.004</td>
</tr>
<tr>
<td>Sold Homes</td>
<td>.029</td>
<td>.006</td>
</tr>
<tr>
<td>Salt Lake City</td>
<td>.030</td>
<td>.004</td>
</tr>
<tr>
<td>1997 Listings</td>
<td>.030</td>
<td>.005</td>
</tr>
<tr>
<td>1998 Listings</td>
<td>.030</td>
<td>.005</td>
</tr>
<tr>
<td>1999 Listings</td>
<td>.030</td>
<td>.005</td>
</tr>
<tr>
<td>2000 Listings</td>
<td>.030</td>
<td>.005</td>
</tr>
<tr>
<td>2001 Listings</td>
<td>.030</td>
<td>.005</td>
</tr>
<tr>
<td>2002 Listings</td>
<td>.029</td>
<td>.006</td>
</tr>
<tr>
<td>2003 Listings</td>
<td>.029</td>
<td>.006</td>
</tr>
<tr>
<td>2004 Listings</td>
<td>.029</td>
<td>.006</td>
</tr>
<tr>
<td>2005 Listings</td>
<td>.029</td>
<td>.006</td>
</tr>
<tr>
<td>2006 Listings</td>
<td>.029</td>
<td>.005</td>
</tr>
<tr>
<td>2007 Listings</td>
<td>.030</td>
<td>.004</td>
</tr>
<tr>
<td>2008 Listings</td>
<td>.030</td>
<td>.004</td>
</tr>
<tr>
<td>Total</td>
<td>.029</td>
<td>.005</td>
</tr>
</tbody>
</table>
Table 3.4: Reproduction of results from Hsieh and Moretti (2003)


<table>
<thead>
<tr>
<th></th>
<th>( \Delta ) Number of Agents</th>
<th>( \Delta \bar{\pi} )</th>
<th>( \Delta \log(Q/N) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta(P/w) )</td>
<td>0.917</td>
<td>0.064</td>
<td>-0.646</td>
</tr>
<tr>
<td></td>
<td>(0.078)</td>
<td>(0.041)</td>
<td>(0.069)</td>
</tr>
<tr>
<td>Observations</td>
<td>Not Reported</td>
<td>282</td>
<td>282</td>
</tr>
<tr>
<td>R(^2)</td>
<td>Not Reported</td>
<td>0.05</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Table 3.5: Regression: Productivity on Price

Regression of the change in average productivity of agents (\( \log(Q/N) \)) on the change in house price (\( \log(P) \)). The regression estimates, while imprecise, are illustrative of the arithmetic relationship between the regression of the ratio and the regressions of the individual variables (\( \log(Q) \) and \( \log(N) \)). Specifically, \( 0.099 = -0.587 - (-0.686) \).

<table>
<thead>
<tr>
<th>Variables</th>
<th>( \log(Q/N) )</th>
<th>( \log(Q) )</th>
<th>( \log(N) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in log Price</td>
<td>0.099</td>
<td>-0.587</td>
<td>-0.686</td>
</tr>
<tr>
<td></td>
<td>(0.187)</td>
<td>(0.630)</td>
<td>(0.641)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.001</td>
<td>-0.274</td>
<td>-0.275</td>
</tr>
<tr>
<td></td>
<td>(0.089)</td>
<td>(0.246)</td>
<td>(0.248)</td>
</tr>
<tr>
<td>Observations</td>
<td>65</td>
<td>65</td>
<td>65</td>
</tr>
<tr>
<td>R(^2)</td>
<td>0.002</td>
<td>0.013</td>
<td>0.016</td>
</tr>
</tbody>
</table>

** p<0.01, * p<0.05
Table 3.6: Regression: Number of Agents on Price and Quantity
Regression of the change in the log number of agents for 65 Utah cities from 1997 to 2007. The very small sample size leads to imprecise estimates of the relationships, but it is clear that the number of agents is more strongly correlated with the number of homes sold than with the average house price. Any estimation that neglects this relationship is likely to suffer from an omitted variable bias.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Change in log Price</th>
<th>Change in log Quantity</th>
<th>Constant</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.686 (0.641)</td>
<td>0.962** (0.0861)</td>
<td>-0.275 (0.248)</td>
</tr>
<tr>
<td></td>
<td>-0.123 (0.215)</td>
<td>0.959** (0.0891)</td>
<td>0.00212 (0.0867)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-0.0123 (0.0797)</td>
</tr>
</tbody>
</table>

Observations: 65 65 65
R²: 0.016 0.858 0.858

** p<0.01, * p<0.05
Table 3.7: Regression: Wage on Price, Quantity, and Number of Agents

Regression of the change in log mean (and median) wage for 65 Utah cities from 1997 to 2007. Again, the small sample size leads to imprecise estimation, but the potential for omitted variable bias is illustrated - consider whether changes in price are correlated with changes in wages.

In column 1, where other market variables are omitted, I fail to reject the hypothesis that changes in price covary with changes in wage. Once I include the omitted market quantity and number of agents (columns 2 and 3), we see that, in fact, wages and prices move together in the data.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Δ Mean Wage</th>
<th></th>
<th></th>
<th>Δ Median Wage</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in log Price</td>
<td>0.113</td>
<td>0.737**</td>
<td>0.758**</td>
<td>-0.205</td>
<td>0.204</td>
</tr>
<tr>
<td></td>
<td>(0.320)</td>
<td>(0.254)</td>
<td>(0.242)</td>
<td>(0.271)</td>
<td>(0.208)</td>
</tr>
<tr>
<td>Change in log Quantity</td>
<td>0.311**</td>
<td>0.0562</td>
<td></td>
<td>0.204**</td>
<td>-0.0643</td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.163)</td>
<td></td>
<td>(0.021)</td>
<td>(0.183)</td>
</tr>
<tr>
<td>Change in log Number of Agents</td>
<td>0.234</td>
<td></td>
<td>0.246</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.153)</td>
<td></td>
<td>(0.174)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.591*</td>
<td>0.0931</td>
<td>0.0895</td>
<td>0.317</td>
<td>-0.00978</td>
</tr>
<tr>
<td></td>
<td>(0.245)</td>
<td>(0.080)</td>
<td>(0.092)</td>
<td>(0.159)</td>
<td>(0.080)</td>
</tr>
<tr>
<td>Observations</td>
<td>65</td>
<td>65</td>
<td>65</td>
<td>65</td>
<td>65</td>
</tr>
<tr>
<td>R²</td>
<td>0.002</td>
<td>0.674</td>
<td>0.693</td>
<td>0.012</td>
<td>0.479</td>
</tr>
</tbody>
</table>

** p<0.01, * p<0.05
Table 3.8: Fixed Effects

Fixed effects estimation of the relationship between various measures of agent productivity and the real estate and brokerage markets. The measures of productivity are: 1. The total number of transactions the agent engaged in (including unsold listings) 2. The total number of listings that sold and 3. The total number of times the agent represented a buyer in a successful purchase. Fixed effects are allowed for the individual agents. Also included (but not reported) are dummy variables for the markets and year.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Transactions</th>
<th>Sold</th>
<th>Bought</th>
<th>Transactions</th>
<th>Sold</th>
<th>Bought</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home Price</td>
<td>0.117*</td>
<td>-0.060</td>
<td>-0.074</td>
<td>0.129*</td>
<td>-0.010</td>
<td>-0.039</td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td>(0.058)</td>
<td>(0.052)</td>
<td>(0.057)</td>
<td>(0.058)</td>
<td>(0.051)</td>
</tr>
<tr>
<td>Homes Sold</td>
<td>0.214**</td>
<td>0.213**</td>
<td>0.186**</td>
<td>0.283**</td>
<td>0.397**</td>
<td>0.320**</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.022)</td>
<td>(0.020)</td>
<td>(0.029)</td>
<td>(0.034)</td>
<td>(0.0284)</td>
</tr>
<tr>
<td>Number of Agents</td>
<td></td>
<td></td>
<td></td>
<td>-0.096**</td>
<td>-0.227**</td>
<td>-0.164**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.027)</td>
<td>(0.030)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.753</td>
<td>0.804</td>
<td>1.170</td>
<td>-0.854</td>
<td>0.220</td>
<td>0.746</td>
</tr>
<tr>
<td></td>
<td>(0.711)</td>
<td>(0.726)</td>
<td>(0.639)</td>
<td>(0.705)</td>
<td>(0.720)</td>
<td>(0.638)</td>
</tr>
<tr>
<td>Observations</td>
<td>81609</td>
<td>55176</td>
<td>66085</td>
<td>81609</td>
<td>55176</td>
<td>66085</td>
</tr>
<tr>
<td>R²</td>
<td>0.025</td>
<td>0.023</td>
<td>0.031</td>
<td>0.025</td>
<td>0.025</td>
<td>0.032</td>
</tr>
</tbody>
</table>

** p<0.01, * p<0.05
Table 3.9: Fixed Effects - Incumbents Only

Fixed effects estimation of the relationship between various measures of agent productivity and the real estate and brokerage markets this estimation is performed solely on the subsample of agents that are *incumbents* in a given year - they were active in the previous year. The measures of productivity are: 1. The total number of transactions the agent engaged in (including unsold listings) 2. The total number of listings that sold and 3. The total number of times the agent represented a buyer in an successful purchase. Fixed effects are allowed for the individual agents. Also included (but not reported) are dummy variables for the markets and year.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Transactions</th>
<th>Sold</th>
<th>Bought</th>
<th>Transactions</th>
<th>Sold</th>
<th>Bought</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home Price</td>
<td>-0.041</td>
<td>-0.143*</td>
<td>-0.150**</td>
<td>-0.028</td>
<td>-0.090</td>
<td>-0.115*</td>
</tr>
<tr>
<td></td>
<td>(0.061)</td>
<td>(0.063)</td>
<td>(0.057)</td>
<td>(0.061)</td>
<td>(0.062)</td>
<td>(0.057)</td>
</tr>
<tr>
<td>Homes Sold</td>
<td>0.100**</td>
<td>0.200**</td>
<td>0.160**</td>
<td>0.154**</td>
<td>0.400**</td>
<td>0.294**</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.024)</td>
<td>(0.022)</td>
<td>(0.036)</td>
<td>(0.036)</td>
<td>(0.032)</td>
</tr>
<tr>
<td>Number of Agents</td>
<td>-0.069*</td>
<td>-0.247**</td>
<td>-0.162**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.032)</td>
<td>(0.0285)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>1.992**</td>
<td>2.065**</td>
<td>2.453**</td>
<td>1.850*</td>
<td>1.444</td>
<td>2.036**</td>
</tr>
<tr>
<td></td>
<td>(0.754)</td>
<td>(0.783)</td>
<td>(0.711)</td>
<td>(0.755)</td>
<td>(0.775)</td>
<td>(0.708)</td>
</tr>
<tr>
<td>Observations</td>
<td>59090</td>
<td>47208</td>
<td>50579</td>
<td>59090</td>
<td>47208</td>
<td>50579</td>
</tr>
<tr>
<td>R²</td>
<td>0.053</td>
<td>0.031</td>
<td>0.059</td>
<td>0.054</td>
<td>0.033</td>
<td>0.060</td>
</tr>
</tbody>
</table>

** p<0.01, * p<0.05
Figure 3.1: Histogram - Commission

Histogram of the commission offered to the buyer’s broker in the Wasatch Front Regional MLS.

Figure 3.2: Scatter Plot - Commission vs. Sales Price

Scatter plot of the commission offered to the buyer’s agent against the sales price of the home. Data presented are for a randomly selected sample of 1000 listings. Since commission is discrete (and, indeed, mostly equal to 3%), the data points have been jittered by adding some white noise to each variable.
Figure 3.3: Histogram - Listing Commission

Histogram of the list agents reported commission from the Wasatch Front Regional MLS. This is not a required field in the MLS and is observed for less than 2% of all observations.

Figure 3.4: Scatter Plot - Listing Commission vs. Sales Commission

Scatter plot of the commission offered to the buyer’s broker against the listing commission reported in the MLS. Again, the data are jittered by adding white noise to each variable so the circle centered at (3%, 3%) is the result of jittering the many observations there.
Figure 3.5: Model Equilibrium

A graphical representation of the equilibrium in the model of real estate brokerage with heterogeneous agents. Individual types lie along the x axis and the two lines represent the agents’ outside wages and earnings as brokers. Brokerage earnings cross the outside wage line once (from below) at point $a$. The intersection bisects the agents - all those above $a$ enter the brokerage industry.
Figure 3.6: Time Series - Price and Quantity
Figure 3.7: Time Series - Price and Number of Agents

Figure 3.8: Time Series - Number of Agents and Quantity
Figure 3.9: Time Series - Price and Wages
Figure 3.10: Time Series - Wages

Figure 3.11: Time Series - Productivity
Figure 3.12: Time Series - Price and Productivity

Figure 3.13: Time Series - Number of Agents and Productivity
Figure 3.14: Time Series - Quantity and Productivity